

32ND EUROPEAN SYMPOSIUM ON COMPUTER AIDED PROCESS ENGINEERING

ESCAPE-32

Edited by
LUDOVIC MONTASTRUC
STEPHANE NEGNY



COMPUTER-AIDED CHEMICAL ENGINEERING, 51

32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING

VOLUME 1

COMPUTER-AIDED CHEMICAL ENGINEERING, 51

**32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING**

VOLUME 1

Edited by

Ludovic Montastruc

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

ludovic.montastruc@ensiacet.fr

Stephane Negny

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

stephane.negny@ensiacet.fr



ELSEVIER

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

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

Copyright © 2022 Elsevier B.V. All rights reserved.

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

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

Notices

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

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

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

British Library Cataloguing in Publication Data

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

Library of Congress Cataloging-in-Publication Data

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

ISBN (Volume 1): 978-0-443-18631-8

ISBN (Set) : 978-0-323-95879-0

ISSN: 1570-7946

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



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Candice Janco
Acquisition Editor: Anita Koch
Editorial Project Manager: Lena Sparks
Production Project Manager: Paul Prasad Chandramohan
Designer: Mark Rogers

Typeset by STRAIVE

Contents

Preface	xv
Theme Coordinators	xvii
T1: Modelling and Simulation	
1. Periodic Oscillations in Methane Reactor: Effects of the Main Operating Parameters <i>Piero Bareschino, Alberto E. Cutillo, Claudio Tregambi, Francesco Pepe, Gaetano Continillo, Erasmo Mancusi</i>	1
2. Modeling and Simulation of Non-Isothermal Ceramic Drying <i>Achilleas L. Arvanitidis, Margaritis Kostoglou, Michael C. Georgiadis</i>	7
3. Numerical study on the solid phase residence time distribution in a counter-current screw extractor <i>Annemarie Lehr, Gábor Janiga, Andreas Seidel-Morgenstern, Dominique Thévenin</i>	13
4. Dynamic modelling of light and temperature effects on biomass growth and biohydrogen production by the photosynthetic bacterium <i>Rhodospseudomonas palustris</i> <i>Bovinille Anye Cho, Brandon Sean Ross, Jan-Pierre du Toit, Robert William McClelland Pott, Ehecatl Antonio Del Rio-Chanona, Dongda Zhang</i>	19
5. Reducing the experimental effort to design pharmaceutical tablet lubrication by model-based design of experiments <i>Francesca Cenci, Gabriele Bano, Charalampos Christodoulou, Yuliya Vueva, Simeone Zomer, Massimiliano Barolo, Fabrizio Bezzo, Pierantonio Facco</i>	25
6. A model-based approach to predict the flowability of directly compressed pharmaceutical blends from individual components <i>Gabriele Bano, Magdalini Aroniada, Yuliya Vueva</i>	31
7. Purification of MIBA by Continuous Distillation <i>Pavan Veldandi, Ralph Cos, David Earp</i>	37
8. Modelling of Organophilic Pervaporation for Separation of Acetone-Butanol-Ethanol Mixture <i>Andras Jozsef Toth</i>	43

- 9. Development of Deep Learning Architectures for Forecasting Distillation Columns Dynamic Behavior of Biobutanol Purification**
Abraham Rodarte de la Fuente, Eduardo Sanchez-Ramirez, Martha Patricia Calderon-Alvarado, Juan Gabriel Segovia-Hernandez, Esteban A. Hernández-Vargas **49**
- 10. Impact of Methanol Synthesis Kinetics on Bulk Production Prediction: an In-Silico Assessment**
Filippo Bisotti, Matteo Fedeli, Carlo Pirola, Giulia Bozzano, Flavio Manenti **55**
- 11. Optimal layout of modular multi-floor process plants using MILP**
P Wrigley, P Wood, S O'Neill, R Hall, S Marr, D Robertson **61**
- 12. Paving the way to multi-case optimization of a steam Rankine cycle for cogeneration in nuclear power plants**
Guilherme Vescovi, Nicolas Alpy, David Haubensack, Catherine Azzaro-Pantel, Pascal Stouffs **67**
- 13. Dynamic modelling of non-isothermal open-cell foam catalyst packings: selective sugar hydrogenation to sugar alcohols as a case study**
Catarina G. Braz, Ali Najarnezhadmashhadi, Vincenzo Russo, Kari Eränen, Henrique A. Matos, Tapio Salmi **73**
- 14. Optimising a wind farm with energy storage considering curtailment and uncertainties**
Flora A. V. Biggins, Jude O. Ejeh, Diarmid Roberts, Aaron S. Yeardley and Solomon F. Brown **79**
- 15. Optimisation of Biofuel and Kerosene Fuel Blends to Support Sustainable Aviation**
Ridab Khalifa, Mohammad Alherbawi, Adel Elomri, Tareq Al-Ansari **85**
- 16. Dynamic Surrogate Modeling for Continuous Processes Control Applications**
Alessandro Di Pretoro, Andrea Tomaselli, Flavio Manenti, Ludovic Montastruc **91**
- 17. Probabilistic machine learning based soft-sensors for product quality prediction in batch processes**
Max Mowbray, Aaron Hicks, Harry Kay, Sam Kay, Amanda Lane, Cesar Medonza, Philip Martin and Dongda Zhang **97**
- 18. Property Estimation Method for Cannabinoids and Terpenes Using Machine Learning**
Laura A. Vergara, Hector J. Hortúa, Gustavo A. Orozco **103**

- 19. Kinetic modelling of γ -linolenic acid production by *Cunninghamella echinulata***
Ziqi Song, Alexander Rogers, Bovinille Anye Cho, Keju Jing, Dongda Zhang **109**
- 20. Techno-economic and environmental analysis of pyrolysis process simulation for plastic (PET) waste**
Muhammad Shahbaz, Ahmed AlNouss, Gordon Mckay, Hamish Mackey, Tareq-Al Ansari **115**
- 21. Monte Carlo Simulation of the Mechanical Processing of Bulk Materials with Fluctuating Compositions - Compositional Probability Density**
Karim Khodier, Tobias Krenn, Lisa Kandlbauer, Lisa Tatschl, Renato Sarc **121**
- 22. Determination of the burst pressure of pillow plates using finite element methods**
Alexander Zibart, Bernhard Spang, E.Y. Kenig **127**
- 23. A hybrid multi effect distillation and double reverse osmosis system for most economical brackish water desalination**
O.M.A. Al-hotmani, Mudhar A. Al-Obaidi, Y.M. John, Raj Patel, and Iqbal M. Mujtaba **133**
- 24. Computational intelligence applied to the mathematical modeling of enzymatic syntheses of biosurfactants**
Alice de C. L. Torres, Rafael A. Akisue, Lionete N. de Lima, Paulo W. Tardioli, Ruy de Sousa Júnior **139**
- 25. A Practical Guide to Coffee Roaster Modelling**
Cameron E. Bolt and Philip L. de Vaal **145**
- 26. Modeling and simulation of anoxic-aerobic algal-bacterial photobioreactor for nutrients removal**
Irina Bausa, Raúl Muñoz, Smaranda Podar, César de Prada **151**
- 27. Gaussian-Process based inference of electrolyte decomposition reaction networks in Li-ion battery failure**
Dr Peter J. Bugryniec, Aaron Yeardley, Aarjav Jain, Nicholas Price, Dr Sergio Vernuccio and Dr Solomon F. Brown **157**
- 28. Energy evaluation of processes for the production of hydrogen from biomass biodigestion under Aspen Plus**
Lokmane Abdelouahed, Fatma Kourdourli, Bechara Taouk, Lionel Estel **163**

- 29. Life cycle optimization of energy systems integrated with carbon capture and utilization**
Iasonas Ioannou, Alex Mercandetti, Gonzalo Guillén Gosálbez 169
- 30. A comparative study of swarm intelligence and artificial neural networks applications in modeling complex reaction processes**
Min Wu, Ulderico Di Caprio, Furkan Elmaz, Bert Metten, Dries De Clercq, Olivier Van Der Ha, Siegfried Mercelis, Peter Hellinckx, Leen Braeken and M. Enis Leblebici 175
- 31. Combined optimization of start-up shutdown and grade transition of a multistage continuous crystallization process**
Jiaxu Liu, Brahim Benyahia 181
- 32. Systematic dynamic modelling of heat exchanger network**
Bertrand Zitte, Isabelle Pitault, Boussad Hamroun and Françoise Couenne 187
- 33. Technical and economic assessment of a castor bean biorefinery to produce renewable aviation fuel: a computer-aided design**
Araceli Guadalupe Romero-Izquierdo, Claudia Gutiérrez-Antonio, Fernando Israel Gómez-Castro, Salvador Hernández 193
- 34. SiCN fibers as advanced materials for electromagnetic shielding in X-band: experiments and computational modelling and simulation**
Heloisa Ramlow, Liangrid Lutiani da Silva, Bráulio Haruo Kondo Lopes, Maurício Ribeiro Baldan, Ricardo Antonio Francisco Machado 199
- 35. Continuous-Time Surrogate Models for Data-Driven Dynamic Optimization**
Burcu Beykal, Nikolaos A. Diangelakis, Efstratios N. Pistikopoulos 205
- 36. Parameter estimation in dynamic metabolic models applying a surrogate approximation**
Rafael D. de Oliveira, Dielle P. Procópio, Thiago O. Basso and Galo A.C. Le Roux 211
- 37. A Benchmark Model to Generate Batch Process Data for Machine Learning Testing and Comparison**
Margarida L.C. Vicente, José F.O. Granjo, Ruomu Tan and Franz D. Böhner 217
- 38. Global warming impact of electric city buses in Chile: Critical stages of their fabrication and use**
Guillermo Valenzuela-Venegas, Daniel Peña-Torres, Franco Lizama-Valenzuela and Melanie Colet-Lagrille 223

39. Sustainable Ammonia Production via Electrolysis and Haber-Bosch Process <i>Gbemisola Ojo, Kyle Camarda</i>	229
40. Optimization for sustainable hydrogen production path <i>Patience B. Shamaki, G.A.C Le Roux</i>	235
41. Development of lactic acid evaporation process model with multi effect evaporator and mechanical vapor recompression system <i>Sujin Cha, Jonghun Lim, Yurim Kim, Hyungtae Cho, Il Moon, Junghwan Kim</i>	241
42. Optimization-Based Framework for Robust Modeling and Design of Kinetic Systems <i>Eduardo Sánchez-Ramírez, Brenda Huerta-Rosas, Juan José Quiroz-Ramírez, Víctor Alejandro Suárez-Toriello, Juan Gabriel Segovia-Hernández</i>	247
43. Optimal Control System For Products Quality from a Deethanizer Column <i>Cristian Patrascioiu, Nawwar A. Rahman, Marian Popescu</i>	253
44. Importance of detailed experimentation in the model based design and scale up of pharmaceutical spray dryers for heat sensitive products <i>P. Martin-Salvador, T. De Beer, A. Kumar</i>	259
45. Process modelling of Direct Air Capture (DAC) of CO₂ using solid amine sorbents <i>So-mang Kim, Grégoire Léonard</i>	265
46. Mechanistic modelling for thrips incidence in organic banana <i>Jean C. Campos, José Manrique-Silupú, William Ipanaqué, Bogdan Dorneanu and Harvey Arellano-García</i>	271
47. A novel approach to modelling trickle bed reactors <i>Bogdan Dorneanu, Norbert Heinzelmann, Klaus Schnitzlein, Harvey Arellano-García</i>	277
48. Alkaline Water Electrolysis Model to Purify GMP grade NaOH Solutions for Biopharmaceutical Manufacturing Processes <i>Yeonghyun Kim, Youngjin Kim, Jae Hyun Cho, Il Moon</i>	283
49. Effect of air dynamics on the discharge of a pharmaceutical powder using the discrete element method <i>L. Naranjo, I. Nopens, T. De Beer, A. Kumar</i>	289

50. Biorefinery modelling is in tatters, and here is why <i>Robert Pujan and Heinz A. Preisig</i>	295
51. Simulation of a Fischer-Tropsch reactor for jet fuel production using Aspen Custom Modeler <i>Alejandro Morales, Gregoire Leonard</i>	301
52. Techno-economic-environmental analysis of a microbial oil production integrated into a bioethanol sugarcane biorefinery <i>Andreza Aparecida Longati, Felipe Fernando Furlan, Roberto de Campos Giordano, Everson Alves Miranda</i>	307
53. Modeling the hydrodynamic sizing and rating of reactive packing in Aspen Plus <i>Andressa Neves Marchesan, Ingrid Lopes Motta, Rubens Maciel Filho and Maria Regina Wolf Maciel</i>	313
54. Simplified Model-based Design of Plate-fin Microdevices with Uniform Flow Distribution at High Flow Rates <i>Osamu Tonomura, Kaori Maenaka, Shinji Hasebe</i>	319
55. CFD-based study of fluid flow and transport phenomena in fixed bed compact reactors <i>Osamu Tonomura, Akihiro Kitagawa, Kazuki Kato, Taisuke Maki, Ken-Ichiro Sotowa</i>	325
56. Multicomponent, nonisothermal VOC adsorption modelling for pharmaceutical effluent purification: effect of operating conditions on bed performance <i>Vasiliki E. Tzanakopoulou, Alexandra Costa, Daniel Castro-Rodriguez, Dimitrios Gerogiorgis</i>	331
57. Hydrogen Separation via Continuous Hydrate Formation <i>Marcelino Artur L. Fernandes, Mariana G. Domingos, Fernando G. Martins, Isabel S. Fernandes, Marcelo F. Costa, Ricardo J. Santos, José Carlos B. Lopes</i>	337
58. Combined particle model and experimental approach for predicting pyrolysis with palm kernel shells <i>Andres Chico-Proano, Michelle Romero, Ricardo A. Narváez C., Boris G. German, Daniel Rivadeneira, George Manos, Lazaros G. Papageorgiou and Eric S. Fraga</i>	343
59. Methanation of CO₂ byproduct from an ammonia plant with green hydrogen <i>Samuel Asante, Mark W. Hlawitschka, Robert Schlesinger</i>	349

60. Uncertainty analysis applied to distillation columns calculations <i>Leonardo O. S. Santana, Ewerton E. S. Calixto, Fernando L. P. Pessoa</i>	355
61. Gas diffusion channels in Li-O₂ batteries: a comparison of wet and flooded electrodes <i>Jean Felipe Leal Silva, Gustavo Doubek, Rubens Maciel Filho</i>	361
62. Techno-economic analysis of flexible AP-X LNG production process under risks and uncertainties <i>Noor Yusuf, Rajesh Govindan, Tareq Al-Ansari</i>	367
63. Modelling and Parameter Fitting of the Dosage of Hydrogen Peroxide in a Photo-Fenton Process <i>Kourosch Nasr Esfahani, Montserrat Pérez-Moya, Moisés Graells</i>	373
64. Optimising the performance of the condensate stabilisation unit in LNG Processes <i>Abdul Aziz Shaikh, Ahmed AlNouss, Tareq Al-Ansari</i>	379
65. Data-driven modelling of full batch distillation cycles based on recurrent neuronal networks <i>Gerardo Brand-Rihm, Erik Esche and Jens-Uwe Repke</i>	385
66. Stackelberg Game Design and Operation of a Non-Cooperative Bi-Level H₂ Supply Chain Under Cournot Equilibrium <i>Jose M. FLORES-PEREZ, Catherine AZZARO-PANTEL, Antonin PONSICH and Alberto A. AGUILAR LASSERRE</i>	391
67. Modelling of Heat-Driven Water Treatment Systems: Multi-Effect Distillation (MED) model in Modelica <i>Miguel Castro Oliveira, Pedro Coelho, Muriel Iten, Henrique A. Matos</i>	397
68. Continuous operation of a solar photobioreactor with linearizing control – A simulation study <i>Joris Sébile-Meilleroux, Mariana Titica, Jérémy Pruvost</i>	403
69. Dynamic Inherent Safety Analysis of a Distillation Column under Simultaneous Design and Control <i>Denis Su-Feher, Efstratios N. Pistikopoulos</i>	409
70. Modeling of Phosphates Slurry Pipelines Through Dynamic Non-Newtonian Fluid Model With Modelica <i>Fatima Ez-Zahra El Hamra, Radouan Boukharfane, Saad Benjelloun, Ahmed Ja and Jean-Michel Ghidaglia</i>	415

71. Techno-Economic Analysis of the Conversion of Waste Plastics to Hydrogen Fuel <i>Ali A. Al-Qadri, and Usama Ahmed</i>	421
72. Evaluating the flexible operation of vacuum-pressure swing adsorption for CO₂ capture from modern gas turbines <i>Mathew Dennis Wilkes, Solomon Brown</i>	427
73. Exergoeconomic assessment of the optimized vapour-recompression assisted column for palm-based fatty acid fractionation <i>Norul M. Sidek, Mohamad R. Othman,</i>	433
74. A Reduced Population Balance Model for Coupled Hydrodynamics and Mass Transfer in Shallow Bubble Column Reactors <i>Menwer Attarakih, Ferdaous Al-Slaihah, Armin Fricke, Hans-Jörg Bart</i>	439
75. Modeling of the crystallization of gypsum produced in the digestion tank of an industrial phosphoric acid manufacturing process <i>Ilias Bouchkira, Abderrazak M. Latifi, Lhachmi Khamar and Saad Benjelloun</i>	445
76. Development of a whole-body physiologically-based pharmacokinetic model for high-dose methotrexate <i>Giuseppe Pesenti, Dario Massari, Marco Foppoli, Davide Manca</i>	451
77. Analysis of an industrial adsorption process based on ammonia chemisorption: model validation <i>Cristian Cardenas, Abderrazak M. Latifi, Cécile Vallières and Stéphanie Marsteau</i>	457
78. CFD modeling and simulation of an ammonia adsorption process <i>Flora Esposito, Cristian Cardenas, Abderrazak M. Latifi and Stéphanie Marsteau</i>	463
79. Polygeneration from sugarcane industries enhanced by functionalizing novel cultivars and excess thermal energy <i>Shoma Fujii, Yuichiro Kanematsu, Yasunori Kikuchi</i>	469
80. Economic optimization of a reactive distillation column with multiple reactive sections for silane production <i>Alcántara-Maciel, Francisco D., Victor E. Casillas-Céspedes, J. Armando López-García, Julián Cabrera-Ruiz, César Ramírez Márquez, J. Rafael Alcántara-Ávila</i>	475

81. A multiscale model of proliferating and quiescent cell populations coupled with cell cycle dynamics <i>Iqra Batoool and Naim Bajcinca</i>	481
82. Multi-objective optimization of the food chain as a support for the development of agricultural sector <i>Jan Drofenik, Klavdija Zirngast, Bojan Pahor, Zdravko Kravanja, Zorka Novak Pintarič</i>	487
83. Optimal CO₂ allocation for enhanced oil recovery operations within carbon utilisation networks in Qatar <i>Razan Sawaly, Ikhlas Ghiat, Abdulkarim Mohamed, Ahmad Abushaikha, Tareq Al-Ansari</i>	493
84. Optimal operation of an evaporator for the controlled production of titania nanoparticles <i>Filippo Tamagnini and Sebastian Engell</i>	499
85. A model-based approach for the prediction of banana rust thrips incidence from atmospheric variables <i>Carlos A. Estrada, José Manrique-Silupú, William Ipanaque, Bogdan Dorneanu and Harvey Arellano-García</i>	505
86. Comparison between 3D numerical simulations and experimental results of a lab-scale liquid-solid fluidized bed <i>Almir G. S. L. Ritta, Renaud Ansart and Olivier Simonin</i>	511
87. Non-process elements in kraft bleach plants: adsorption equilibrium aiming at reducing water consumption <i>Ana M. Sousa, Carolina T. Pinheiro, José Granjo, Licínio M. Gando-Ferreira, Lino O. Santos, Margarida M. J. Quina</i>	517
88. Re-design and scheduling of dairy thermal treatment processes for continuous operation <i>Wei-Fu Tseng and Sandro Macchietto</i>	523
89. Xflow modelling for investigation of fluid structure interaction of artificial reef: application to burial effect <i>Baptiste Oudon, Chanez Belaidi, Jessica Salaun, Arnaud Coutu</i>	529
90. Metamodeling of chemical engineering unit operations using Kriging and prediction error estimation <i>Thibault Delage, Sanaa Zannane and Thibaut Neveux</i>	535

- 91. On the integration of process engineering with metabolomics for the production of muconic acid: the case for *Saccharomyces Cerevisiae***
Stefanos Xenios, Daniel Weilandt, Hatzimanikatis Vasilis, Ljubisa Miskovic, Antonis Kokosis **541**
- 92. Modelling and simulation of a residual lignocellulosic biomass pyrolysis pilot plant**
Nezly Martelo, Antonio Gagliano, Alberto Fichera, Rosaria Volpe, Mirari Antxustegi, Rodrigo Llano-Ponte, Maria Gonzalez Alriols **547**
- 93. Techno-economic-environmental analysis of biodiesel production by magnetic nanoparticles CLEAs of eversa® transform**
Alves, E.S., Miranda, L.P., Guimarães, J.R., Tardioli, P.W., Giordano R.C., Furlan F.F. **553**
- 94. Numerical Analysis of Impellers Hydrodynamics Performance in a Bioreactor CSTR with SPH**
R. Murrieta-Dueñas, C.E. Alvarado-Rodríguez, J. Cortez-González, R. Gutiérrez-Guerra **559**
- 95. On the use of embedded models and advanced analytics to model complex processes in the cement industry**
Alexandros Pyladarinos, Antonis Kokossis, Ioannis Marinos, Thanassis Gentimis **565**
- 96. Aspen Plus® modeling approach of beechwood gasification in a fluidized bed reactor using biochar as bed material**
Leonela Martes Hernández, Minda Loweski Feliz, Luis Reyes Alonzo, Lokmane Abdelouahed, Bechara Taouk **571**
- 97. Mathematical modeling of the diffusion-limited (DLA) aggregation accompanied by particles swarming in reactors**
Leila Musabekova, Sabira Akhmetova, Assel Sailau, Arnold Brener **577**
- 98. Verification of Neural Network Surrogates**
Joshua Haddad, Michael Bynum, Michael Eydenberg, Logan Blakely, Zachary Kilwein, Fani Boukouvala, Carl D. Laird and Jordan Jalving **583**
- 99. Sustainable Analysis of Recent Acid Gas Treatment Schemes for LNG production**
Ahmed AlNouss, Saad Al-Sobhi **589**

Preface

This volume of the Computer-Aided Chemical Engineering series puts together a selection of the contributions presented at the 32th European Symposium on Computer Aided Process Engineering (ESCAPE), held in Toulouse, France, from June 12th to 15th, 2022.

This 32th event of the ESCAPE series is a continuation of the conferences under the auspices of the CAPE Working Party of the European Federation of Chemical Engineering (EFCE), and the Société Française de Génie des Procédés (SFGP).

The ESCAPE series serves as a forum to bring together scientists, researchers, managers, engineers, and students from academia and industry, who are interested in CAPE and Process Systems Engineering (PSE). The scientific aim of the symposium is to present and review the latest developments in CAPE and/or PSE. The conference has been organized since 1992, starting with two meetings in 1992 in Denmark and France, and since then having one event annually. Hosting countries to the conference have been Austria (1993, 2018), Ireland (1994), Slovenia (1995, 2016), Greece (1996, 2011), Norway (1997), Belgium (1998), Hungary (1999, 2014), Italy (2000, 2010, 2020), Denmark (1992, 2001, 2015), The Netherlands (2002, 2019), Finland (2003, 2013), Portugal (2004), Spain (2005, 2017), Germany (2006), Romania (2007), France (1992, 2008), Poland (2009), United Kingdom (2012) and Turkey (2021).

The main focus for ESCAPE-32 is on the methodical approaches in process systems engineering with emphasis on uncertainty towards sustainability. The themes of ESCAPE-29 have been selected after a comprehensive discussion with the CAPE Working Party members and the scientific community. The particular topics within these overarching themes have been formulated to allow researchers from CAPE-related sciences to present their results and exchange valuable knowledge and experience. The themes include:

Modelling and Simulation

Coordinators : Iqbal Mujtaba, Jena-Pierre Belaud and Ludovic Montastruc

Product/Process Synthesis and Design

Coordinators: Grégoire Léonard and Laurent Cassayre

Large Scale Design and Planning/Scheduling

Coordinators: Antonio Espuna and Catherine Azzaro-Pantel

On Line Model Based Applications and Control

Coordinators: Miroslav Fikar and Nataliya Shcherbakova

Concepts, Methods and Tools

Coordinators: André Bardow and Pascal Floquet

Digitalization and Artificial Intelligence

Coordinators: Norbert Aspirion, Rachid Ouaret and Stéphane Negny

CAPE Applications Addressing Societal Challenges

Coordinators: Ana Barbosa-Povoa, Raphaële Thery-Hetrex and Marianne Boix

Education in CAPE and Knowledge Transfer

Coordinators: Eric Schaer and Vincent Gerbaud

ESCAPE-32 attracted 467 contributions from four continents (Europe, Americas, Africa, Asia). The papers have been reviewed and 281 selected for publication by the International Scientific Committee together with the help of Theme Coordinators. The selection process involved review of abstracts, review of manuscripts and final selection of the revised manuscript. We are deeply thankful for timely and careful reviews by these Scientists, as well as their invaluable help.

As editors of this special volume, we hope that the contributions in this edition of Computer Aided Process Engineering are excellent illustrations of the current state of the art in their respective field, that it will contribute to the progress in computer aided process and product engineering.

March 2022

Ludovic Montastruc Stéphane Negny

National Organizing Committee

- Prof. **Stéphane Negny**, Toulouse INP
- Prof. **Ludovic Montastruc**, Toulouse INP
- Prof. **Xavier Joulia**, Toulouse INP
- Prof. **Xuan Meyer**, Toulouse INP

Scientific Committee

- Prof. **Ludovic Montastruc**, Toulouse INP
- Prof. **Stéphane Negny**, Toulouse INP
- Prof. **Xavier Joulia**, Toulouse INP
- Prof. **David Boogle**, University College London
- Prof. **Flavio Manenti**, Politecnico Di Milano
- Prof. **Jean Marc Le Lann**, Toulouse INP
- Dr **Ing. Yann Le Moullec**, EDF
- Prof. **Abderrazak Latifi**, Université de Lorraine

International Scientific Committee

- A KARIMI I
- A PREISIG Heinz
- ABILDSKOV Jens
- AGHA Mujtaba Hassan
- AHMETOVIĆ Elvis
- ALMEIDA-RIVERA
Cristhian
- ANIBAL SANTOS DE
MATOS Henrique
- ANTONIO CARRILLO
LE ROUX Galo
- ARPENTINIER Philippe
- ARSENYEVA Olga
- ASPRION Norbert
- AZZARO PANTEL
Catherine
- BARBOSA-POVOA Ana
- BARDOW André
- BAUDOUIN Olivier
- BELAUD Jean Pierre
- BIEGLER Lorenz
- BOGLE David
- BOIT Baptiste
- BOIX Marianne
- BOUNACEUR Roda
- CABALLERO Jose

- CARVALHO Ana
- CASSAYRE Laurent
- CASTRO Pedro
- CAUDRON Didier
- CHEN Xi
- COMMENGE Jean-Marc
- COURTOIS Francis
- CREMASCHI Selen
- DIAZ Soledad
- DIRION Jean-Louis
- EDEN Mario Richard
- ENGELL Sebastian
- ESPUNA Antonio
- FAIRWEATHER
Michael
- FERRASSE Jean-Henry
- FIKAR Miroslav
- FLOQUET Pascal
- GABRIEL SEGOVIA-
HERNANDEZ Juan
- GEORGIADIS Michael
- GERBAUD Vincent
- GIULIA Bozzano
- GOUNARIS Chrysanthos
- GRAELLS Moises
- GROSSMANN Ignacio
- GUILLÉN GOSÁLBEZ
Gonzalo
- HARJUNKOSKI Iiro
- HENNING Gabriela
Patricia
- HOFMANN Rene
- IGOR PLAZL
- JALLUT Christian
- JOBSON Megan
- KAPUSTENKO Petro
- KIENLE Achim
- KLEMES Jiri Jaromir
- KOKOSSIS Antonis
- KONDILI Emilia
- KRASLAWSKI Andrzej
- LATIFI Abderrazak M.
- LÉONARD Grégoire
- LOGIST Filip
- MANCA Davide
- MANENTI Flavio
- MANSOURI Seyed
Soheil
- MARAVELIAS Christos
- MARECHAL Francois
- MARTELLI Emanuele
- MARTÍN Mariano
- MHAMDI Adel
- MIZSEY Peter
- MONTASTRUC Ludovic
- MOON Il
- MUJTABA Iqbal
- NEGNY Stéphane
- NEVEUX Thibaut
- OLIVEIRA Fabricio
- OLIVEIRA Nuno
- OUARET Rachid
- PAPADOKONSTANTA
KIS Stavros
- PITAULT Isabelle
- PONSICH Antonin
- PUIGJANER Luis
- RADL Stefan
- RAMOS Manuel
- REKLAITIS Gintaras V
- RENEAUME Jean-
Michel
- SEABRA DOS REIS
Marco
- SHAER Eric
- SHCHERBAKOVA
Nataliya
- SIMON Levente
- SIN Gurkan
- SRINIVASAN
Rajagopalan
- SUGIYAMA Hirokazu

- TELES DOS SANTOS
Moises
- THEODOROPOULOS
Constantinos
- THERY-HETREUX
Raphaële
- TRUSOVA Marina
- TULA Anjan
- TÜRKAY Metin
- VARELA Tânia Rute
Xavier De Matos
- VENKATASUBRAMAN
IAN Venkat
- YAMASHITA Yoshiyuki
- YOU Fengqi
- ZONDERVAN Edwin

Periodic Oscillations in Methane Reactor: Effects of the Main Operating Parameters

Piero Bareschino^{a,b}, Alberto E. Cutillo^a, Claudio Tregambi^{a,b}, Francesco Pepe^a, Gaetano Continillo^a, Erasmo Mancusi^a

^a*Dipartimento di Ingegneria, Università degli Studi del Sannio, Benevento 82100, Italy.*

^b*Istituto di Scienze e Tecnologie per l'Energia e la Mobilità Sostenibili - Consiglio Nazionale delle Ricerche, Italy*
erasmo.mancusi@unisannio.it

Abstract

Biogas is widely considered as one of the most promising renewable energy resources and the most environmental-friendly energy source. Biogas produced through anaerobic digestion contains many impurities and a high percentage of CO₂. Thus, upgrade and purification of the raw biogas by capturing CO₂ before its application are necessary. In this work a catalytic methanation process of biogas was proposed and numerically analyzed. It appears that sustained periodic oscillations occur in a wide range of operating parameters. A detailed nonlinear analysis is performed, and the information produced can be useful for effective plant design and adequate plant control and operation.

Keywords: Biomethane, Power-to-Methane, fixed bed reactor, periodic oscillations, non-linear dynamics.

1. Introduction

Due to the continuous increase in energy demand and to mitigate environmental problems related to greenhouse gas emissions, the research towards non-fossil and renewable energy sources is continuously increasing (Kapoor et al. 2019). In this respect, production of biomass-derived biofuels has emerged as one of the most promising non-conventional energy resources (Tursi 2019, Mancusi et al. 2021). Among biofuels, biogas from the anaerobic digestion of organic wastes stands out as an attractive way of reducing landfilling while producing energy. However, the CH₄ content in biogas usually reaches about 70% at most, so that it needs to be purified (removal of trace components) and upgraded (removal of CO₂) before utilization. The most used carbon dioxide separation technologies are based on absorption, adsorption, cryogenic distillation, and membrane separation, all of them being highly energy consuming (Zhang et al. 2020). In this work, the upgrading process by direct methanation of biogas is analyzed. Particularly, we envision using surplus electrical energy from renewable sources to produce, via electrolysis, the required hydrogen to be fed, along with the biogas, into the methanation reactor (Bareschino et al. 2020). Methanation enables the conversion of H₂ and CO₂ into methane ranking among the power-to-gas technologies that represent the best solution for the energy storage. Several studies show that methanation has a very high CO₂ conversion degree, close to the removal efficiencies of traditional upgrading techniques (e.g. Mhadmhan et al. 2022). In the present work, the proposed methanation process involves simultaneous biogas upgrading and methane enrichment of the leaving gas. The process is carried out in an adiabatic fixed bed reactor with a nickel-based catalyst, and a recycle loop is used for diluting the inlet reactants concentration, to limit the maximum adiabatic

temperature increase (Bareschino et al. 2021). The effects on the feed temperature and recycling ratio (R) are considered. Periodic oscillations are predicted over a wide range of the investigated parameters. The main cause of the sustained periodic oscillations can be found in the feedback from the mass recycle coupled with the typical phenomena of fixed bed reactors, i.e. the inverse response (Luyben, 2007). Although the maximum temperature reached during the oscillations is limited by the thermodynamic equilibrium, the system temperature oscillates remarkably, with a period of about 10 minutes and in a range of $\pm 150^\circ\text{C}$. This behavior must be avoided to prevent catalyst damage. Non-linear analysis is performed to characterize the stability range of periodic regimes and to identify domains of coexistence of multiple stable regimes. This information can be useful for effective plant design and adequate plant control and operation (Mancusi et al., 2007).

2. Mathematical Model

The biogas upgrading process is carried out in a single adiabatic fixed bed reactor with a nickel-based catalyst. Figure 1 reports a schematic layout of the methanation process under study.

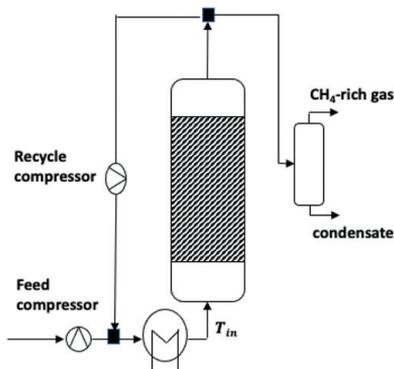


Figure 1 Adiabatic fixed bed methanation with recycle.

The biogas fed to the reactor is produced by anaerobic digestion of food waste. Purified from all impurities (H_2S), it contains CH_4 and CO_2 at 65% and 35% respectively (Tursi, 2019). This stream is enriched with H_2 produced by renewable sources according to a stoichiometric $\text{CO}_2:\text{H}_2$ ratio of 1:4 (R_3 in Tab. 1). The complete CO_2/CO methanation reaction mechanisms for syngas methanation over Ni-based catalyst proposed by Xu and Froment (Xu and Froment 1989) is considered. For readers' convenience, Table 1 summarizes the adopted reactions and associated enthalpy variations.

Table 1 – Reactions scheme and associated standard enthalpies of reactions.

Reaction	ΔH_{298} (kJ.kmol ⁻¹)	
$\text{CO} + 3\text{H}_2 \rightleftharpoons \text{CH}_4 + \text{H}_2\text{O}$	-206	R_1
$\text{CO}_2 + \text{H}_2 \rightleftharpoons \text{CO} + \text{H}_2\text{O}$	41	R_2
$\text{CO}_2 + 4\text{H}_2 \rightleftharpoons \text{CH}_4 + 2\text{H}_2\text{O}$	-165	R_3

To limit temperature increase due to the strong exothermicity of the methanation reactions, the water content in the feed is increased by adopting gas recirculation (Rönsch

et al. 2016). While the presence of water in the feed stream due to recycle shifts the chemical balance towards the products, it is essential to keep the temperature below 600° C and reduce carbon formation (Rönsch et al. 2016). Moreover, since the process is exothermic and reactions R_1 and R_3 take place with increasing volumes, the process benefits from high pressures and low temperatures.

The fixed bed reactor is modeled by a 1D pseudo-homogeneous and adiabatic model (Bareschino et al. 2021). The material balances for each gas component ($i=CH_4, CO, CO_2, H_2, H_2O$) and the energy balance are reported in Table 2 with the initial and boundary conditions:

Table 2 – Governing equations.

Mass Balance	$\varepsilon_g \frac{\partial c_i}{\partial t} = -u_{sg} \frac{\partial c_i}{\partial z} - (1 - \varepsilon_g) \rho_c r_i$
Heat Balance	$(\varepsilon_g \rho_g c_{pg} + (1 - \varepsilon_g) c_{pc}) \frac{\partial T}{\partial t} = -u_{sg} \rho_g c_{pg} \frac{\partial T}{\partial z} - \rho_c \sum_{j=1}^3 \Delta H_{Rj} R_j$
Initial conditions	$c_i(z, 0) = 0, T(z, 0) = T_0$
Boundary conditions	$c_i(0, t) = c_{i,in}, T(z, 0) = T_{in}$

where z is the axial position along each reactor belonging in $[0, L]$, and r_i the rate of consumption or formation of i -species ($i=CH_4, CO, CO_2, H_2, H_2O$) determined by summing up the reaction rates of those species in all the reactions R_j (see Table 1) according to the stoichiometric coefficients (ν) as follows:

$$r_i = \sum_{j=1}^3 \nu_{i,j} R_j \quad (1)$$

The gas superficial velocity (u_{sg}) is calculated as follows:

$$u_{sg}(z, t) = \frac{PM_{in}}{PM} u_{sg,in} \quad (2)$$

where PM is the molecular weight and the subscript in represents inlet conditions. Operating conditions, reactor volumes, and catalyst properties used in the simulations are reported in Table 3.

Table 3 – Parameters values used in the simulations.

Parameter	Value	Parameter	Value
P , bar	15.0	$y_{CO_2,in}$	0.145
T_{in} , °C	280	$y_{H_2,in}$	0.585
L , m	1.5	c_{pc} , J.kg ⁻¹ .K ⁻¹	1100
d_t , m	0.225	ρ_c kg.m ⁻³	2350
$y_{CH_4,in}$	0.27	ε_g	0.4

The method of lines is applied to solve the partial differential equations (PDEs) in two steps: the spatial derivatives are approximated by finite differences over a uniform grid of 200 discretization nodes, and the resulting system of 1200 ODEs is integrated in the

initial value variable (Bareschino et al. 2021). The size of the ODE set is large, but model order reduction techniques are available for future studies (Bizon and Continillo, 2021).

3. Results

The non-linear features of the fixed bed reactor give rise to phenomena such as the inverse response (a temporary behavior contrary to that expected). Figure 2 reports the reactor output according to the legend, for a step decrease (10°C) of the inlet temperature. It is apparent that the first effect observed at the reactor exit is a temperature increase. Since higher temperatures decrease the equilibrium conversion, the temperature increase is followed by an increase in the reactant concentrations. Thus, although the inlet heat exchanger cancels out temperature variations, recycle still provides a feedback.

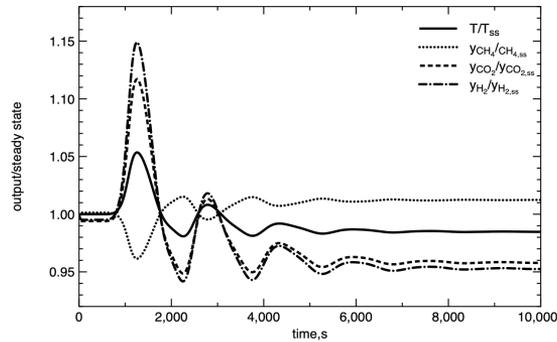


Figure 2 Reactor response to a step decrease of inlet temperature. Subscript ss refers to unperturbed steady-state. Recycle ratio $R=1.6$ whereas the other parameters are those reported in Tab. 3.

The combination of this feedback due to mass recycle and the inverse response can result in sustained oscillatory behavior (Luyben, 2007). Numerical simulations using non-linear dynamical model reproduce periodic oscillations in the reactor. Figure 2 reports a simulation showing time series of the outlet gas temperature and composition.

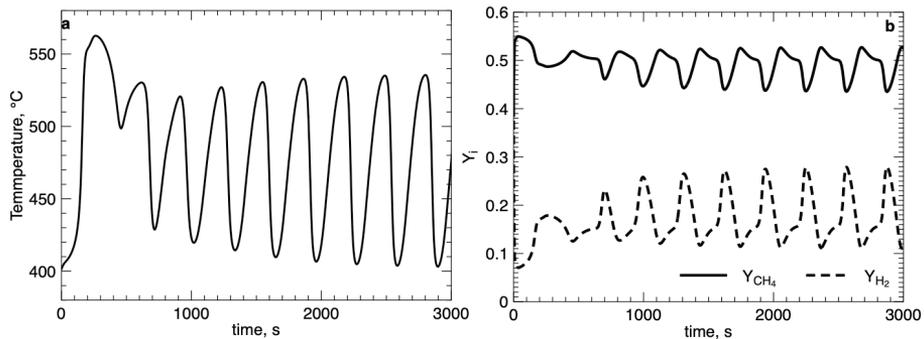


Figure 3 (a) Temperature at reactor exit vs time, (b) CH_4 and H_2 molar fraction at reactor exit vs time for $R=1.6$, whereas the other parameters values are those reported in Tab. 3.

Although the maximum temperature is limited by thermodynamic equilibrium, the system, as it can be seen, oscillates with a period of about 10 minutes and in the range of

$\pm 150^\circ\text{C}$. The amplitude of temperature oscillations is wide, and this could cause thermal stress of the catalyst. We now use a linear analysis to study more carefully the cause of the instability induced by the recycle operation. Close to an operating point, the dynamics of a system are well described by its linearized model. The model of the reactor was linearized numerically at this operating point, yielding a standard linear state space model with 1200 state variables in the form:

$$\frac{dx}{dt} = Jx \quad (3)$$

where J is the Jacobian matrix calculated at the steady state solutions. Computing steady state solution implies solving the nonlinear system of equations as reactor parameters vary (Mancusi et al., 2015). Once the steady state solution was calculated, the eigenvalues are computed by employing the *eig* function in MATLAB.

The largest eigenvalues of J for several values of R were calculated and depicted in Fig. 4. The reactor with no recycle ($R=0$) has a stable stationary response, i.e., all the eigenvalues have a negative real part. As R increases, the real part of the eigenvalues increases until $R=1.5$, where a couple of eigenvalues cross the imaginary axis. In these conditions the stable static regime solution becomes unstable, and periodic oscillations occur due to a Hopf bifurcation (Kuznetsov 1998).

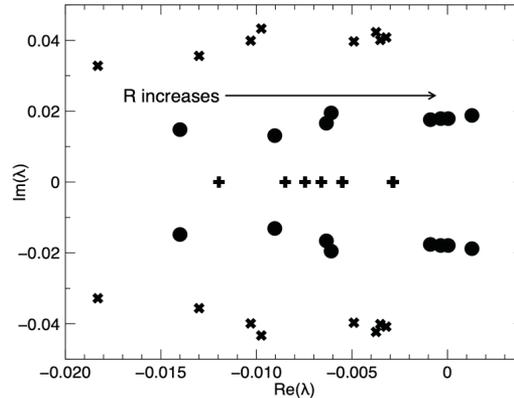


Figure 4 Imaginary and real part of the largest eigenvalues for several values of R .

The effect of the inlet temperature and recycle ratio is addressed by a bifurcation diagram (Fig. 5) where the locus of all Hopf bifurcation points is reported in the plane $R-T_{in}$. In this plot, the lines partition the parameter space into regions characterized by qualitatively similar phase portraits, that is the region characterized by stable steady state and region in which stable periodic oscillations exist (Kuznetsov 1998).

4. Conclusions

The paper reports a dynamical study of an adiabatic fixed bed reactor for the catalytic methanation of biogas. The complete dynamical characterization of a model is very useful to study the existence of periodic regimes and its influence on plant design, control, and operation. For the problem at hand, complex periodic regimes are due to the interaction of feedback induced by the mass recycle and by the inverse response typical of fixed bed reactors. The emergence of periodic oscillations is due to a Hopf bifurcation. The effect

of both the inlet temperature and the recycle ratio was investigated. The knowledge of the global dynamics can be helpful in designing an effective control strategy.

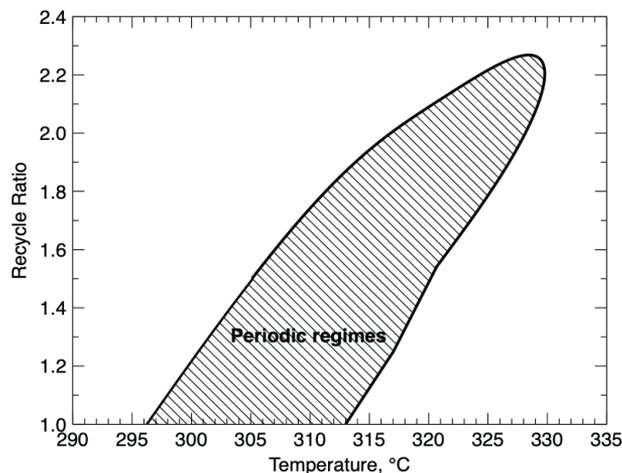


Figure 5 The bifurcation diagram in the R - T_{in} plane.

References

- P. Bareschino, E. Mancusi, M. Urciuolo, A. Paulillo, R. Chirone, F. Pepe, 2020, Life cycle assessment and feasibility analysis of a combined chemical looping combustion and power-to-methane system for CO₂ capture and utilization, *Renew. Sust. Energ. Rev.* 130, 109962
- P. Bareschino, E. Mancusi, C. Tregambi, F. Pepe, M. Urciuolo, P. Brachi, G. Ruoppolo, 2021, Integration of biomasses gasification and renewable-energies-driven water electrolysis for methane production, *Energy*, 230, 120863.
- K. Bizon, G. Continillo, 2021, Efficient optimization of a multifunctional catalytic fixed-bed reactor via reduced-order modeling approach, *Chem. Eng. Res. Des.* 165, 214-229.
- S. Mhadmhan, C. Ngamcharussrivichai, N. Hinchiranan, P. Kuchonthara, Y. Li, S.g Wang, Pr. Reubroycharoen, 2022 *Fuel*, 310, 122289.
- E. Mancusi, P. Bareschino, P. Brachi, A. Coppola, G. Ruppolo, M. Urciuolo, F. Pepe, 2021, Feasibility of an integrated biomass-based CLC combustion and a renewable-energy-based methanol production systems, *Renew. Energy*, 179, 29-36.
- E. Mancusi, L. Acampora, F.S. Marra, P. Altimari, 2015, Hysteresis in autothermal methane reforming over Rh catalysts: Bifurcation analysis. *Chem. Eng. J.*, 262, 1052-1064.
- E. Mancusi, L. Russo, A. Brasiello, S. Crescitelli, M. di Bernardo, 2007, Hybrid modeling and dynamics of a controlled reverse flow reactor, *AIChE J.*, 53(8), 2084-2096.
- W.L. Luyben, 2007, *Chemical Reactor Design and Control*, Nnew Jersey, John Wiley & Sons.
- R. Kapoor, P. Ghosh, M. Kumar, V.K. Vijay, 2019, Evaluation of biogas upgrading technologies and future perspectives: a review. *Environ Sci. Pollut. Res.*, 26, 11631-11661.
- Y.A. Kuznetsov, 1998. *Elements of applied bifurcation theory*. New York: Springer.
- S. Rönsch, J. Schneider, S. Matthischke, M. Schlüter, M. Götz, J. Lefebre, P. Prabhakaran, S. Bajoh, 2015. Review on methanation – From fundamentals to current projects. *Fuel* 9, 82–102.
- A. Tursi, 2019, A review on biomass: importance, chemistry, classification, and conversion *Biofuel Res. J.*, 22, 962-979.
- J. Xu, G. Froment, 1989. Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics. *AIChE J* 35, 85–96.
- X. Zhang, J. Witte, T. Schildhauer, C. Bauer, 2020, Life cycle assessment of power-to-gas with biogas as the carbon source. *Energy Fuels* 4, 1427–1436.

Modeling and Simulation of Non-Isothermal Ceramic Drying

Achilleas L. Arvanitidis,^a Margaritis Kostoglou,^b Michael C. Georgiadis^{a,*}

^a*Department of Chemical Engineering, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*

^b*Department of Chemistry, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*
mgeorg@auth.gr

Abstract

The main objective of this work is to model, simulate and assess the drying behavior of a single ceramic tile which is exposed to the convective flux of air of known humidity, velocity and temperature. The proposed mathematical model is inspired by a moving boundary model which was originally presented by Adrover et al. (2019). More specifically, 1-D mass and heat transfer is assumed along the thickness of the tile. The mechanism of mass transfer inside the material is a combination of diffusion and capillary motion, whereas at the surface moisture is removed through forced convection by the blowing air. Based on this assumption, shrinkage takes place only on one dimension. The model is used to accurately capture the drying behavior of a parallelepiped roof tile for two distinct case studies. Namely, the ideal-shrinkage case and the no-shrinkage case. The proposed modeling approach leads to high quality results with low computational costs. Furthermore, the developed modeling framework can provide the basis for modeling drying for a wide variety of operating conditions and various material properties.

Keywords: Process Modeling, Ceramic Drying, Moving-Boundary Problem

1. Introduction

Drying constitutes an essential step in the ceramic industry consuming large amounts of energy. During this process most of the water that was added in a previous molding step is removed. This step is characterized by significant complexity since heat and mass transfer are coupled processes and take place simultaneously. According to Geankoplis (1993), at first, the wet green body is heated by the blowing air and thus the evaporation initiates. Afterwards, the body's temperature reaches an equilibrium value which implies that the drying rate is constant. At some point, the moisture content at the surface becomes insufficient to maintain the constant rate drying and the drying rate starts decreasing. Accounting to the nature of the process, drying may also be accompanied by dimensional variations which occur due to the moisture removal. Specifically, the total volume variation of the material, in an ideal shrinkage case, should be equal to the removed water's volume. In reality though, this phenomenon is often antagonized by an increase in the body's porosity which is caused by the replacement of the evaporated water by an equal volume of air. Thus, products with differing shrinkage than the ones expected are usually rejected as they often do not meet the required quality standards or there is a high probability of breakage in the subsequent firing step. Furthermore, it is possible that breakage phenomena occur during the drying process itself, especially when the drying

rates are not properly controlled. Consequently, drying is a product quality defining process which makes the knowledge of the evolution of moisture content and linear shrinkage imperative.

Due to the complex nature of the process, drying has been studied extensively through the years. Scherer (1990) presented a comprehensive review in which he explains the mechanisms as well as the various phenomena that take place during drying of porous material. Jarque et al. (2016) combined experimental results and modeling techniques and showed that drying of ceramic roof tiles is a non-isothermal process and that the assumption of an instantaneous increase of the body's surface temperature to the drying air's temperature does not stand well. More recently, a numerical model was developed to describe the drying kinetics of ceramic green bodies by using moisture dependent thermophysical properties (Lauro et al., 2021). In this work a non-isothermal model is proposed to predict the water distribution and shrinkage of a ceramic tile over time.

2. Mathematical description of the model

Air of uniform velocity is assumed to flow over a rectangular ceramic roof tile. The temperature, T_g , and relative humidity, RH , of the blowing air are fixed. The mass and heat transfer inside the tile are assumed to be one dimensional and are described by the corresponding diffusion equations, which in this case, also consider the local shrinkage of the material via a local shrinkage velocity variable, v_s (Adrover et al., 2019). The mass conservation equation is written in terms of the water mass concentration, C_w , and the energy equation is given in terms of the tile temperature T .

$$\frac{\partial C_w}{\partial t} = \frac{\partial}{\partial z} \left[D_{eff} \frac{\partial C_w}{\partial z} - v_s C_w \right] \quad (1)$$

$$\rho C_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left[k \frac{\partial T}{\partial z} - v_s \rho C_p T \right] \quad (2)$$

Where D_{eff} is the effective diffusion coefficient of the moisture in the porous body and k is the thermal conductivity of the wet tile. Both properties have constant values for the present study.

The local shrinkage velocity, v_s , accounts for the rate at which a specific point of the green body moves and is surmised to be proportional to the diffusive flux of water inside the body.

$$v_s = a \frac{D_{eff}}{\rho_{H_2O}} \frac{\partial C_w}{\partial z} \quad (3)$$

Where a is the shrinkage factor, which is a parameter that lies between 0 to 1. These two marginal values correspond to the no-shrinkage and the ideal-shrinkage scenario respectively. However, since shrinkage is considered, the sample thickness, L , is not fixed and should be measured at any time according to equation (4).

$$\frac{dL}{dt} = v_s |_{z=L(t)} \quad (4)$$

In addition, bearing in mind that volume variations occur during drying and that the shrinkage velocity is a pointwise variable, the local concentration of the solid phase, C_s should be described by:

$$\frac{\partial C_{solid}}{\partial t} = - \frac{\partial [v_s C_{solid}]}{\partial z} \quad (5)$$

Given that, the local porosity of the green body, ε , can be derived from the equation (6).

$$1 = \varepsilon + \frac{C_{solid}}{\rho_{intrinsic}} + \frac{C_w}{\rho_{H_2O}} \quad (6)$$

Where $\rho_{intrinsic}$ is the particle density of the solid and does not include the volume of the pores.

The density and the thermal conductivity of the wet porous body are moisture dependent properties and are calculated through equations (7) and (8) respectively.

$$\rho = C_w + C_s + \varepsilon \rho_{air} \quad (7)$$

$$C_p = \frac{C_{p,H_2O} C_w + C_{p,s} C_s + C_{p,air} \varepsilon \rho_{air}}{\rho} \quad (8)$$

Furthermore, it is assumed that the water evaporation takes place only on the surface of the tile. Hence, the boundary conditions for the mass conservation equation (1) is:

$$-D_{eff} \left. \frac{\partial C_w}{\partial z} \right|_{z=L(t)} = h_m \rho_{da} (H_g^{eq} - H_g) \quad (9)$$

Where h_m is the mass transfer coefficient, ρ_{da} is the dry air density and H_g is the humidity ratio of air which is given by the equation (10).

$$RH = \frac{H_g}{H_g^{sat}} \quad (10)$$

The term H_g^{sat} is the saturation humidity ratio and it is calculated by equation (11).

$$H_g^{sat} = 0.622 \frac{P_{H_2O}^{sat}}{P - P_{H_2O}^{sat}} \quad (11)$$

Where P is the pressure of the air and $P_{H_2O}^{sat}$ is the saturation pressure of water.

The boundary condition of the energy balance equation (2) is given by:

$$-k \left. \frac{\partial T}{\partial z} \right|_{z=L(t)} = h_T (T|_{z=L(t)} - T_g) - \lambda_v D_{eff} \left. \frac{\partial C_w}{\partial z} \right|_{z=L(t)} \quad (12)$$

Where h_T is the heat transfer coefficient and λ_v is the latent heat of vaporization of water. Finally, the specific moisture, W , at the tile - air interface correlates with the equilibrium relative humidity through the Henderson desorption isotherm (Murugesan et al., 2001):

$$RH = 1 - \exp(-17W^{0.6}) \quad (13)$$

The water concentration can be converted to specific moisture through the following transformation:

$$W = C_w / C_{solid} \quad (14)$$

3. Results and Discussion

The above system consists of ordinary differential equations coupled with partial differential equations with moving boundaries. The numerical solution of the model in this form is not trivial. This makes the need of a new, fixed in time spatial variable which exempts the partial differential equations of the moving boundaries and allows the model to predict the position of the boundary as part of the solution. The boundary immobilization technique requires the dimensionalization of the model by utilizing the new fixed-in-time spatial variable $y = z/L(t)$, which lies between 0 and 1.

The transformed model is implemented in the gPROMS™ modeling environment. The following data were used as inputs:

Table 1: Model Data

Model Inputs					
k	1.23 W/(mK)	λ_v	2500 kJ/kg	T_g	273.15 + 60 K
D_{eff}	$6 \cdot 10^{-8} m^2/s$	$\rho_{intrinsic}$	2826 kg/m ³	T_{tile}	273.15 + 25 K
RH	0.50	ρ_{H_2O}	1000 kg/m ³	$C_w _{t=0}$	414 kg _{H₂O} /m ³
h_m	0.01 m/s	ρ_{air}	1.063 kg/m ³	$C_{solid} _{t=0}$	2070 kg/m ³
h_T	60 W/(m ² K)	P	101325 Pa	$L _{t=0}$	0.0187 m

In this work two marginal shrinkage scenarios were examined and compared. In the ideal shrinkage case, where the a parameter is equal to unity, the tile's volume variation from its initial state corresponds to the total volume of the liquid water that has been removed until that time. In other words, no change in the porosity takes place. On the other hand, in the no-shrinkage case, where the a parameter is equal to 0, the moisture migrates towards the body's surface where it evaporates leaving empty pores inside the green body.

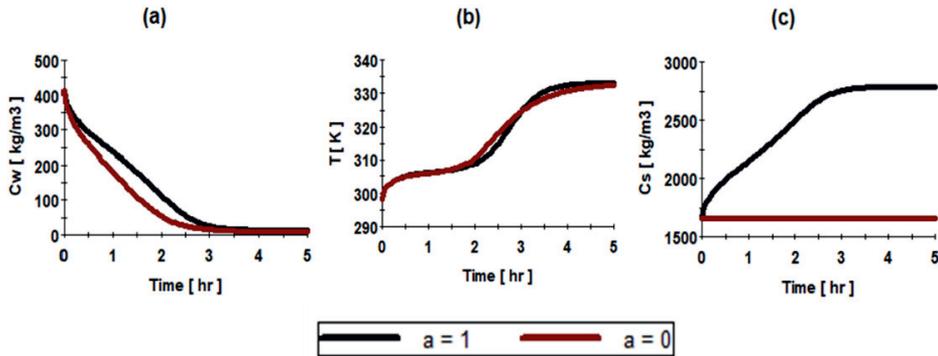


Figure 1: (a) Simulated surface moisture concentration vs time for the two cases (b) Simulated tile surface temperature vs time for the two cases (c) Surface solid concentration vs time for the two cases

Figures 1 and 3 illustrate the comparison between these two case scenarios and how the body shrinkage affects the drying kinetics of the body.

Figure 2 shows the tile shrinkage with respect to time, for the ideal-shrinkage case. It can be observed that the shrinkage stops when drying exits its constant rate period, as stated in the literature (Scherer, 1990; Lauro et al., 2021). As far as the no-shrinkage case is concerned, the tile thickness is fixed to its initial value since no volume variations occur.

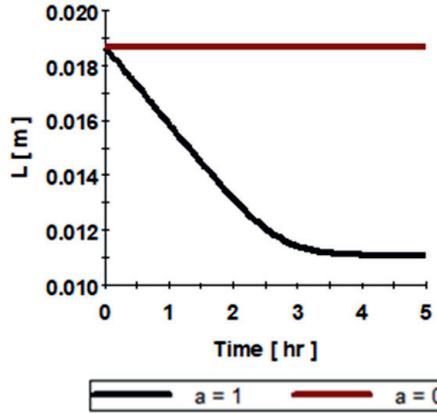


Figure 2: Tile thickness vs time for the ideal-shrinkage case and the no-shrinkage case

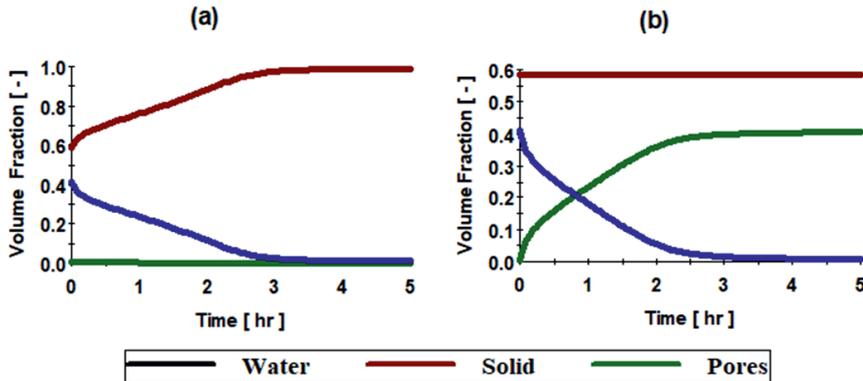


Figure 3: Surface volume fractions of the moisture, solid and the pores for (a) the ideal-shrinkage case and (b) for the no-shrinkage case

As it was stated earlier, when $a = 1$, the green body's volume is decreased by the volume of the evaporated water. In that case, an increase in the local solid concentration is expected along the tile. For the ideal-shrinkage scenario, since no pores are created as drying proceeds, the solid concentration approaches the solid's intrinsic density asymptotically. On the other hand, when $a = 0$, pores with an equal volume as the removed water's are created along the tile. Hence, the solid concentration is not affected. This phenomenon can be displayed in Figures 1c and 3.

Figure 3 presents the graphical illustration of the terms of the right-hand side of equation (6). For the ideal-shrinkage scenario (see Figure 3a), the local solid volume fraction increases with respect to time while the moisture volume fraction decreases. No pores are created. When it comes to the no-shrinkage scenario (see Figure 3b), the solid volume fraction remains fixed to its initial value through the process as no volume variations take place on the body. At any time, the volume of the removed moisture is replaced by an equal volume of air. Thus, the product's porosity increases as drying proceeds.

4. Conclusions

In this work, a 1-D modeling framework is proposed to simulate the drying behavior of a shrinking roof tile. Two marginal shrinkage scenarios were examined. These two scenarios were distinguished by the value of the shrinkage parameter a , which illustrates a key role in the overall analysis. The simulated results confirm the robustness of the model. It is important to note that the above framework can be used to capture the phenomena that accompany drying of a wet tile even for more realistic shrinkage scenarios, i.e., when $a \in (0,1)$, in which the tile's thickness decreases while its porosity increases.

Acknowledgments

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

References

- A. Adrover, A. Brasiello, G. Ponso, 2019, A moving boundary model for food isothermal drying and shrinkage: General setting, *Journal of Food Engineering*, 244, 178 -191
- C. J. Geankoplis, 1993, *Transport Processes and Unit Operations*, 3rd edn., Prentice Hall, New Jersey
- J. C. Jarque, C. Segarra, V. Cantavella, R. Mondragón, 2016, Non-Isothermal modeling of drying kinetics of ceramic tiles, *Drying Technology*, 34, 761 - 772
- N. Lauro, S. Oummadi, A. Alzina, B. Nait-Ali, D. S. Smith, 2021, Computer model of drying behaviour of ceramic green bodies with particular reference to moisture content dependent properties, *Journal of European Ceramic Society*, 41, 7321 – 7329
- K. Murugesan, H. N. Suresh, K. N. Seetharamu, P. A. Aswatha Narayana, T. Sundararajan, 2001, A theoretical model of brick drying as a conjugate problem, *International Journal of Heat and Mass Transfer*, 44, 4075 - 4086
- G. W. Scherer, 1990, Theory of Drying, *Journal of American Ceramic Society*, 73, 3 – 14

Numerical study on the solid phase residence time distribution in a counter-current screw extractor

Annemarie Lehr,^a Gábor Janiga,^a Andreas Seidel-Morgenstern,^b Dominique Thévenin^a

^a*Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*

^b*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*
annemarie.lehr@ovgu.de

Abstract

The application of a highly efficient continuous counter-current extraction to extract Artemisinin from *Artemisia Annu*a leaves is desirable since Artemisinin is efficiently used as anti-malaria drug. The residence time distribution (RTD) of the solid and liquid phases are most important for understanding this process since they influence the reaction efficiency. This work is devoted to the numerical investigation of the solid-phase RTD in a fully-filled screw extractor by using computational fluid dynamics (CFD). The solid phase is considered as a liquid Eulerian phase with a high viscosity. To track it, the commonly used species model has been implemented in a frozen quasi-steady-state simulation. Validation experiments have been performed by using dry leaves. A very good agreement between numerical and experimental residence times can be observed, with a relative error lower than 12 %. As a next step the model will be extended to predict the RTD in a multiphase flow model including the liquid solvent.

Keywords: computational fluid dynamics (CFD), extraction, residence time distribution

1. Introduction

The implementation of counter-current extraction processes provides advantages compared to the direct current method, since higher final concentrations of the target substance can be achieved in the solvent. Especially in the field of natural product extraction high yields of plant substances are required for drug preparation (Lack, 1985). In this work the solid-liquid counter-current extraction process of *Artemisia Annu*a leaves to gain Artemisinin is investigated. Derivatives of Artemisinin (e.g. artesunate) are increasingly used as efficient anti-malaria drugs (Gilmore et al., 2014). The extraction efficiency of this leaching process is influenced by the contact times between the liquid solvent and the solid plant material. Consequently, the knowledge of the residence time distribution (RTD) of both phases is of particular interest. The RTD is directly linked to the apparatus performance as it is influenced by various process parameters, such as throughput, the rotation speed of the screw, or the process temperature. Its characterization is highly important for many industrial processes such as the continuous production of chemicals, food and pharmaceutical products (Gao et al., 2012). Experimental measurements are the most common way to study the RTD in an apparatus. However, such measurements can be very expensive, especially when considering scale up or plant design. Here, computational fluid dynamics (CFD) may provide a deeper

understanding of the multiphase flow and of the controlling interactions. Chen et al. (2019) used CFD simulation to measure the particle residence time distribution in a fluidized bed by using the species method and the multi-solid method for tracer injection, which are the most common Eulerian methods. While the species method only solves a species transport equation for a numerical tracer, the multi-solid method describes two or more solids as Eulerian phases by considering interactions. Both methods can in principle deliver accurate predictions of the RTD, but the species method needs prior information concerning the diffusion coefficient. Nevertheless, the associated computational time is lower compared to the multi-solid method. This is the reason why several additional studies used the species method for getting the RTD behavior of multiphase flows, e.g. Adeosum et al. (2009), Deshmukh et al. (2009), and Zhang et al. (2015).

In this study the species method is used as well to describe the RTD distribution of the solid phase in a fully-filled screw extractor. The diffusion coefficient is derived from prior experiments. Finally, the numerical results are validated by comparison with experimental data.

2. Mathematical Model

2.1. Mean residence time and variance

The mean residence time and variance can be determined from the residence time distribution (RTD), which is a function of the length of the observed extruder. For measuring the RTD in a device, the stimulus response technique with pulse or stepwise input of tracer is typically used. The tracer concentration is measured by an appropriate device at the outlet of the extruder, leading to concentration curves named $E(t)$ and $F(t)$ curves, also known as normalized residence time distribution function and cumulative exit age distribution function, respectively. Both functions can be transformed into each other following Levenspiel (1999):

$$F(t) = \int_0^t E(t) dt \quad (1)$$

or conversely

$$E(t) = \frac{dF(t)}{dt} \quad (2)$$

The mean residence time τ and the variance σ^2 can be derived from $E(t)$.

$$\tau = \int_0^{\infty} t \cdot E(t) dt \quad (3)$$

$$\sigma^2 = \int_0^{\infty} (t - \tau)^2 \cdot E(t) dt \quad (4)$$

2.2. Dispersion coefficient

The dispersion model is frequently applied to describe back-mixing effects in tubular reactors. For this purpose, an additional axial dispersion coefficient D_{ax} is introduced into a second-order term in the mass balance:

$$\frac{\partial c_{\alpha}}{\partial t} = -u_{ax} \frac{\partial c_{\alpha}}{\partial z} + D_{ax} \frac{\partial^2 c_{\alpha}}{\partial z^2} \quad (5)$$

Implementing the species method in CFD requires the provision of the axial dispersion coefficient as input parameter. It can also be expressed via the dimensionless Bodenstein number Bo , which includes the axial velocity u_{ax} and the length of axial coordinate L .

$$Bo = \frac{2\tau^2}{\sigma^2} \quad (6)$$

$$D_{ax} = \frac{u_{ax}L}{Bo} \quad (7)$$

3. Experiments and Simulation

3.1. Experimental Setup

For validating the CFD simulations three experiments with equal operating conditions have been performed. The employed screw extractor with a total length of 320 mm is shown in Figure 1. The milled dry leaves are introduced as powder on the left-hand side until total filling of one segment is reached. When starting the screw rotation with 1.3 rpm the leaves are transported towards the right side of the device over a total flow length of 260 mm. The RTD during the transport process is controlled by a sampling of leave material at four different locations, marked 1 to 4 in Figure 1. Since only the dry leaves transport in the screw extruder has been considered no compression inside the device occurs. Consequently, the pressure equals atmospheric pressure.

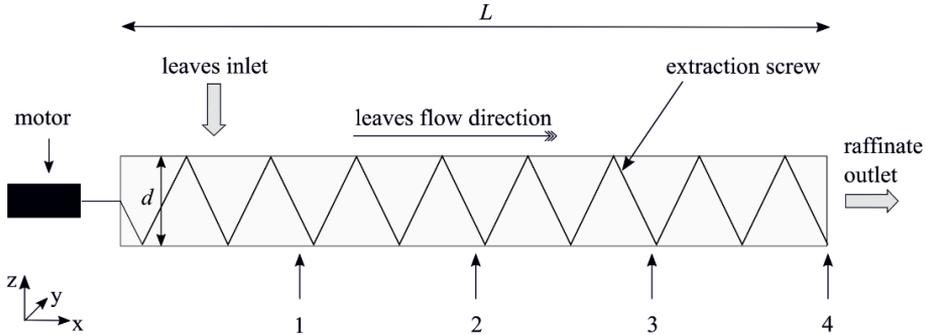


Figure 1: Counter-current screw extractor used for solid RTD experiments with equally distributed sampling points 1-4, $L = 320$ mm, $d = 29.7$ mm

3.2. Numerical Setup

The flow investigation was performed by using the commercial CFD software package StarCCM+ (Simcenter STAR-CCM+ 2019.2.1 Build 14.04.013). The computational geometry corresponds to the configuration used in the experimental setup (Figure 1). A block-structured mesh has been generated with a total number of 830.931 finite-volume cells. According to a previously performed mesh independence study this resolution provides the most accurate results by reducing numerical diffusion. Since only the screw rotates, the domain has been separated into a rotating and a stationary domain. Refinements at the interface between the domains were implemented to accurately capture the occurring leakage flows. As the numerical simulation of solid phases is very complex, the *Artemisia annua* leaves have been modelled as a Newtonian fluid with a very high viscosity of $\eta = 1,000$ Pa·s. Due to the slow rotation of the screw (1.3 rpm) and the high viscosity the resulting flow is laminar. A slip boundary for all walls has been implemented. The CFD simulation considers a completely filled system. The real experiments naturally involve a thin layer of gas phase (air) at the top, but it does not play any significant role for the extraction process. It is therefore neglected in CFD to reduce the computational effort. To adapt the inflow velocity of the leaves in CFD to the experiments an average axial velocity \bar{u}_{ax} has been calculated from experimental data.

$$\bar{u}_{ax} = \frac{L}{\tau_{exp}} \quad (8)$$

To measure the residence time distribution, the species method has been implemented by using the experimentally determined axial diffusion coefficient at the outlet (sampling point 4). First, the flow has been solved for 30 s of physical time until velocities and pressure converged to a specific value to ensure a quasi-steady state. Subsequently, the resulting flow has been frozen, and the species has been initialized and solved as single transport equation in the domain. Further modeling details can be taken from Table 1.

Table 1: Physical models and boundary conditions

Models	Parameter	Value	Unit
Laminar	Density	260	kg/m ³
Liquid	Viscosity	1000	Pa·s
Segregated Flow	Velocity Inlet	4.5×10^4	m/s
2nd order Implicit Unsteady	Time Step	0.5	s
Three Dimensional	Screw rotation	1.3	rpm
	Gravity z	-9.81	m/s ²

4. Results and Discussion

The residence time distributions obtained from the CFD simulation at four sampling points are shown in Figure 2a and compared with the residence time distributions of the first experiment. The RTD curves are similar, but the differences in maxima increase with time. Figure 2b shows the resulting mean residence times of experiments and CFD depending on the sampling points. The numerical results show slightly higher values than the experiments resulting in an error of 7.05 % for sampling point 2, up to 14.65 % for sampling point 4 (Table 2). These errors correspond to a difference in residence time between 0.31 and 1.32 min. Nevertheless, during the independent repetitions of the same experiment, some deviations of the RTD curves have also been observed. They occur due to minute changes of particle size distributions, particle-particle interactions, but also small variations in the screw rotation speed. As the experimental values are not perfectly repeatable a second comparison of numerical residence times is performed by involving three experimental realizations, considering the mean values and their deviations to each other (Figure 3). Here, the numerical mean residence times fit the experimental residence time for sampling points 2 and 3. For sampling point 1 the experimental residence time is slightly higher than the numerical one, by 12 % (Table 3). For sampling point 4 the numerical residence time is outside the experimental range by only 2.48 %. Consequently, the numerical predictions are very close to the experimentally measured values, proving that the CFD model can appropriately predict the residence time.

Table 2: Comparison of numerical results and experimentally measured residence times ($D_{ax} = 1.41 \times 10^{-6} \text{ m}^2/\text{s}$)

Sampling Point	τ_{CFD} (min)	$\tau_{exp,1}$ (min)	error (%) (vs. exp.)
1	1.54	1.74	11.49
2	4.71	4.40	7.05
3	7.88	6.99	13.31
4	10.33	9.01	14.65

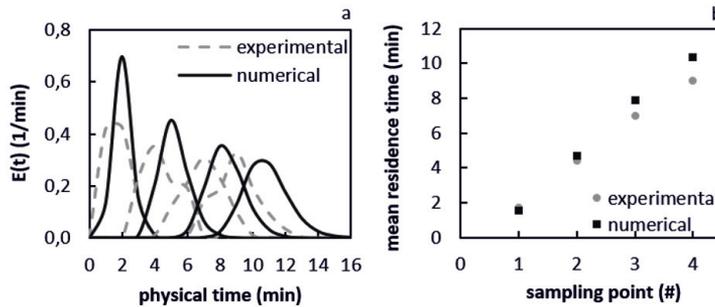


Figure 2: Results of the solid RTD study and comparison between experiments and numerical simulation regarding a) RTD curves at four sampling points, b) mean residence times

Table 3: Numerical mean residence time compared to the deviations observed when repeating the experiments ($D_{ax} = 1.41 \times 10^{-6} \text{ m}^2/\text{s}$)

Sampling Point	τ_{CFD} (min)	$\bar{\tau}_{\text{exp,all}}$ (min)	$\sigma_{\text{exp,all}}$ (min)	error (%) (vs. range)
1	1.54	1.89	± 0.14	12.00
2	4.71	4.94	± 0.47	-
3	7.88	7.77	± 0.68	-
4	10.33	9.58	± 0.50	2.48

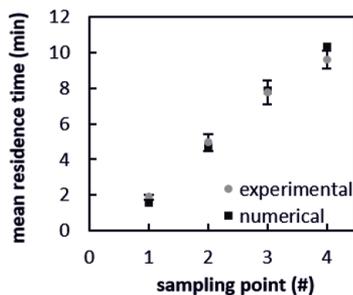


Figure 3: Comparison of the numerical mean residence time with three experimental runs

5. Conclusion and Outlook

In this study the solid residence time in a counter-current screw extractor has been analyzed by using CFD simulations. Since solid-phase RTD experiments show deviations in the resulting mean residence times, the numerical simulations have been first compared to one single experimental trial, then to the average value involving three measurement campaigns. The results show that the CFD mean residence time fits in the range of the experimental observations. The first sampling point shows the largest deviation, with an error of only 12 % (12.6 s). As this is considered an acceptable range, the numerical model with a fully-filled Eulerian phase can appropriately predict the solid RTD of *Artemisia Annua* leaves. Using the frozen quasi-steady-state flow allows short computational times for the implemented species method. This will enable future multiphase flow simulations with liquid and solid phase even closer to the experimental reality.

The numerical representation of the RTDs will then be used to parametrize compartment models (CM) in which the reaction kinetics can be considered with acceptable computational times. This mixed CFD/CM model will be used for the final optimization.

References

- E. Lack, 1985, Kriterien zur Auslegung von Anlagen für die Hochdruckextraktion von Naturstoffen, PhD Thesis, Graz, Austria
- J. Adeosun, A. Lawal, 2009, Numerical and experimental studies of mixing characteristics in a T-junction microchannel using residence time distribution, *Chemical Engineering Science*, 64, 10, 2422-2432
- K. Chen, P. Bachmann, A. Bück, M. Jacob, E. Tsotsas, 2019, CFD simulation of particle residence time distribution in industrial scale horizontal fluidized bed, *Powder Technology*, 345, 129-139
- K. Gilmore, D. Kopetzki, J. W. Lee, Z. Horvath, D. T. McQuade, A. Seidel-Morgenstern, P. H. Seeberger, 2014, Continuous synthesis of artemisinin-derived medicines, *Chemical Communications*, 50, 12652-12655
- O. Levenspiel, 1999, *Chemical Reaction Engineering*, 3rd. ed. John Wiley & Sons, 1999
- S. Deshmukh, M.J. Sathé, J.B. Joshi, S.B. Sudhir, 2009, Residence time distribution and flow patterns in the single-phase annular region of annular centrifugal extractors, *Industrial & Engineering Chemistry Research*, 48, 1, 37-46
- Y. Gao, F.J. Muzzio, M.G. Ierapetritou, 2012, A review of the Residence Time Distribution (RTD) applications in solid unit operations, *Powder Technology*, 228, 416-423
- Y. Zhang, Z. Wang, Y. Jin, Z. Li, W. Yi, 2015, CFD simulation and experiment of residence time distribution in short contact cyclone reactors, *Advanced Powder Technology*, 26, 4, 1134-1142

Dynamic modelling of light and temperature effects on biomass growth and biohydrogen production by the photosynthetic bacterium *Rhodospseudomonas palustris*

Bovinille Anye Cho,^a Brandon Sean Ross,^b Jan-Pierre du Toit,^b Robert William McClelland Pott,^b Ehecatl Antonio Del Rio-Chanona,^c Dongda Zhang^{a,c}

^a*The University of Manchester, Centre for Process Integration, Manchester, M1 3AL, United Kingdom*

^b*Stellenbosch University, Department of Process Engineering, Stellenbosch, 7599, South Africa*

^c*Imperial College London, Centre for Process System Engineering, London, SW7 2AZ, United Kingdom*

email of the corresponding author: dongda.zhang@manchester.ac.uk

Abstract

Background: Parallel to the impending hydrogen economy, the simultaneous treatment of organic waste and biohydrogen synthesis by the photosynthetic bacterium *Rhodospseudomonas palustris* is a promising renewable energy technology. However, studies so far have been mostly laboratory based with upscaling of the bioprocess still being an open challenge. Therefore, this study investigates two different photobioreactors (PBRs): schott bottle-based and vertical tubular-based PBRs and presents three original contributions to facilitate the biotechnology transfer across PBR scales and configurations.

Study Design: Firstly, a dynamic model is constructed to simulate the complicated influences of light intensity, light attenuation, and temperature, previously not unified for any photosynthetic bacteria to the best of our knowledge. Secondly, perturbation analysis was exploited to identify critical parameters influencing the model accuracy and reliability for across the scale extrapolations. Thirdly, two model parameters: *effective light coefficient* and *biohydrogen enhancement coefficient*, both linked to the PBR's transport phenomena were proposed for recalibrations during bioprocess upscaling predictions.

Major results: By comparing against experimental data, the upscaling prediction accuracy was thoroughly verified for the two investigated PBR scales. As well, the enhancement of biohydrogen production rate by improved culture mixing and gas removal was mechanistically described.

Conclusion: This provides important advances for the efficient design of novel PBRs and future online optimisation for biohydrogen production.

Keywords: Photobioreactor, Biohydrogen production, kinetic modelling, Purple non-sulfur bacteria, Upscaling.

1. Introduction

Renewable biohydrogen is a key biofuel identified as one promising alternative to the conventional fossil-based fuels for providing energy meant for: (i) electricity generation,

(ii) transportation, and (iii) heating [1, 2]. Amongst the several species of photosynthetic microorganisms which can synthesis biohydrogen, the purple-non-sulfur photosynthetic bacterium *Rhodospseudomonas palustris* (hereinafter referred to as *R. palustris*) has been identified as a promising candidate due to: (i) the absence of oxygen-induced nitrogenase repression; (ii) a versatile metabolic repertoire with capabilities of degrading a wide variety of organic substrates including those toxic to other microorganisms; as well as (iii) the continuous synthesis of biohydrogen during all the growth phases, including the stationary phase which is observed to last significantly longer under anaerobic conditions than in cyanobacteria and microalgae species [2]. Despite these benefits, literature investigations have been mostly carried out at the laboratory scale and to a lesser extent at the pilot scale but remains pending industrial scale applications. Although mathematical models can help facilitate the experimental design and process upscaling, two of the most important cultivating factors, namely temperature and light intensity (plus light attenuation herein - the decrease in light transmission due to cellular absorption and scattering within PBRs) have never been unified/coupled in a biokinetic model of any photosynthetic bacteria in general and *R. palustris* in particular. Therefore, this paper aims to: (i) investigate the photoheterotrophic biomass growth and biohydrogen production of *R. palustris* in two different Photobioreactors (PBRs) configuration and scales: schott bottle-based (0.5 L) and vertical tubular-based (1 L) PBRs, (ii) construct an accurate mechanistic model of the bioprocess under the unified influences of temperature, and light intensity, and light attenuation, and (iii) evaluate model's applicability for simulating process dynamics over different scales and configurations of PBRs.

2. Methodology

2.1. Mechanistic model construction

The photo-heterotrophic biomass growth, substrate consumption and biohydrogen production under the influences of light intensity, light attenuation, and temperature are simulated with Equations (1) to (3). The biomass growth (Equation 1) assumes negligible cell death and replete amount of substrate (i.e., > 20 mM at the end of each batch) with the rate of substrate consumption described in Equation (2). The biohydrogen production model (Equation 3) is an extension of the Luedeking-Piret model with the incorporation of temperature and light intensity influences on the linear dependence of biomass growth rate and instantaneous biomass concentration. The first square brackets on the Right-Hand-Side of Equations (1) and (3) simulates the monotonic increase of microbial activity by temperature which does not exceed the optimal operational value (i.e., 313.15 K) for *R. palustris* biomass growth [3]. Assuming the absence of photoinhibition for light intensities not higher than 200 W m^{-2} [4], photolimitation and photosaturation are the main photo-mechanisms captured from the Aiba model as represented by the second square brackets on the Right-Hand-Side of Equations (1) and (3).

2.2. Dynamic parameter estimation

Generally, the model parameter estimations were solved by weighted nonlinear least-square regression. However, due to the high nonlinearity, the entire process was decoupled into two steps with step I: estimating biomass growth and substrate consumption model parameters, and step II: estimating biohydrogen model parameters while fixing the optimal solution of step I. Due to stiffness and nonlinearity, orthogonal collocation over finite elements in time was used to discretise the differential equations, then solved with the interior point solver, IPOPT through the open-source interface Pyomo, within Python version 3.7 programming environment.

$$\frac{dX}{dt} = \mu_{max} \cdot \left[A \cdot \exp\left(-\frac{E_a}{RT}\right) \right] \cdot \left[\frac{\frac{I_0}{(\tau \cdot X) \cdot L} \cdot (1 - e^{-(\tau \cdot X) \cdot L})}{\frac{I_0}{(\tau \cdot X) \cdot L} \cdot (1 - e^{-(\tau \cdot X) \cdot L}) + k_s} \right] \cdot X \quad (1)$$

$$\frac{dS}{dt} = -Y_{XS} \cdot \frac{dX}{dt} - m \cdot X \quad (2)$$

$$\frac{dH_2}{dt} = \alpha_{max} \cdot \left[A_{H_2} \cdot \exp\left(-\frac{E_{a,H_2}}{RT}\right) \right] \cdot \left[\frac{\frac{I_0}{(\tau \cdot X) \cdot L} \cdot (1 - e^{-(\tau \cdot X) \cdot L})}{\frac{I_0}{(\tau \cdot X) \cdot L} \cdot (1 - e^{-(\tau \cdot X) \cdot L}) + k_{s,H_2}} \right] \cdot X \quad (3)$$

where X , S and H_2 , are the biomass concentration (g L^{-1}), substrate concentration (mM) and biohydrogen production (mL) respectively, μ_{max} and α_{max} are the maximum specific growth rate (h^{-1}) and maximum specific production rate (h^{-1}) respectively, $\mu_{max} \cdot A = A'$ (h^{-1}) and $\alpha_{max} \cdot A_{H_2} = A'_{H_2}$ (h^{-1}) are the pre-exponential factors which are lumped for identifiability purposes, E_a and E_{a,H_2} are the activation energies (J mol^{-1}), R the is universal gas constant ($8.3145 \text{ J mol}^{-1}\text{K}^{-1}$), T is the absolute temperature (K), I_0 is the incident light intensity (Wm^{-2}), τ ($\text{m}^2 \text{ g}^{-1}$) is the light absorption coefficient, L (m) is the light path length, Y_{XS} and m as the substrate yield coefficient (mmol g^{-1}) and maintenance coefficient ($\text{mmol g}^{-1} \text{ h}^{-1}$), respectively.

2.3. Parameter perturbation analysis

The parameters in the mechanistic model are grouped into three classes: Class I (i.e., Y_{XS} and m) for reaction conversion related parameters, Class II (i.e., k_s , k_{s,H_2} and τ) for light intensity associated parameters, and Class III (i.e., A' , E_a , A'_{H_2} and E_{a,H_2}) for temperature associated parameters based on expert knowledge (e.g., Class I are well-known to great accuracy from microbiological studies meanwhile the same is not true for Class III). Input parameter uncertainties are assigned to each Class, then probabilistic sampling with Latin Hypercube Sampling was carried out, and the effects on the model's prediction uncertainty was propagated by Monte Carlo simulations.

2.4. PBR large scale simulations

The frequencies of local light/dark cycles are the only factor influenced by the PBR's transport phenomena on biomass growth rate. This effect was simulated in our recent study [5] by introducing one additional parameter (i.e., *effective light intensity coefficient*, η) which is larger than 1 if the PBR's culture mixing is intensified thereby leading to Equation (4). As the biohydrogen production rate is known to be a function of the PBR's biohydrogen partial pressure since this reaction is reversible, another parameter (i.e., *biohydrogen enhancement coefficient* ϕ) is added into Equation (6) to affect the overall hydrogen production rate as per the influence of biohydrogen removal rate and gas-liquid mass transfer.

$$\frac{dX}{dt} = A' \cdot \exp\left(-\frac{E_a}{RT}\right) \cdot \left(\frac{\frac{I_0}{(\tau \cdot X) \cdot L_L} \cdot (1 - e^{-(\tau \cdot X) \cdot L_L})}{k_s \left(\frac{1}{\eta}\right) + \left(\frac{I_0}{(\tau \cdot X) \cdot L_L} \cdot (1 - e^{-(\tau \cdot X) \cdot L_L})\right)} \right) \cdot X \quad (4)$$

$$\frac{dS}{dt} = -Y_{XS} \cdot \frac{dX}{dt} - m \cdot X \quad (5)$$

$$\frac{dH_2}{dt} = \phi \cdot A'_{H_2} \cdot \exp\left(-\frac{E_{a,H_2}}{RT}\right) \cdot \left(\frac{\frac{I_0}{(\tau \cdot X) \cdot L_L} \cdot (1 - e^{-(\tau \cdot X) \cdot L_L})}{k_{s,H_2} \left(\frac{1}{\eta}\right) + \left(\frac{I_0}{(\tau \cdot X) \cdot L_L} \cdot (1 - e^{-(\tau \cdot X) \cdot L_L})\right)} \right) \cdot X \quad (6)$$

3. Results and discussion

3.1. Dynamic parameter estimation results

Table 1 lists the obtained parameter values while the orange line in figure 1 shows the model fitting result after the step I of the parameter estimation for 100 Wm^{-2} and 200 Wm^{-2} at 35°C and 30°C , respectively with the remainder (i.e., 200 Wm^{-2} at 35°C and 40°C) not shown due to fitting similarities.

Table 1: Model parameter estimates and literature validation.

Parameters	Estimated	Parameters	Estimated
Step-one optimisation		Step-two optimisation	
$\mu_{T=303.15 \text{ K}}$ (h^{-1})	0.159	$\alpha_{T=308.15 \text{ K}}$ (h^{-1})	81.0
$\mu_{T=308.15 \text{ K}}$ (h^{-1})	0.185	$\alpha_{T=313.15 \text{ K}}$ (h^{-1})	106.0
$\mu_{T=313.15 \text{ K}}$ (h^{-1})	0.225	k_{s,H_2} (Wm^{-2})	500.0
k_s (Wm^{-2})	500.0	A' (h^{-1})	8.30×10^3
τ ($\text{mm}^2 \text{ g}^{-1}$)	90.8	E_a (J mol^{-1})	2.74×10^4
Y_{XS} (mmol g^{-1})	9.66	A'_{H_2} (h^{-1})	1.01×10^{10}
m ($\text{mmol g}^{-1} \text{ h}^{-1}$)	0.0140	E_{a,H_2} (J mol^{-1})	4.78×10^4
$\alpha_{T=303.15 \text{ K}}$ (h^{-1})	56.9		

3.2. Results of parameter perturbation analysis

Figure 1 shows the uncertainty bands and average percentage uncertainty over the process trajectory to be lowest and highest for Class I parameters (not shown) and Class III parameters respectively. This suggests that the model is highly sensitive to changes of Class III parameters but is less responsive to Class I parameters. Therefore, it is necessary to design more experiments for the accurate identification of temperature associated parameters meanwhile it is unsafe to update these Class III parameters during model-based process online optimization considering the high level of prediction uncertainty that can be introduced into the model output. Conversely, Class II parameters only showed a similar magnitude of model uncertainty when their input uncertainty was around 20% thereby indicating the model to be robust in their mild changes. Hence, Class II parameters are the ideal candidates to be re-estimated during online operation if the model is used for dynamic process optimisation.

3.3. Results of PBR large scale simulation

Figure 2 (a) to (c) shows that the model can well predict substrate consumption and hydrogen production in the large scale PBR with the calibrated transport phenomena associated parameters being $\eta = 4.515$ and $\phi = 1.945$. The η was observed to be higher than that reported in the literature [5] (i.e., 0.25 to 3.625) thereby implying a better light utilisation efficiency in the upscaled PBR which can be associated to the enhanced transport phenomena (i.e., better culture mixing that promotes a more frequent light/dark cycling). The obtained ϕ implies that the maximum specific H_2 production rate was almost doubled due to a decrease in the PBR's partial pressure which facilitates biohydrogen synthesis. Similar observations were reported in the literature [6] whereby a 12% increase in the biohydrogen productivity of *Rhodobacter sphaeroides* ZX-5 was attained when the total pressure at the PBR's headspace was decreased from 1.082×10^5 to 0.944×10^5 Pa.

Figure 2 (d) to (f) shows the biomass, substrate and biohydrogen models to be sensitive to the transport phenomena associated parameters. This implies: (i) they tune all of the state variables during any re-calibration process for the prediction of a different PBR scale and configuration, and (ii) can be updated during process online optimisation for scenarios such as a sudden disturbance of the PBR operation (e.g., random failure event

of the mixing pumps in the PBR) since their sensitivities were relatively low in comparison to Class III parameters.

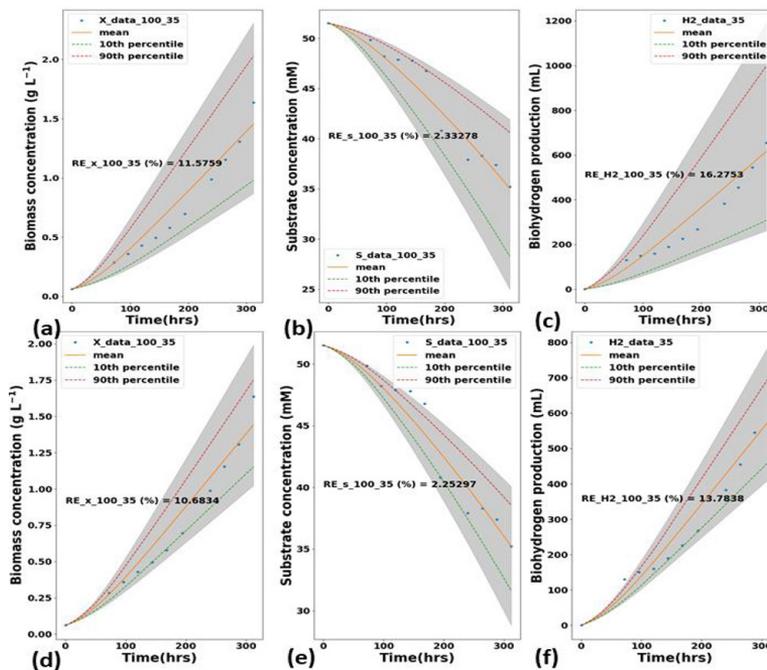


Figure 1: Perturbation analysis for input uncertainties of (a) to (c) 4% for class III, (d) to (f) 20% for class II. The mean, 10th and 90th percentile, and %RE for 100 Monte-Carlo simulations. Also, (a) to (c) and (d) to (f) are state variables corresponding to the biomass, substrate and biohydrogen models respectively

4. Conclusion

In this paper, a mechanistic model capable of simulating the photoheterotrophic biomass growth and biohydrogen production by the photosynthetic bacterium *R. palustris* under the unified influences of light intensity, light attenuation and temperature was developed. Model-based upscaling capabilities via two transport phenomena dependent parameters, η and ϕ embedded into the mechanistic model was possible after re-calibration. By using experimental data from a small (0.5 L) and large (1 L) scale photobioreactors (PBRs), the model prediction accuracy was thoroughly verified for both intra and across-scale predictions. Whilst perturbation analysis revealed temperature, light and transport phenomena associated parameters to be sensitive, the first should be identified and fixed during online model-based process optimisation meanwhile the second and last are more suitable to be updated. Amongst the investigated PBR scales, the enhanced biohydrogen production rate in the large scale PBR was attributed to the superior culture mixing and gas removal performance, and was mechanistically described for the first time. This provides important advances for the efficient design of novel PBRs and future online optimisation for biohydrogen production.

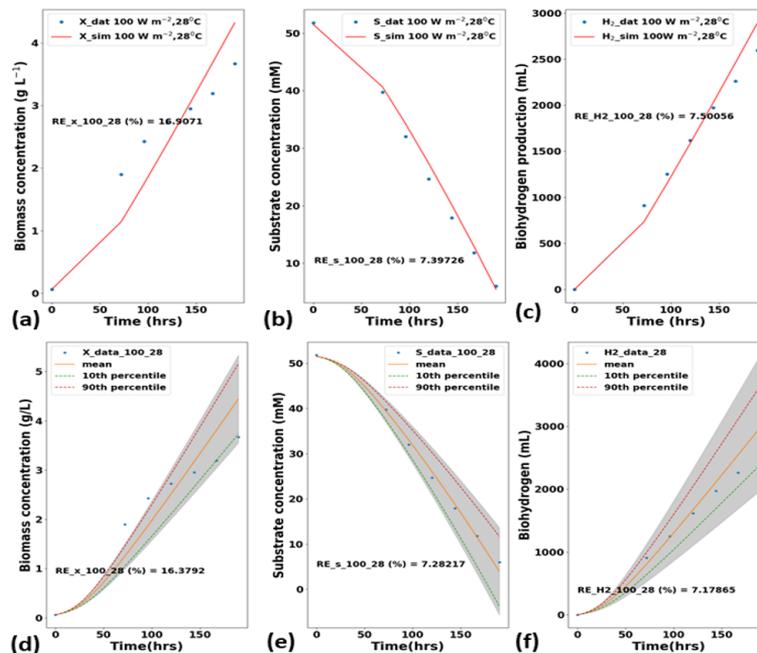


Figure 2: Simulation result of the large scale PBR: (a) to (c) and perturbation analysis (d) to (f) of 20% for transport phenomena associated parameters. The mean, 10th and 90th percentile, and %RE for 100 Monte-Carlo simulations. Each fitting is accompanied by the percentage relative error (%RE). Also, (a) to (c) and (d) to (f) are state variables corresponding to the biomass, substrate and biohydrogen models respectively.

References

- [1] P. Dechatiwongse, S. Srisamai, G. Maitland, and K. Hellgardt, "Effects of light and temperature on the photoautotrophic growth and photoinhibition of nitrogen-fixing cyanobacterium *Cyanothece* sp. ATCC 51142," *Algal Res.*, vol. 5, no. 1, pp. 103–111, Jul. 2014.
- [2] B. Anye Cho, B. S. Ross, J. P. du Toit, R. W. M. C. Pott, E. A. del Río Chanona, and D. Zhang, "Dynamic modelling of *Rhodospseudomonas palustris* biohydrogen production: Perturbation analysis and photobioreactor upscaling," *Int. J. Hydrogen Energy*, vol. 46, no. 74, pp. 36696–36708, 2021.
- [3] J.-P. du Toit and R. W. M. Pott, "Heat-acclimatised strains of *Rhodospseudomonas palustris* reveal higher temperature optima with concomitantly enhanced biohydrogen production rates," *Int. J. Hydrogen Energy*, vol. 46, no. 21, pp. 11564–11572, Mar. 2021.
- [4] H. Koku, I. Eroğlu, U. Gündüz, M. Yücel, and L. Türker, "Kinetics of biological hydrogen production by the photosynthetic bacterium *Rhodobacter sphaeroides* O.U. 001," *Int. J. Hydrogen Energy*, vol. 28, no. 4, pp. 381–388, Apr. 2003.
- [5] B. Anye Cho, M. Á. Carvalho Servia, E. A. del Río Chanona, R. Smith, and D. Zhang, "Synergising biomass growth kinetics and transport mechanisms to simulate light/dark cycle effects on photo-production systems," *Biotechnol. Bioeng.*, vol. 118, no. 5, pp. 1932–1942, May 2021.
- [6] X. Li et al., "Effects of light/dark cycle, mixing pattern and partial pressure of H₂ on biohydrogen production by *Rhodobacter sphaeroides* ZX-5," *Bioresour. Technol.*, vol. 102, no. 2, pp. 1142–1148, Jan. 2011.

Reducing the experimental effort to design pharmaceutical tablet lubrication by model-based design of experiments

Francesca Cenci,^a Gabriele Bano,^b Charalampos Christodoulou,^b Yuliya Vueva,^b Simeone Zomer,^b Massimiliano Barolo,^a Fabrizio Bezzo,^a Pierantonio Facco^{a*}

^aCAPE-Lab – Computer-Aided Process Engineering Laboratory,
Department of Industrial Engineering, University of Padova,
via Marzolo 9, 35131 Padova, Italy

^bGlaxoSmithKline, Ware, United Kingdom

*pierantonio.facco@unipd.it

Abstract

In oral solid-dosage manufacturing through direct compression, lubrication is used to enhance powder flowability and the ejection of the tablet from the die. However, lubrication can negatively impact tablet quality attributes such as tablets hardness or dissolution. In order to facilitate the selection of an appropriate lubrication extent, different models describing the relation between compaction performance and process conditions may be used. In particular, the extension of the Kushner and Moore model proposed by Nassar et al. (Nassar et al., 2021, *Int. J. Pharm.*, **592**, 119980) allows predicting tensile strength over a wide range of tablets solid fraction and powder blending time values. The main drawback of this model is that it requires a large number of experiments for parameter estimation. This results into a significant consumption of active pharmaceutical ingredient (API), which may be scarce and considerably expensive. In this study, model-based design of experiments is used to reduce the required experimental effort for the identification of the model parameters. We propose a novel procedure that is able to reduce parameters uncertainty while minimizing the number of required experiments. Results based on a simulated case-study demonstrate the effectiveness of the approach.

Keywords: model-based design of experiments; pharmaceutical engineering; quality by design; oral solid-dosage manufacturing

1. Introduction

In tablet manufacturing, while dry or wet granulation facilitates compaction and flow properties by converting fine powders into agglomerates (Šantl et al., 2011), direct compression improves powder manufacturability only through lubrication. Accordingly, lubrication is a crucial step: an excessive usage of lubricant may degrade tablets properties, like disintegration and dissolution, thus prejudicing the correct absorption when they are ingested by the patients. Many lubrication models have been proposed to aid the product development in tablets production through direct compression. Kushner and Moore (2010) developed a model which relates tensile strength and lubrication at the fixed solid fraction of 0.85. Recently, Nassar et al. (2021) introduced two additional parameters to the abovementioned model to describe the effect of solid fraction when it varies in a wide range of values. Despite the satisfactory predictive power, a considerable amount of experiments is needed for model parameters identification. In the industrial practice the experiments are usually carried out in a suboptimal, and consequently

ineffective, manner, based on trial-and-error approaches. The typical experimental procedure goes through the preparation of up to 7-9 blends with different lubrication extents, in which one powder blend is first prepared with a specific lubrication extent, and then a compression profiling is performed to produce tablets with different solid fractions. While changing compression pressure to have different solid fractions is not an issue, preparing a considerable number of blends (i.e., a considerable number of lubrication extents) leads to an excessive usage of API, which is expensive and may not be available in the required quantity during drug development.

In this work we adopt model-based design of experiments (MBD_{oE}; Asprey and Macchietto, 2000), which is a science driven method to select optimal experimental conditions yielding the maximum information content for the purpose of parameters estimation. In order to minimize the number of blends to be prepared, a novel approach is proposed based on a two-step optimization: first the optimal solid fractions are found for every possible value of lubrication extent (i.e., blend); then, the lubrication extent providing the most informative experiment is selected.

2. Materials and methods

The efficacy of the proposed MBD_{oE} procedure is demonstrated through numerical simulations (Figure 1): (1) a proper model describing the lubrication process is selected, together with preliminary parameters guesses; (2) based on model equations and parameters, MBD_{oE} provides the design of the most informative experiments; (3) optimal experiments are executed in the process, which in this study is simulated through a digital model; (4) optimal data are available to update parameters estimates. This procedure is iterated until satisfactory parameters precision and model predictive power are obtained.

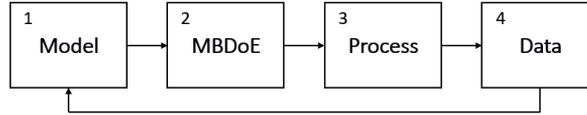


Figure 1 Workflow to validate the MBD_{oE} procedure through numerical simulations.

2.1. Lubrication model

The compression performance in the tablet press is quantified in terms of tablets tensile strength (ts , MPa), which is related to powder lubrication extent and to tablets solid fraction (sf , -) through the extended Kushner and Moore equation proposed by Nassar et al. (2021):

$$ts/ts_{sf=0.85,0} = (1 - \beta) + \beta \exp(-\gamma k) \quad (1)$$

$$ts_{sf=0.85,0} = a_1 \exp(b_1(1 - sf)) \quad (2)$$

$$\beta = a_2(1 - sf) + b_2 \quad (3)$$

where Eq. (1) represents the original Kushner and Moore (2010), while Eqs. (2) and (3) introduce the dependence of tensile strength from solid fraction.

Five parameters (θ) must be identified in this model: a_1 [MPa], a_2 [-], b_1 [-], b_2 [-] and γ [dm^{-1}]. Two variables can be manipulated in the experiments to estimate those parameters, namely solid fraction and lubrication extent.

2.2. Model-based design of experiments and global sensitivity analysis

MBD_{oE} aims at finding the optimal experiment conditions $\phi_{\text{opt}} = [sf_{\text{opt}}, k_{\text{opt}}]^T$ in terms of solid fraction and lubrication extent in such a way as to estimate the model parameters

with minimum uncertainty. This is performed by solving an optimization problem which maximizes the information obtained by the experiment, which is evaluated through the Fisher information matrix (FIM), $\mathbf{H}_{\hat{\theta}}$ (Box and Lucas, 1959):

$$\mathbf{H}_{\hat{\theta}}(\hat{\theta}, \phi) = \frac{1}{\sigma_y^2} \mathbf{S}(\hat{\theta}, \phi)^T \mathbf{S}(\hat{\theta}, \phi), \quad (4)$$

where \mathbf{S} is the sensitivity matrix whose elements are:

$$s_{sf,k}(\hat{\theta}, \phi) = \left[\frac{\partial ts}{\partial a_1}(\hat{\theta}, \phi), \frac{\partial ts}{\partial b_1}(\hat{\theta}, \phi), \frac{\partial ts}{\partial a_2}(\hat{\theta}, \phi), \frac{\partial ts}{\partial b_2}(\hat{\theta}, \phi), \frac{\partial ts}{\partial \gamma}(\hat{\theta}, \phi) \right]_{sf,k}, \quad (5)$$

the decorator symbol ($\hat{\cdot}$) indicates estimated quantities, ϕ is the design vector which collects inputs variables, i.e. $\phi = [sf, k]^T$, and $\hat{\theta}$ are the estimated parameters. In this study, the information content is maximized through a D-optimal strategy (Pukelsheim, 1993) which minimizes the volume of the parameters uncertainty region:

$$\phi_{\text{opt}} = \underset{\phi}{\text{argmax}} \{ \det[\mathbf{H}_{\hat{\theta}}(\hat{\theta}, \phi)] \} \quad (6)$$

Note that, to avoid the dominance of some variables or parameters due to different scales, data are pre-treated through a division by a reference value (e.g., the mean value of each variable) and the parameters are estimated with scaled data and indicated with capital letters, i.e. $\hat{\theta} = [\hat{A}_1, \hat{B}_1, \hat{A}_2, \hat{B}_2, \hat{\Gamma}]^T$.

Considering that MBDofE calculations rely on sensitivity indices and that parameters with negligible influence on the response tend to have higher uncertainty, a preliminary global sensitivity analysis is performed. Sobol's method (Saltelli et al., 2008) is applied, estimating first-order ($S_{\text{main},i}$, $i=A_1, \dots, \Gamma$) and total-order ($S_{T,i}$, $i=A_1, \dots, \Gamma$) sensitivity indices with a latin-hypercube sampling strategy. In particular, 8000 resamplings are performed in the entire parameters domain whose ranges are shown in Table 1.

Table 1 Lower and upper bounds for model parameters used in global sensitivity analysis.

Parameter	Lower bound	Upper bound
A_1	0	5
B_1	-7	-2
A_2	0	5
B_2	0	5
Γ	0	5

2.2.1. Two-step MBDofE

With a classic MBDofE, input variables are freely varied inside their domain; therefore, in each iteration of the MBDofE procedure ϕ_{opt} may have different k_{opt} , which corresponds to the preparation of different powder blends after one compression point. However, the number of blends that should be prepared (i.e., the changes in powder lubrication extent) should be minimized for operational reasons. To reach this goal, we adapt the MBDofE procedure by developing a two-step optimization. At first, experimental domains of the input variables are specified: sf varies in the continuous range [0.65, 0.90], while k can take on integer values in [90, 2000] dm. Then, the following optimizations are performed:

- a possible value of k is fixed and the optimization (6) is solved to get N_{SF} optimal values of sf ; the set of N_{SF} solid fractions for the same lubrication extent is named "profile"; this procedure is repeated for every possible value of k ;
- among all calculated profiles, the one maximizing the objective function in (6) is selected as the optimal experiment to be executed ($\phi_{\text{opt}} = [sf_{\text{opt},1}, \dots, sf_{\text{opt},N_{SF}}, k_{\text{opt}}]^T$).

In general, the same values of $sf_{\text{opt},i}$ and k_{opt} can be obtained by solving (6) multiple times; however, since the objective is to reduce the experimental burden, replications of similar optimization results are avoided by imposing that the optimal values differ by at least 150 dm for k and 0.04 for sf . This allows to better explore the experimental domain, which is advantageous in case of process-model mismatch.

In order to have sufficient experimental data for model calibration, the two optimization steps are iterated $N_K = 4$ times, thus obtaining N_K profiles. For each profile, $N_{SF}=5$ optimal solid fractions are calculated.

In both optimization steps, results depend on the quality of the parameters estimates $\hat{\theta}$ used in (6), but usually “true” parameters estimates are not available at the beginning of the experimental campaign and parameters guesses $\hat{\theta}_{\text{guess}}$ must be used. In this study, we choose guesses based on historical data. Indeed, the datasets of five different formulations named A, B, C, D and E (for details, see Nassar et al., 2021) are available. Since they comprise a considerable amount of data points, the estimated parameters are treated as “true values” ($\theta_A, \theta_B, \theta_C, \theta_D, \theta_E$ respectively) and reliable initial guesses are obtained by setting $\hat{\theta}_{\text{guess}} = \text{mean}(\theta_A, \theta_B, \theta_C, \theta_D, \theta_E)$.

2.3. Process

In this study, we represent the process through a digital model of the historical formulation A, which is made of equations (1)-(3) and “true” parameters θ_A . Noise is added to the in-silico data to mimic experimental and measurement errors. In particular, given that specific solid fractions are difficult to be achieved in the physical system, pseudo-random noise is added to $sf_{\text{opt},i}$, $i = 1, \dots, N_{SF}$ to reproduce experimental errors (the mean sf error is 0.01). Furthermore, pseudo-random noise is added also to the values of ts generated by the digital model to reproduce a ts error variance of 0.02.

2.4. Global sensitivity analysis

First-order and total-order sensitivity indices are calculated in the selected domain of sf values for three different given lubrication extents, namely 90 dm, 800 dm, 2000 dm. Results (Figure 2) show that:

- at low lubrication extents ($k=90$ dm in Fig 2.a), the most influential parameters are \hat{A}_1 , \hat{B}_2 and \hat{B}_1 , while at medium and high lubrication extents ($k=800$ dm and $k=2000$ dm in Fig 2.a) the influence of \hat{B}_2 on the response is dominant;
- parameter \hat{A}_2 has always little influence on the response (Fig 2.b); therefore, it is expected to have higher uncertainty when the model is calibrated with experimental data. However, S_{T,\hat{A}_2} is not constantly equal to zero, thus this parameter cannot be set to a nominal value;
- the summation of the total effects (omitted for sake of conciseness) is almost always higher than 1, suggesting a high level of interaction among parameters at every sf and k value, which in turn complicates the unique identification of parameters.

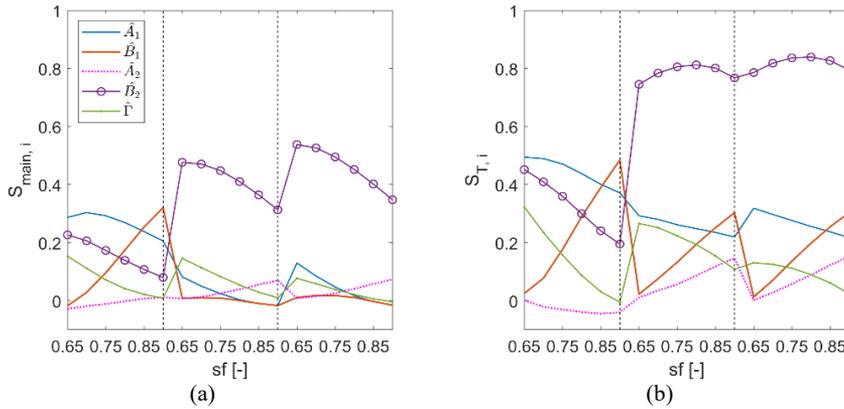


Figure 2 Global sensitivity analysis on model (1)-(3). Two types of sensitivity indices are shown: (a) main effects; (b) total effects. They are calculated for three different lubrication extents: 90 dm, 800 dm and 2000 dm (separated by vertical black dotted lines).

2.5. Design of optimal experiments

By applying the two-step MBDofE procedure, the following optimal lubrication extents are obtained: 90 dm, 2000 dm, 718 dm, 1849 dm. Since the process is affected by experimental and measurement errors, four replicates are generated in silico for every optimal solid fraction level.

Optimal profiles are used to identify model parameters, whose precision is assessed through a t-test (Table 2). Moreover, model validation is performed by calculating the tensile strength absolute error (*ts* AE) with the historical dataset A in the domain of the MBDofE optimization, i.e. with *k* in [90, 2000] dm. As suggested by the process experts, the desired model predictive power is a percentage of at most 5% of data points with a *ts* AE exceeding 0.25 MPa.

Table 2 Results of model calibration performed with an increasing amount of optimal profiles generated in silico. Parameters estimates are shown together with 95% confidence intervals ($\hat{\theta}_i, \pm 100(1-\alpha)\%CI$); CI are not calculated when the FIM is ill-conditioned. Parameters precision is assessed through a t-test; estimates not satisfying the t-test are indicated with an asterisk (*).

no. k_{opt}	$\hat{\theta}_i, \pm 100(1-\alpha)\%CI$				
	\hat{A}_1	\hat{B}_1	\hat{A}_2	\hat{B}_2	$\hat{\Gamma}$
2	0.14*	-4.82*	0.92*	1.00*	0.63*
3	0.15±0.03	-4.81±0.41	0.72±0.44*	0.79±0.23	1.26±0.58
4	0.15±0.03	-4.81±0.40	0.71±0.37	0.80±0.19	1.21±0.53

The t-test for parameters precision show that two optimal profiles do not allow to obtain statistically sound parameters estimates; a considerable improvement is achieved by adding the third profile, since all parameters except for \hat{A}_2 pass the t-test. The higher uncertainty of this parameter is likely due to the scarce influence on the response, as revealed by GSA (Section 2.4). The addition of the fourth optimal profile significantly improves the quality of \hat{A}_2 , which passes the t-test; moreover, and the fourth optimal profile slightly improves the precision of parameters \hat{B}_2 and $\hat{\Gamma}$ (their 95% CI are smaller) and brings negligible differences in terms of parameters estimates and intervals for \hat{A}_1 and \hat{B}_1 . Moreover, three optimal profiles allow to attain the desired model predictive

performance (Fig. 3), leading to a reduction of the experimental effort by more than 50% with respect to the standard industrial practice.

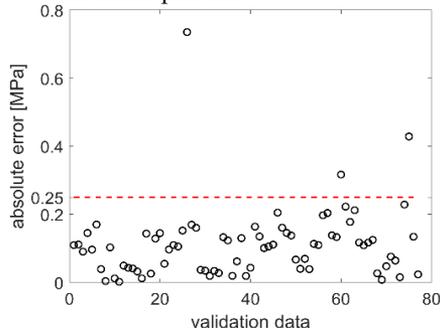


Figure 3 Tensile strength absolute error when the model is calibrated with three optimal profiles.

3. Conclusions

Lubrication models are used in the pharmaceutical industry to aid the design of lubrication processes, which in turn are crucial to facilitate tablets manufacturability and to ensure the desired product quality. Precise model parameters are required to the scope, but they are usually obtained with an excessive amount of experimental data, which implies a considerable API usage (thus, high costs) and time-consuming experimental campaigns. We solve this issue by proposing a novel MBDoE procedure that allows to select the most informative experiments for the purpose of parameters estimation. Moreover, the proposed MBDoE approach is able to minimize the change in lubrication extent, thus the number of blends to be prepared, through a two-step optimization: first, solid fraction values are optimized for every lubrication extent, thus providing multiple profiles of optimal sf ; among them, the profile maximizing parameters precision is selected as the best experiment to be performed. We demonstrate that it is efficient in reducing the experimental effort by more than 50% with respect to the standard industrial practice, with considerable savings in terms of API and time for the experimentation. Indeed, three optimal blends instead of the typical 7-9 are sufficient to obtain both statistically sound parameters estimates and the desired prediction quality. However, with three optimal profiles a higher uncertainty has to be accepted for parameter \hat{A}_2 , which is likely due to the small influence of this parameter on the model response.

References

- Asprey, S. P., Macchietto, S. (2000). Statistical Tools for Optimal Dynamic Model Building. *Comput. Chem. Eng.*, **24**, 1261-1267.
- Box, G. E. P., Lucas, H. L. (1959). Design of experiments in non-linear situations. *Biometrika*, **46**, 77-90.
- Kushner, J., Moore, F. (2010). Scale-up model describing the impact of lubrication on tablet tensile strength. *Int. J. Pharm.*, **399**, 19-30.
- Nassar, J., Williams, B., Davies, C., Lief, K., Elkes, R. (2021). Lubrication empirical model to predict tensile strength of directly compressed powder blends. *Int. J. Pharm.*, **592**, 119980.
- Šantl, M., Ilić, I., Vrečer, F., Baumgartner, S. (2011). A compressibility and compactibility study of real tableting mixtures: The impact of wet and dry granulation versus a direct tableting mixture. *Int. J. Pharm.*, **414(1-2)**, 131-139.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer*; Wiley. Hoboken, NJ, USA.

A model-based approach to predict the flowability of directly compressed pharmaceutical blends from individual components

Gabriele Bano*, Magdalini Aroniada, Yuliya Vueva

GlaxoSmithKline, 45 Priority St., Ware SG1 20DE, United Kingdom

**gabriele.x.bano@gsk.com*

Abstract

In oral solid dosage (OSD) forms manufacturing, understanding and characterizing the flowability of pharmaceutical formulations is pivotal since product and process performance directly depend on this property. Therefore, whenever a formulation change occurs during drug development, the impact on powder flowability must be quantified to address any potential effects on the manufacturability or critical quality attributes of the final product. Current industrial practice relies on extensive experimentation to address this problem, resulting into a significant consumption of active pharmaceutical ingredient (API), which is expensive and scarce during the early stages of drug development.

In this study, we propose a mixing rule model that can be used to predict the flowability of pharmaceutical blends from the flowability of the individual components. The model is validated with 35 different powder blends consisting of common APIs and excipients used for directly compressed pharmaceutical formulations. Results show that, by measuring the flowability of the pure API and only 6 binary mixtures, the mixing rule model can be used to predict the flowability of directly compressed pharmaceutical blends containing the API without further experiments. For a new pharmaceutical product, the reduction in experimental effort using this modelling approach can translate to ~80% reduction of API consumption, depending on the formulated product and the API physical properties.

Keywords: quality by design; pharmaceutical manufacturing; mixing rules; OSD

1. Introduction

Understanding powder flowability is critical to the success of several unit operations involved in the manufacture of oral solid dosage (OSD) forms (Prescott and Barnum, 2000). The performance of feeders, blenders, tablet presses and capsule fillers depends, among other factors, on the flow properties of the formulation = (Osorio and Muzzio, 2013). Nevertheless, measuring and predicting powder flowability from the individual components of the formulation is notoriously a difficult task (Seville et al., 2000), especially for powders with small particle size and uncontrolled size and shape distributions (Leung et al., 2017).

Despite this complexity, attempts to predict the flowability of multi-component mixtures of solids have appeared in the literature. These attempts can broadly be categorized in two groups: empirical approaches based on statistical models (Hildebrandt et al., 2019; Barjat et al., 2021) and approaches based on the granular Bond number (Capece et al., 2015; Giraud et al., 2021). In the former case, powder flowability was predicted from both bulk and particle properties of the individual components using multivariate regression techniques, latent variable models, support vector regression models or a combination thereof (Barjat et al., 2021). In the latter case, specific flow properties of the powder mixture such as the flow function coefficient (ffc) were correlated to bulk and particle properties of the individual components via the Bond number, which quantifies the relative balance between gravitational forces and attractive interparticle forces such as the van der Waals forces.

Both these approaches have limitations. From one side, the prediction fidelity of statistical models depends on the size and the range of composition and bulk/particle properties included in the calibration dataset. The models currently available cover only a small range of powder flow characteristics observed in pharmaceutical OSD forms, and therefore have limited applicability for industrial applications, independently of the manufacturing platform considered (direct compression, dry or wet granulation). On the other side, the approaches based on the granular Bond number, although promising, require some input measurements that are not routinely collected in an industrial environment (e.g., particle surface energy), and often show a low prediction fidelity. Their applicability is therefore limited, since the additional effort to retrieve the required input measurements does not often justify their resulting low prediction fidelity.

In this study, we develop a pragmatic mixing rule model to predict the flowability of pharmaceutical blends from the individual components. We constrain our study to formulations that are suitable to direct compression as the chosen manufacturing platform for the final OSD product. Under this assumption, we identify a restricted set of components and composition ranges for the formulations which are relevant in an industrial context, and we quantify the prediction fidelity of the proposed modelling approach with experimental data.

2. Materials and experimental methods

2.1. Materials

Experiments were performed using blends composed of the following materials:

- microcrystalline cellulose (MCC) as Avicel PH102 (FMC Corporation, USA);
- anhydrous lactose as lactose Supertab 21AN (DFE Pharma, Germany);
- croscarmellose-sodium as Ac-Di-Sol (FMC Corporation, USA);
- magnesium stearate as LIGAMED MF-2-V (Peter Greven, Germany)
- monohydrate crystalline lactose as Pharmatose 200M (DFE Pharma, Germany) (used as surrogate API)
- micronized paracetamol (Mallinckrodt Pharmaceuticals, USA).

All materials were used as received by the vendors. Binary and ternary mixtures containing Pharmatose 200M as API (formulation A) and micronized paracetamol (formulation B) were prepared spanning the range of composition reported in Table 1. These components and composition ranges were selected based on current state-of-the art for directly compressed pharmaceutical formulations in an industrial environment.

Table 1. Materials used in this study and composition ranges used for the ternary mixtures.

Material	Grade	Range (%w/w)
API	n/a	[5,40]
MCC	Avicel PH102	[10, 90]
Lactose anhydrous	SuperTab 21AN	[15, 70]
Croscarmellose sodium	Ac-Di-Sol	Fixed at 3%
Magnesium stearate	LIGAMED MF-2-V	Fixed at 1%

Binary blends and ternary blends were prepared for the two formulations at different compositions according to the ternary plots shown in Figure 1.

2.2. Blend preparation (pre-lubrication)

2.3. The blends were prepared using 3L bin blender (Sino Pharmaceutical Equipment Development Co, Ltd). All excipients and the API were transferred into the blender after screening through 1.0 mm sieve, and then mixed at 20 rpm for 20 min.

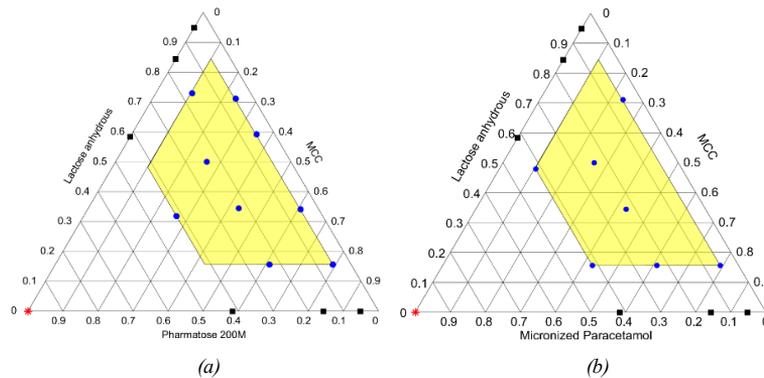


Figure 1. Powder blends prepared for this study using (a) Pharmatose 200M as surrogate API and (b) micronized paracetamol as API. Red stars are pure components, black squares binary blends and blue circles ternary blends. The shaded area represents the formulation design space of interest.

2.4. Blend lubrication

Magnesium stearate (MgSt) was added to the pre-lubricated blends in the amount of 1% w/w in the 5L bin blender. The mass of the non-lubricated blend required to achieve 40% head space and the lubrication time were calculated respectively from the blend bulk density, and according to the relationship proposed by Kushner (2012), by setting a target lubrication extent equal to 100 dm.

2.5. Flowability measurements

Powder flowability was measured using a Brookfield shear cell tester (Amtrek, USA). Prior to the analysis, all blends were pre-conditioned at 20 °C/50% RH for 24h. Flow function tests were performed for each blend using a standard (volume = 266 cm³) cell.

3. Proposed modelling approach

We aimed to construct a mixing rule model that could predict the following properties of the powder blend:

- slope and intercept of the flow locus;
- fill bulk density

using these properties of the individual components as inputs. From the slope and intercept of the flow locus, we derived the flow function coefficient at a consolidation endpoint = 3 kPa. This value is often used in industry to assess the flowability behaviour of the blend according to the classification system reported in Schulze (2008).

Most of the APIs and excipients that are used in pharmaceutical development are fine, dry and uncharged powders. In this scenario, previous studies (Capece et al., 2015; Giraud et al., 2021) have shown that the flowability of the mixture is dominated by inter-particle cohesive forces such as the van der Waals forces. Based on this assumption, we used statistical-mechanical arguments developed for mixture of fluids to propose a functional relationship between the property of the solid mixture and the individual components.

For a mixture whose pair intermolecular potential energy as a function of distance between the pair ($u_{ij}(r)$) can be described as:

$$u_{ij}(r) = \epsilon_{ij} f\left(\frac{r}{\sigma_{ij}}\right), \quad (1)$$

the following mixing rules can be derived to describe the molecular volume σ^3 and the molecular energy ϵ of the mixture:

$$\sigma^3 \propto \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}^3 \quad (2)$$

$$\epsilon \propto \frac{\sum_{i=1}^N \sum_{j=1}^N w_i w_j \epsilon_{ij} \sigma_{ij}^3}{\sigma^3}. \quad (3)$$

Here, σ_{ij} is intermolecular interaction distance between molecules of component i and j , ϵ_{ij} is the interaction energy parameter, N is the number of components in the mixture and w_i is the mass fraction of the i -th component. For a mixture of solids, we assumed that the property of the mixture can be described as a sum of two terms: 1) a term that depends on the molecular volume of the mixture σ^3 (zero-interaction term), and 2) a term that depends on the product (molecular energy) \times (molecular volume) $\epsilon \sigma^3$ (first-order interaction term). Leveraging on the mixing rules developed by Van der Waals for fluids and their statistical mechanics interpretation (Van der Waals, 1873; Kwak and Mansoori, 1986), and considering the analogy with our assumptions, we formulated the following mixing rule for a generic property ϕ_m of the mixture:

$$\phi_m = \sum_{i=1}^N w_i \phi_i + \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sqrt{\phi_i \phi_j} (1 - k_{ij}) \quad (4)$$

where k_{ij} is a binary interaction parameter (BIP) for the pair of components i and j that needs to be fitted from experimental data of the binary mixture, and ϕ_i is the property of the i -th component of the powder blend. The first term on the right-hand side of Eq. (4) is the zero-interaction term, while the second term is the first-order interaction term. Note the analogy between the first and second term of Eq. (4) and the "b" and "a" term in the van Der Waals equation of state for mixture of fluids.

Eq. (4) sets the structural relationship between the property of the mixture and the individual components. In order to use it, binary interaction parameters must be estimated for each pair of components of the mixture. For the formulation considered in Table 1, this requires estimation of the binary interaction parameters $k_{API,MCC}$, $k_{API,Lac}$ and $k_{MCC,Lac}$, with $MCC =$ Avicel PH102 and $Lac =$ Lactose SuperTab21AN. Note that, for a specific property of the mixture, only the first two parameters depend on the API, and therefore need to be estimated from experimental data when a new product is considered. The binary interaction parameter $k_{MCC,Lac}$ can be estimated only once and then used for any formulation containing those two components.

3.1. Software

The models presented in this study were coded in MATLAB R2020b. All simulations were performed on an Intel Core i7-5600U CPU@2.60GHz processor with 16.0 GB RAM.

4. Results

The BIPs were estimated from the binary mixtures reported in Figure 1 (black square points) for each binary system API-excipient or excipient-excipient. The parameter estimation was performed using a maximum likelihood estimator (Johansen and Juselius, 1990). The estimated values of the BIPs for the slope and intercept of the flow locus and for the fill bulk density are reported in Table 2. Model predictions were validated against the experimental data obtained for the ternary mixtures reported in Figure 1 (blue circles). Note that these mixtures were chosen in order to cover the design space of mixtures of industrial relevance (yellow area in Figure 1). The same plot shows the classification ranges for the flow behavior of the mixture according to the ffc values reported in Schulze (2008). Good agreement between the model predictions and the experimental observations was obtained for both formulations.

Table 2. Estimated value of the binary interaction parameters for the two formulations considered in this study. PT = Pharmatose 200M. MCC = Avicel PH102. Lac = Lactose SuperTab21AN. APAP=micronized paracetamol.

Binary interaction parameter	Value for slope of the flow locus [-]	Value for intercept of the flow locus [-]	Value for fill bulk density [-]
$k_{PT,MCC}$	3.648	0.085	0.753
$k_{PT,Lac}$	-0.861	1.756	0.830
$k_{APAP,MCC}$	-9.575	4.695	1.061
$k_{APAP,Lac}$	-8.897	3.497	2.387
$k_{MCC,Lac}$	3.107	0.720	1.306

The classification of the flow behavior of the blend was predicted correctly by the model for all blends apart from a single blend with micronized paracetamol (blend #3), where the blend was categorized as cohesive rather than very cohesive. However, the model prediction was still within the experimental error for the specific blend.

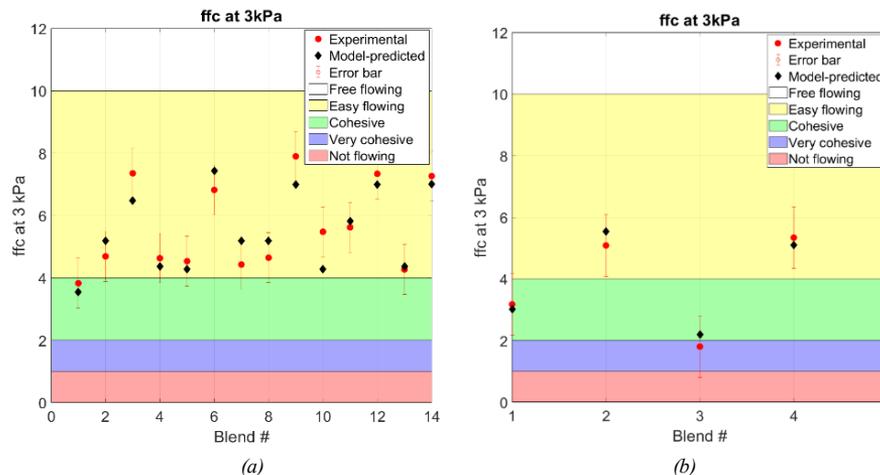


Figure 2. Model predictions Vs experimental values of the flow function coefficient at 3kPa consolidation endpoint for (a) formulation A (API: Pharmatose 200M) and (b) formulation B (API: micronized paracetamol).

Figure 3 shows the parity plots predicted vs experimental for the fill bulk density. Good prediction fidelity was obtained for both formulations; however, the model seemed to systematically underestimate the bulk density of the formulation containing micronized paracetamol. Investigation on this behavior is ongoing and the study will be further corroborated with additional experimental data for model validation. Overall, the model was able to consistently predict the behavior of the ternary blends with just binary and pure API data used as model inputs. Whenever a minor or major change of the formulation occurs, the model can be used to assess any impact on the flow behavior without additional experiments.

5. Conclusions

In this study, we developed a simple, yet effective, mixing rule model to predict the flowability of multi-component pharmaceutical powder blends that are suitable for direct compression. The model requires an upfront experimental characterization of the flowability of the pure API and 6 binary mixtures, and can then be used to predict the flowability of any ternary blend whose ingredient compositions fall within pre-defined ranges relevant to a direct compression

manufacturing route in an industrial environment. Whenever a minor or major change in the formulation composition occurs during the drug development process, the model can be used to predict the impact on the blend flowability without requiring any further shear cell experiments.

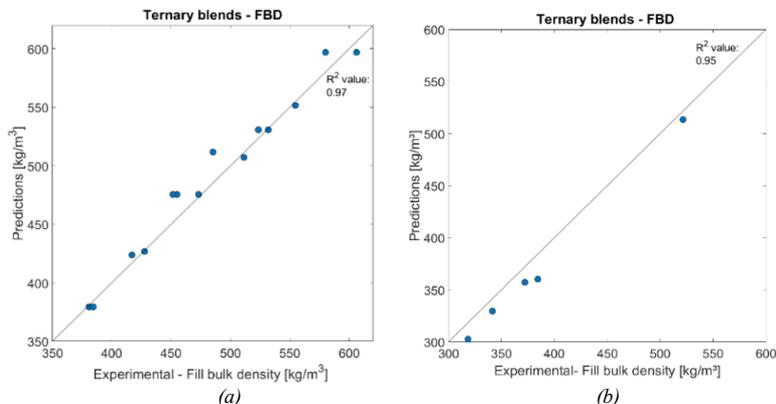


Figure 3. Model predictions Vs experimental values of fill bulk density for (a) formulation A (API: Phartamose 200M) and (b) formulation B (API: micronized paracetamol).

For some OSD products, where several alterations of the formulation composition might occur, this can translate into a reduction of up to 80% of API consumption for blend flowability assessment. Future work will focus on testing the methodology with additional APIs and excipients (e.g., mannitol) with different particle properties in order to define the range of validity of the underlying assumptions of the proposed approach.

Acknowledgments

We gratefully acknowledge the support of Innovate UK through the funding of the “Digital Design Accelerator Platform to Connect Active Material Design to Product Performance” (Grant No. 10581).

Conflict of interest

G.B., M.A. and Y.V. are employees of the GSK group of companies. The authors declare no competing conflicts of interest.

References

- Barjat, H., Checkley, S., Chitu, T., Dawson, N., Farshchi, A., Ferreira, A., Gamble, J., Leane, M., Mitchell, A., Morris, C. and Pitt, K. (2021) Demonstration of the Feasibility of Predicting the Flow of Pharmaceutically Relevant Powders from Particle and Bulk Physical Properties. *J. Pharm. Innov.*, 16, 181-196.
- Capece, M., Ho, R., Strong, J., & Gao, P. (2015). Prediction of powder flow performance using a multi-component granular Bond number. *Powder Technol.*, 286, 561-571.
- Giraud, M., Vaudez, S., Gatumel, C., Nos, J., Gervais, T., Bernard-Granger, G., & Berthiaux, H. (2021). Predicting the flowability of powder mixtures from their single components properties through the multi-component population-dependent granular bond number; extension to ground powder mixtures. *Powder Technol.*, 379, 26-37.
- Hildebrandt, C., Gopireddy, S. R., Fritsch, A. K., Profitlich, T., Scherließ, R., & Urbanetz, N. A. (2019). Evaluation and prediction of powder flowability in pharmaceutical tableting. *Pharm. Dev. Technol.*, 24, 35-47.
- Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxf. Bull. Econ. Stat.*, 52, 169-210.
- Kushner, J. (20212) Incorporating Turbula mixers into a blending scale-up model for evaluating the effect of magnesium stearate on tablet tensile strength and bulk specific volume. *Int. J. Pharm.*, 429, 1-11
- Kwak, T. Y., & Mansoori, G. A. (1986). Van der Waals mixing rules for cubic equations of state. Applications for supercritical fluid extraction modelling. *Chemical engineering science*, 41(5), 1303-1309
- Osorio, J. G., & Muzzio, F. J. (2013). Effects of powder flow properties on capsule filling weight uniformity. *Drug Develop. Ind. Pharm.*, 39, 1464-1475.
- Prescott, J. K., & Barnum, R. A., 2000. On powder flowability. *Pharm. Technol.*, 24, 60-85.
- Schulze, D. (2008). Powders and bulk solids. Behaviour, characterization, storage and flow. Springer, 22.
- Seville, J. P. K., Willett, C. D., & Knight, P. C. (2000). Interparticle forces in fluidisation: a review. *Powder Technol.*, 113(3), 261-268.
- Van der Waals, D., 1873. Dr. Thesis, Leiden University.

Purification of MIBA by Continuous Distillation

Pavan Veldandi^a, Ralph Cos^b, David Earp^c

^a*Cognizant Technology Solutions, Hyderabad, India*

^{b,c}*AVEVA Group plc, United Kingdom*

david.earp@aveva.com

Abstract

MIBA (Methyl Iso Butyryl Acetate) is an intermediate used in the pharmaceutical industry and is generally produced in small scale batch processes making it expensive to produce. This study was to look at the potential for a continuous process for the manufacture of MIBA, with the aim to improve yields and hence reduce costs., with the initial stage of the work focused on the purification of the MIBA product, this being required at high levels of purity, typically in excess of 99 weight %. Vacuum distillation column models were built using both AVEVA's PRO/II Simulation and AVEVA Process Simulation process modeling tools with the feasibility of separations assessed using PRO/II's Ternary VLE tool. This study investigated achieving high purities of MIBA by using different column configurations modelled in AVEVA PRO/II Simulation and AVEVA Process Simulation. This paper summarizes the methods and findings of this study.

Keywords: MIBA, Distillation, Simulation, modeling, Ternary Plots.

1. Introduction

The component Methyl Iso Butyryl Acetate (MIBA) has good demand in Pharmaceutical Industry across the world. Most of the manufactures are synthesizing MIBA in batch processes and as a result the cost of the product is high. So, in this work we tried to design a single continuous distillation unit. This is a challenging problem because the properties of the components involved in the reaction are not well known. It is a general method that MIBA is synthesized by the reaction of Methyl Acetoacetate with Iso Butyryl Chloride in the presence of Calcium Hydroxide and Ammonium Chloride. In the reaction the main byproducts are Acetamide and water. In this work the reaction part of the process is not considered for simulation. Rather, the focus is on removing Acetamide and water to produce 99.9 wt% MIBA by means of a continuous distillation process.

2. Simulation of continuous distillation of MIBA

2.1. Simulation using AVEVA PRO/II Simulation Software

In this simulation the following feed composition and conditions were used:

Table 1. Feed stream composition and condition

Stream Name		FEED
Temperature	C	80
Pressure	MM HG	133.488
Total Molar Rate	g-mol/hr	1384.401
Total Mass Rate	KG/HR	125
Stream Phase		Liquid
Thermodynamic system		NRTL
Total Weight Comp Fractions		
ACETAMIDE		0.0739
MIBA		0.8561
WATER		0.07
Total Weight Comp. Rates		
	KG/HR	
ACETAMIDE		9.2323
MIBA		107.0159
WATER		8.7518

A rigorous distillation column unit operation with 13 trays, reboiler and condenser is used to model the separation. The feed enters at the bottom. This is to make the distillation equivalent to the configuration in a batch distillation. The pressure of the column is maintained at 4 mmHg. The thermodynamic method used is Non-Random Two Liquid (NRTL). When the simulation model is run the top product purity achieved was 92.7 wt% MIBA, but this purity is not of sufficient quality for use in practical applications. The objective is to reach 99.9 wt% purity. An attempt has been made to improve the purity by changing several parameters for the column, but none of them achieved the desired purity. The MIBA composition profile versus the tray number, as illustrated by Figure 1, reveals a peak composition MIBA in the upper stages below the condenser. Resulting from this observation a side draw was added at tray 3 and this produced 99.9 wt% pure MIBA stream. The stream details of the top product and side draw are shown in Table 2.

2.2. Simulation using AVEVA Process Simulation Software

The same approach as mentioned in 2.1 has been followed to simulate the Distillation operation using AVEVA Process Simulation Software with identical results. The initial results without side draw are shown in Figure 2., and the process flow diagram with side draw is shown in Figure 3.

2.3 Application of Ternary plots

The separation has been analyzed with PRO/II's Ternary Plot tool confirming the obtained results through a conceptual analysis. The ternary plot analysis of the initial process flowsheet with 92 wt% pure MIBA is shown in Figure 4. The ternary plot analysis of the modified process flowsheet with 99.9 wt% pure MIBA is shown in Figure 5.

Table 2. Product stream composition and condition

Stream Name		SIDEDRAW	TOP PRODUCT
Temperature	C	56.501	0.653
Pressure	MM HG	5.246	4
Total Molar Rate	g-mol/hr	650	576.954
Total Mass Rate	KG/HR	92.853	22.643
Stream Phase		Liquid	Liquid
Thermodynamic system		NRTL	NRTL
Total Weight Comp. Fractions			
ACETAMIDE		0.0008	0
MIBA		0.9981	0.6182
WATER		0.0012	0.3818
Total Weight Comp. Rates			
ACETAMIDE	KG/HR	0.074	0.0002
MIBA		92.672	13.9982
WATER		0.107	8.6447

Figure 1. Weight fraction of MIBA from top tray to bottom tray

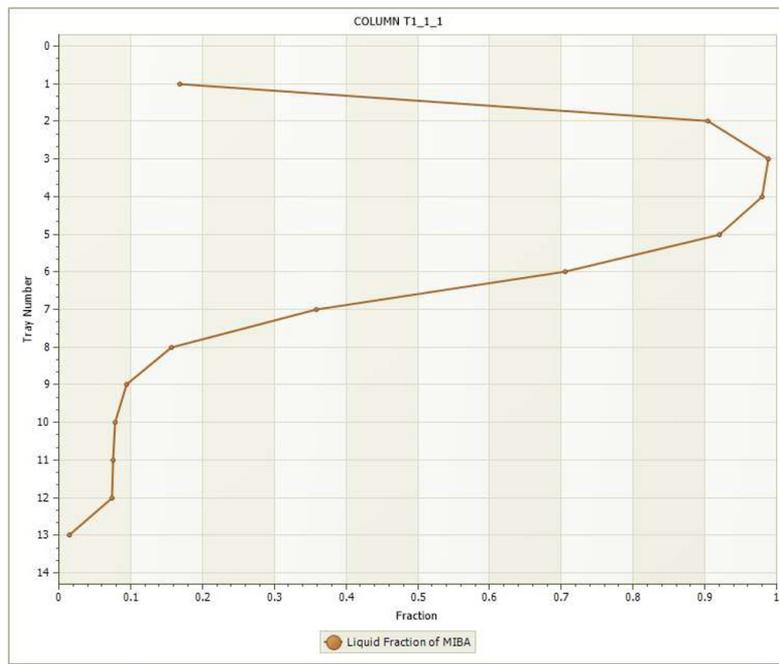


Figure 2. MIBA synthesis initial process flowsheet using AVEVA Process Simulation software

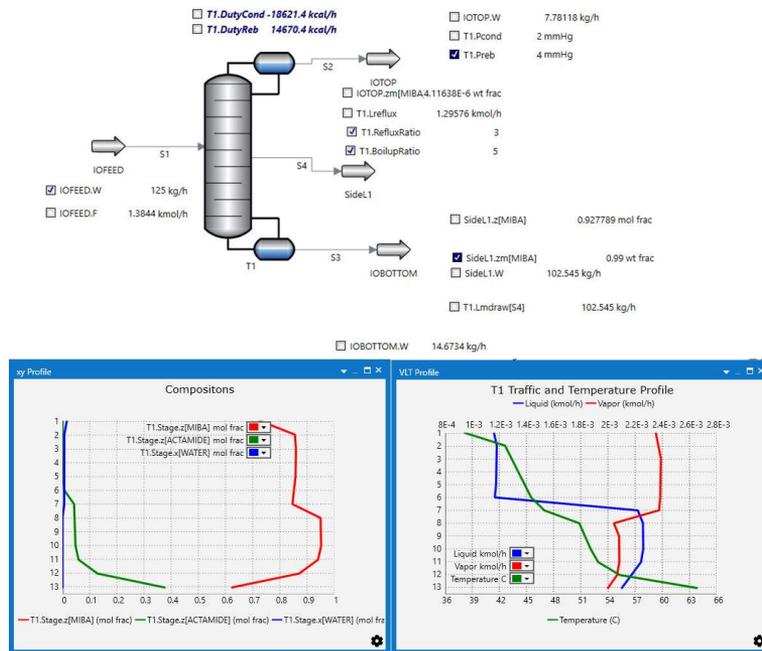


Figure 3. 99.9 wt% MIBA modified process flowsheet using AVEVA Process Simulation software

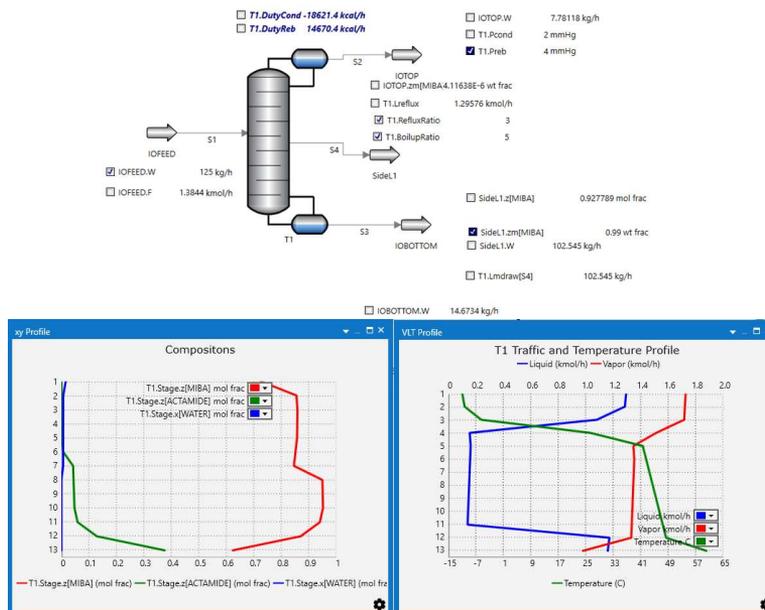
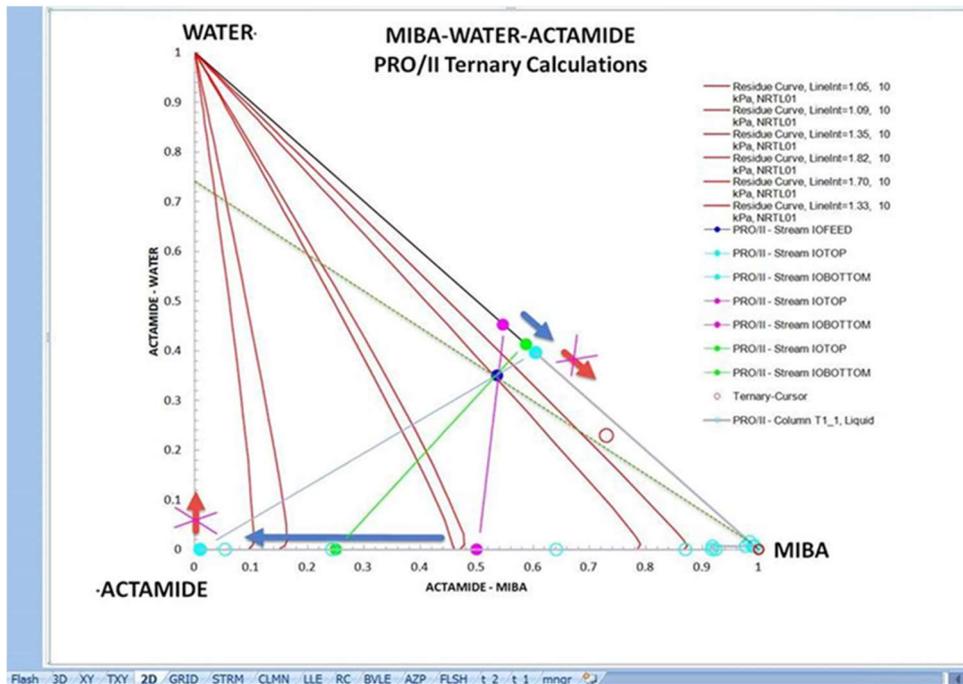


Figure 4. Ternary Plot for Base Flow diagram



The intention is to gain a 99 wt% pure MIBA distillate. But this is not achievable, it is infeasible. As the feed is a ternary system composed of MIBA, ACETAMIDE and water, we used the PRO/II Ternary Plot tool that allows to analyze the thermodynamic behavior along with the operational process specifications.

The filled dots represent stream compositions, the marine circles represent the column profile for the extreme solution when 60 mol% MIBA is obtained. This allows to analyze the separation statement. We can see the residue curves in red. According to separation theory for ternary systems, two requirements describe the solution:

- Distillate, bottoms and feed compositions form a straight line.
- Distillate and bottoms compositions must lie on the same residue curve.

This defines the separation region for a given feed. The chart below shows the feed as the blue dot. If we want the blue feed to separate into a pure MIBA composition then the bottoms must be resulting somewhere along the dotted green curve, it would be somewhere in the upper left section. This is due to requirement a).

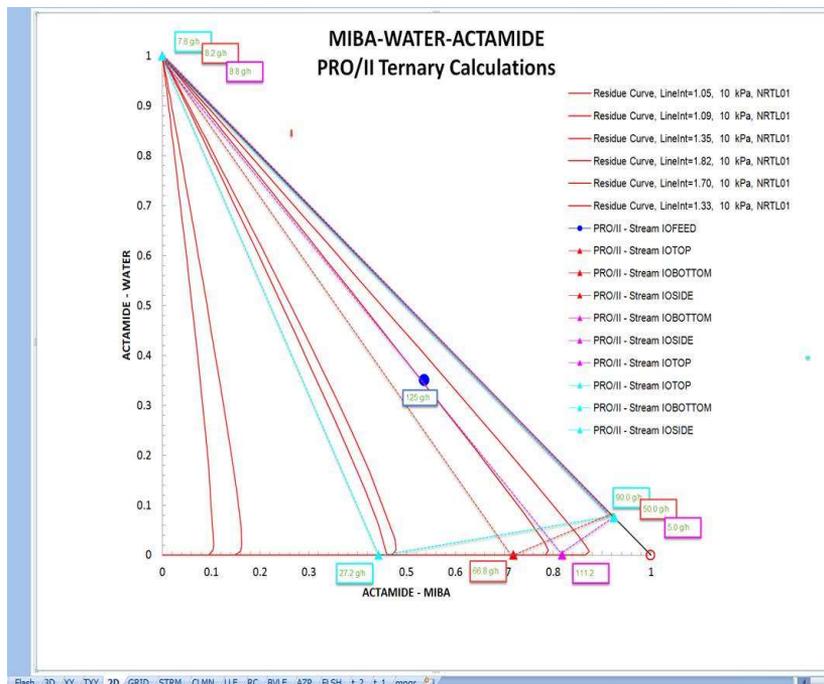
Requirement b) shows that a bottoms composition in that section is infeasible:

The column can be solved for the distillate/bottom composition pairs in pink, green, and marine. The arrow shows that as we move the specification for the distillate's MIBA from pink to marine, we also push the bottom's ACETAMIDE towards 100 wt%.

The point is that the residue curve does not extend upwards (the crossed out red arrow on lower left corner). Hence, there is no way a bottoms product can be gained that would be in the upper left section. The pure ACETAMIDE node (the chart's lower left corner) limits the MIBA distillate to the marine blue fraction of 0.6, i.e. 60 mol%.

Below is a plot with various side draw rates (90, 50, 5 g/h), all at 99 wt% MIBA (ca. 92 mol%), and the resulting top and bottom purities and rates. In case of side draws, the feed is located in the gravity center of the triangle that can be spanned by the top, side and bottom composition (the dotted lines). It is a simple mass balance.

Figure 5. Ternary Plot for 99.9 wt% MIBA Flow diagram



Conclusion

In this work a continuous distillation process has been proposed for purifying MIBA. Simulation has been carried out using AVEVA PRO/II Simulation software and AVEVA Process Simulation Software. The results from the standalone flowsheet have been verified with ternary plots feature available in both the software. The location of side draw plays an important role in achieving 99.9 wt% MIBA. The data used and the predicted performance of the distillation columns should be validated by experimental analysis.

References

- [1] Pai, G.G.; Kishore, N.; Chaudhari, N.P.; Anjaneyulu, N.; Ghogare, B.N. 2011, U.S. Patent No. 7,872,154 B2.
- [2] Islam M.; Sarkar, M.S. Detection of impurities in a tablet of Atorvastatin, KYAMCH J. 2011, 1(2), 43-47.
- [3] Sagyam, R.; Himabindhu, V.; Reddy, P.P.; Reddy, M.G. An efficient synthesis of highly substituted pyrrole and bis pyrrole derivatives. J. Het. Chem.,2007, 44, 923-926
- [4] Eranki, R.J.V.; Inki, G.; Jayaramn, V.; Vidiyala, S.R.; Ramulu, J.S. New stability indicating method for qualification of impurities in Amlodipine and Atorvastatin calcium tablet by validated HPLC. Int. J. Pharmaceut. Res. Scholar.,2013, 2(4), 458-471
- [5] De Lange, B. 2013, U.S. Patent No. 20130041162 A.

Modelling of Organophilic Pervaporation for Separation of Acetone-Butanol-Ethanol Mixture

Andras Jozsef Toth, ^{*,a}

^a*Department of Chemical and Environmental Process Engineering, Műegyetem rkp. 3., Budapest, 1111, Hungary
andrasjozsefth@edu.bme.hu*

Abstract

The motivation of the research is based on separation problem from fermentation industry, that is Acetone-Butanol-Ethanol solvent residue removal from its aqueous solutions. Pervaporation is a novel, promising technology for the separation of complex, azeotropic mixtures. Within pervaporation membrane category, organophilic pervaporation was investigated. In this work, laboratory experiments were performed with commercially available test membranes at different temperatures conditions. The separation factor and total organic fluxes were determined as well. It was found that, the separation factor and fluxes were inversely proportional to the feed butanol concentration. The laboratory results were consistent with literature studies. Using partial fluxes, semiempirical pervaporation models were fitted. Exponentially improved model version of Rautenbach pervaporation model (Szilagyi and Toth, 2020) was investigated in the case of binary mixtures. The aim of this work was to extend the observations of organophilic pervaporation model to more complex mixture. It can be observed that, the exponentially Rautenbach model describes accurately the transport process of organophilic pervaporation. Thus, it is possible to implement further studies in process simulator environment.

Keywords: Parameter estimation, Organophilic pervaporation, Acetone-Butanol-Ethanol mixture.

1. Introduction

During the ABE (A: Acetone, B: N-butanol, E: Ethanol) fermentation, Acetone, N-butanol and Ethanol are mainly produced, other components may be e.g. organic acids (butyric acid, acetic acid) (Kollarik, 2018). There are several possibilities for the separation of fermentation products. The most common solutions are distillation adsorption, liquid-liquid extraction, stripping, reverse osmosis, pervaporation (Valentinyi et al., 2018). The latter two methods fall within the scope of membrane procedures. The advantages of membrane operations are that they are generally energy efficient, flexible operations and do not require the addition of foreign (organic) substances to improve the separation. Any fouling and scaling during the process must be prevented. Furthermore, the membrane operations are environmentally beneficial because they do not produce significant additional waste (Haaz and Toth, 2018).

Pervaporation is a membrane operation in which the mixture to be separated evaporates to the low-pressure side of the membrane and the separation between the components through the membrane is carried out by the principle of sorption-diffusion. Vacuum pump is mostly used to achieve the low vapor pressure on the permeate side. Depending on the

permeable component, two types of pervaporation are distinguished: hydrophilic and organophilic pervaporation (Baker, 2012). The “sorption-diffusion” model is the most accurate to describe pervaporation in bilayer composite membranes, so the separation process of pervaporation can be derived from the sorption-diffusion principle (Wijmans and Baker, 1995). The industrial application of pervaporation has become more widespread in recent decades due to its lower energy consumption compared to traditional separation techniques (distillation, absorption, etc.). In the case of the products of ABE fermentation, the actual task is the problem of separating the quaternary mixture of Acetone - N-butanol - Ethanol - Water. Only organophilic pervaporation is possible, not hydrophilic because dilute aqueous solutions must be separated. Table 1 introduces some organophilic pervaporation membranes treating the products of ABE fermentation.

Table 1 Organophilic membranes separating ABE fermentation products (Feed, total organic flux and separation factor values)

Membrane type	Feed [wt%]			Total organic flux [kg/m ² h]	Separation factor [-]			Reference
	A	B	E		A	B	E	
PDMS/ceramic	0.6	1.2	0.2	1.21	34.6	20.0	6.5	Liu et al., 2014
PDMS	1.1	1.0	10.2	58.60	4.7	11.6	2.7	Liu et al., 2005
PEBA	0.6	1.9	0.7	0.03	5.1	12.4	3.5	Liu et al., 2005
PDMS filled with s. (60wt%)	7.0	15–20	1.0	0.91	7.0	20.0	1.0	Huang and Meagher, 2001
PDMS/ceramic	0.6	1.1	0.2	1.05	30.0	18.0	5.0	Liu et al., 2011
PDMS filled with s. (65wt%)	0.5	1.0	0.2	0.28	40.0	70.0	10.0	Zhou et al., 2011
PDMS filled with s. (50wt%)	3.0	10.0	1.0	0.12	44.0	70.0	5.2	Qureshi et al., 2001
TOA LM with PP s.	0.8	1.5	0.5	0.02	220.0	275.0	80.0	Thongsukmak and Sirkar, 2007

Studying Table 1, it can be seen that PDMS-based membranes are the most common in practical application. The aim of this research work is to investigate the quaternary Acetone - N-butanol - Ethanol - Water test mixture and the parameter estimation for semiempirical pervaporation model.

2. Material and methods

The measurements were performed on CM Celfa Membrantechnik AG P-28 membrane apparatus with the effective membrane area of 28 cm². On the permeate side, constant 13.33 mbar vacuum was provided with VACUUMBRAND PC2003 VARIO vacuum pump. Measurements were performed under isothermal conditions at three different temperatures (303, 318 and 328K) with an organophilic PDMS type membrane (Sulzer PERVAP 4060). The starting compositions of the 500 mL feed mixtures were 0.4 wt% Acetone, 0.8 wt% N-butanol and 0.1 wt% Ethanol. The rest of the test mixtures were water. The permeate was collected in liquid nitrogen-cooled traps and the compositions were analyzed by SHIMADZU GCMS-QP2010 gas chromatograph and Hanna HI 904 Karl Fischer coulometric titrator (Toth, 2015).

Appropriate computer modelling is an essential tool for designing and optimizing separation processes, which requires models that describe the processes as well as possible. Among the pervaporation models found in the literature, the used model

(Szilagyí and Toth, 2020) was based on Rautenbach's sorption-diffusion model. The model defines the process of pervaporation in the following steps.

- adsorption of the component in the selective layer of the membrane,
- diffusion of the component through the membrane material,
- desorption of the target component on the steam side.

The basic equation of this model is (J : partial permeate flux) (Szilagyí and Toth, 2020):

$$J_i = \bar{D}_i^* \exp \left[\frac{E_i}{R} \left(\frac{1}{T^*} - \frac{1}{T} \right) \right] \left(\frac{p_{i1} - p_{i3}}{p_{i0} \bar{y}_i} \right) \exp(x_{i1}^B) \quad i = (1, \dots, k) \quad (1)$$

First the parameters of the model must be estimated based on measured data. These parameters are the activation energies (E_i) and the reference transport coefficients (\bar{D}_i^*), and the B parameters that shows the concentration dependencies of the transport coefficients. Transport coefficient depends on the temperature in an Arrhenius type exponential way. The liquid activity coefficients can be calculated with different vapor-liquid equilibrium models or with the Wilson equation (Haaz and Toth, 2018). The estimations are completed with the STATISTICA® program environment. The verification can be obtained with objective function, that is minimized the deviation of the modelled and the measured values (Toth et al., 2018).

$$OF = \sum_{i=1}^n \left(\frac{J_{i,measured} - J_{i,modelled}}{J_{i,measured}} \right)^2 \quad (2)$$

Partial pressures (p_{i0}) are calculated according to the Antoine equation (Haaz and Toth, 2018).

3. Results and discussion

The most valuable component during the separation is N-butanol, so the parameters evaluating the separation was plotted as a function of this. Figure 1 shows the separation factors at 303K. Total organic fluxes at 303, 318 and 328K can be seen in the Figure 2.

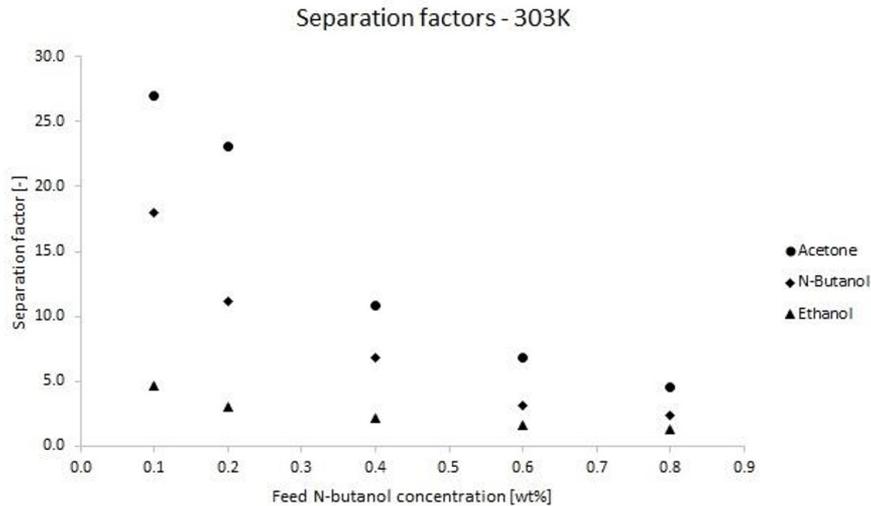


Figure 1. Separation factors of ABE mixtures in a function of feed N-butanol content in weight percent with Sulzer PERVAP™ 4060 membrane

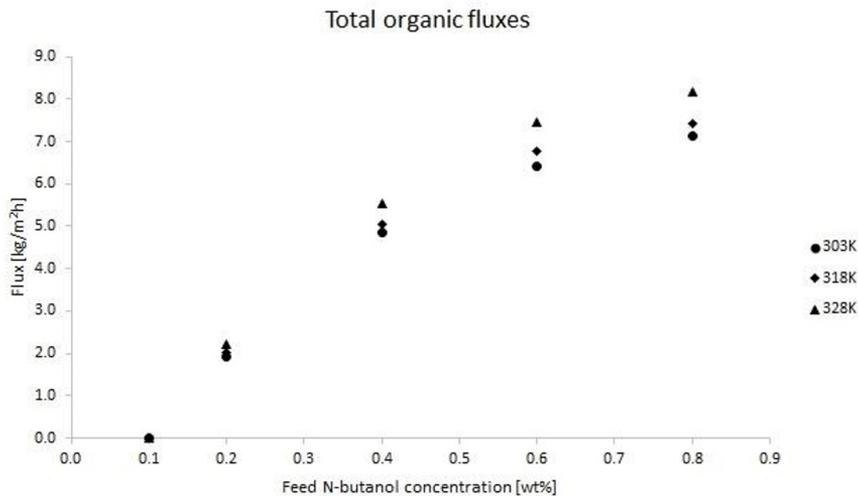


Figure 2. Measured total organic fluxes of ABE mixtures in a function of feed N-butanol content in weight percent with Sulzer PERVAP™ 4060 membrane

Figure 1 and Figure 2 show that the fluxes are inversely related to the separation factors as a function of the N-butanol feed concentration. Higher temperatures resulted in higher flux and separation factor values. In order of magnitude, the order is: Ethanol < N-butanol < Acetone, which is in agreement with the literature (Kujawska et al., 2015). Figure 3, Figure 4 and Figure 5 show the comparison of the measured partial fluxes with the model at 303K.

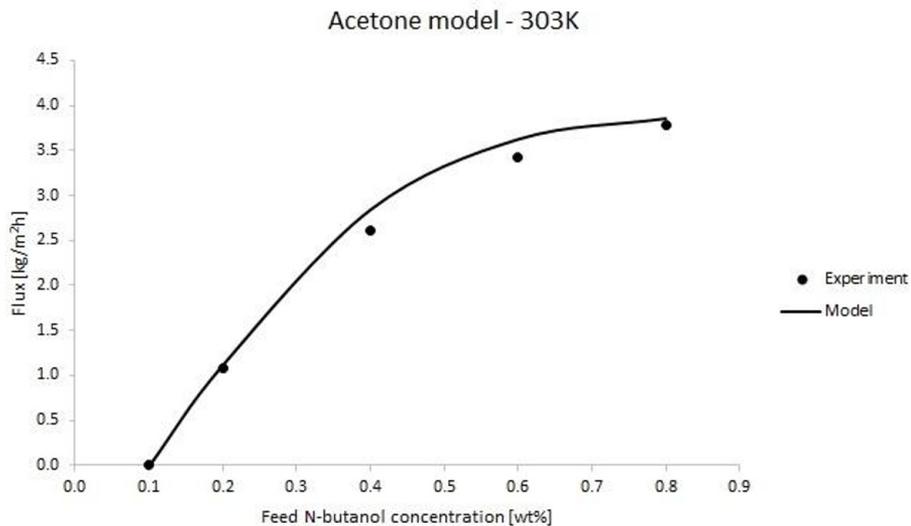


Figure 3. Measured partial Acetone fluxes of ABE mixtures compared to Acetone fluxes calculated with pervaporation model at 303K in a function of feed N-butanol content in weight percent

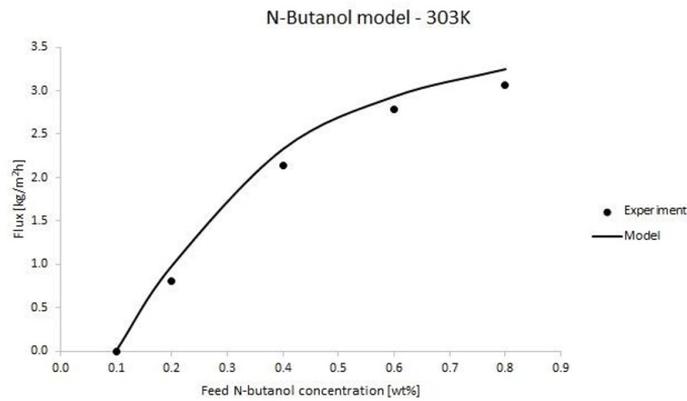


Figure 4. Measured partial N-butanol fluxes of ABE mixtures compared to N-butanol fluxes calculated with pervaporation model at 303K in a function of feed N-butanol content in weight percent

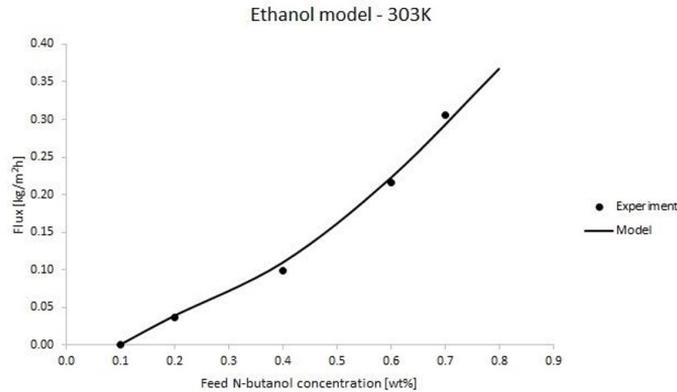


Figure 5. Measured partial Ethanol fluxes of ABE mixtures compared to Ethanol fluxes calculated with pervaporation model at 303K in a function of feed N-butanol content in weight percent

It can be seen that, in the case of the pervaporation model good agreement is found at the lower and also at the higher feed N-butanol content between the measured and calculated values. The reason for the better fit of the supplemented model probably lies in the concentration dependence of the transport coefficient. The results of the laboratory measurements suggest that the transport coefficient is also concentration dependent. The minimized objective functions and estimated values for transport coefficients, activation energies and B parameters of the pervaporation model are shown in Table 2. It can be seen, the low OF values also confirm the accuracy of the model.

Table 2 Estimated parameters for Acetone - N-butanol - Ethanol - Water mixture with Sulzer PERVAP™ 4060 membrane and objective functions

	Acetone	N-Butanol	Ethanol	Water
\bar{D}_i [kmol/m ² h]	0.034	0.193	0.003	65821
E_i [kJ/kmol]	18075	34298	43943	35637
B [-]	257.12	-30.45	624.12	-15.42
OF [-]	0.012	0.031	0.025	0.106

4. Conclusions

In this study, organophilic pervaporation works in the literature related to the separation of ABE fermentation products was presented. Laboratory experiments were performed at three different temperatures. It was found that the separation factors and the fluxes are inversely proportional to the N-butanol feed concentration. The results are in good accordance with literature studies. It can be observed that the Rautenbach model with exponential exponent precisely describes the transport process of organophilic pervaporation. This publication was supported by NTP-NFTÖ-21-B-0014, MEC 140699, OTKA 128543, OTKA 131586 and TKP2020 National Challenges Subprogram.

References

- R. W. Baker, 2012, *Membrane Technology and Applications*, 3rd edition ed. Wiley.
- E. Haaz, A. J. Toth, 2018, Methanol dehydration with pervaporation: Experiments and modelling, *Sep Purif Technol* 205, 121-129.
- J. Huang, M. M. Meagher, 2001, Pervaporative recovery of n-butanol from aqueous solutions and ABE fermentation broth using thin-film silicalite-filled silicone composite membranes, *J Membr Sci* 192, 231-242.
- C. Kollarik, 2018, Investigation of quaternary mixtures - Separation of ABE fermentation products by pervaporation, MSc Thesis, Budapest University of Technology and Economics, Budapest.
- A. Kujawska, J. Kujawski, M. Bryjak, W. Kujawski, 2015, ABE fermentation products recovery methods—A review, *Renew Sustain Energy Rev* 48, 648-661.
- F. Liu, L. Liu, X. Feng, 2005, Separation of acetone–butanol–ethanol (ABE) from dilute aqueous solutions by pervaporation, *Sep Purif Technol* 42, 273-282.
- G. Liu, L. Gan, S. Liu, H. Zhou, W. Wei, W. Jin, 2014, PDMS/ceramic composite membrane for pervaporation separation of acetone–butanol–ethanol (ABE) aqueous solutions and its application in intensification of ABE fermentation process, *Chem Eng Process* 86, 162-172.
- G. Liu, W. Wei, H. Wu, X. Dong, M. Jiang, W. Jin, 2011, Pervaporation performance of PDMS/ceramic composite membrane in acetone butanol ethanol (ABE) fermentation–PV coupled process, *J Membr Sci* 373, 121-129.
- N. Qureshi, M. M. Meagher, J. Huang, R. W. Hutkins, 2001, Acetone butanol ethanol (ABE) recovery by pervaporation using silicalite–silicone composite membrane from fed-batch reactor of *Clostridium acetobutylicum*, *J Membr Sci* 187, 93-102.
- B. Szilagyi, A. J. Toth, 2020, Improvement of Component Flux Estimating Model for Pervaporation Processes, *Membranes* 10, 418.
- A. Thongsukmak, K. K. Sirkar, 2007, Pervaporation membranes highly selective for solvents present in fermentation broths, *J Membr Sci* 302, 45-58.
- A. J. Toth, 2015, *Liquid Waste Treatment with Physicochemical Tools for Environmental Protection*, PhD Thesis, Budapest University of Technology and Economics, Budapest.
- A. J. Toth, E. Haaz, S. Solti, N. Valentinyi, A. Andre, D. Fozér, T. Nagy, P. Mizsey, 2018, Parameter estimation for modelling of organophilic pervaporation, in: A. Friedl, J. J. Klemeš, S. Radl, P. S. Varbanov, T. Wallek, (Eds.), *Comput Aided Chem Eng*, Elsevier, pp. 1287-1292.
- N. Valentinyi, G. Marton, A. J. Toth, E. Haaz, A. Andre, P. Mizsey, 2018, Investigation of process alternatives for the separation of ethanol, n-butanol and water ternary mixture, *Chem Eng Trans* 69, 607-612.
- J. G. Wijmans, R. W. Baker, 1995, The solution-diffusion model: a review, *J Membr Sci* 107, 1-21.
- H. Zhou, Y. Su, X. Chen, Y. Wan, 2011, Separation of acetone, butanol and ethanol (ABE) from dilute aqueous solutions by silicalite-1/PDMS hybrid pervaporation membranes, *Sep Purif Technol* 79, 375-384.

Development of Deep Learning Architectures for Forecasting Distillation Columns Dynamic Behavior of Biobutanol Purification

Abraham Rodarte de la Fuente^a, Eduardo Sanchez-Ramirez^a, Martha Patricia Calderón-Alvarado^a, Juan Gabriel Segovia-Hernandez^a, Esteban A. Hernández-Vargas^b

^a *Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, México. gsegovia@ugto.mx*

^b *Instituto de Matemáticas, UNAM, Unidad Juriquilla, Blvd. Juriquilla 3001, Juriquilla, Queretaro, C.P. 76230, México
gsegovia@ugto.mx; esteban@im.unam.mx*

Abstract

In recent years, Artificial Neural Networks (ANNs) have received special attention to a widespread application in the field of engineering, biology, energy, and finance. Within the ANN design, several factors play a major role in the correct prediction of the process, such as the number of internal layers, the number of neurons, the number of used features, the training algorithm, the activation function, the number of epochs, among many others. Considering datasets of an intensified distillation column generated by Aspen Plus Dynamics at different operation conditions, here we bring clarity in the field with different architectures of ANNs to abstract the dynamics of both an intensified and conventional distillation process that separates an effluent coming from fermentation producing acetone, butanol, and ethanol (ABE) for spark-ignition purposes. Our results highlight that a one-layer neural network can represent the dynamics of an intensified column to forecast the concentration of acetone, butanol, and ethanol. Remarkably, the linear activation function overperforms the tangent hyperbolic as activation functions. Ultimately, we found that the reflux ratio and reboiler duty are key features to reconstruct the full dynamics of the intensified column.

Keywords: Deep Learning, ABE purification, LSTM, Biobutanol.

1. Introduction

The use of liquid biofuels such as ethanol and butanol has been presented as an important advance due to their origin from biomass fermentation and their low emission levels (Ribeiro et al., 2007). One of the processes by which it can be obtained is ABE fermentation from biomass, this process produces a mixture composed of acetone, butanol, and ethanol which creates a challenge in trying to reduce the energy in the separation process (González-Bravo et al., 2016). The use of new design approaches for the control, modeling, and simulation of chemical processes has allowed the development of intensified processes, seeking a radical change in the unit operations used to meet current needs under the development of a sustainable process (Sánchez-Ramírez et al., 2017). The dividing wall column (DWC) allows to reduce energy and capital costs due to the reduction of equipment since the separation of the mixture would be carried out in one column and not in two, in addition to the fact that shorter piping and electrical currents

are required (Gómez-Castro et al., 2008). Dynamic properties are an important issue in DWC since apparently dynamic behavior can be greatly reduced. Although it has been shown that control sequences can be better than conventional distillation sequences so you can have total annual cost savings as good dynamic performance (Rewagad & Kiss, 2012). Furthermore, there is not any preliminary work reported involving the processing of mixtures with a high degree of complexity relative to thermodynamic modeling. Deep learning models can learn extremely complicated patterns from a large amount of data without much manual expertise so they can be used for a large number of applications where their structure is constituted by several hidden layers that allow transforming the input data several times before producing the output, so it is possible to manipulate their architecture to obtain better output values without overfitting (López-Tapia et al., 2021). The use of Long short-term memory (LSTM) ANNs in time series prediction may offer more efficient and effective alternatives for highly complex multivariate systems (Ookura & Mori, 2020). The main objective of this study is to optimize the ANN architecture that models the dynamics behavior of a DWC comparing the activation function, the optimizers that minimize the Mean Squared Error (MSE), and analyzing the increase in the number of hidden layers based on the AIC (Akaike Information Criterion) value. Finally, the manipulable variables that have a greater weight on the system modeling while maintaining a good prediction of the output data will be identified.

2. Methodology

Data generation. The datasets are obtained from a simulation software Aspen Dynamics, this simulation is performed in a closed-circuit test where a set point was implemented in the composition of each component of the mixture to be separated (Acetone-Butanol-Ethanol), in this way three setpoint changes were performed, and tuned at the same time, thus producing time-varying operational datasets. Simulations of 100 hours were performed using a sampling time of 0.4 hours. The feed stream considers a mixture of acetone, butanol, ethanol, and water in proportions of 0.3018, 0.1695, 0.0073, and 0.5214 wt%, respectively. The datasets were subdivided into the first 24 hours for training and the rest for testing. The data manipulation is done in the Python 3.8 programming language where use is made of the PANDAS library which is ideal for data analysis.

Neural Network basic elements. An artificial neural network (ANN) is a distributed computing scheme inspired by the structure of the human nervous system. The architecture of a neural network is formed by connecting multiple elementary processors, being an adaptive system that has an algorithm to adjust its weights to meet the performance requirements of the problem based on representative samples. Eq. 1 and 2 represent the equivalent model of the synaptic connections in a k neuron.

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

where the vector x_j is the set of input signals, w_{kj} is the set of synaptic weights of neuron k , u_k is the linear combination of the weighted inputs, b_k is the polarization and y_k is the output signal of the neuron. The activation function φ serves the purpose of limiting the output range of the neuron and can be linear or nonlinear. In this study, we compare the performance of the functions Linear, ReLU, and Tanh. To obtain a good minimization of the loss function, the choice of the optimizer will be compared from a defined activation function to obtain a good prediction of the time series thanks to its convergence speed and its generalization speed (Manickam et al., 2021). The optimizers compared are

Adaptive Moment Estimation (ADAM), Root Mean Square Propagation (RMSP), and Stochastic Gradient Descent (SGD).

LSTM modeling. Recurrent Neural Networks (RNN) are distinguished in that they have at least one or more feedback loops that have a profound impact on the learning capability of the neural network. The LSTM is a special type of RNN that can learn long-term dependent information making considerable progress in problems related to time series analysis. This type of network needs a higher computational power because it has feedback based on a time step (Hochreiter & Schmidhuber, 1997). In this study, a step of 5 is used for the feedback, with which it is possible to have a good prediction for a large amount of data.

Model Selection. Optimal architectures of the neural network model can be found by minimizing the MSE concerning the variation of the number of hidden neurons (Shin et al., 2020).

$$MSE = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n} \quad (3)$$

where n is the number of data points, y_i is the actual value and \bar{y}_i is the value estimated by the ANN. Given a collection of models for a dataset, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a value for model selection. It deals with the trade-off between the goodness of fit of the model and the complexity of the model. AIC does not provide a test of a model in the sense of testing a null hypothesis, so it can tell nothing about the quality of the model in an absolute sense. If all the candidate models fit poorly, AIC does not give any warning of that. The formula for AIC depends upon the statistical model. A lower AIC value means that a given model describes the data better than other models with higher values.

$$AIC = N \log(MSE) + \frac{2MN}{N - M - 1} \quad (4)$$

N is the number of data points, M is the number of unknown parameters.

The normalization was performed using Eq. 5 to transform all feature values into an interval of $[0,1]$, mathematically speaking y_i refers to the data values, $\min(y_i)$ is the minimum value and $\max(y_i)$ refers to the maximum value. Increasing the efficiency of the algorithm by reducing fluctuations.

$$y_{scaled} = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)} \quad (5)$$

3. Results

From the DWC system, were considered as manipulable variables the reflux ratio, side streamflow, and reboiler heat duty. In the Linear activation function and the RMSP optimizer, an analysis of the impact of these manipulable variables was performed; first, the ANN was fed with the three manipulable variables the result is shown in blue color line in Figure 1(a), then the reboiler duty input to the ANN was inactivated and the predictive capability of the ANN was measured with the input only of the reflux ratio and the side stream flow the result is shown in orange color line Figure 1(a), also the analysis of inactivating the reflux ratio to the ANN and feeding the ANN reboiler duty and the Side streamflow the result is shown in Figure 1(a) green color line and finally inactivating the side-stream flow feed to the ANN and feeding the ANN with the Reboiler duty and ratio the results are shown black line Figure 1(a). According to these results, it can be observed that the one that showed the worst behavior was the one that deprived the Reboiler duty feeding, and the ones that showed the best results were on the one hand the

one that feed the three manipulable variables and the ones that deprived the side streamflow feeding and the reflux ratio. The one that showed the best result for 5 neurons was the one feed with the reflux ratio and the reboiler duty.

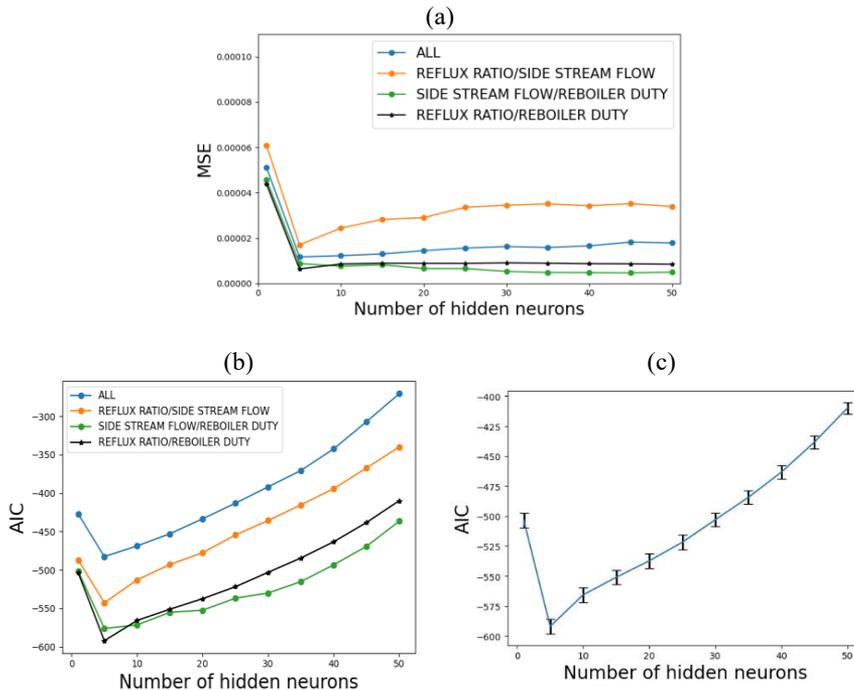


Figure 1. ANNs with one single hidden layer (a) Effects of the input manipulable variables on MSE; (b) AIC value comparison for input manipulable variables; (c) Best value of AIC is for Reflux ratio/Reboiler duty.

The ANN RMSP optimizer is analyzed for the three manipulable variables; Reflux ratio side streamflow and reboiler duty. The black-colored line in Figure 1(b) was the one that showed the lowest value of the AIC for 5 hidden neurons which is the one that was fed only with the Reflux ratio and the Reboiler heat duty. An error analysis was performed only for the one showing the lowest AIC shown in Figure 1(c).

According to the analysis performed to the different hyperparameters, the final topology of the ANN shown in Figure 2(b) was configured, where the manipulable variables that showed the greatest influence in the modeling of the DWC dynamics, as well as the perturbed variables, are feed to the ANN and only one layer of neurons is used, achieving a good prediction in the output variables, the composition profiles of Acetone, Ethanol, and Butanol. With the architecture shown in Figure 2(b), a good prediction was achieved for the system shown in the diagram in Figure 2(a), the results obtained by ANN show a good fit to the data simulated by Aspen Dynamics for the mass fraction of Acetone, Ethanol, and Butanol (see Figure 3). Table 1 summarizes the methodology that gave the best results in the ANN configuration.

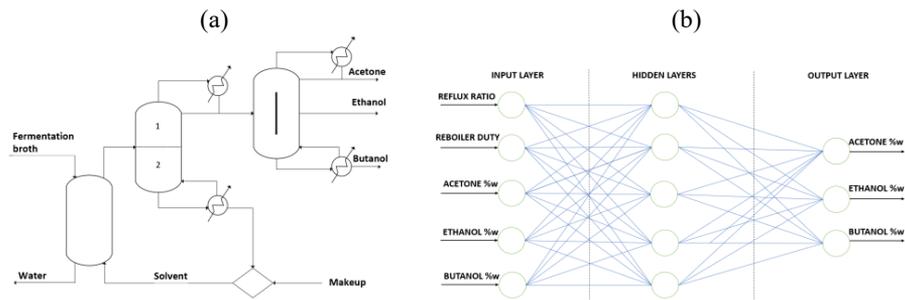


Figure 2. Artificial neural network final architecture. (a) Study case; (b) Final ANNs architecture.

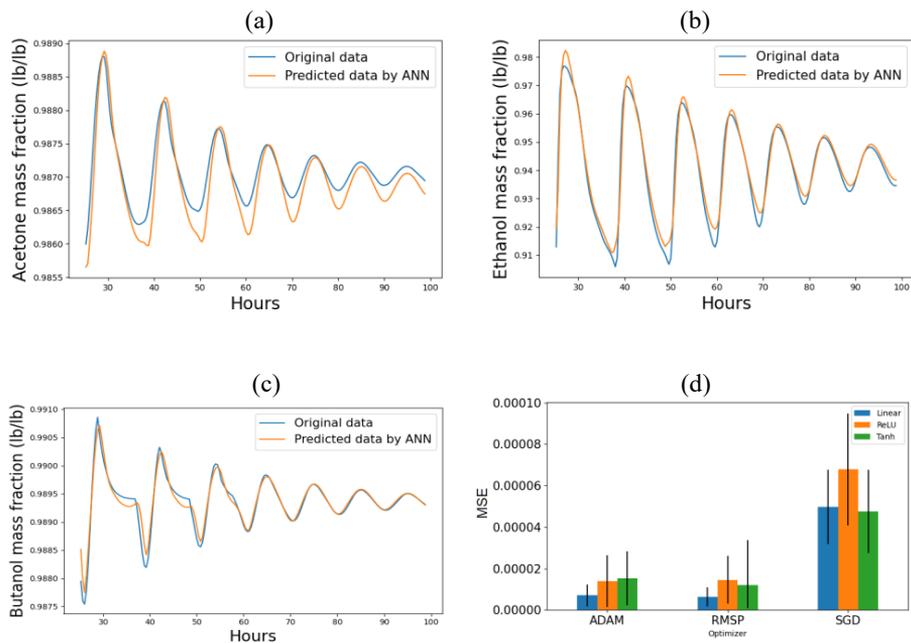


Figure 3. Comparison between the predicted and actual a) acetone, b) ethanol and c) butanol mass fraction; d) MSE values for activation functions and optimizers.

Table 1. ANN Configuration Summary.

Percentage of training data	24%	Activation function	Linear
Percentage of test data	76%	Optimizer	RMSP
Feedback step	5	Key Features	Reflux ratio/ Reboiler heat duty
Number of input features	5	MSE value	6.396095×10^{-6}
Type of neurons	LSTM	AIC value	-591.96
Number of hidden layers	1	Activation function	Linear
Number of hidden neurons	5	Optimizer	RMSP

4. Conclusions

The architecture of the ANN was optimized to obtain good forecasting of Acetone, Butanol, and Ethanol in a dividing wall column. A major result in our study is that we found that to have a good data prediction it is not necessary to consider the three manipulable variables. Results show that it is only necessary to consider two: the reflux ratio and the reboiler heat duty. This would have an impact on future work by implementing only two controllers for the three perturbed output variables of the DWC system, which would reduce the operating cost considerably by having fewer controllers and with a low number of neurons in the ANN structure.

References

- Gómez-Castro, F. I., Segovia-Hernández, J. G., Hernández, S., Gutiérrez-Antonio, C., & Briones-Ramírez, A. (2008). Dividing Wall Distillation Columns: Optimization and Control Properties. *Chemical Engineering & Technology*, 31(9). <https://doi.org/10.1002/ceat.200800116>
- González-Bravo, R., Sánchez-Ramírez, E., Quiroz-Ramírez, J. J., Segovia-Hernández, J. G., Lira-Barragán, L. F., & Ponce-Ortega, J. M. (2016). Total Heat Integration in the Biobutanol Separation Process. *Industrial & Engineering Chemistry Research*, 55(11). <https://doi.org/10.1021/acs.iecr.5b03168>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8). <https://doi.org/10.1162/neco.1997.9.8.1735>
- López-Tapia, S., Molina, R., & Katsaggelos, A. K. (2021). Deep learning approaches to inverse problems in imaging: Past, present, and future. *Digital Signal Processing*. <https://doi.org/10.1016/j.dsp.2021.103285>
- Manickam, A., Jiang, J., Zhou, Y., Sagar, A., Soundrapandiyar, R., & Dinesh Jackson Samuel, R. (2021). Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures. *Measurement*, 184. <https://doi.org/10.1016/j.measurement.2021.109953>
- Ookura, S., & Mori, H. (2020). An Efficient Method for Wind Power Generation Forecasting by LSTM in Consideration of Overfitting Prevention. *IFAC-PapersOnLine*, 53(2). <https://doi.org/10.1016/j.ifacol.2020.12.1008>
- Rewagad, R. R., & Kiss, A. A. (2012). Dynamic optimization of a dividing-wall column using model predictive control. *Chemical Engineering Science*, 68(1). <https://doi.org/10.1016/j.ces.2011.09.022>
- Ribeiro, N. M., Pinto, A. C., Quintella, C. M., da Rocha, G. O., Teixeira, L. S. G., Guarieiro, L. L. N., do Carmo Rangel, M., Veloso, M. C. C., Rezende, M. J. C., Serpa da Cruz, R., de Oliveira, A. M., Torres, E. A., & de Andrade, J. B. (2007). The Role of Additives for Diesel and Diesel Blended (Ethanol or Biodiesel) Fuels: A Review. *Energy & Fuels*, 21(4). <https://doi.org/10.1021/ef070060r>
- Sánchez-Ramírez, E., Quiroz-Ramírez, J. J., Hernández, S., Segovia-Hernández, J. G., & Kiss, A. A. (2017). Optimal hybrid separations for intensified downstream processing of biobutanol. *Separation and Purification Technology*, 185. <https://doi.org/10.1016/j.seppur.2017.05.011>
- Shin, Y., Smith, R., & Hwang, S. (2020). Development of model predictive control system using an artificial neural network: A case study with a distillation column. *Journal of Cleaner Production*, 277. <https://doi.org/10.1016/j.jclepro.2020.124124>

Impact of Methanol Synthesis Kinetics on Bulk Production Prediction: an In-Silico Assessment

Filippo Bisotti^a, Matteo Fedeli^a, Carlo Pirola^b, Giulia Bozzano^a, Flavio Manenti^{a*}

^a *Politecnico di Milano, Dept. CMIC “Giulio Natta”, Piazza Leonardo da Vinci, 32, Milan, 20133, Italy*

^b *Università degli Studi di Milano, Chemistry Dept., Via Golgi, 19, Milan, 20133, Italy*

* corresponding author - flavio.manenti@polimi.it

Abstract

Kinetic modelling covers a key role in process simulation and design. Recently the methanol sector is assisting a remarkable enhancement due to its applications as fuel, solvent, and precursor as shown in Bozzano and Manenti (2016). The increasing number of patents, the market prospects, and recent research witness this renewed interest. However, despite this rate in developing and improving technologies, the kinetics modelling does not follow these trends. The methanol synthesis chemical paths, the intermediates, and the real role of the active sites are nowadays still unclear. However, process engineering requires reliable models to estimate the methanol synthesis rate, hence, to design and size the reactor and downstream equipment. Currently, the most used kinetics are Graaf and Vanden Bussche - Froment's models which in any case show some shortcomings and weaknesses. Starting from these premises, the need for updated kinetics is clear. This work aims at comparing and highlighting the impact of different kinetic models (1) original Graaf (or-GR), (2) Vanden Bussche - Froment (VBF), and (3) refitted Graaf (ref-GR) on the methanol synthesis configuration for different feedstocks through an in-silico assessment. The general simulation flowsheet includes the single-stage PFR for the methanol synthesis, the condensation step, and recycle loop for the unreacted syngas. The comparison with industrial data proves that the ref-GR model predicts better than the original Graaf model as in Graaf et al. (1988), while the VBF, Vanden Bussche and Froment (1997), tends to overestimate methanol production. The validation exploits industrial data published in the literature.

Keywords: kinetics comparison, methanol reactor, process simulation, industrial comparative case studies, Lurgi and ICI technologies

1. Introduction

As highlighted in recent publications by Bisotti (2021), and Bozzano and Manenti (2016), the methanol molecule has been gaining increasing interest for twenty years as in Olah's wishes for the methanol economy anticipating the energy transition period. Although the technology is shifting towards milder operating conditions decreasing pressure and optimizing catalyst formulation, the kinetic steps, species role and interaction with active sites, and consequently modelling appear to lack a comprehensive catalytic path description for CO and CO₂ hydrogenation over CZA catalysts. The different kinetic models (Table 1) reflect this fragmented framework where carbon source and kinetic scheme are not uniform. The methanol synthesis is essentially limited to three different reactions: CO and CO₂ hydrogenation, and (reverse) water-gas shift reaction. Looking at

the table, with few exceptions, these models were proposed more than 20 years ago, hence, a kinetic parameters robust refit would be beneficial as demonstrated in Bisotti et al. (2021) for the or-GR model. They decided to refit the Graaf model since several authors use regardless Graaf's and VBF's models showing good accuracy with industrial and/or experimental observations. Hence, or-GR and VBF models are considered as industrial benchmark for methanol synthesis modelling, as proven in Mäyrä and Leiviskä (2018). In this work, adopting the ref-GR model, we would show the impact and the different predictions in methanol synthesis for or-GR, ref-GR and VBF models furtherly corroborating observations/conclusions proposed in previous work, Bisotti et al. (2021).

Table 1 – Main kinetic models available in the literature.

Model (year)	Source	Model reactions
Villa et al. (1985)	CO	$\text{CO} + 2\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH}$ $\text{CO} + \text{H}_2\text{O} \rightleftharpoons \text{CO}_2 + \text{H}_2$
Klier et al (1982) McNeil et al. (1989) Ma et al. (2009)	CO and CO ₂	$\text{CO} + 2\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH}$ $\text{CO}_2 + 3\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH} + \text{H}_2\text{O}$
Graaf et al. (1988) Park et al. (2014) Seidel et al. (2018) Slotboom (2020) Bisotti et al. (2021)		$\text{CO} + 2\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH}$ $\text{CO}_2 + 3\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH} + \text{H}_2\text{O}$ $\text{CO} + \text{H}_2\text{O} \rightleftharpoons \text{CO}_2 + \text{H}_2$
Skrzypek et al. (1991) Askgaard et al. (1995) Vanden Bussche-Froment (1996) Kubota et al. (2001)	CO ₂	$\text{CO}_2 + 3\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH} + \text{H}_2\text{O}$ $\text{CO}_2 + \text{H}_2 \rightleftharpoons \text{CO} + \text{H}_2\text{O}$

2. Methods

The in-silico assessment consists of two different steps: (1) comparison of the methanol production and reactants conversions using three different kinetic models, syngas quality, and operating conditions; (2) industrial case studies. For the first task, the analysis is performed in Aspen Hysys[®] V11 using a PFR followed by a cooling step (up to 25°C) and flashing unit where liquid products and light gas are separated. The recovered syngas is then recycled back to the reactor feed after a re-compression as in Figure 1. The fresh syngas make-up and recycled syngas temperatures are set to 225°C. Pressure drop is estimated using the Ergun model already implemented in Aspen Hysys[®]. The PFR length is subdivided into 100 segments (the default value is 20). The recycle function settings impose low errors in the molar flow and stream composition, the adjust guarantees the residual error for the feed stream (FEED) lower than 0.02%. Since, we aim at comparing kinetic models for industrial applications, we choose conventional feed compositions or at least close to industrial among the one proposed in Leonzio (2020). For the second task, reactor design/configuration and flowsheet are assigned in accordance with the analyzed industrial technologies (i.e., Lurgi and ICI). For the industrial plant data, we refer to Chen et al. (2011) and Froment et al. (1995) which report the process schemes and equipment details for the Lurgi and ICI technologies, respectively. We focused our attention on such technologies since they cover almost 90% the installed reactor technology worldwide, as proven in Bozzano and Manenti (2016).

3. Sensitivity analysis to the feed composition and pressure

The sensitivity analysis aims at comparing the methanol production using the three kinetics under different operating conditions: 50-75 bar is a typical operating window for the MegaMethanol technology and 70-90 bar range is a standard in other technologies such as Lurgi and ICI. To further test the kinetics, different feed compositions are considered. Such compositions are petro-syngas, COG, and bio-syngas in Leonzio (2020) and reported also in Bisotti et al. (2021). The petro-syngas is representative for conventional syngas from natural gas (i.e., steam methane reforming), the COG for syngas from coke gasification that in China is still the main carbon source to chemicals, and bio-syngas for emerging biosources such as biogas and biomasses. The implemented Lurgi reactor specifications, including feed temperature, are the same reported in Chen et al. (2011). As purge ratio, we assumed that 10% of the recycle loop is removed. Figure 1 depicts the implemented reactor flowsheet, while Figure 2 gathers the simulations results.

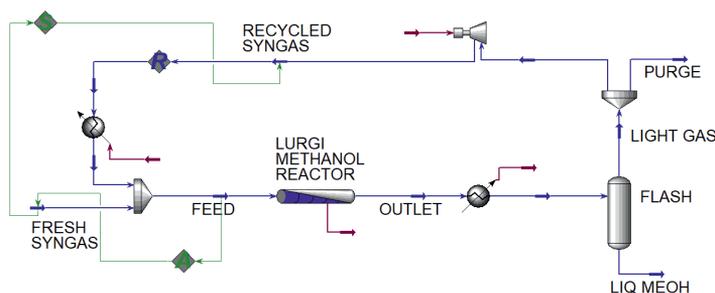


Figure 1 - methanol synthesis complete configuration.

The results show that the ref-GR is like the VBF model in predicting the methanol production and reactant conversions meaning that the robust refit procedure modified the kinetic parameters forcing the or-GR to move closer to the VBF. Since the or-GR tends to underestimate the methanol productivity, the refit induced an increment in the CO₂ and H₂ conversion as depicted in Figure 2 (2A)-(2C). Furtherly, it is noticeable that VBF and ref-GR predicts larger methanol production with regards to the or-GR which never overcomes 3 ton/h. This is one of the main Graaf's model shortcoming already emerged. Moreover, the ref-GR exhibits a strong pressure dependence in the methanol production, on average moving from 50 bar to 90 bar there is +35% increment, while or-GR and VBF is almost flat regardless the operating pressure. Generally, the ref-GR behavior is comprised the VBF and or-GR predictions (except for the petro-syngas feedstock). These have been already discussed in Bisotti et al. (2021). However, considering the order of magnitude of the simulation results is possible to state that the ref-GR and VBF predictions are quite similar both in terms of methanol production (4.5 ton/h vs 3.7 ton/h respectively). Instead, it is quite evident that the discrepancies are magnified looking at COG and bio-syngas feedstocks. Under such conditions, using the VBF model, the methanol production is one order of magnitude larger than the one predicted with ref-GR as in Figure 2 (1B)-(1C). Finally, the reactants conversions point out that the VBF implies larger H₂ and CO₂ consumptions. This is due to: (1) the VBF tends to overestimate the methanol production as proven in Bisotti et al. (2021) and (2) the CO₂ is the only direct hydrogenation path considered, hence, as expected, higher CO₂ conversion are strictly related to the suppressed CO hydrogenation which push the water-gas-shift equilibrium towards CO₂ production which is then converted into methanol.

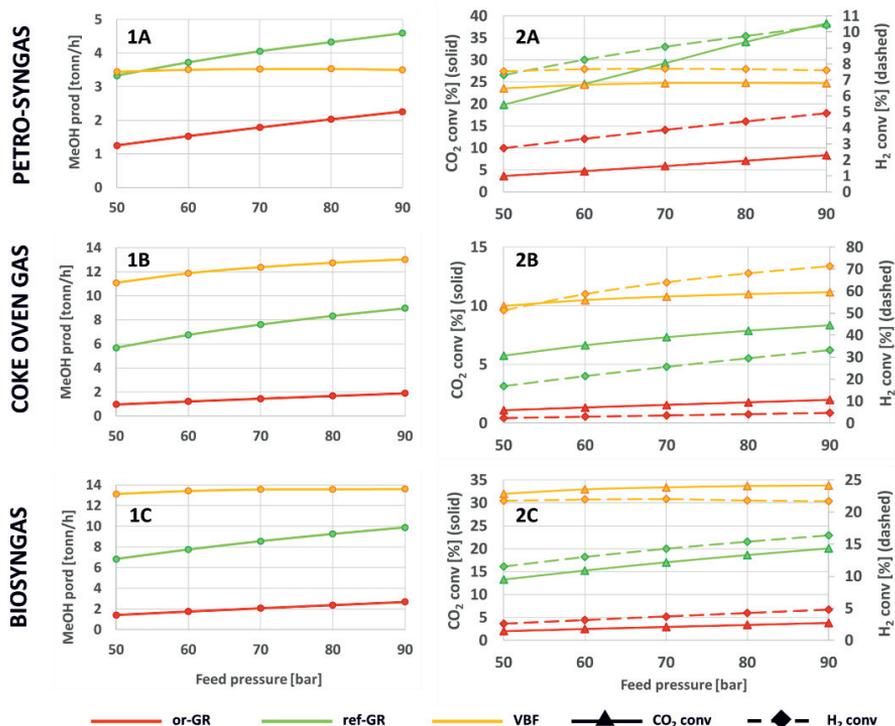


Figure 2 - sensitivity analysis results: (1) methanol production and (2) CO₂ and H₂ conversions for different feedstocks (A) petro-syngas, (B) COG, and (C) bio-syngas. For the conversion chart CO₂ conversion is the solid line (left y axis) and H₂ conversion is the dashed one (right y axis).

4. Industrial case studies

The industrial case studies include the Lurgi Boiling Water Reactor (BWR) and the fixed adiabatic beds gas-quenched reactor (formerly ICI, nowadays Johnson Matthey). The Lurgi and ICI technologies are described in Chen et al. (2011) and Froment et al. (1995) and depicted in Figure 3 and Figure 4 respectively. The cited works report further details such as the reactor specifications, feed composition, and operating conditions. Simulations results are graphically depicted in Figure 5 (BWR Lurgi) and Figure 6 (ICI).

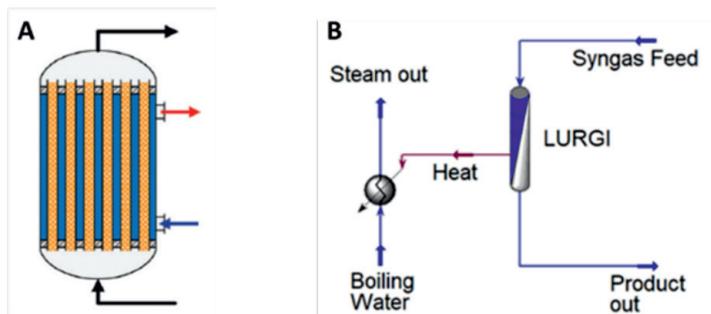


Figure 3 - Lurgi BWR technology: (A) reactor configuration and (B) simulation flowsheet

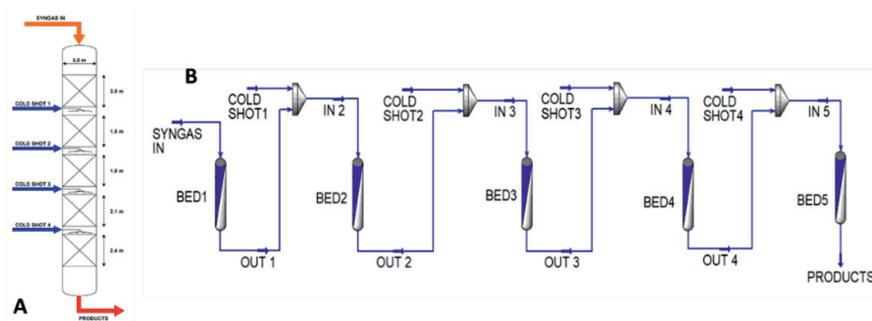


Figure 4 - ICI technology: (A) reactor configuration and (B) simulation flowsheet

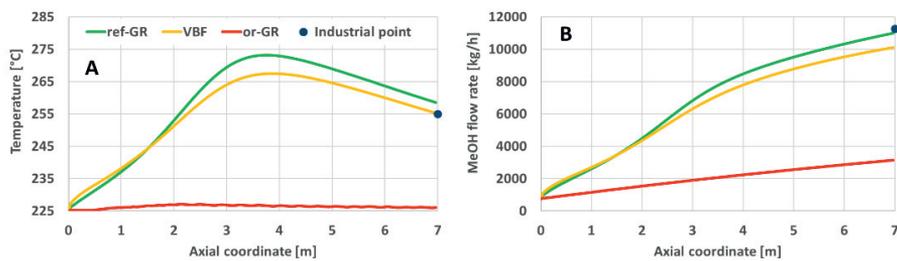


Figure 5 - Lurgi BWR reactor results: (A) temperature profile and (B) methanol production using or-GR (red solid line), ref-GR (green), and VBF (yellow). The blue dots are the experimental data.

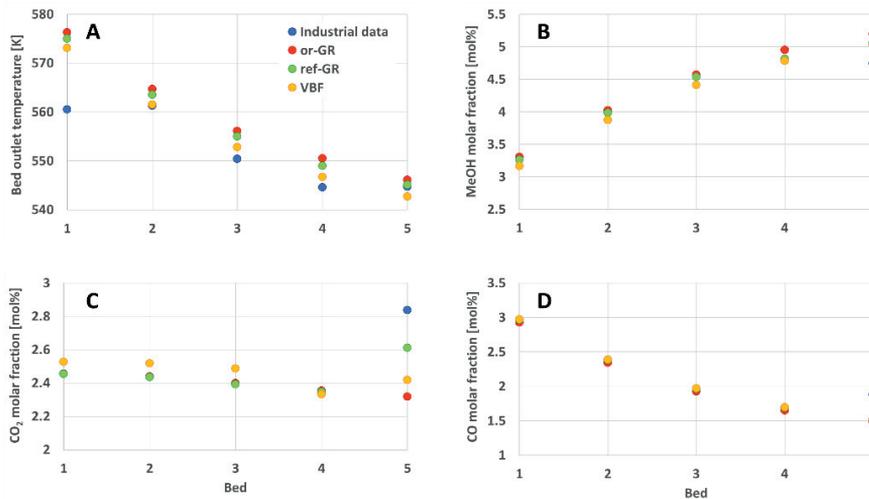


Figure 6 - ICI reactor results: (A) bed outlet temperatures, (B) methanol, (C) CO₂, and (D) CO molar fractions in the outlet mixture using ref-GR (green dots), or-GR (red) and VBF (yellow). The blue dots are the industrial data.

The industrial case studies results show that the ref-GR kinetic better predicts the methanol production in the Lurgi reactor while the predicted outlet temperature is close to the industrial value. Similarly, also for the ICI technology it is possible to appreciate that the ref-GR is based on an accurate kinetic model since it enables to catch both methanol and CO_x molar fractions, and the outlet temperatures. Specifically, VBF and ref-GR have similar deviations with regards to the industrial data, however, the ref-GR model is the only one catching the final CO₂ amount. In Figure 6, the hydrogen content is not reported since industrial data are not available. Concerning the ICI technology neither of the considered kinetic models can correctly predict the products mixture and outlet temperature values for the first and second adiabatic beds. This may be due to the presence of mass transfer limitation which may occur where temperature is larger than 275°C. It is not possible to characterize this aspect using simulation software such as Aspen Hysys. The VBF model presents similar accuracy even though it is not as accurate as the ref-GR in predicting the methanol production, for instance, in the Lurgi technology VBF underestimates the methanol production for almost 1.0 ton/h. The or-GR exhibits the worst accuracy meaning that it does appear as a suitable kinetic model for the methanol synthesis reactor simulation and validation.

5. Conclusions

This work demonstrated that kinetics affect the reactor predictions. Furtherly, it proved that in process simulation environment (Aspen Hysys[®]) the or-GR is not a suitable model to properly design the methanol synthesis reactor. Conversely, the ref-GR guarantees accuracy, and it is more precise of the VBF in the analyzed case studies. Hence, it is possible to state that ref-GR solved the shortcomings emerged in the or-GR and it is a potential candidate as reliable alternative for industrial and simulation purposes.

References

- A.M. Al-Fadli, M.A. Soliman, G.F. Froment, 1995, Steady State Simulation of a Multi-Bed Adiabatic Reactor for Methanol Production, *Journal of King Saud University-Engineering Sciences*, 7 (2), 101–132, [https://doi.org/10.1016/S1018-3639\(18\)31053-5](https://doi.org/10.1016/S1018-3639(18)31053-5)
- F. Bisotti, M. Fedeli, K. Prifti, A. Galeazzi, A. Dell'Angelo, M. Barbieri, C. Pirola, G. Bozzano, F. Manenti, 2021, Century of Technology Trends in Methanol Synthesis: Any Need for Kinetics Refitting?, *Industrial Engineering & Chemistry Research*, 60 (44), 16032–16053, <https://doi.org/10.1021/acs.iecr.1c02877>
- G. Bozzano, F. Manenti, 2016, Efficient methanol synthesis: Perspectives, technologies and optimization strategies, *Progress in Energy and Combustion Science*, 56, 71-105, <https://doi.org/10.1016/j.peccs.2016.06.001>
- L. Chen, Q. Jiang, Z. Song, D. Posarac, 2011, Optimization of Methanol Yield from a Lurgi Reactor, *Chemical Engineering & Technology*, 34(5), 817-822, <https://doi.org/10.1002/ceat.201000282>
- G.H. Graaf, E.J. Stamhuis, A.A.C.M. Beenackers, 1988, Kinetics of Low-Pressure Methanol Synthesis, *Chemical Engineering Science*, 43 (12), 3185–3195, [https://doi.org/10.1016/0009-2509\(88\)85127-3](https://doi.org/10.1016/0009-2509(88)85127-3)
- G. Leonzio, 2020, Mathematical modelling of a methanol reactor by using different kinetic models, *Journal of Industrial and Engineering Chemistry*, 85, 130-140, <https://doi.org/10.1016/j.jiec.2020.01.033>
- O. Mäyrä, K. Leiviskä, 2018, Chapter 17 - Modeling in Methanol Synthesis; in *Methanol: Science and Engineering*, Editors A. Basile, F. Dalena; Elsevier, 475–492, <https://doi.org/10.1016/B978-0-444-63903-5.00017-0>
- K.M. Vanden Bussche, G.F. Froment, 1996, A Steady-State Kinetic Model for Methanol Synthesis and the Water Gas Shift Reaction on a Commercial Cu/ZnO/Al₂O₃Catalyst, *Journal of Catalysis*, 161 (1), 1–10, <https://doi.org/10.1006/jcat.1996.015>

Optimal layout of modular multi-floor process plants using MILP

P Wrigley,^a P Wood,^a S O'Neill^a, R Hall^b, S Marr^b, D Robertson^c

^a*School of Computing and Engineering, University of Derby, DE13HD, UK*

^b*Nuclear Advanced Manufacturing Research Centre, University of Sheffield, UK*

^c*Rolls-Royce Plc, UK*

p.wrigley@derby.ac.uk

Abstract

Off-Site Modular Construction (OSMC) research has been a growing research area over the past two decades driven by low productivity levels in construction (Bock 2015). Some of the chemical plant industry has started analyzing small modular chemical plants. Some work has been achieved in automating earlier parts of the small modular chemical plant design process such as automated database creation and selection of equipment. However, a layout optimization methodology has not been applied to OSMC industrial process plants. Plant layout is an important step in the plant design process (Moran 2017). Methods have been made to proposed to optimize the plant for different requirements. Requirements have been modeled in a mathematical programming model as constraints such as connectivity, pumping, safety, pipe routing, multi-floor arrangements, etc., to establish the optimal cost and safe plant layouts.

This paper proposes to develop and utilize a MILP mathematical layout optimization model to help design and construct modular process plant. The main considerations are the module sizes for transport requirements and factory handling. Data from previous research was utilized and run through the new modular optimization model. The previous research layout results were compared to the new modular layouts process plant optimization to compare how modularization may affect the design of industrial process plants. The results demonstrate that building a plant in road transportable, factory built could enable equipment to be located closer together due to advanced factory manufacturing processes as assembly and tools are more accessible than building stick built plants.

Keywords: MILP, mathematical layout optimization, off site modular construction, industrial process plants,

1. Introduction

Off-Site Modular Construction (OSMC) research has been a growing research area over the past two decades driven by low productivity levels in construction (Bock 2015). Productivity is higher in factories when compared to a stick-built site due easier access to superior tools, methods and learning. This has spawned the development of small, factory built, rapidly deployable and flexible process plants (Seifert et al. 2012) to take advantage of the gains in OSMC productivity. This is a rapidly growing area in Chemical process plant research (Bielenberg and Palou-Rivera 2019). Research has shown that OSMC can provide 20% savings in cost and up to 50% savings in scheduling, providing reductions in risk and finance (Mignacca et al. 2018). The same work performed in a factory may be

8x more efficient and cost effective than performing work in situ according to a study in shipbuilding (Barry 2009).

Plant layout is an important step in the plant design process (Moran 2017) where requirements are taken into consideration for layouts. Requirements can take the form of process flow connectivity, pumping, safety, pipe routing, multi-floor arrangements, etc., to establish the optimal cost and safe plant layouts. Three core approaches have been applied to optimising costs of plant layouts: Heuristics and Metaheuristics, Mathematical optimisation along with Rule based expert systems. Some work has been achieved in automating earlier parts of the small modular chemical plant design process such as automated database creation and selection of equipment, however, a layout optimization methodology has not been applied to OSMC industrial process plants (Eilermann et al. 2018).

In plant layout literature, heuristics were first utilised in the plant design problem to arrange equipment. Mathematical approaches were then developed using Mixed integer (Nonlinear) Programming (MINLP) and (MILP) and techniques (MINLP) or graph partitioning (Ejeh, Liu, and Papageorgiou 2019b) for pumping and floor construction costs, financial risk and safety. Pipe routing and other safety methods were developed (Ejeh, Liu, and Papageorgiou 2019b) along with methods to solve larger problems. Ejeh et al., expanded the optimisation model with considerations for tall equipment items that span across floors and the availability of predefined production sections (Ejeh, Liu, and Papageorgiou 2018) and an updated model accounting for fire and explosion risk (Ejeh, Liu, and Papageorgiou 2019b). In another work, three improvements are added over the previous methods: an extension for multi-floor equipment items to extend above the maximum possible number of levels, the choice of an available number of floors fewer than the maximum amount essential for any equipment item and Integer cuts to improve the efficiency (Ejeh, Liu, and Papageorgiou 2019a).

This work builds upon the previous work of (Patsiatzis and Papageorgiou 2002) by extending the problem to consider process plant systems for OSMC and road transportable modules.

2. Problem Description

This work aims to obtain the optimal system layout for equipment in modules constrained for OSMC and transport. The factory-built road transportable requirement for this work is a key requirement. Requirements for EU transport are outlined in Table 1, (Barrot 2019).

	No permit (1)	Long term permit (2)	Corridor (3) (except Germany)
Width	3 m	3,5 m	4,5 m
Overall length	24 m	30 m	40 m
Overall height	Directive 96/53/EC	4,2 m	4,4 m
Weight	Directive 96/53/EC	80 tonnes	100 tonnes
Axle load: Beam axle	Directive 96/53/EC 12 tonnes	12 tonnes 15 tonnes	12 tonnes 15 tonnes

Table 1 – EU road transport requirements (Barrot 2019)

In this section, the extension of the optimisation model will be considered with the addition of module cost. An outline of the modular process plant layout problem and the fundamental assumptions are provided in this section for the mathematical model formulation. If equipment items extend above the available module height, these are identified a suitable vertical module and located outside the module stack. Equipment items are assumed as a cuboid. The geometrical centres are utilized for rectilinear distances between equipment in the x-y axis. A position on the equipment height is determined for the vertical distances/connections. Every equipment item is permitted to rotate in 90 angles in the x-y plane unless constrained. Equipment items could be built on platforms to increase the work off site. A trade off study would be required to assess these criteria. The problem description is as follows

Provided to the Model:

- A set of N process plant equipment items, i, j, size: (width, height, depth)
- Directed connections between items
- Module size (width, height, depth)
- NF number of modules
- a set of K available modules for layout with module height, MH;
- Connection points height on items h
- Costs of connections and modules connection, Cci j, pumping (horizontal, Chi j, and vertical, Cvi j), land purchase (LC) and construction (FC1, FC2) cost data;
- Additional margin space between items
- Module Positions (fixed)
 - i j, pumping (horizontal, Ch
 - i j, and vertical, Cv
 - i j), land purchase (LC) and construction (FC1, FC2) cost data;

Determined by the Model

- Positions of items
 - Rotations of items
 - Number of modules used
 - Cost of connections
 - floor area;
- so as to minimise the total plant layout cost

2.1. Model Formulation

Binary variables V_{ik} and Z_{ik} , Integer variable NF, and Parameter H from (Patsiatzis and Papageorgiou 2002) are altered to represent modules rather than floors.

Assuming construction and piping installation costs can be reduced 20% by constructing modules in factories (Barry 2009)(Mignacca et al. 2018), FC1&FC2 and connection costs C_{ij} are reduced by 20%.

2.2. Non overlapping of Items with module bounds

A lower and upper bound must be imposed on the equipment coordinates to ensure they remain within the module dimensions depending on its current rotation. Assuming that all modules are positioned at the origin and distances are calculated by taking their relative position into account:

$$x_v + l_i \leq W_k + M(1 - m_k), \quad \forall v \in V, \forall k \in K \quad \text{Equation 1}$$

$$y_v + d_i \leq H_k + M(1 - m_k), \forall v \in V, \forall k \in K \quad \text{Equation 2}$$

Considering if items can move in the z axis:

$$z_v + dv \leq D_k + M(1 - m_k), \forall v \in V, \forall k \in K \quad \text{Equation 3}$$

and $(x_{uv}, y_{uv}, z_{uv}) = (0, 0, 1)$ respectively.

2.3. Rectilinear distance between connected items

The rectilinear distance in the x-axis between item i and j can be calculated by the absolute value of D_{xij} and is considered for the set of connected items, $f(i; j)$ 2 Eg, where D_{xij} is defined by:

$$R_{ij} - L_{ij} = x_i + \sum_{k \in K} V_{ik} m_{xk} - x_j - \sum_{k \in K} m_{jk} m_{xk}, \forall (i, j) \in f_{ij} \quad \text{Equation 4}$$

$$A_{ij} - B_{ij} = y_i + \sum_{k \in K} m_{ik} m_{yk} - y_j - \sum_{k \in K} m_{jk} m_{yk}, \forall (i, j) \in f_{ij} \quad \text{Equation 5}$$

Considering the rectilinear in the z-axis:

$$U_{ij} - D_{ij} = [z_i + \sum_{k \in K} m_{ik} m_{zk}] - [z_j + \sum_{k \in K} m_{jk} m_{zk}], \forall (i, j) \in f_{ij} \quad \text{Equation 6}$$

2.4. Objective function

Given that the unit cost of a connection between i and j is c_{ij} and the cost of a module is given by g_k , minimise the following objective function,

$$\begin{aligned} Of = & \sum_{(i,j) \in E} c_{ij} TD_{ij} + C_{ij}^v D_{ij} + C_{ij}^h (R_{ij} + L_{ij} + A_{ij} + B_{ij}) + C1 \cdot NF + FC2 \cdot NF \cdot FA \\ & + LC \cdot FA \end{aligned} \quad \text{Equation 7}$$

All continuous variables in the formulation are defined as non-negative. Total layout cost (eq. 21), subject to floor constraints (eqs. 1 - 3, 5 - 9), multi-floor equipment constraint (eq. 4), distance constraints (eqs. 10 - 14), area constraints (eqs. 15 - 20), equipment orientation, non-overlapping and layout constraints

3. Results and Discussion

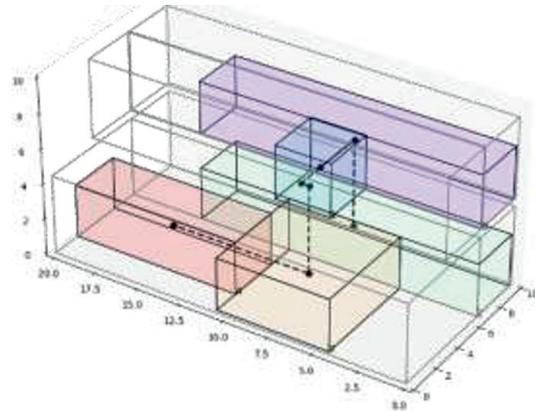
For the case study, the five-unit instant coffee process (Patsiatzis and Papageorgiou 2002) was chosen.

3.1. 2 Module stack

The results for the 3 layer module stack can be seen in Figure 1.

	x	y	z
Percolator	7.9	7.85	5
Cyclone	7.9	4.7	5
Spray Dryer	7.9	7.85	0
Press	7.45	3.15	0
Drier	15.3	3.15	0

Figure 1 - Locations of equipment for 2 floor module



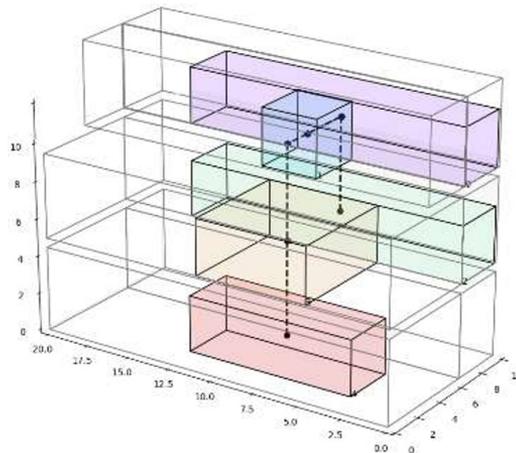
Assuming construction and piping installation costs can be reduced 20% by constructing modules in factories (Barry 2009)(Mignacca et al. 2018), construction of the module stack (FC1&FC2) reduces to 26624rmu and connection costs decrease to 12374rmu. The total plant cost is 77184rmu a 6.3% decrease compared to the original stick-built site. Land costs stay the same at 13340rmu. However, pumping costs increase to 24846rmu a 3.6% increase of the total original costs of 82366rmu.

3.2. 3 Module stack

The results for the 3-layer module stack can be seen in Figure 2.

	x	y	z
Percolator	7.9	7.85	10
Cyclone	7.9	4.7	10
Spray Dryer	7.9	7.85	5
Press	7.9	3.15	5
Drier	7.9	3.15	0

Figure 2 - Locations of equipment for 3 floor modules



Again, assuming 20% reduction in construction and piping installation, construction of the module stack (FC1&FC2) increases to 39936rmu. Connection costs are 13863rmu. In this instance, pumping costs decrease to 12854rmu, a 41% decrease on the original pumping costs of 21909rmu. Land costs stay the same at 13340rmu. The total plant cost is 79992rmu, a 2.9% decrease compared to the original stick-built site.

Although the construction costs have increased by 13312rmu (a 50% increase) the reduction in pumping costs is 11992rmu. The operators of this plant would therefore see significant savings if running the plant for over a year.

4. Conclusions

A layout optimisation model was introduced to consider off site manufacture and transport of process plant systems. Although operational pumping costs may increase slightly due to the requirement to locate equipment items in modules (more constrained), construction costs for structures and piping costs can be decreased. The results showed that this reduction provides an overall benefit when comparing OSMC to stick built process plants on the smaller scale plants analysed here.

Future work would be to increase the process plant system size to see how it copes with larger models. A more detailed analysis is required on the construction of modules, this would require the involvement and collaboration with civil engineering and infrastructure experts.

Acknowledgement

This work was sponsored by Rolls-Royce Plc.

References

- Barrot, Jacques. 2019. *European Best Practice Guidelines for Abnormal Road Transports*. EU. https://ec.europa.eu/transport/road_safety/sites/default/files/vehicles/doc/abnormal_transport_guidelines_en.pdf.
- Barry, K. 2009. "Modularization of Equipment for New Nuclear Applications (1021178)." <https://www.epri.com/research/products/1021178>.
- Bielenberg, James, and Ignasi Palou-Rivera. 2019. "The RAPID Manufacturing Institute – Reenergizing US Efforts in Process Intensification and Modular Chemical Processing." *Chemical Engineering and Processing - Process Intensification* 138 (February): 49–54. <https://doi.org/10.1016/j.cep.2019.02.008>.
- Bock, Thomas. 2015. "The Future of Construction Automation: Technological Disruption and the Upcoming Ubiquity of Robotics." *Automation in Construction* 59: 113–21. <https://doi.org/10.1016/j.autcon.2015.07.022>.
- Eilermann, Martin, Christian Post, Heiko Radatz, Christian Bramsiepe, and Gerhard Schembecker. 2018. "A General Approach to Module-Based Plant Design." *Chemical Engineering Research and Design* 137: 125–40. <https://doi.org/10.1016/j.cherd.2018.06.039>.
- Ejeh, Jude O., Songsong Liu, and Lazaros G. Papageorgiou. 2018. "Optimal Multi-Floor Process Plant Layout with Production Sections." *Chemical Engineering Research and Design* 137: 488–501. <https://doi.org/10.1016/j.cherd.2018.07.018>.
- . 2019a. "Optimal Layout of Multi-Floor Process Plants Using MILP." *Computers & Chemical Engineering* 131 (December): 13. <https://doi.org/10.1016/j.compchemeng.2019.106573>.
- Ejeh, Jude O, Songsong Liu, and Lazaros G Papageorgiou. 2019b. "An MILP Model for Safe Multi-Floor Process Plant Layout." In *Computer Aided Chemical Engineering*, 46:379–84. Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-818634-3.50064-3>.
- Mignacca, Benito, Giorgio Locatelli, Mahmoud Alaassar, and Diletta Colette Invernizzi. 2018. "We Never Built Small Modular Reactors SMRs but What Do We Know About Modularization in Construction." In *26th International Conference on Nuclear Engineering*. London: ASME.
- Moran, Seán. 2017. "Planning of Layout Activities." In *Process Plant Layout*. <https://doi.org/10.1016/b978-0-12-803355-5.00005-6>.
- Patsiatzis, Dimitrios I., and Lazaros G. Papageorgiou. 2002. "Optimal Multi-Floor Process Plant Layout." In *Computers and Chemical Engineering*, 26:575–83. [https://doi.org/10.1016/S0098-1354\(01\)00781-5](https://doi.org/10.1016/S0098-1354(01)00781-5).
- Seifert, Tim, Stefan Sievers, Christian Bramsiepe, and Gerhard Schembecker. 2012. "Small Scale, Modular and Continuous: A New Approach in Plant Design." *Chemical Engineering and Processing: Process Intensification* 52: 140–50. <https://doi.org/10.1016/j.cep.2011.10.007>.

Paving the way to multi-case optimization of a steam Rankine cycle for cogeneration in nuclear power plants

Guilherme Vescovi,^{a,b} Nicolas Alpy,^a David Haubensack,^a Catherine Azzaro-Pantel,^b Pascal Stouffs^c

^a*DES/IRENE/DER/SESI, Bat 1222, CEA Commissariat à l'Energie Atomique Cadarache, 13108 Saint-Paul-lez-Durance, France*

^b*Laboratoire de Génie Chimique, Université Toulouse, CNRS, INPT, UPS, 31029 Toulouse, France*

^c*Université de Pau et des Pays de l'Adour, E2S UPPA, LaTEP, 64000 Pau, France*

Abstract

The current global climate situation requires adequate energy technologies and systemic solutions to reduce CO₂ emissions to reach the carbon neutrality target by 2050. Cogeneration in nuclear power plants can provide a low-carbon source of process heat. These plants are in most cases optimized only for electricity production, which is the case for Pressurized Water Reactors in France. A significant asset of cogeneration units is that they can efficiently switch from electricity-only production to hybrid heat and power production. In this context, the objective of this work is to assess the benefits of performance optimization for a multi-case operation applied to Rankine cycle system design. The key for such an optimization is to minimize the exergetic losses that change from one operational mode to another. As a proof of concept of the multi-case optimization relevancy for a steam-water cycle, a seasonal variation of the heat sink temperature is considered as it induces several operating modes for the cycle, even for electricity-only production. The model of the system developed in the Modelica environment with the ThermoSysPro library is first presented. The formulation of the optimization problem involves dimensional parameters as optimization variables to maximize the global efficiency of the cycle. Three cases are then simulated: minimal condenser pressure, maximal condenser pressure and a seasonal variation profile of condenser pressure. Multi-case optimization allows improving the mean operating efficiency of the cycle in the considered heat-sink temperature range, compared to an optimization focused on a single operating point. The relative efficiency gain obtained for a narrow condenser temperature range is about 0.5 %. While the gain is modest, this demonstrates the interest for the concept of partial regimes modelling in support of a multi-case optimization, which should be rather emphasized for a cogenerating Rankine cycle, for which operational modes will be much more different. Further developments on the model and the study of a cogeneration case constitute a natural perspective of this work.

Keywords: Modelica simulation, Optimization, Cogeneration, Rankine cycle, CO₂ reduction.

1. Introduction

The energy transition has become a cornerstone of the global response to climate change. In order to avoid nature's irreversible turning point, extremely challenging goals are being set for reducing greenhouse gas emissions to zero by 2050, and limiting the rise in global temperatures to 1.5 °C (IEA, 2021). The role of nuclear energy is very relevant in that context, since nuclear power plants (NPP) do not emit CO₂ during electricity production. Alongside with the growth of renewables, which have an intrinsic variability, the need for flexibility in power systems increases. The combined production of electricity and heat in power systems contributes to grid stability and security, dealing with supply-demand fluctuations and load following transients. The heat produced during a decrease in electricity supply has several potential applications depending on the available temperature levels, such as hydrogen production, thermal storage (for further shifted electricity production when the demand becomes higher), and district heating. Hence, cogeneration in power plants has a key role in decarbonisation of the energy systems in the long term (Taibi et al. 2018).

In this work, a seasonal temperature variation is used as a proof of concept for the multi-case optimization approach, which could be useful in the design of cogenerating systems. The dimensional characteristics of the cycle are optimized considering the whole temperature profile, instead of a single operation case with the extreme or average temperature values. The assessment of this case study will provide better insights in a real cogeneration case, with a larger range in temperature variations to which the cycle is submitted, so that more significant efficiency improvements could be achieved.

2. Rankine cycle model

The case study considered is a Nuclear Power Plant (NPP) equipped with a 540 MW Small Modular Reactor (SMR). In the secondary circuit, saturated steam exits the steam generator at 45 bar, is expanded in the turbine group, releasing heat through the condenser to the cooling circuit before being pumped back to the steam generator as liquid water. In the cooling system, sea water is pumped to the condenser, where it extracts heat and flows back to the sea in an open system configuration.

The secondary circuit of the NPPs is usually a Rankine cycle containing a reheater and several preheaters. However, the reference case for the present work contains only one preheater and no reheat. Even if this cycle is significantly simplified, this choice facilitates the calculations and numerical convergence of the model, and is not restrictive for the goal of this work.

Figure 1 shows the case study, which consists of the secondary circuit of the SMR-PWR. The system was modeled using Dymola software (Dassault Systèmes 2021) and the ThermoSysPro library (EDF 2021). These tools are capable of simulating system models comprising physical components from several engineering domains (Brück 2018).

The cycle in Figure 1 was designed with the standard ThermoSysPro components. In order to validate the Rankine model developed in Dymola, the CYCLOP (Cycle Optimization) tool was used (Haubensack, Thévenot, and Dumaz 2004). This tool is being developed by the CEA and its accuracy has already been proven with real operation data of NPPs. Table 1 shows the results obtained with Dymola, compared to CYLOP, for the steady-state design operation for two values of the condenser pressure, i.e. 50 and 100 mbar.

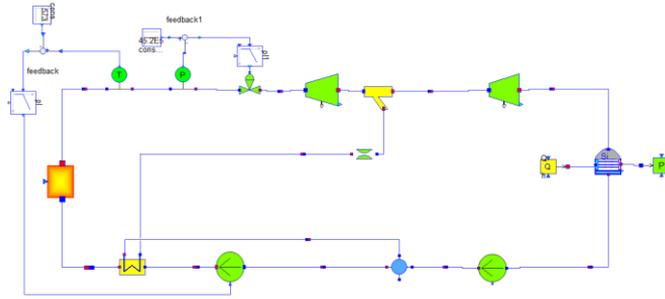


Figure 1: Rankine cycle model, designed with ThermoSysPro library in Dymola.

For these two design conditions, the results show that Dymola model is capable of obtaining optimized values very similar to CYCLOP results for different nominal conditions, hence validating a single case optimization process with Dymola.

Table 1: Validation of Dymola cycle in design conditions. Where η is the cycle efficiency, $T_{in,sg}$ is the steam generator inlet temperature, P_{ph} and Q_{ph} are the pressure and mass flow of steam entering the preheater, respectively. The relative error between cases is presented.

Design Value	50mbar			100mbar		
	CYCLOP	Dymola	Rel. Error	CYCLOP	Dymola	Rel. Error
η	0.3306	0.3306	0	0.3141	0.3140	-3.2E-04
$T_{in,sg}$ (°C)	150.00	150.01	6.7E-05	150.00	150.01	6.7E-05
P_{ph} (mbar)	5436	5436	0	5436	5437	1.8E-04
Q_{ph} (kg/s)	46.99	47.02	6.4E-04	42.77	42.79	4.7E-04

3. Seasonal temperature variation of cold source

In a water-cooled condenser, typical of a Pressurized Water Reactor (PWR), the pressure is defined by the available temperature of the cold source. Consequently, temperature variations of the cooling water in the vicinities of the NPP affect the electricity production, since the efficiency of the Rankine cycle (secondary circuit) is dependent on the condenser pressure. This effect is shown in Figure 2.

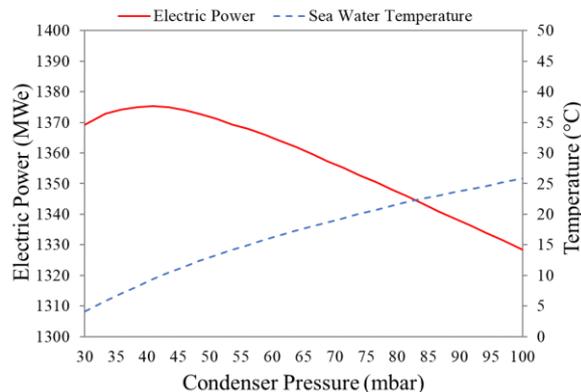


Figure 2: Electrical power production sensibility to cold source temperature. Adapted from (Grand 2014). Courtesy of the author.

As the condenser pressure decreases, due to a reduction in cooling water temperature, the production of electricity increases, as expected by the Carnot efficiency. However, there is an inflexion point beyond which electrical power begins to drop with a further decrease in condenser pressure. This effect is highly influenced by the increase in irreversibilities along the turbine during off-design operation. Indeed, the increase in condenser pressure causes an increase of turbine outlet swirl velocities, generating losses, which degrade cycle efficiency, reinforcing the Carnot tendency. In addition, the reduction in condenser pressure increases turbine outlet volumetric flow, which increases kinetic losses, contradicting the ideal Carnot cycle efficiency trend.

Temperature variations occur seasonally, and their amplitude depends on the geographic location of the plant. Table 2 shows the monthly temperature averages for sea water used as a cold source in the nuclear power plant located in Blayais, France.

Table 2: Average values of sea temperature, with corresponding condenser pressures, in Blayais, France (Weather Spark 2021).

Month	T_{sea} (°C)	P_{cond} (mbar)	Month	T_{sea} (°C)	P_{cond} (mbar)
January	11	44,99	July	20	73,88
February	11	44,99	August	21	77,91
March	11	44,99	September	20	73,88
April	13	50,38	October	17	62,85
May	15	56,32	November	15	56,32
June	18	66,36	December	13	50,38

The water used for cooling the three 900 MW reactors of this NPP undergoes a temperature variation from 11 to 21°C, with a yearly average of 15.4°C. In order to assess the electricity production efficiency and optimize the cycle for the seasonal temperature variation, the exhaust loss effect must be considered in the turbine model.

In order to account for off-design conditions, the turbine model available in ThermoSysPro library was then modified to include the exhaust loss as a function of last stage exit velocity (1) adapted from (Spencer, Cotton, and Cannon 1974). This loss has been chosen since it has considerable influence on the cycle efficiency in part-load operation.

$$\Delta h = 3.901 \cdot 10^{-8} V_s^4 - 4.515 \cdot 10^{-5} V_s^3 + 1.954 \cdot 10^{-2} V_s^2 - 3.447 V_s + 2.28 \cdot 10^2 \quad (1)$$

In this expression, Δh is the enthalpy loss due to the exhaust-pressure loss, and V_s is the velocity of steam leaving the last stage blades. In order to assess the operation at off-design conditions, a pressure variation was imposed to the condenser, in a cycle designed for a condenser pressure of 57.63 mbar. For comparison, the same cycle was re-optimized at every pressure with CYCLOP tool, which does not account for the irreversibility effects. The results from the Dymola simulations in Figure 3 include the inflexion point seen in Figure 2, which is influenced by the exhaust loss in the turbine. As expected, the cycles simulated at rated conditions with CYCLOP follow the Carnot efficiency trend, which increases as the condenser temperature (and pressure) decreases. These results support the Rankine model developed in Dymola regarding the operation in off-design conditions.

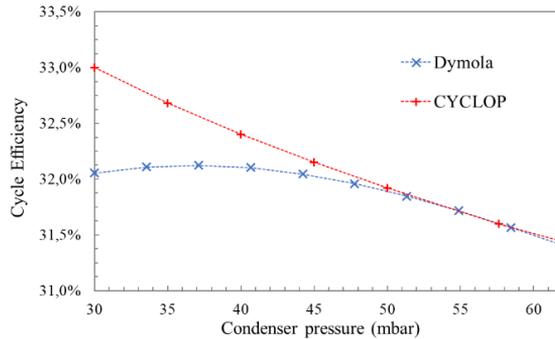


Figure 3: Efficiency of Rankine cycle with exhaust loss in non-rated conditions (Dymola) and the design efficiency of each condition (CYCLOP).

4. Optimization

The optimization module available in Dymola was then used to optimize the Rankine model for the multi-case scenario given by the temperature profile of Table 2. The Simplex method was used with six optimization variables, given by dimensional characteristics: the Stodola coefficients of the turbines (Cooke 1984), the exhaust area of the turbine’s last stage and a coefficient for pump head characteristic curve (for both the high-pressure and low-pressure sections). The multi-case optimization was compared to the performance of cycles optimized for the minimal and maximal temperatures (Figure 4).

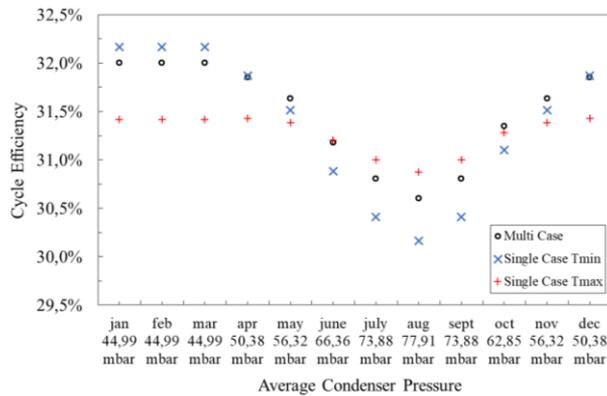


Figure 4: Multi-case optimization result for the Rankine cycle submitted to a seasonal variation of cold source temperature. A comparison is made with the cycle optimized for the minimal and maximal temperatures.

The multi-case results do not give the best efficiency values at minimal and maximal pressures. However, it increases the average efficiency of the cycle with respect to the single-case extreme-temperature optimizations. Table 3 shows the respective average efficiency.

The multi-case optimization presented the highest annual average efficiency for the Rankine cycle. This demonstrates the interest in considering the different operation scenarios during the optimization, instead of a single-case approach. The results obtained indicate that even if the cycle is optimized for the average temperature (of the seasonal

variation), the efficiency obtained is lower than the multi-case performance. This methodology allowed a relative gain of 0.41%, 0.67% and 0.06% with respect to the single case optimizations for the minimal, maximal and average temperatures, respectively.

Table 3: Average values of Rankine cycle efficiency for the different optimization approaches.

Average Efficiency (12 months)	Single Case	Single Case	Single Case	Multi
	Tmin	Tmax	Tavg	Case
	0,3135	0,3127	0,3146	0,3148

5. Conclusion

The presented work is a proof of concept of the methodological feasibility of the multi-case optimization approach. First, the case study was presented and the optimization was validated using two nominal conditions as a reference. The turbine model was then modified to include the exhaust loss effect, which has a considerable influence on cycle efficiency. A literature example was then used as a reference for supporting the behavior of the modified turbine model. For a seasonal temperature variation, relative efficiency gains of 0.41% and 0.67% were obtained through the multi-case optimization of the Rankine cycle, in comparison to the cycle optimized only for the minimal or maximal temperatures. When the cycle was optimized for the average value of the seasonal variation, the efficiency was still slightly lower than with the multi-case optimization result. In a cogeneration situation, with temperature variations that might be much more significant, this methodology should further improve the efficiency of combined heat and power production.

Acknowledgements

Authors would like to acknowledge CEA/DES/SMR/IDNES project, for its support.

References

- Brück, Dag. 2018. *Dymola Referential*. www.dymola.com.
- Cooke, D.H. 1984. "On Prediction of Off-Design Multistage Turbine Pressures by Stodola's Ellipse." *Proceedings of ASME*: 1–6.
- Dassault Systèmes. 2021. "Ingénierie Système DYMOLA Modélisation et Simulation Multi-Ingénierie Basées Sur Modelica et FMI." <https://www.3ds.com/fr/produits-et-services/catia/produits/dymola/> (October 12, 2021).
- Electricité de France. 2021. "ThermoSysPro." <https://thermosyspro.com/> (October 18, 2021).
- Grard, Hubert. 2014. EDP Sciences INSTN *Physique, Fonctionnement et Sécurité Des REP: Le Réacteur En Production*. Collection.
- Haubensack, David, Caroline Thévenot, and Patrick Dumaz. 2004. "The Copernic / Cyclop Computer Tool : Pre-Conceptual Design of Generation 4 Nuclear Systems." In *2nd International Topical Meeting on HIGH TEMPERATURE REACTOR TECHNOLOGY*, Beijing, 1–11.
- International Energy Agency. 2021. *Net Zero by 2050 A Roadmap for the Global Energy Sector*. www.iea.org/t&c/.
- Spencer, R. C., K C Cotton, and C. N. Cannon. 1974. *A Method for Predicting the Performance of Steam Turbine-Generators 16500 KW and Larger*.
- Taibi, Emanuele et al. 2018. *Power System Flexibility for the Energy Transition: Part 1, Overview for Policy Makers*. www.irena.org/publications.
- Weather Spark. 2021. "Climat et Moyennes Météorologiques Tout Au Long de l'année Pour Blaye." <https://fr.weatherspark.com/y/43659/Météo-moyenne-à-Blaye-France-tout-au-long-de-l'année> (October 19, 2021).

Dynamic modelling of non-isothermal open-cell foam catalyst packings: selective sugar hydrogenation to sugar alcohols as a case study

Catarina G. Braz,^a Ali Najarnezhadmashhadi,^{b,c} Vincenzo Russo,^d Kari Eränen,^b Henrique A. Matos,^a Tapio Salmi^b

^a*Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, PT-1049-001 Lisboa, Portugal*

^b*Laboratory of Industrial Chemistry and Reaction Engineering (TKR), Johan Gadolin Process Chemistry Centre (PCC), Åbo Akademi University, Aurum, Henriksgatan 2, FI-20500 Turku/Åbo, Finland*

^c*Department of Chemical Engineering, KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden*

^d*Chemical Sciences Department, Università degli Studi di Napoli ‘Federico II’, Complesso Universitario Monte S. Angelo, IT-80126 Napoli, Italy*
tapio.salmi@abo.fi

Abstract

A comprehensive multiphase model was developed for a trickle bed reactor with solid foam packings. Three-dimensional dynamic mass and energy balances in the three phases of heterogeneously catalysed reaction systems were implemented, and the mass and heat transfer resistances in the gas-liquid and liquid-solid phases and inside the pores of the catalyst were included in the model. Hydrogenation of arabinose and galactose mixtures on a ruthenium catalyst supported by carbon-coated aluminium foams was applied as an industrially relevant case study for the multiphase model. The kinetic parameters were estimated with confidence intervals within 10% error, indicating a good accuracy of the parameters, and the model results present a good adjustment to the experimental values. Finally, a sensitivity analysis on several model parameters demonstrated that the model could be applied to industrially sized reactors and various multiphase catalytic systems.

Keywords: Open-cell foam catalyst packing, Non-isothermal trickle bed reactor, Reaction kinetics, Mass transfer, gPROMS.

1. Introduction

In the context of increasing energy costs and future stringent environmental regulations for industrial production, structured catalysts play an essential role in designing more energy-efficient chemical reactors. In recent years, several advances in this field have been taken, with diversified structured catalysts being invented and studied in detail, such as monoliths, fibres, solid foams as well as structures prepared by 3D printing.

For three-phase catalytic systems (solid catalyst, gas phase, and liquid phase), open-cell foams have been investigated as suitable alternatives for catalytically active reactor packings because of their advantageous structural properties. The structures of pores and struts in open-cell foams provide high porosity (75-95%) and high specific surface area, allowing radial liquid flow and high local turbulence, which result in enhanced mass and

heat transfer. Axial and radial mixing are improved by the high pore tortuosity of solid foams, ensuring the even distribution of the gas and liquid phases, which is critical for a maximum product yield (Stemmet *et al.* (2008), Mohammed *et al.* (2014), Twigg *et al.* (2002)).

In this work, an advanced mathematical model of a three-phase catalytic tubular reactor was developed for solid foam packings and implemented in the *gPROMS[®] ModelBuilder 7.0* software. The two-dimensional gas, liquid and solid phase mass and energy balances include individual terms such as internal diffusion, gas-liquid and liquid-solid mass transfer, and intrinsic kinetics. Furthermore, the gas and liquid flows are described by axial and radial dispersion terms along with liquid hold-up and pressure drop expressions.

The hydrogenation of arabinose and galactose mixtures on a ruthenium catalyst supported by carbon-coated aluminium foams was applied as an industrially relevant case study. The kinetic parameters of the reaction were estimated using experimentally obtained concentrations of arabinose, galactose, arabitol and galactitol in a continuously operating tubular reactor (Figure 1) with a diameter of 11 mm and bed length of 33 mm, equipped with an open-cell aluminium foam with a pore density of 40 PPI. The experiments were conducted at varying operating conditions, with temperatures between 90 and 120 °C and arabinose-galactose molar ratios of 1:1 and 1:2. The experimental methods are explained in detail in a previous study by Najarnezhadmashhadi *et al.* (2020).

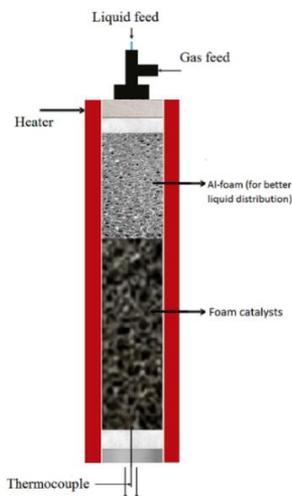


Figure 1 – Continuous tubular reactor system: trickle bed with cylindrical solid foam packing.

2. Model development

Modelling of trickle bed reactors is a very challenging and complex problem. In addition to the description of the flow pattern and the intrinsic reaction kinetics, mass and heat transfer effects must be accounted for accurately calculate the concentrations and temperatures inside the reactor tube. Heat and mass transfer resistances can appear at the gas-liquid and liquid-solid interfaces, as well as inside the pores of the catalyst layer. These resistances are more pronounced for lower fluid velocities (lower Reynolds, Sherwood, and Nusselt numbers) of the fluids inside the reactor. This is the case for many organic reactions applied in the production of fine and speciality chemicals and

ingredients in dietary products, where reactions are slow and long residence times are necessary to obtain high reactant conversions. On the other hand, internal mass and heat transfer resistances can be significant too, depending on the ratio between the reaction and diffusion rates, and, particularly, the thickness of the catalyst layer. Typically, internal mass and heat transfer resistances in the catalyst pores are negligible in structured reactors with thin catalyst layers such as 50 μm or less. However, these resistances might become prominent for rapid reactions and slow diffusion rates.

It was assumed that the reactions proceeded in the porous catalyst layer exclusively. Hence, the reactants diffuse first into the solid catalyst, after which they react on the active centre on the surface of the solid catalyst. Figure 2 illustrates the scheme of the system considered in the present work.

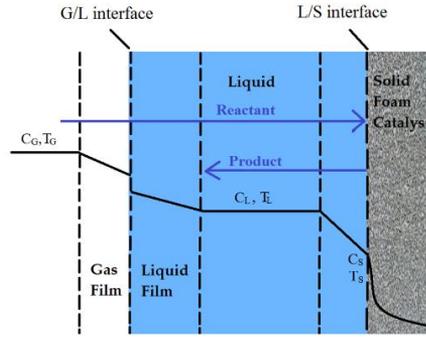


Figure 2 – Interaction of kinetic, mass and heat transfer effects in the three-phase system: gas, liquid, and solid foam catalyst. Adapted from Najarneshadimashhadi (2021).

2.1. Advanced model for continuous trickle bed reactors with open-cell foam packings

The model for the tubular reactor enables the calculation of the axial and radial concentration and temperatures profiles in the gas and liquid phases, as well as within the catalyst active layer by solving the dynamic mass and energy balances simultaneously for the gas, liquid, and solid phases. The model considers intrinsic kinetics, gas-liquid mass and heat transfer, internal diffusion inside the catalyst pores, heat conduction within the reactor and inside the catalyst pores, and the flow description by axial and radial dispersion. These essential features are combined with up-to-date correlations for liquid hold-up, mass transfer coefficients, axial and radial dispersion coefficients and pressure drop. The energy balances for the tubular trickle bed reactor are presented in equations (1)-(3). Because the catalyst layers are very thin ($< 10 \mu\text{m}$), and the reactor and particle coordinates are quite different in the model, a dimensionless coordinate, $x = r_p/R_c$, is used in the mass and heat balances of the solid phase, where R_c is the catalyst layer thickness, and r_p is the distance from the catalyst centre to its surface. The mass balance equations and the correlations implemented in the reactor model are described in detail by Najarneshadimashhadi (2021).

Gas phase

$$\begin{aligned} \frac{dT_G(t, z, r)}{dt} = & - \frac{h_{G,L} a (T_G(t, z, r) - T_L(t, z, r))}{\varepsilon_G \rho_G(z, r) C_{p,G}} - \frac{\partial(u_G(z, r) T_G(t, z, r))}{\partial z} + \left(\frac{\lambda_{z,G}}{\varepsilon_G \rho_G(z, r) C_{p,G}} + D_{z,G} \right) \frac{\partial^2 T_G(t, z, r)}{\partial z^2} \\ & + \left(\frac{\lambda_{r,G}}{\varepsilon_G \rho_G(z, r) C_{p,G}} + D_{r,G} \right) \frac{1}{r} \frac{\partial T_G(t, z, r)}{\partial r} + \left(\frac{\lambda_{r,G}}{\varepsilon_G \rho_G(z, r) C_{p,G}} + D_{r,G} \right) \frac{\partial^2 T_G(t, z, r)}{\partial r^2} \end{aligned} \quad (1)$$

Liquid phase

$$\begin{aligned} \frac{dT_L(t, z, r)}{dt} = & + \frac{h_{G,L} a (T_G(t, z, r) - T_L(t, z, r))}{\varepsilon_L \rho_L C_{p,L}} - \frac{u_L \partial T_L(t, z, r)}{\partial z} + \left(\frac{\lambda_{z,L}}{\varepsilon_L \rho_L C_{p,L}} + D_{z,L} \right) \frac{\partial^2 T_L(t, z, r)}{\partial z^2} \\ & + \left(\frac{\lambda_{r,L}}{\varepsilon_L \rho_L C_{p,L}} + D_{r,L} \right) \frac{1}{r} \frac{\partial T_L(t, z, r)}{\partial r} + \left(\frac{\lambda_{r,L}}{\varepsilon_L \rho_L C_{p,L}} + D_{r,L} \right) \frac{\partial^2 T_L(t, z, r)}{\partial r^2} - \frac{\lambda_p}{\varepsilon_L \rho_L C_{p,L} R_c^2} \left(\frac{\partial T_L(t, z, r, x)}{\partial x} \right) \Big|_{x=1} \end{aligned} \quad (2)$$

Solid catalyst phase

$$\frac{dT_S(t, z, r, x)}{dt} = \frac{\lambda_p}{\varepsilon_p \rho_p C_{p,p} R_c^2} \left(\frac{\partial^2 T_S(t, z, r, x)}{\partial x^2} + \frac{s}{x} \frac{\partial T_S(t, z, r, x)}{\partial x} \right) + \frac{1}{C_{p,p}} \sum r_j(t, z, r, x) (-\Delta H_r) \quad (3)$$

The boundary conditions for the energy balances are presented below. The closed-closed boundary conditions of Danckwerts were applied at the reactor inlet and outlet in gas and liquid phases. In the solid phase, the energy balance is coupled to the bulk liquid balance through the boundary condition, which states that the heat flux through the liquid film is equal to the heat flux in/out to/from the catalyst particle.

	Entrance	Outlet	Center	Wall
Gas phase	$T_G^in = T_G(t, r) \Big _{z=0} - \frac{D_{z,G}}{u_G(z, r)} \frac{\partial T_G(t, r)}{\partial z} \Big _{z=0}$	$\frac{\partial T_G(t, r)}{\partial z} \Big _{z=L} = 0$	$\frac{\partial T_G(t, z)}{\partial r} \Big _{r=0} = 0$	$\frac{\partial T_G(t, z)}{\partial r} \Big _{r=R} = 0$
Liquid phase	$T_L^in = T_L(t, r) \Big _{z=0} - \frac{D_{z,L}}{u_L} \frac{\partial T_L(t, r)}{\partial z} \Big _{z=0}$	$\frac{\partial T_L(t, r)}{\partial z} \Big _{z=L} = 0$	$\frac{\partial T_L(t, z)}{\partial r} \Big _{r=0} = 0$	$-\lambda_L \frac{\partial T_L(t, z)}{\partial r} \Big _{r=R} = h_w (T_L(t, z) \Big _{r=R} - T_w)$
	$T_S^in = T_S(t, r, x) \Big _{x=0}$	$\frac{\partial T_S(t, r, x)}{\partial z} \Big _{z=L} = 0$	$\frac{\partial T_S(t, z, x)}{\partial r} \Big _{r=0} = 0$	$-\lambda_S \frac{\partial T_S(t, z, x)}{\partial r} \Big _{r=R} = h_w (T_S(t, z, x) \Big _{r=R} - T_w)$
Solid phase	Particle center		Particle surface	
	$\frac{\partial T_S(t, z, r)}{\partial x} \Big _{x=0} = 0$		$\frac{\partial T_S(t, z, r)}{\partial x} \Big _{x=1} = \frac{h_{L,S} R_c}{\lambda_S} (T_L(t, r) - T_S(t, z, r) \Big _{x=1})$	

3. Modelling results and discussion

The dynamic model for trickle bed reactors equipped with cylindrical solid foam packings was applied to describe the catalytic hydrogenation of arabinose and galactose mixtures as a representative case study. The kinetic parameters of the highly selective arabinose (A) and galactose (G) hydrogenation to arabitol and galactitol are presented in Table 1. The activation energies and pre-exponential factors for the process were estimated with a 95 % confidence interval below 10 % error, indicating a good accuracy of the estimated parameters.

Table 1 – Parameter estimation results.

Parameters	Estimated Value	95 % Confidence interval
<i>Ea_A</i>	5.32×10^4	0.36×10^4
<i>Ea_G</i>	5.48×10^4	0.37×10^4
<i>k_{ref,A}</i>	5.24×10^{-7}	0.31×10^{-7}
<i>k_{ref,G}</i>	5.31×10^{-7}	0.30×10^{-7}

Figure 3 compares the experimental and calculated conversion values, of both arabinose and galactose, as a function of the reaction temperature. The model is able to predict the sugar conversions successfully, the average relative deviations being maximally about 16 %. The experimentally recorded sugar conversions displayed in the figures are averages from multiple samples. The selectivity toward sugar alcohols was always very high, typically exceeding 95%.

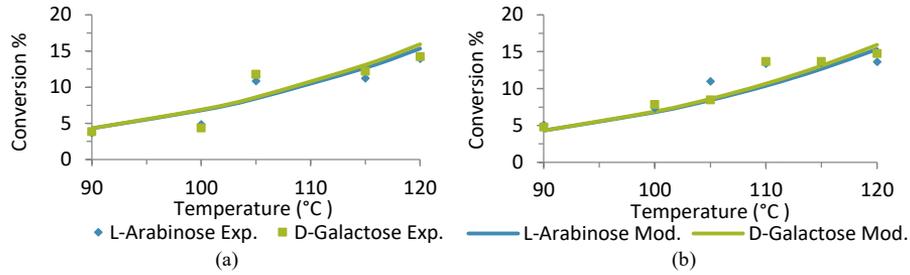


Figure 3 – Experimental and model values for the conversion of sugar mixtures as a function of the temperature. Inlet molar ratio (a) 1:1, (b) 1:2.

Figure 4 (a) and (b) present the arabinose concentration and the liquid-phase temperature axial and radial profiles for different reactor lengths. The conversion increases in line with the reactor length. The radial profiles in the conversion of arabinose were not observable because local conversions are very small. The temperature inside the reactor tube does not vary significantly because the reactor radius is very small. Higher temperature profiles are observed for higher reactor radiuses, Figure 4 (c).

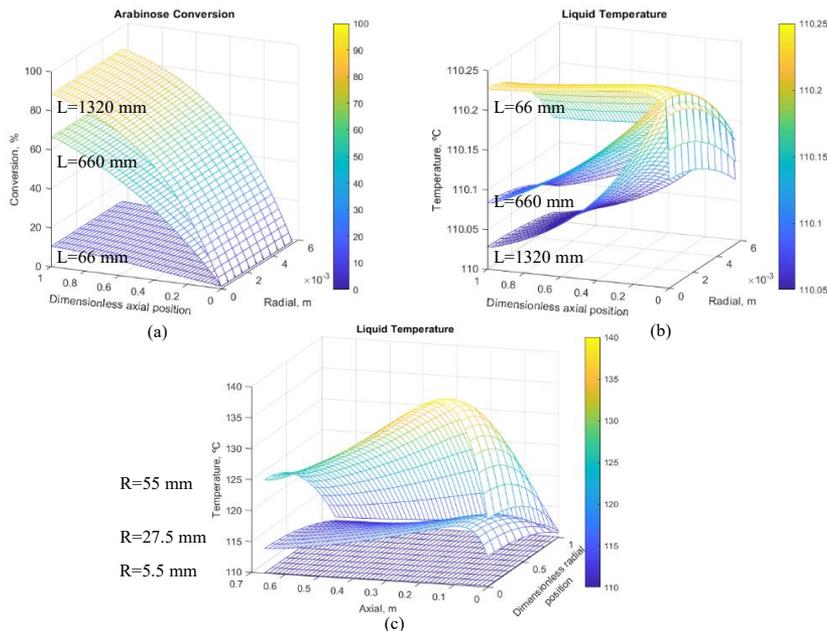


Figure 4 – Arabinose conversion (a) and liquid temperature (b) as a function of the reactor length, and liquid temperature as a function of the reactor radius for L=1320 mm (c).

The capabilities of the model proposed were tested through a sensitivity analysis on the effect of the catalyst layer thickness on the concentrations (galactose) and temperature profiles inside the catalyst pores, for $k_{ref,i} = 25 \times k_{ref,i}^{estimated}$. Analysing Figure 5, it is possible to conclude that, as the washcoat thickness increases, so do the concentration gradients and temperature in the catalyst layer. These results also demonstrate the flexibility of the generalised model, which can simulate both ideal cases, where there are no internal transfer limitations, and systems dominated by high intraparticle transport

limitations. These properties demonstrate that this model is suitable for use in other chemical systems.

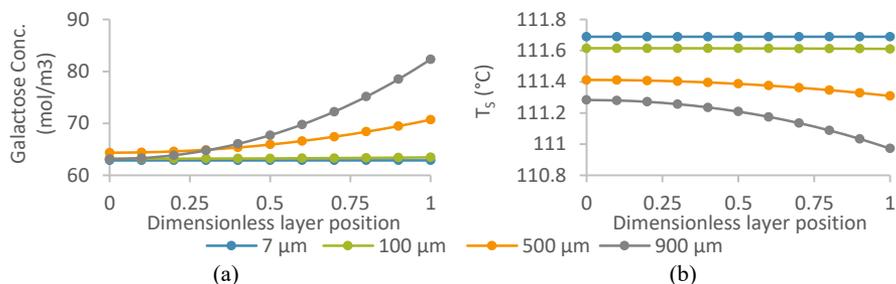


Figure 5 – (a) Galactose and (b) temperature profiles in the washcoat in the middle of the reactor ($z = 0.033\text{m}$), for $k_{ref,i} = 25 \times k_{ref,i}^{estimated}$

4. Conclusions

A comprehensive dynamic model for three-phase catalytic tubular reactors with solid foam packings considering all the main chemical and transport effects occurring in the system was developed and implemented successfully in the very modern software. The kinetic parameters of the model were estimated with good accuracy by non-linear regression analysis. The sensitivity analysis demonstrated that the model could predict the effect of different reactor dimensions, as well as the kinetic and transport phenomena included in the advanced multiphase reactor model. It is possible to conclude that this modelling approach can be applied to industrially sized reactors and various reactive multiphase catalytic systems, which are of fundamental and industrial interest.

Acknowledgements

The present work was financed by Finnish Cultural Foundation, Neste-Fortum Foundation and Rector of Åbo Akademi University (Ali Najarnezhadmashhadi), Academy of Finland, Academy Professor grant 319002 (Tapio Salmi) and CERENA, strategic project FCT-UIDB/04028/2020 (Henrique Matos and Catarina Braz). The economic support is gratefully acknowledged.

References

- I. Mohammed, T. Bauer, M. Schubert, R. Lange, 2014, Liquid–solid mass transfer in a tubular reactor with solid foam packings, *Chemical Engineering Science*, 108, 223–232.
- A. Najarnezhadmashhadi, K. Eränen, S. Engblom, A. Aho, D. Murzin, T. Salmi, 2020, Continuous Hydrogenation of Monomeric Sugars and Binary Sugar Mixtures on a Ruthenium Catalyst Supported by Carbon-Coated Open-Cell Aluminum Foam, *Industrial & Engineering Chemistry Research*, 59, 30, 13450–13459.
- A. Najarnezhadmashhadi, 2021, Development of structured catalyst and reactor technologies for biomass conversion- Continuous production of sugar alcohols, Doctoral thesis, Åbo Akademi University, Turku/Åbo, Finland.
- C.P. Stemmet, 2008, Gas-liquid solid foam reactors: hydrodynamics and mass transfer, PhD dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands.
- M.V. Twigg, J.T. Richardson, 2002, Theory and Applications of Ceramic Foam Catalysts, *Chemical Engineering Research and Design*, 80, 2, 183–189.

Optimising a wind farm with energy storage considering curtailment and uncertainties

Flora A. V. Biggins^a, Jude O. Ejeh^a, Diarmid Roberts^a, Aaron S. Yeardley^a and Solomon F. Brown^{a,*}

^a*Department of Chemical Biological Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom
s.f.brown@sheffield.ac.uk*

Abstract

In this work, we present a scenario-based stochastic optimisation (SBSO) model to schedule a wind farm with battery storage (BS) and a hydrogen electrolyser (HE) considering curtailment and uncertainties in generation and market prices. We compare cases with BS only, HE only, and a combination of the two. We apply Markov Chain (MC) and Gaussian Process (GP) techniques to generate wind curtailment and electricity price scenarios, respectively, capturing their inherent uncertainties. The model then assesses the economic benefits of incorporating BS and/or HE alongside wind generation and their scheduling as a function of curtailed and non-curtailed wind. The results can be used to determine the suitability of such systems for the purposes of maximising profits and making optimal use of curtailed generation. Results show that HE increases mean income and curtailed wind utilisation significantly more than BS. However, by combining the two, wind curtailment can be reduced by 95%.

Keywords: Wind farm, Battery storage, Hydrogen electrolysis, Curtailment, Stochastic optimisation

1. Introduction

Energy storage technologies (EST) can facilitate the decarbonisation of energy systems and lead to more sustainable futures. Battery storage (BS) has been found to improve power quality in electrical grids (Das et al., 2018) – particularly with high renewable energy penetration – and hydrogen storage (HS) can also replace fossil fuels in heating, industry and shipping (Gielen et al., 2019). Operating these technologies alongside renewables allows for the adoption of variable electricity sources (IRENA, 2019) and a means to use otherwise curtailed generation. However, in order to do so optimally, the scheduling of these ESTs must further take into account uncertainties relating to renewable generation, curtailment and market prices due to their unpredictable nature.

There are a number of recent studies optimising the scheduling of renewable energy - energy storage systems under uncertainties, specifically wind-hydrogen systems. Xiao et al. (2020) consider uncertainties in wind generation and electricity price and present a scenario-based stochastic optimisation (SBSO) model which evaluates financial risk. They find that a hydrogen electrolyser (HE) can increase the value of a wind system, the extent of which depends on hydrogen price. Yu et al. (2019) and Mirzaei et al. (2019) present SBSO models which minimise operation costs of a system with wind generation and HS, the latter study also considers demand response (DR). Both of these papers consider uncertainties in wind generation, whilst Yu et al. (2019) also considers uncertainties in demand. The efficacy of these models at reducing the risk of uncertainties

is demonstrated. Finally, Cao et al. (2020) present a SBSO model to minimise operation costs of an intelligent parking lot with HS and renewable generation. They present a Pareto set of solutions for different levels of risk aversion. These studies highlight the value of SBSO models for scheduling wind-hydrogen systems under uncertainties. However, they do not consider curtailed wind, which is an important issue as renewable penetration increases, nor do they consider other forms of energy storage, such as BS.

Several studies address using curtailed wind for an HE. For example, Yan et al. (2018) explore different approaches for handling curtailed wind. They find that investing in an HE is both a profitable and environmentally friendly approach. However, they do not consider uncertainties in wind power or electricity price. Shams et al. (2021) present a machine learning model to predict curtailed power which is used for an HE and BS. However, they optimise from a system operator point of view rather than that of an investor. On the other hand, Jiang et al. (2019) present a chance-constrained model that optimises the size of a wind-hydrogen system from an investor's perspective. Their methodology allows flexibility for the decision variables to not satisfy the constraints at a given probability level; thus adverse conditions can be accounted for. However, they do not model different curtailment or electricity price scenarios nor do they incorporate BS.

In this work we consider an investor's point of view, and present an SBSO which schedules a wind farm with HE and BS. We optimise their usage to maximise income, considering curtailment and uncertainties in wind generation, curtailment and electricity price. We compare case studies with HE and BS, HE only, BS only and no storage (NS). From this we determine the optimal choice for a wind farm owner to maximise income and utilise the maximum amount of curtailed wind.

2. Model Description

Wind curtailment occurs when generation exceeds demand, and generators are instructed to reduce, or sometimes halt, power export. At time t , in scenario, i , total wind generation can be divided into two categories: curtailed wind, $w_{t,i}^c$, which cannot be exported, and non-curtailed wind, $w_{t,i}^n$, which is available to export. The electrolyser can be powered using curtailed, $e_{t,i}^c$, or non-curtailed wind, $e_{t,i}^n$. Likewise the battery can be charged using curtailed, $c_{t,i}^c$, or non-curtailed wind $c_{t,i}^n$. The discharged power from the battery can be exported to the grid $d_{t,i}^n$ or curtailed $d_{t,i}^c$. It is assumed that at times when wind generation is not curtailed, $w_{t,i}^c = 0$, discharged battery is also not curtailed, $d_{t,i}^c = 0$. However where there is wind curtailment, it is assumed that discharged power cannot be exported to the grid, $d_{t,i}^n = 0$, and is also curtailed.

The objective function is given in Equation 1 and maximises revenue due to selling non-curtailed power in the day-ahead market (first term), selling hydrogen (second term) and minimises losses due to using curtailed wind (third term). The day-ahead price at time, t , and scenario, i , is $p_{t,i}^{da}$, p^h is hydrogen price and η^e is electrolyser hydrogen conversion efficiency. The cost of using curtailed wind, p^c , is neglected in most models, which assume that curtailed wind is free. This assumption is overly simplistic and not realistic, hence we consider p^c here.

$$\max \sum_{t,i=0}^{T,I} (w_{t,i}^n + d_{t,i}^n - c_{t,i}^n - e_{t,i}^n) p_{t,i}^{da} + (e_{t,i}^n + e_{t,i}^c) \frac{p^h}{\eta^e} - (e_{t,i}^c + c_{t,i}^c) p^c \quad (1)$$

The constraints are given in Equations 2-10. Equation 2 sets the lower limits on the battery charging and discharging powers and the power going to the electrolyser, where \underline{e} represents the minimum power required for hydrogen production. Equations 3 and 4 set the upper limits; the upper bound on $d_{t,i}^n$ is set such that discharged power can only be exported when there is no wind curtailed. Equation 5 prevents the sum of curtailed and non-curtailed powers exceeding the maximum limits.

$$c_{t,i}^n, c_{t,i}^c, d_{t,i}^n, d_{t,i}^c \geq 0, \quad e_{t,i}^n, e_{t,i}^c \geq \underline{e} \quad \forall t, i \quad (2)$$

$$c_{t,i}^n, c_{t,i}^c \leq \bar{c}, \quad d_{t,i}^c \leq \bar{d}, \quad e_{t,i}^n, e_{t,i}^c \leq \bar{e} \quad \forall t, i \quad (3)$$

$$\text{if } w_{t,i}^c = 0: d_{t,i}^n \leq \bar{d}, \quad \text{else: } d_{t,i}^n \leq 0 \quad \forall t, i \quad (4)$$

$$c_{t,i}^n + c_{t,i}^c \leq \bar{c}, \quad d_{t,i}^n + d_{t,i}^c \leq \bar{d}, \quad e_{t,i}^n + e_{t,i}^c \leq \bar{e} \quad \forall t, i \quad (5)$$

Equations 6 and 7 set limits on the battery's capacity, $x_{t,i}$, and ensure that it is equal to the capacity at the previous time period plus any charging/discharging in the current time period, respectively. The charging and discharging efficiencies are η^c and η^d , respectively, and are equal to 90%.

$$\underline{x} \leq x_{t,i} \leq \bar{x} \quad \forall t, i \quad (6)$$

$$x_{t,i} = x_{t-1,i} + (c_{t,i}^n + c_{t,i}^c)\eta^c - \frac{d_{t,i}^n + d_{t,i}^c}{\eta^d} \quad \forall t, i \quad (7)$$

Equation 8 prevents the battery from being simultaneously charged and discharged. In Equation 9 the curtailed generation and discharge is greater than or equal to the curtailed power used for charging and powering the electrolyser. Equation 10 ensures that when there is no wind curtailed there is also no curtailment of discharged battery. M is a very large positive co-efficient which allows curtailed discharge to take on any value, satisfying previous constraints, when there is non-zero curtailed wind.

$$(c_{t,i}^n + c_{t,i}^c)(d_{t,i}^n + d_{t,i}^c) = 0 \quad \forall t, i \quad (8)$$

$$w_{t,i}^c + d_{t,i}^c \geq c_{t,i}^c + e_{t,i}^c \quad \forall t, i \quad (9)$$

$$Mw_{t,i}^c - d_{t,i}^c \geq 0 \quad \forall t, i \quad (10)$$

3. Scenario Generation

A range of scenarios are generated to represent possible outcomes of the uncertain parameters, in this case wind generation, curtailment and electricity price. Three wind power profiles are randomly generated from Drax Power Ltd (2021), by adding noise from a Gaussian centred around each data point with a mean equal to that point and a standard deviation $0.25 * \text{data point}$. This wind data is scaled such that the farm has a maximum output of 20 MW. Five curtailment profiles are then generated using a MC with probabilities of moving between states 'curtailed' and 'not-curtailed' determined using historic data, and initial state 'not-curtailed'. When the state is 'curtailed', the percentage of wind power curtailed is determined by randomly selecting from historic data. The 'not-curtailed' wind profile is the difference between this and the original wind profile. Three price profiles are created using Gaussian Process (GP) techniques as described in Yeardley et al. (2021); the first profile is an ordinary GP, and the other two are created using a novel hybridisation method which combines Gaussian Processes with K-means clustering and hierarchical cluster.

A summary of the scenario generation procedure is shown in Figure 1. Each of these scenarios is input into our SBSO model which optimises the scheduling our of energy storage for cases with HE and BS, HE only, BS only, and no storage (NS). The input parameters for the storage for each of these cases is shown in Table 1. Finally, it is assumed that hydrogen can be sold at a price of £3.50/kg and the cost of curtailed wind is £0.01/kWh.

Table 1: Input storage parameters for each case considered.

	\underline{e} (MW)	\bar{e} (MW)	\bar{c}, \bar{d} (MW)	\underline{x} (MWh)	\bar{x} (MWh)
HE + BS	0.04	2	2	1.6	8
HE	0.04	2	0	0	0
BS	0	0	2	1.6	8
NS	0	0	0	0	0

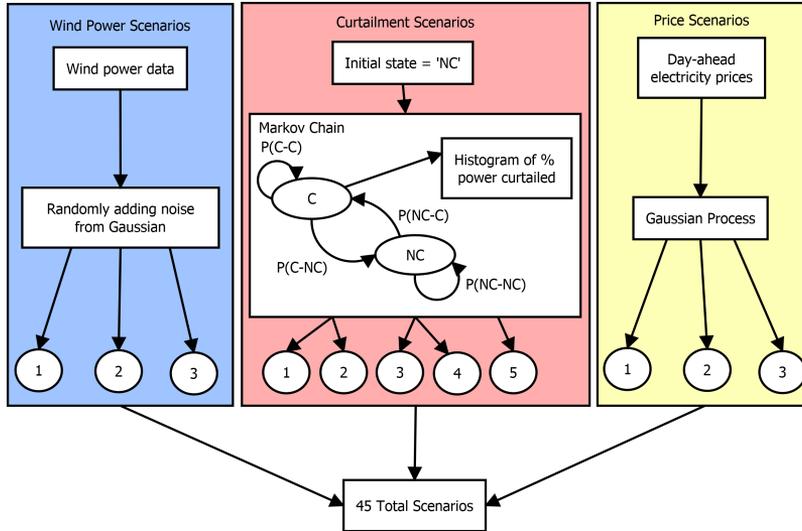


Figure 1: Scenario generation diagram

4. Results and Discussion

In Table 2 the mean income and percent of curtailed wind utilisation, across all scenarios, are shown for the different case studies, along with their standard deviations. It can be seen that the inclusion of storage increases both the mean expected income and curtailed wind utilisation. In particular, the HE is able to increase the values of these more than the BS. However, the combination of both is the most effective of the case studies presented here. Additionally, the inclusion of storage reduces the standard deviation of mean expected income. This is since storage adds flexibility; for instance, when wind generation is low and curtailment is high, additional revenue can still be achieved due to selling hydrogen and discharged power from the BS.

Figure 2 shows the optimised daily power profiles for each case study for Scenario 6. This scenario was chosen because there is a large amount of wind curtailment, occurring between 4:00 and 14:00 (8 and 28 in Figure 2), and shows how the scheduling of the storage responds to this. Wind generation is indicated by the red lines; wind power that is directly imported or curtailed is shown by a solid area, power used for the HE: a dashed area, and power used for or discharged by the BS: a dotted area. Non-curtailed wind may be exported, along with non-curtailed discharge from the BS (indicated by a dark blue area); alternatively, it may be consumed by the storage (green area). Curtailed wind (and BS discharge) is used to power the HE and/or BS (although the BS cannot simultaneously charge and discharge) and is indicated by a white area.

In the case of NS, we can see that all curtailed wind is wasted. By adding BS, we are able to use some of the curtailed wind, however, once the BS is fully charged we cannot use it anymore. The

Table 2: Mean and standard deviation of daily income and percentage of curtailed wind utilised across all scenarios for each case study.

	Mean income (£)	Standard deviation (£)	Curtailed wind usage (%)	Standard deviation (%)
HE + BS	4989	222	0.949	0.029
HE	4870	300	0.679	0.034
BS	4311	629	0.165	0.037
NS	4208	689	0.0	0.0

BS also allows a greater amount of power to be exported in the evening when electricity prices are typically higher (34 - 38 in Figure 2). By adding HE we are able to use a greater proportion of the curtailed wind. Furthermore, under the conditions specified here, it is economical to self-consume and import power for the HE. By combining BS and HE we are able to utilise the most curtailed wind and maximise power used for the HE; at 10:30 and 12:00 (21 and 24 in Figure 2), there are two peaks above the red line which indicate curtailed BS discharge powering the HE. As shown in Table 2 this case generates the highest mean income across the different scenarios. Hence we conclude that of the cases explored here, a combination of BS and HE is optimal for both maximising income and utilising the maximum curtailed wind.

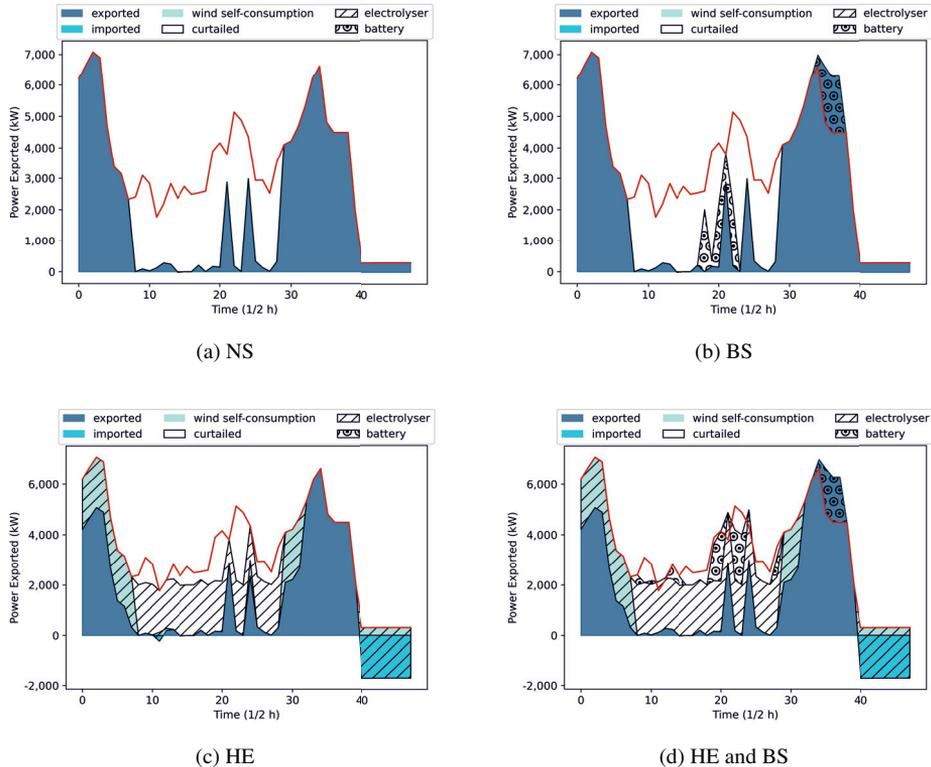


Figure 2: Daily power profile for Scenario 6, for each case study. Red line indicates total wind generation.

5. Conclusion

A scenario-based stochastic optimisation (SBSO) model is presented to schedule a wind farm with battery storage (BS) and a hydrogen electrolyser (HE) under uncertain conditions and considering curtailment. We generate wind curtailment and electricity price scenarios using Markov Chain (MC) and Gaussian Process (GP) techniques, respectively, to model a range of possible outcomes. We compare daily mean predicted income and utilisation of curtailed wind with BS only, HE only, both BS and HE, and no storage (NS).

We find that HE increases mean income and curtailed wind utilisation significantly more than BS. However, by combining HE and BS curtailed wind utilisation increases from 68% to 95%, compared with HE alone. At times when curtailed wind is greater than the HE maximum power, it can also be used to charge the BS; then at times when curtailed wind is lower than this maximum power, it can be additionally powered by discharging the BS. Future work will consider capital and operational costs of these technologies, as well as varying their sizes, ratios and hydrogen price.

6. Acknowledgements

This work was supported by the UK Engineering and Physical Science Research Council via grant EP/L016818/1 which funds the Centre for Doctoral Training in Energy Storage and its Applications. The authors are also grateful to Drax Power Ltd. for their financial support.

References

- Y. Cao, J. Du, X. Qian, S. Nojavan, K. Jernsittiparsert, 2020. Risk-involved stochastic performance of hydrogen storage based intelligent parking lots of electric vehicles using downside risk constraints method. *International Journal of Hydrogen Energy* 45 (3), 2094–2104.
- C. K. Das, O. Bass, G. Kothapalli, T. S. Mahmoud, D. Habibi, 2018. Overview of energy storage systems in distribution networks: Placement, sizing, operation, and power quality. *Renewable and Sustainable Energy Reviews* 91, 1205–1230.
- Drax Power Ltd, 2021. Private communication.
- D. Gielen, E. Taibi, R. Miranda, 2019. Hydrogen: A reviewable energy perspective: Report prepared for the 2nd hydrogen energy ministerial meeting in tokyo, japan.
- IRENA, 2019. Innovation landscape brief: Utility-scale batteries.
- Y. Jiang, Z. Deng, S. You, 2019. Size optimization and economic analysis of a coupled wind-hydrogen system with curtailment decisions. *International Journal of Hydrogen Energy* 44 (36), 19658–19666.
- M. A. Mirzaei, A. S. Yazdankhah, B. Mohammadi-Ivatloo, 2019. Stochastic security-constrained operation of wind and hydrogen energy storage systems integrated with price-based demand response. *International Journal of Hydrogen Energy* 44 (27), 14217–14227.
- M. H. Shams, H. Niaz, J. Na, A. Anvari-Moghaddam, J. J. Liu, 2021. Machine learning-based utilization of renewable power curtailments under uncertainty by planning of hydrogen systems and battery storages. *Journal of Energy Storage* 41, 103010.
- P. Xiao, W. Hu, X. Xu, W. Liu, Q. Huang, Z. Chen, 2020. Optimal operation of a wind-electrolytic hydrogen storage system in the electricity/hydrogen markets. *International Journal of Hydrogen Energy* 45 (46), 24412–24423.
- X. Yan, X. Zhang, C. Gu, F. Li, 2018. Power to gas: addressing renewable curtailment by converting to hydrogen. *Frontiers in Energy* 12 (4), 560–568.
- A. S. Yeardeley, D. Roberts, R. A. Milton, S. F. Brown, 2021. Robust probabilistic electricity price forecasting using a hybridisation of gaussian processes and clustering.
- D. Yu, J. Wang, D. Li, K. Jernsittiparsert, S. Nojavan, 2019. Risk-averse stochastic operation of a power system integrated with hydrogen storage system and wind generation in the presence of demand response program. *International journal of hydrogen energy* 44 (59), 31204–31215.

Optimization of Biofuel and Kerosene Fuel Blends to Support Sustainable Aviation

Ridab Khalifa, Mohammad Alherbawi, Adel Elomri, Tareq Al-Ansari*

College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar.

*talansari@hbku.edu.qa

Abstract

The aviation industry supports the world economy, contributing US\$ 2.7 trillion to global the gross domestic product. However, aviation raises environmental concerns, where the industry has a 12% share of CO₂ emissions within the transportation sector. Therefore, the International Civil Aviation Organization (ICAO) suggested the implementation of Carbon Offsetting Scheme for International Aviation (CORSIA) as a market-based measure to mitigate CO₂. The CORSIA scheme may increase the operational costs by setting a carbon price on every extra tonne of CO₂ beyond the baseline limits. In order to reduce operators' obligations, the integration of reduction measures such as Sustainable Aviation Fuels (SAF) may reduce the cost associated with offsetting requirements. As such, a multi-objective optimization model is presented in this study to identify optimal blending ratios of jet biofuels with conventional kerosene fuel for multiple aircrafts and destinations. The model considers the operators fuel cost, carbon price and renewable credit under CORSIA; aiming to minimize the total fuels' associated costs. In addition, the model is implemented in a case study considering three fuel categories. The results indicate that Jatropa-based jet fuel, within the current tested data, is a preferable synthetic fuel to be blended with Jet-A at a maximum margin of 50%. Fuel prices highly influenced the results of the model. Whereas other factors including carbon prices, fuels' lifecycle emissions, and supplied fuel quantity may directly or indirectly impact the process of incorporating SAF as an integrated mitigation tool under CORSIA.

Keywords: CORSIA, Carbon Policy, Aviation, Sustainability, Mitigation.

1. Introduction

The airline industry has not only influenced global mobility, it has also contributed US\$3.5 trillion to the global gross domestic product (GDP). In 2019, the aviation sector created over 87.7 million jobs and facilitated 4.5 billion boarding passengers on the world's airlines. During the pandemic, the economic contribution from the aviation sector reduced to US\$1.7 trillion, eliminating 46 million jobs supporting the industry (ATAG, 2020). As a mode of transportation, the aviation sector contributes 12% of the total CO₂ emissions produced from transportation (IEA, 2019). It is reported that in the next 15 years, the demand for transportation will double (Airbus, 2019), thus increasing CO₂ emissions from the sector. As a consequence, the IATA suggested four CO₂ mitigation techniques: market-based measures (MBM), alternative fuels, technological enhancement, and operational modifications. The International Civil Aviation Organization ICAO proposed the "Carbon Offsetting Scheme for International Aviation

(CORSIA)" as an applicable methodology of global MBM techniques to support aviation environmental goals.

1.1. CORSIA

During the 39th assembly, the ICAO announced CORSIA as a global offsetting scheme for international flights, with an applicable methodology of a combinatorial approach, beginning with a 100% sectorial approach and moving to a 100% individual approach (IATA, 2019). Fuels classified as CORSIA's Eligible Fuel (CEF) should have the capability of achieving a minimum 10% emission reduction compared to Jet-A, determined by life cycle assessment (LCA) of the fuel. The integrability of SAF, technological improvement, and operational enhancement within a CORSIA framework can help minimize the cost of carbon, support IATA's goal of reducing carbon emissions, and invest in research or through the implementation of SAF (Staples et al., 2018). The acceleration on developing drop-in fuels seems the most suitable substitute for Jet-A, in which its implementation may require minimum or no design modification of the engine or fuel system. These fuels can be synthesized from biomass for example, through production pathways of Fischer-Tropsch (FT), Hydroprocessed Ester Fatty Acid (HEFA), Hydrothermal Liquefaction (HTL), Alcohol to Jet (ATJ), and Direct Sugars to Hydrocarbons (DSHC) or commonly known as Synthesized Iso-paraffins (SIP) (Stratton et al., 2010).

1.2. CORSIA Associated Models

Optimization techniques can predict the influence of SAF in minimizing carbon emissions and support managerial decisions. Jiang & Yang (2021) discussed the correlation between SAF and SAF's policies from an operational decisions perspective, comparing carbon tax and SAF quota. Sharma et al. (2021) investigated CORSIA by examining technologically innovative factors and the related emissions through regression techniques. Sharma et al. (2021) developed a minimization model of aviation emissions using Vensim. Chao et al. (2019) modified the Fleet-level Environmental Evaluation Tool (FLEET) and created multiple emissions scenarios. The modified model was integrated with the SAF lifecycle assessment to evaluate SAF development economically and environmentally. Whereas, this study presents a multi-objective optimization model to select optimal blending ratios based on carbon price, fuel price and lifecycle emissions. The model is applicable to multiple data entries considering different types of aircrafts and multiple destinations through the incorporation of key parameters, such as, aircraft fuel consumption rate, number of passengers and distance travelled.

2. Methodology

2.1. Mathematical optimization model

A mathematical model was developed to select and set the blending ratios of viable SAF's and Jet-A under CORSIA as presented in Table 1. The model was solved using CPLEX OPL considering two assumptions. The model assumed that the Trips Number ($TN_{k,j}$) is a ready-prepared aircraft schedule and that the blending ratio does not vary for different trips on the same route. The model aimed to minimize the cost of investing and operating fleets using certain SAF and the cost of emissions under CORSIA framework. It also considers the reward provided to operators utilizing SAF.

Table 1: Mathematical formulation of blend ratio selection of biofuel and Jet-A.

Sets, Decisions Variables, and Parameters	
Sets	Parameters
k: is the aircraft type	Pr: Price of fuel f (\$/MJ.)

j: is the routes f: is the fuel	CPrice: Price of carbon (\$/kgCO ₂) Ef_f: Amount of CO ₂ produced per MJ of fuel f. LCA_f: Lifecycle Analysis value of fuel f.
Decision Variable BR_{f,k,j}: Blending ratio of fuel f on aircraft k and route j.	TN_{k,j}: Number of trips of aircraft type k on route j. Pax_{k,j}: Number of passengers using aircraft type k on route j. CR_k: Consumption rate using aircraft k (M.J./Pax.km). LHV_f: Lower Heating Value of fuel f (M.J./liter) Supply_f: The amount of supplied fuel f (liter) FCF: Fuel Conversion Factor (0.0718 kgCO ₂ /MJ for Jet-A). GF: Growth factor of carbon (20%). NFC: Net cost of jet fuel. CTax: Carbon taxing cost. RCredit: Reduction credit, representing a reclaimed value of reduced emissions from the offsetting requirements.
Objective Function	
Min NFC + CTax – RCredit	
Where:	
NFC = $\sum_f \sum_k \sum_j FuelConsumption_{f,k,j} * P_f$	
CTax = $CPrice * GF * \sum_f \sum_k \sum_j OperatorsEmissions_{f,k,j}$	
RCredit = $CPrice * FCF * \sum_f \sum_k \sum_j FuelConsumption_{f,k,j} * (1 - \frac{LCA_f}{89})$	
OperatorsEmissions = $\sum_f \sum_k \sum_j FuelConsumption_{f,k,j} * Ef_f$	
FuelConsumption = $\sum_f \sum_k \sum_j Pax_{k,j} * TN_{k,j} * CR_k * D_j * BR_{f,k,j}$	
Constraints	
$\sum_f BR_{f,k,j} = \begin{cases} 1 & TN_{k,j} \neq 0 \\ 0 & TN_{k,j} = 0 \end{cases} \quad \forall k, j$	The blending ratio of fuels f utilized by assigned aircraft k on route j should add up to 100%.
$BR_{f=1,k,j} \geq 0.5 \quad TN_{k,j} \neq 0, \quad \forall k, j$	A minimum blending ratio of 50% should be allocated to Jet-A utilized by aircraft k on route j.
$\sum_k \sum_j Pax_{k,j} TN_{k,j} CR_{k,j} BR_{f,k,j} / LHV_f \leq Supply_f \quad \forall f$	For each fuel, the total amount of the consumption by all aircraft k and all route j should not exceed
$\sum_{f=2}^F \sum_k \sum_j FuelConsumption_{f,k,j} * FCF * (1 - \frac{LCA_f}{89}) \geq 0.1 \sum_{f=1}^F \sum_k \sum_j FuelConsumption_{f,k,j} * Ef_f$	The total SAF fuels utilized by all aircraft k associated with route j should pertain to a minimum of 10% emissions reduction from the total fuel consumption.

2.2. Case study

Aside from the conventional Jet-A, data of two biofuels including jatropha and algae-based fuels, were obtained to test and validate the developed model. Previous literature referenced a well-to-wheel (WTW) approach to account for GHG emissions of jatropha and algae fuels as studied by Alherbawi et al. (2021a) and Fortier et al. (2014), respectively. CORSIA specifies that these values shall be compared to the lifecycle emissions of 89 gCO₂/MJ of conventional jet fuel, as referenced by the ICAO (IATA, 2019). Jatropha lifecycle emissions values were collected from Alherbawi et al. (2021a). Whereas Fortier et al. (2014) performed an LCA using a WTW approach to study algae-based jet fuel's environmental performance. The study analyzed two pathways: the Hydrothermal Liquefaction (HTL) and the wastewater treatment plant (WWTP) HTL. The study also compared the results to previous LCA values of bio-jet fuels. Table 2 shows the average emissions for the two selected fuels. Fuel prices were collected to evaluate the additional costs associated to clean fuels for airlines operators.

Algae and jatropha jet fuel prices were obtained based on the minimum selling price of jet biofuels (Alherbawi et al., 2021b). The fuel prices were presented in USD per energy equivalent units of (MJ) based on the lower heating value (LHV) and density characteristics for each biofuel (Azami & Savill, 2017), and for jet-A as specified by Chevron Products Company (2007). In 2019, biojet production was estimated with 140 million liters per year, the production is assumed to be shared between Jatropha and Algae as in Table 2 (Renewable Energy Agency, 2021).

Table 2: Selected fuels' price, quantity, and characteristic.

Fuels	Price (\$/MJ)	Supplies (Million L)	LHV (MJ/L)	LCA emissions (gCO_{2eq}/MJ)
Jatropha	0.026	70.000	39.0	37.95
Algae	0.031	70.000	38.0	61.38
Jet-A	0.015	87.625*	34.7	89.00

* Jet-A supplied fuel amount was assumed.

The associated direct emissions of Jet-A, jatropha, and algae synthesized fuels are defined based on their emissions factors of 3.16, 2.81 (Carels et al., 2012), and 3.7 (kgCO₂/kgfuel) (Ponnusamy et al., 2014), respectively. Implementing CORSIA requires setting a growth factor value by the ICAO to specify the increase in the emissions compared to a baseline of 2020 for each country. The tested scenario assumed a growth factor of 20% and a carbon price tax of 3 \$/kgCO₂.

Table 3: Selected aircraft types and fuel consumption rates.

Type of Aircraft	Consumption Rate (MJ/Pax/km)
Airbus A320-200	0.7814
Airbus A330-200	1.0802
Airbus A380	1.1357

The scenario was developed with three aircraft types as summarized in Table 3 along with their consumption rates. It was assumed that the operators have a ready flight schedule associated with a specific aircraft type and an average number of passengers on board. The number of trips and passengers per aircraft on each route and the associated traveled distance are illustrated in Table 4 and Table 5.

Table 4: Trips number per an aircraft type on a specific route

Type of Aircraft	Trips Number			Passengers		
	Route	Route	Route	Route	Route	Route
	1	2	3	1	2	3
Airbus A320-200	0	0	46	0	0	180
Airbus A330-200	77	0	0	220	0	0
Airbus A380	0	22	0	0	320	0

Table 5: Trips destinations and distances.

Route	Doha-Kuwait	Doha-London	Doha- Miami
Distance (km)	572	5240	12340

3. Results

The results illustrated Figure 1 indicate that the model selected kerosene and jatropha to create an optimal blend. Algae is excluded from the blend in the current scenario, which is possibly due to its higher price as compared to Jet-A and jatropha fuels. The first route was assigned a 100% jet-A to minimize the overall fuel associated costs, while the SAF blend comes is introduced from the second route to satisfy the minimum emission reduction constraints. However, as the distance travelled increases, the subsequent emissions also increase, which require more SAF to be blended to restrict the emissions within the defined margins. As such the third route was associated with a 50:50 Jet-A and jatropha blended fuels. The initial run of the model suggested a cost of approximately 1.5 million U.S.\$. Along with fuel price, the supplied fuel quantity, fuel emission factor, and emissions reduction capability also influenced the optimal blending ratio and the SAF selection. In addition, an increasing carbon tax is expected to further promote the integration of SAF with higher blending ratios. In such case, airlines can shift to SAF with reasonable blending ratios to satisfy the environmental targets without compromising the economic gains.

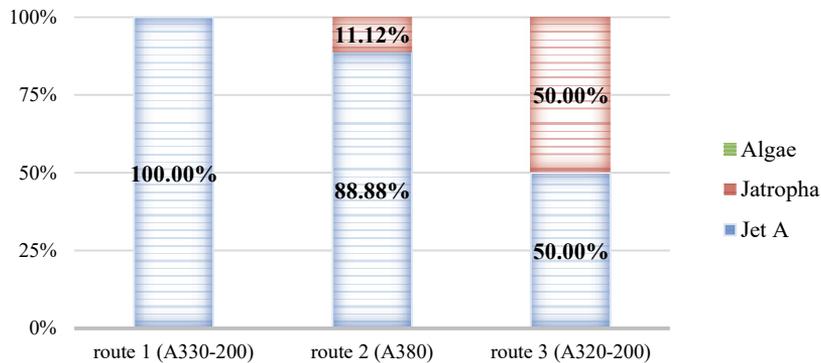


Figure 1: Results of the optimal fuel blending ratios for selected routes.

4. Conclusion

CORSIA promotes the integration of SAF to mitigate the environmental impact of air travel. The policy framework incentivizes operators utilizing SAF by providing a reduction reward function that allows operators to reclaim some of the offsetting carbon emissions. However, a 100% shift can be more challenging due to the high price of SAF as compared to Jet-A, leaving operators required to maintain a higher blending ratio for Jet-A utilization on fleets, as the results reflect a 100%, 88.88 %, and a minimum of 50% Jet-A utilization. It is expected that SAF blending ratios will be further promoted with the expected increase in carbon tax. However, more efforts are required to produce cheaper SAFs and reduced life cycle emissions to reach a full transition to clean fuels.

References

- Airbus. (2019). Global Market Forecast: Cities, Airports, and Aircraft 2019 -2038 . <http://gmf.airbus.com/>
- Alherbawi, M., McKay, G., Mackey, H. R., & Al-Ansari, T. (2021a). Jatropha curcas for jet biofuel production: Current status and future prospects. In *Renewable and Sustainable Energy Reviews* (Vol. 135, p. 110396). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2020.110396>

- Alherbawi, M., McKay, G., Mackey, H. R., & Al-Ansari, T. (2021b). A novel integrated pathway for Jet Biofuel production from whole energy crops: A *Jatropha curcas* case study. *Energy Conversion and Management*, 229, 113662. <https://doi.org/10.1016/J.ENCONMAN.2020.113662>
- ATAG. (2020). *Aviation Beyond Borders*. Atag, 96.
- Azami, M. H., & Savill, M. (2017). Comparative study of alternative biofuels on aircraft engine performance. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 231, 1509–1521. <https://doi.org/10.1177/0954410016654506>
- Carels, N., Sujatha, M., & Bahadur, B. (2012). *Jatropha, challenges for a new energy crop: Volume 1: Farming, economics and biofuel*. *Jatropha, Challenges for a New Energy Crop: Volume 1: Farming, Economics and Biofuel*, 1–599. <https://doi.org/10.1007/978-1-4614-4806-8>
- Chao, H., Agusdinata, D. B., DeLaurentis, D., & Stechel, E. B. (2019). Carbon offsetting and reduction scheme with sustainable aviation fuel options: Fleet-level carbon emissions impacts for U.S. airlines. *Transportation Research Part D: Transport and Environment*, 75, 42–56. <https://doi.org/10.1016/J.TRD.2019.08.015>
- Chevron Products Company. (2007). *Aviation Fuels Technical Review*. <http://www.chevron.com/productservices/aviation/>
- Fortier, M. O. P., Roberts, G. W., Stagg-Williams, S. M., & Sturm, B. S. M. (2014). Life cycle assessment of bio-jet fuel from hydrothermal liquefaction of microalgae. *Applied Energy*, 122(July 2011), 73–82. <https://doi.org/10.1016/j.apenergy.2014.01.077>
- IATA. (2019). *An Airline Handbook on CORSIA*.
- IEA. (2019). *Tracking Transport – Analysis - IEA*. <https://www.iea.org/reports/tracking-transport-2019>
- Jiang, C., & Yang, H. (2021). Carbon tax or sustainable aviation fuel quota. *Energy Economics*, 105570. <https://doi.org/10.1016/J.ENERCO.2021.105570>
- Ponnusamy, S., Reddy, H. K., Muppaneni, T., Downes, C. M., & Deng, S. (2014). Life cycle assessment of biodiesel production from algal bio-crude oils extracted under subcritical water conditions. *Bioresource Technology*, 170, 454–461. <https://doi.org/10.1016/j.biortech.2014.07.072>
- Renewable Energy Agency, I. (2021). *Reaching zero with renewables: Biojet fuels*. www.irena.org
- Sharma, A., Jakhar, S. K., & Choi, T. M. (2021). Would CORSIA implementation bring carbon neutral growth in aviation? A case of US full service carriers. *Transportation Research Part D: Transport and Environment*, 97, 102839. <https://doi.org/10.1016/J.TRD.2021.102839>
- Staples, M. D., Malina, R., Suresh, P., Hileman, J. I., & Barrett, S. R. H. (2018). Aviation CO2 emissions reductions from the use of alternative jet fuels. *Energy Policy*, 114(December 2017), 342–354. <https://doi.org/10.1016/j.enpol.2017.12.007>
- Stratton, R. W., Wong, H. M., & Hileman, J. I. (2010). *Life Cycle Greenhouse Gass Emission from Alternative Fuel*. <http://www.partner.aero>

Dynamic Surrogate Modeling for Continuous Processes Control Applications

Alessandro Di Pretoro^{a,b}, Andrea Tomaselli^b, Flavio Manenti^b, Ludovic Montastruc^{a*}

^a*Laboratoire de Génie Chimique, Université de Toulouse, CNRS/INP, Toulouse, France*

^b*Politecnico di Milano, Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Piazza "L. Da Vinci" 32, 20133 Milano, Italia*

ludovic.montastruc@ensiacet.fr

Abstract

With the increasing amount of data to quickly process, surrogate modeling has become a topic of major interest in process engineering during the last decades. Recently, black-box models have gained renewed interest as effective tools to reduce the computational effort and to enhance the solution of optimization algorithms. Among the surrogate model applications, scheduling and control problems are getting particular attention in the last years. However, when dealing with dynamics, the majority of literature works concern batch processes to be modeled within the corresponding residence time interval. With the purpose to extend the established procedures to continuous operations, in this research work the surrogate modeling procedure of a dynamic non-isothermal CSTR reactor is discussed and analyzed in detail. The modeling phase was carried out for the temperature and concentration with different sampling size and techniques by means of the software ALAMO[®]. After the open-loop model was built and validated, the related closed loop configuration has been tested by means of a conventional PID controller implementation. The controller has been properly tuned on both models in order to compare the performances. The obtained trends show good agreement with those of the rigorous model both for open loop and closed loop performances. More importantly, the implemented surrogate model requires reduced calculation time thanks to the explicit input-output variable correlations. In conclusion, the proposed dynamic surrogate modeling approach for process control applications has proved to be effective and provided reliable results. Moreover, the computational time was reduced by an order of magnitude and the best compromise between the performances and accuracy can be detected.

Keywords: data-driven modeling, design of experiment, process control, ALAMO

1. Introduction

Nowadays, the exponential increase of the available computational performances is leading the entire engineering domain towards a more and more data-driven approach. With respect to phenomenological models, data-driven models sometimes allow a smoother calculation, lower computational effort and better convergence of optimization algorithms. In the process engineering domain, the impact of data processing in the digitalization transition can be detected in particular in the growing interest towards surrogate modeling. Initially conceived to compensate the lack of phenomenological models, surrogate modeling for chemical processes has become a topic of major interest in the Process Systems Engineering and it is studied and exploited to considerably reduce

the computational effort for complex systems modeling as well as for optimization (Bhosekar and Ierapetritou, 2018; McBride and Sundmacher, 2019).

The majority of surrogate modeling related research works addresses steady state models to derive input-output variables correlations without the need to solve all the units of the system. During the last years, the interest towards this approach is nevertheless involving dynamic systems as well (Di Pretoro et al., 2022). When dynamic models are analyzed, the literature studies mainly refers to batch operations to be modeled over the related residence time interval (Shi and You, 2015; Shokry et al., 2020). The purpose of this work is then to extend the surrogate modeling activity to continuous processes and suggest a preliminary approach that could be suitable for this scope. Moreover, once completed the modeling phase, we are interested as well in comparing the rigorous and data-driven model closed-loop performances for a conventional PID feedback control strategy.

The selected case study is a simple non-isothermal CSTR reactor and its thorough description is addressed in the next section as well as the control strategy one. Section 3 refers to the different approaches proposed for the surrogate modeling procedure and the obtained results are then discussed in section 4. Finally, some conclusions of general validity and possible perspectives and developments are commented in the last section.

2. Case study

This section introduces first the selected case study, i.e. a non-isothermal CSTR reactor, along with its model equations and parameter values. The second part of the section is focused instead on the feedback control strategy and on the parameters tuning approach.

2.1. The non-isothermal CSTR reactor

The continuous process selected for the surrogate modeling procedure is a conventional non-isothermal CSTR reactor, with constant inlet and outlet overall flowrates, where a first order reaction occurs according to the kinetic scheme:



$$k = k_0 \cdot \exp\left(-\frac{E_a}{R \cdot T}\right) \quad (2)$$

The set of differential equations required for the phenomenological modeling of the system consists of mass and heat balances as follows:

$$\frac{dC_A}{dt} = \frac{Q}{V} \cdot (C_{A,in} - C_A) - k_0 \cdot \exp\left(-\frac{E_a}{R \cdot T}\right) \cdot C_A \quad (3)$$

$$\frac{dC_B}{dt} = -\frac{Q}{V} \cdot C_B + k_0 \cdot \exp\left(-\frac{E_a}{R \cdot T}\right) \cdot C_A \quad (4)$$

$$\frac{dT}{dt} = \frac{Q}{V} \cdot c_p \cdot (T_{in} - T) - \frac{\Delta H_r}{\rho \cdot c_p} \cdot k_0 \cdot \exp\left(-\frac{E_a}{R \cdot T}\right) \cdot \frac{UA}{\rho \cdot V \cdot c_p} \cdot (T_j - T) \quad (5)$$

All physical properties and system parameters are listed in Table 1.

Process parameter	Value	Unit
Q/V	1	1/s
$C_{A,in}$	0.98	mol/m ³
k_0	$7.2 \cdot 10^{10}$	1/s
E_a/R	8750	K
T_{in}	304.2	K
ΔH_r	$5 \cdot 10^4$	J/mol
$U \cdot A$	$5 \cdot 10^4$	W/K
$\rho \cdot c_p$	239	J/(m ³ ·K)

Table 1 – Process operating conditions and reactor parameters

2.2. Control strategy and tuning

This section introduces the control strategy employed for the aforementioned case study. As already explained, given the constant inlet and outlet flowrates, no level control is accounted for while the reactor temperature is controlled by manipulating the coolant temperature in the jacket. For this purpose, a conventional PID control was used according to the characteristic equation:

$$u(t) = u_s + K_C \cdot \left(e(t) + \frac{1}{\tau_I} \cdot \int_0^t e(\tau) \cdot d\tau + \tau_D \cdot \frac{de(t)}{dt} \right) \quad (6)$$

Where $e(t)$ is the error, K_C is the proportional gain, τ_I and τ_D the integral and derivative time constant respectively. As concerns these controller parameters, both for phenomenological and data-driven models, they were tuned so that they minimize the objective function given by:

$$F_{obj} = w_1 \cdot CA_{max} + w_2 \cdot ISE + w_3 \cdot OS + w_4 \cdot ST \quad (7)$$

where w_i is the relative weight of each term, CA_{max} is the maximum control action, ISE is the Integral Square Error, OS is the OverShoot and ST is the Settling Time.

3. Methodology

The surrogate modeling procedure was carried out based on different approaches better detailed in the following subsections.

In any case, the dataset used for temperature and concentration dynamic trends is the same and it was derived by means of the Latin Hypercube Sampling (LHS) function already implemented in MatLab®. The key difference between this approach for control of continuous processes and those available in the literature is that sampling should be performed over the [time]x[manipulated variable] domain for a time interval long enough to achieve the new steady state conditions.

Figure 1 a and b show respectively the results obtained for reactor temperature and reactant concentration for coolant temperature step perturbation of different magnitude.

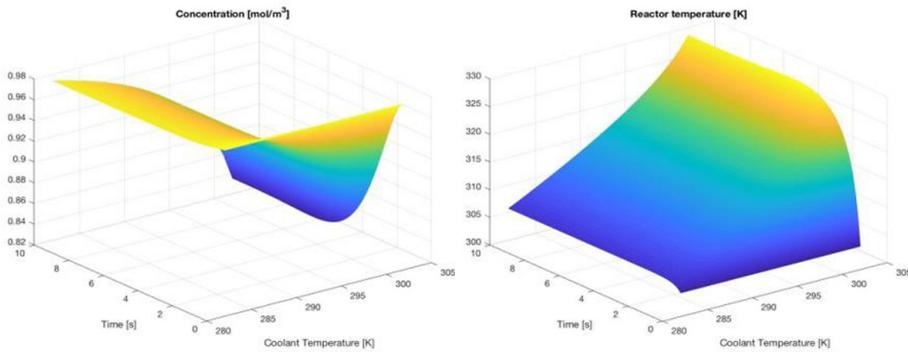


Figure 1 – a) Concentration and b) reactor temperature response vs coolant temperature and time

A time interval equal to 10 seconds was found in order to achieve the new steady state conditions for both variables and every step perturbation size.

3.1. Surrogate modeling with ALAMO[®]

The first two approaches for surrogate modeling exploit the dedicated software ALAMO[®]. This tool performs surrogate modeling based on Response Surface Methodology by using selected interpolating base functions. In this study, mixed polynomial, logarithmic and exponential functions were selected for the regression. Datasets with different size have been imported and the related performances in terms of model accuracy and computational time required to solve the dynamic model over the 10s time interval have been recorded as later discussed in the results section.

The main difference between the two is related to the sample density. In particular, as pointed out in the next section, the model obtained with the conventional LHS sampling shows some discrepancies at the proximity of the domain boundaries and is not able to exactly represent the steady state behaviour. That is why, in the second approach, the sample density was doubled in proximity of the 10 s boundary, so that the steady state points have a higher weight on the objective function minimized by the software.

3.2. First order response regression

The third approach is based on the pre-selection of a polynomial function for the controlled variable whose coefficients are obtained by means of a regression minimizing the relative error with respect to the data sampling. Given the temperature behaviour (cf Figure 1b) a suitable candidate as response surface is a first order response function as:

$$T_r = f(T_j) \cdot (1 - A \cdot e^{-B \cdot t}) + C \quad (7)$$

The advantage of using this expression lies on the fact that, for high values of the time variable, it exhibits a steady state behaviour that is exactly what is experienced by the system under study. This procedure was implemented with a MatLab[®] dedicated script. The results for these three approaches, along with the related advantages and inconveniences are discussed in the next section.

4. Results

4.1. Modeling

This first subsection addresses the performances of the modeling procedure from a computational point of view. Although results are showed for the first approach, the same remarks both from a qualitative and a quantitative perspective are valid for the others.

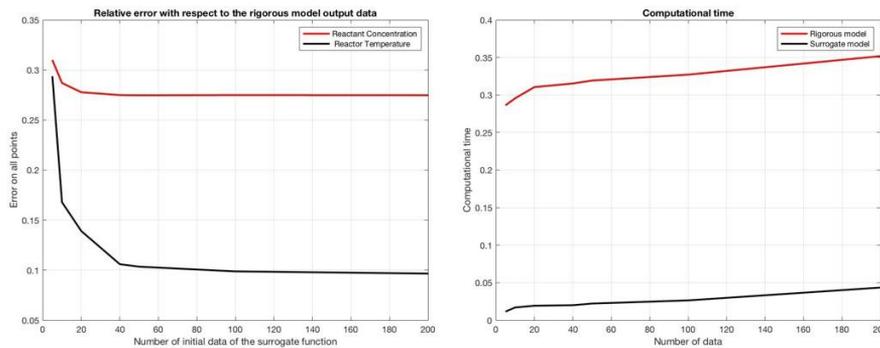


Figure 2 – a) Relative error [%] and b) computational time [s] vs dataset sample size

Figure 2a presents the mean relative error for concentration and temperature while Figure 2b compares the solution of the dynamic model over the 10 s time span for the rigorous and surrogate model respectively. As it can be noticed, for this specific case study, the relative error decreases almost exponentially with the sample size and achieves a stable value of about 0.1 % for a dataset with more than 40 points. On the other hand, the computational time, that is one order of magnitude lower than that of the rigorous model, follows an almost linear trend with respect to the sample size.

From this first analysis, we can observe the considerably higher performance of surrogate model with respect to ODE system from a computational perspective and that a good compromise between accuracy and computational time can be obtained for a number of points in the range 40-60.

4.1.1. Closed-loop response

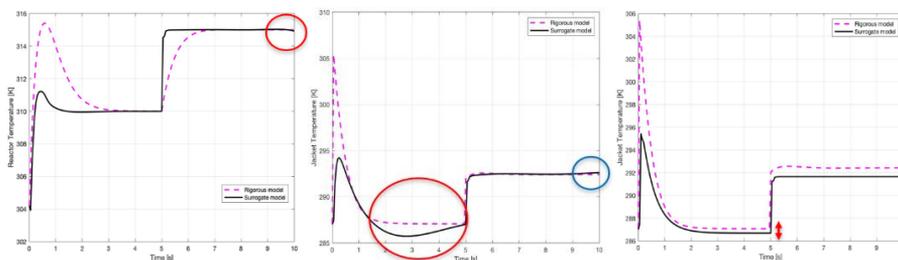


Figure 3 – Controlled/manipulated variable for the three approaches

Once the surrogate models according to each of the three approaches have been derived, their closed-loop response to two sequential 5 K perturbations on the inlet reactor temperature have been tested. Figure 3 shows, for each of them, the controlled or manipulated variable behaviour according to the one that is more significant for the model performances.

As it can be noticed, in the first case, the model is able to reproduce with some inaccuracies the dynamic trend by respecting the settling time and the overshoot for the second perturbation in particular. However, in proximity of the second steady state, the curve starts dropping down to the impact of the exponential term of the obtained analytical function. In order to mitigate this behaviour, the second approach increases the number of points in proximity of the domain boundary. In this case, the steady-state issue has been fixed but the cost to pay is a higher inaccuracy in the central part of the time domain, in particular for the manipulated variable trend as showed in Figure 3b.

Finally, the first order regression performances were tested. The obtained trend exhibits good agreement with the rigorous model one both in terms of transient quality and steady state behaviour. However, an offset with respect to the actual steady state values can be noticed due to the fact that the functional form of the surrogate model has been fixed and includes, as a function of time, the exponential term only.

5. Conclusions

As a first result, the study shows that surrogate modeling to describe continuous processes dynamics is possible and effective by means of the currently available tools. In particular, the obtained model is able to describe the system performances both in open- and closed-loop behaviour. However, although the most critical aspects such as transient peaks and steady state value are correctly described, the accuracy of all points in the perturbation time interval is not always ensured. To deal with these inaccuracies different modeling and sampling approaches can be employed according to the specific parameter of interest. As concerns the computational aspects, with respect to the set of differential equation, the model obtained by means of the proposed procedure exhibits a solution time over the defined time span that is lower by an order of magnitude and a relative error lower than the 0.5 %. Moreover, it allows to detect the best compromise between solution accuracy and computational time in terms of sample size for the specific case study.

In conclusion, the research work shows good potential for continuous process modeling by means of data-driven approach and it is worth a deeper investigation in terms of both modeling strategies and sampling in order to become a well-established tool in the process systems engineering domain and also further applications in different research fields.

References

- A. Bhosekar, M. Ierapetritou, 2018. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Computers & Chemical Engineering* 108, 250–267.
- A. Di Pretoro, B. Bruns, S. Negny, M. Grünewald, J. Riese, 2022. Demand Response Scheduling Using Derivative-Based Dynamic Surrogate Models. *Computers & Chemical Engineering*, 107711.
- K. McBride, K. Sundmacher, 2019. Overview of Surrogate Modeling in Chemical Process Engineering. *Chemie Ingenieur Technik* 91, 228–239.
- H. Shi, F. You, 2015. Adaptive surrogate-based algorithm for integrated scheduling and dynamic optimization of sequential batch processes. 54th IEEE Conference on Decision and Control (CDC) proceedings, Osaka, 7304-7309.
- A. Shokry, P. Baraldi, E. Zio, A. Espuña, 2020. Dynamic Surrogate Modeling for Multistep-ahead Prediction of Multivariate Nonlinear Chemical Processes. *Ind. Eng. Chem. Res.* 59, 15634–15655.

Probabilistic machine learning based soft-sensors for product quality prediction in batch processes

Max Mowbray^a, Aaron Hicks^a, Harry Kay^a, Sam Kay^a, Amanda Lane^b, Cesar Mendoza^b, Philip Martin^a and Dongda Zhang^{a,*}

^a*Department of Chemical Engineering and Analytical Science, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

^b*Unilever Research Port Sunlight, Quarry Rd East, Bebington, CH63 3JW, UK
dongda.zhang@manchester.ac.uk*

Abstract

Statistical machine learning algorithms have been widely used to analyse industrial data for batch process monitoring and control. In this study, we develop a three-step methodology to identify, visualize and systematically reduce data dimensionality for the construction of robust soft-sensors for end-product quality prediction. The approach first employs partial least squares to screen the entire dataset and identify critical time regions and operational variables, then adopts multiway partial least squares to construct a latent space descriptive of the existing batches. Nonlinear estimators are then constructed from the reduced latent space to estimate final product quality, which is able to express model uncertainty. Specifically, in this study, we explore the performance of Gaussian processes. Innovations of this approach include the ease of data visualisation and ability to identify major operational activities within the factory, as well as robustly predict end-quality. To highlight efficiency and practical benefits, an industrial consumer goods product manufacturing process was presented as an example and the soft sensor was successfully constructed and cross validated. Furthermore, the accuracy, reliability, and interpretability of the soft-sensor is discussed, tested and shown to generalise well.

Keywords: Machine Learning, Batch process, Soft-sensor, Dimensionality reduction, Viscosity prediction, Interpretability

1. Introduction

The operation of nonlinear, uncertain batch processes is a well established domain of research within process systems engineering. Batch process data tends to consist of rank-3 tensorial datasets which are highly dimensional. Further, given that the process is batch, it tends to be the case that operation's primary concern is to ensure that end-quality is kept on specification (meaning that one typically only has a single measurement of the desired qualities of a given batch). As a result, the construction of end-quality predictive models for process monitoring and operation is a highly complex process, requiring identification of different dynamical regimes (handling of non-stationarity in the data) and highly dimensional data points that typically exhibit multicollinearity. This is combined with the fact that operational data is often reasonably limited (at least relative to the era of big data) and the identification of physical mechanisms (mathematically) is extremely challenging. Conventionally, the problem has been handled by various dimensionality reduction techniques (see e.g. Wold et al. (2009)). However, proper quantification of the uncertainties associated with latent variable models is challenging as discussed in Zhang and Garcia-Munoz (2009).

As a result, in this work we consider the development of probabilistic latent variable models, which can naturally quantify the uncertainty of their prediction. This is important because not only is the nature of end-quality prediction in batch processing of high uncertainty, but such quantification can help inform operation of the batch to mitigate waste and reduce inefficiencies.

2. Methodology

2.1. Model construction with rank-3 tensorial input data

In this work, we assume the availability of a dataset, $\mathcal{D} = \{\mathcal{X}, Y\}$, composed of N batch runs, with J process variables recorded at T discrete time intervals, such that $\mathcal{X} = [X_1, \dots, X_N]^T \in \mathbb{R}^{N \times J \times T}$ and $Y = [y_1, \dots, y_N] \in \mathbb{R}^{N \times n_y}$, where n_y represents the number of process qualities one would like to monitor. In this work, we would like to identify a predictive mapping expressed as follows:

$$\mathbf{y} = f(X, \theta) \quad (1)$$

where $\theta \in \mathbb{R}^{n_\theta}$ are the parameters of the functional mapping, $f(\cdot, \theta)$, defined generally. The input data, \mathcal{X} , represents a rank-3 tensorial dataset, which generally poses challenge to constructing classical predictive models. A number of techniques exist to handle such data (that all aim to obtain a rank-2 tensorial representation), including convolutional neural networks, and a number of tensorial analysis approaches as outlined in Sun and Braatz (2020). Further, it is common in the paradigm of batch processing for datasets to exhibit multi-collinearity, such that the problem of identifying the mapping (in Eq. 1) is ill-posed. As a result, in this work, we use the method developed in Hicks et al. (2021) with inspiration from Nomikos and MacGregor (1995, 1994) to construct a reduced latent representation of a batch, such that the dataset \mathcal{X} can be well described by $Z = [\mathbf{z}_1, \dots, \mathbf{z}_N]$, where $\mathbf{z}_i \in \mathbb{R}^{n_z}$ and $i \in \{1, \dots, N\}$. This reduced latent representation, Z can be used for model construction as outlined subsequently. Specifically, we propose to identify Z via multi-way projection to latent structures (MPLS), which can be thought as a special case of projection to latent structures (PLS). Essentially, the additional component of MPLS (relative to vanilla PLS) is the transformation (or unfolding) of a rank-3 tensorial dataset via a functional transformation, f_{unf} , to obtain a rank-2 matrix. In this work, we use time-wise unfolding, which may be defined as $f_{unf} : \mathbb{R}^{N \times J \times T} \rightarrow \mathbb{R}^{N \times JT}$. Having obtained an unfolded representation of the dataset, $\hat{X} = f_{unf}(\mathcal{X})$, one can then identify Z that correlates with the variables one desires to predict via PLS:

$$\begin{aligned} \hat{X} &= ZP^T + E \\ Y &= UQ^T + F \end{aligned} \quad (2)$$

where $P \in \mathbb{R}^{JT \times n_z}$ are the *input loadings*, which project \hat{X} to the input latent space; $Q \in \mathbb{R}^{n_y \times n_u}$ are the *output loadings*, which project Y to the output latent space; $U = [\mathbf{u}_1, \dots, \mathbf{u}_N]^T \in \mathbb{R}^{N \times n_u}$ are the *output scores*; and, $E \in \mathbb{R}^{N \times JT}$ and $F \in \mathbb{R}^{N \times n_y}$ quantify the information loss from projecting both \hat{X} and Y to their respective latent spaces. The latent spaces themselves are related via a linear regression, which enables formation of a predictive model (which together with an unfolding mechanism enables identification of a model of the form of Eq. 1).

Although PLS is a linear decomposition, the relationship expressed within the latent spaces can often be nonlinear. This means that PLS may inherently underfit the modelling problem at hand. Further, approximation of the model uncertainties is a difficult problem given that the typical methods for uncertainty estimation do not consider the indirect relationship of model input and output via the latent structures. This was studied in Zhang and Garcia-Munoz (2009). To handle these two facets, we propose to a nonlinear estimation from the latent space, U to the target Y , to identify a model which naturally expresses predictive uncertainty via probabilistic inference. A suitable estimator is discussed in the following.

2.2. Nonlinear Estimators for Probabilistic Inference and Gaussian Processes

Here, we briefly formalise a general introduction to nonlinear estimators which lend naturally to probabilistic inference. Specifically, we can consider probabilistic inference as the construction of a conditional probability distribution function over predictions, \mathbf{y} , given a model input, \mathbf{z} . In this case, we would like to identify a prediction as follows:

$$\mathbf{y} \sim p(\mathbf{y}|\mathbf{z}, \cdot) \quad (3)$$

Specifically, we consider the use of Gaussian processes (GPs), the high level intuition for which is provided in the following. The use of GPs is particularly appealing as the predictive distribution constructed (in Eq. 3) expresses both aleatoric (arising due to the underlying process) and epistemic (due to a lack of information) uncertainties in closed form. In short, GPs operate within a nonparametric, Bayesian inference framework. The idea here is to (instead of identifying a fixed model structure and number of parameters) simply exploit the statistical relationships within the data to identify a function, $f: \mathbb{R}^{n_z} \rightarrow \mathbb{R}$, (i.e. an infinite dimensional weight vector) such that we can make predictions, $\mathbf{f} = [f(\mathbf{z}_1), \dots, f(\mathbf{z}_N)]$, where $f(\mathbf{z}_i) \in \mathbb{R}$, by simply querying the function at given model inputs, Z . Formally, however, GP models are a subset of stochastic process models (SPs). SPs define a probability model over a collection of random variables, such that any finite subset of the random variables have a joint distribution - meaning they are often identified as a distribution over functions. When this joint distribution is assumed Gaussian, one obtains a GP. A GP is fully specified by a mean, $\mathbf{m}(\cdot)$, and covariance function, $k(\cdot, \cdot)$. Many covariance functions exist, but all are constituted by some hyperparameters, which we define generally as $\boldsymbol{\lambda}$, such that the covariance function may be denoted $k(\mathbf{z}, \mathbf{z}'; \boldsymbol{\lambda})$. Selection of the covariance function, and the associated hyperparameters, defines the behaviour of the GP in function space, otherwise known as the *prior*, $p(\mathbf{f}|Z, \boldsymbol{\lambda})$:

$$\begin{aligned} f(\mathbf{z}) &\sim p(\mathbf{f}|Z, \boldsymbol{\lambda}) \\ p(\mathbf{f}|Z, \boldsymbol{\lambda}) &= \mathcal{G} \mathcal{P}(\mathbf{m}(\mathbf{z}), k(\mathbf{z}, \mathbf{z}'; \boldsymbol{\lambda})) \end{aligned} \quad (4)$$

The mention of a prior, leads us nicely into discussion regarding probabilistic inference in GPs. As mentioned, inference leverages a Bayesian framework, which allows us to directly exploit the statistical relationships in data to infer $f(\mathbf{z}_i) \approx y \in \mathbb{R}$ (i.e. a product end-quality). Therefore, we may write a variant of Bayes' rule for inference of a latent coordinate at the training points in GPs as follows:

$$p(\mathbf{f}^j|Y^j, Z, \boldsymbol{\lambda}) = \frac{p(Y^j|Z, \mathbf{f}, \boldsymbol{\lambda})p(\mathbf{f}|Z, \boldsymbol{\lambda})}{p(Y^j|Z, \boldsymbol{\lambda})} \quad (5)$$

where $Y^j \in \mathbb{R}^N$, $j \in \{0, \dots, n_y - 1\}$ denotes the j^{th} column of Y ; $p(\mathbf{f}^j|Y^j, Z, \boldsymbol{\lambda})$ is known as the *posterior* predictive distribution; $p(Y^j|Z, \mathbf{f}, \boldsymbol{\lambda})$ is the likelihood (formalised as conditional to the covariance function hyperparameters); and, $p(Y^j|Z, \boldsymbol{\lambda})$ is the marginal likelihood. Due to the dependence on $\boldsymbol{\lambda}$, it is important to properly identify $\boldsymbol{\lambda}$. This is achieved by maximisation of the marginal log-likelihood, $p(Y^j|Z, \boldsymbol{\lambda})$. If a homoscedastic Gaussian additive noise model is chosen in construction of the likelihood term, and given that Gaussians are closed under both conditioning and marginalisation, we can construct the *posterior* exactly as a Gaussian distribution, such that at a new test point \mathbf{z}^* , a posterior predictive distribution over function values, f^* , may be constructed as:

$$p(f^*|Y^j, Z, \mathbf{z}^*, \boldsymbol{\lambda}) = \mathcal{N}(\mu(\mathbf{z}^*; Y^j, Z, \boldsymbol{\lambda}), \sigma(\mathbf{z}^*; Y^j, Z, \boldsymbol{\lambda})) \quad (6)$$

As GPs are multiple-input, single-output models, if one has an n_y dimensional latent representation of desired process qualities to predict, then n_y separate GP models can be constructed with the

predictions aggregated as follows:

$$\begin{aligned}
 \boldsymbol{\mu}(\mathbf{z}^*; Y, Z, \cdot) &= [\boldsymbol{\mu}_1(\mathbf{z}^*; Y^0, Z, \cdot), \dots, \boldsymbol{\mu}_{n_y}(\mathbf{z}^*; Y^{n_y-1}, Z, \cdot)] \\
 \boldsymbol{\Sigma}(\mathbf{z}^*; Y, Z, \cdot) &= \text{diag}(\boldsymbol{\sigma}_1^2(\mathbf{z}^*; Y^0, Z, \cdot), \dots, \boldsymbol{\sigma}_{n_y}^2(\mathbf{z}^*; Y^{n_y-1}, Z, \cdot)) \\
 p(\mathbf{y}^* | \mathbf{z}^*, \cdot) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\
 \mathbf{y}^* &\sim p(\mathbf{y}^* | \mathbf{z}^*, \cdot)
 \end{aligned} \tag{7}$$

2.3. Contribution

By integrating the concepts outlined in this Section, one can identify a reduced latent representation of the batch trajectory, which correlates to the desired end-qualities to predict. By constructing nonlinear estimators from the latent space, Z , one can make predictions given a new test batch, $X^* \in \mathbb{R}^{1 \times J \times \hat{T}}$, as follows:

$$\mathbf{y}^* \sim p(\mathbf{y}^* | f_{unf}(X^*)P, \cdot) \tag{8}$$

Given, we identify our mapping from the latent space to end-batch quality as a Gaussian process, we can write $\mathbf{y}^* \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, which describes both the expected end-quality given the trajectory information of the batch, as well as a variance which is quantitative of the aleatoric and epistemic uncertainty associated with the prediction.

3. Case Study

3.1. Consumer goods product-quality prediction

The work developed in this paper focuses on product quality prediction in batch processes. The feasibility and performance of the proposed soft-sensors is assessed by predicting the viscosity of a consumer good product with data obtained from a production plant. This data belongs to two different product specifications with similar processing procedure. Specifically, the data was generated from real-time monitoring of the process as well as one off-line measurement of end-viscosity. These measurements are performed with a rheometer, and can take upwards of 20 min to be completed, which can potentially make errors irreversible. As is demonstrated subsequently, a possible solution is to substitute periodical off-line viscosity measurements with on-line final viscosity prediction using live process data.

3.2. Datasets

The first dataset used to develop the soft-sensor belongs to the first product specification. It comprises 30 individual batches that follow the same recipe, and have the same target viscosity. Deviations from the set target are due to operation errors and uncertainty in the final viscosity measurements. This dataset was used to train all the models. In the first instance, model performance was assessed using leave-two-out cross-validation on dataset α . Two further datasets, β and γ , were used for testing. Dataset β originates from the specification as dataset α and has 16 batches, while dataset γ belongs to a different product specification and is composed of 11 batches. Hence datasets β and α represent the same process. As for dataset γ , test predictions assess the generalisation capacity of the model, since γ belongs to a different specification of product.

To ensure that all datasets could be used for prediction, heterogeneity was standardised via a process of batch alignment and critical time regions and process variables were selected to constitute \mathcal{X} (for the respective datasets). This was conducted via the framework outlined in Hicks et al. (2021), which is based on finding datapoints that correlate strongly with the process quality one would like to predict. For all datasets, critical data regions were identified such that $J = 8$ and $T = 300$, reducing the datasets from an original dimensionality of 7000 measurements of 30 different process variables.

3.3. Metrics for Model Selection and Evaluation

The evaluation and comparison of the different models' performance was based on the following metrics that are required to quantify the quality of the predictive distribution constructed by the models. The accuracy of the models was calculated with the mean average percentage error (MAPE) with respect to the real viscosity as seen in Eq. 9:

$$MAPE = \frac{|y - \mu(X^*, \cdot)|}{y} \times 100 \quad (9)$$

where $|\cdot|$ denotes the absolute value; y is the real viscosity measurement and $\mu(\mathbf{z})$ corresponds to the mean of the predictive distribution over viscosity values given a latent coordinate, \mathbf{z} , as predicted by the models constructed in this work. The scaled coefficient of variation, σ_c , is denoted:

$$\sigma_c = \frac{c\sigma(\mathbf{z})}{\mu(X^*, \cdot)} \times 100 \quad (10)$$

where $c \in \mathbb{R}$ is defined to quantify the normalised spread of potential viscosity values observed under the predictive distribution modelled, with a given probability (i.e. $c = 3$ indicates the spread of potential viscosity with probability 0.997); $\sigma(X^*, \cdot) = \sqrt{\Sigma(X^*, \cdot)}$ is the standard deviation of the predictive distribution identified. In the region of the latent spaces that one has data, this should be reflective of the aleatoric uncertainty of the data generation process and e.g. represent variability in the measurement of viscosity and operational error. For the underlying process it has been estimated that this aleatoric uncertainty should be in the range of $\sigma_c = 26\%$. The final metric used in this work, is the coverage probability, CP :

$$CP = \mathbb{P}(|y - \mu(\mathbf{z})| \leq c\sigma(\mathbf{z})) \quad (11)$$

As in Zhang and Garcia-Munoz (2009), CP is defined with c and hence the inequality should be satisfied with a hypothetical probability (i.e. $c = 3$ indicates $CP = 0.997$, as used in this work).

4. Results and Discussion

Given the nonparametric nature of Gaussian process models, determination of model structure primarily considered the number of latent variables used with the MPLS model input latent space. The results of the cross validation for the model are expressed by Fig. 1. From Fig. 1, one can see the optimal number of latent variables (PCs) in the latent space is around 2. At this PC, there is the lowest validation MAPE ($\approx 10\%$), $\sigma_c \approx 26\%$, which well represents the expected variation of the process, and $CP = 0.95$, which considering the finite number of predictions is close to the hypothetical value desired.

The results obtained in model testing are displayed by Table 1. The results demonstrate the ability of the model identified to obtain impressive predictive accuracy (via low MAPE $\approx 10\%$) and identify predictive distributions, which well represent the aleatoric and epistemic uncertainties (quantified via $CP \approx 0.9$). The framework has particular utility, because it is able to identify predictive uncertainties (innately) in closed form. Further, combining Bayesian inference with latent space modeling, interpretability is provided in the form of both the uncertainty prediction, but also clustering and loading type analyses, which together are able to jointly inform operators when the process is displaying dynamical regimes not commonly observed (i.e. when operational error is present) and for what reason. This enables efficient process monitoring and an online indicator if operators should take action and how best to take it.

5. Conclusions

In this work, we have provided a framework for the identification of a robust soft sensor for end-quality prediction in batch processes. The soft sensor is able to well identify a predictive

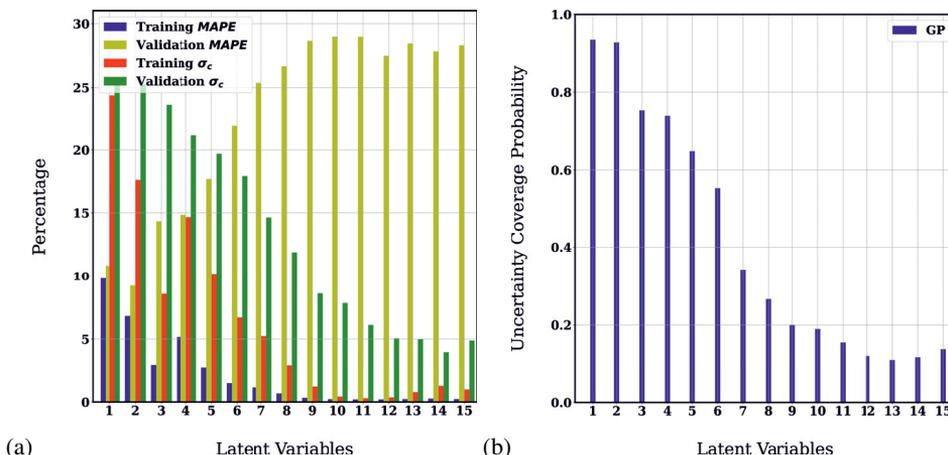


Figure 1: a) Average training and validation MAPE (bar 1 and 2, respectively) and σ_c (bar 3 and 4, respectively) over all cross validation folds as latent variables are added to the MPLS latent space. b) Evolution of coverage probability with the addition of latent variables.

Table 1: Results of predictive tests on datasets β and γ .

Dataset	MAPE (%)	σ_c (%)	CP
β	10.0	22.46	0.88
γ	11.4	22.90	0.83

distribution that quantifies both the expected end-quality (quantified via MAPE in the region of 10% on both validation and test predictions), as well as a closed form uncertainty prediction, which represents the underlying process variation and epistemic uncertainty of the model (quantified via the coverage probability and coefficient of variation). Combining both Bayesian inference and latent space modelling enables interpretability of the predictions, and could help inform operators if and how to take action. We hope to implement this framework to a real process in the scope of process monitoring.

References

- A. Hicks, M. Johnston, M. Mowbray, M. Barton, A. Lane, C. Mendoza, P. Martin, D. Zhang, 2021. A two-step multivariate statistical learning approach for batch process soft sensing. *Digital Chemical Engineering*, 100003.
- P. Nomikos, J. F. MacGregor, 1994. Monitoring batch processes using multiway principal component analysis. *AIChE Journal* 40 (8), 1361–1375.
- P. Nomikos, J. F. MacGregor, 1995. Multi-way partial least squares in monitoring batch processes. *Chemometrics and intelligent laboratory systems* 30 (1), 97–108.
- W. Sun, R. D. Braatz, 2020. Opportunities in tensorial data analytics for chemical and biological manufacturing processes. *Computers & Chemical Engineering*, 107099.
- S. Wold, N. Kettaneh-Wold, J. MacGregor, K. Dunn, 2009. Batch process modeling and mspc.
- L. Zhang, S. Garcia-Munoz, 2009. A comparison of different methods to estimate prediction uncertainty using partial least squares (pls): a practitioner’s perspective. *Chemometrics and intelligent laboratory systems* 97 (2), 152–158.

Property Estimation Method for Cannabinoids and Terpenes Using Machine Learning

Laura A. Vergara,^a Héctor J. Hortúa,^b Gustavo A. Orozco^a

^a*Grupo de Procesos Químicos y Bioquímicos, Departamento de Ingeniería y Ambiental, Universidad Nacional de Colombia. Cra 30#45-03, Bogotá, Colombia*

^b*Semillero de investigación en Data Science, Escuela de Ciencias Básicas, Tecnología e Ingeniería, Universidad Nacional Abierta y a Distancia, Bogotá, Colombia*

Abstract

Nowadays, medical cannabis has great importance as a coadjuvant in the treatment of some pathologies as glaucoma, rheumatoid arthritis, HIV, Alzheimer's, asthma, cancer, chronic pain that is difficult to control, Crohn's disease, epilepsy, multiple sclerosis, insomnia, and Parkinson's; additionally, cannabinoids and terpenes have important properties that need to be considered and studied in a deeper way to understand the way these compounds work. The main goal of this study is to develop a new hybrid methodology based on both contribution groups and machine learning algorithms. In this particular case, the study is focused on estimating properties of cannabinoids and terpenes whose experimental information is really scarce and, in most cases, not yet reported. To do so, a database of hundreds of thousands of molecules which includes different thermodynamic properties have been considered. Our method is able to estimate different properties such as: boiling point, melting point, vapor pressure, viscosity, and vaporization enthalpy. Machine learning algorithms have been used to establish the contribution of every functional group and its prediction capability has been compared with other well-known methods like the Joback and Reid (1987) method, Constantinou and Gani (1994) method, among others. One of the major findings of the present study is the fact that the current estimation methods are not adequate enough for cannabinoids and terpenes, so there is an important need to find new and more precise ones that allow to improve the accuracy in the groups as well as to extend the predictions to a bigger set of chemical groups. This study clearly shows the importance and utility of the machine learning methods in one of the most relevant chemical engineering areas such as the properties estimation one. This study clearly shows the benefits of machine learning techniques in chemical engineering applications.

Keywords: Property Estimation Method, functional groups, Machine Learning, Neural Networks, Cannabinoids.

1. Introduction

In the last years the interest in studying the cannabis plant has increased notoriously in different fields due to the different pharmaceutical properties of several of its chemical compounds, for instance it is known cannabinoids are of great importance as coadjuvants in the treatment of different pathologies as cancer, epilepsy, and multiple sclerosis Abyadeh et al. (2021). Besides cannabinoids, cannabis plants also contain several compounds such as flavonoids, fats, and terpenes. The latter have biological properties that make them potential mechanism against different diseases, including cancer chemo preventive effects, antimicrobial, antifungal, antiviral, and antiparasitic activities Paduch et al. (2007). Despite the great range of applications, the thermodynamics properties of

cannabinoids and terpenes are difficult to find in the literature or in specialized databases. Therefore, as a first cheap and fast alternative predictive tool such as the based ones on group contribution methods can be used in order to estimate them based on experimental data of other molecules, in particular due to their importance in the crystallization and distillation process it is relevant to study properties such as melting and boiling points, densities, viscosities, among others. This work is aimed to study the melting point of several cannabinoids and terpenes using a group contribution strategy along with machine learning techniques. The basic idea when using group contribution methods, once a set of chemical groups is defined, all molecules belonging to the database are split accordingly allowing to establish a system of algebraic equations. Finally, an optimization method is used in order to minimize the error between predictions and experimental data. Several contribution methods have been proposed for different purposes in the property's estimation field. One of the most famous was proposed by Joback and Reid (1987), this method is able to estimate critical points, melting temperature and normal boiling temperature, it is commonly used for relatively simple molecules and is based on 41 molecular groups. Additionally, Constantinou and Gani (1994) made the estimation of critical properties of pure organic compounds, which due to its sophistication promises better results than Joback's method, as the estimation is performed at two levels: the basic level that uses contributions from first-order groups, and the next higher level that uses a small set of second-order groups having the first-order groups as building blocks. The predictions, first and second order approximations, are based on the molecular structure of the compounds, and is able to make the distinction between isomers. Recently, other studies unified the group contribution methods with artificial neural networks. For instance, Valderrama et al. (2015) created this type of hybrid method for the estimation of the density of ionic liquids. For this purpose, an experimental database of 399 data for 100 ionic liquids was used for the network training. In order to discriminate between different isomers, the molecular mass and the structure of the molecule were given as input variables. Additionally, Gharagheizi et al. (2011) proposed an artificial neural network-group (ANN-GC) contribution method in order to determine surface tension of pure compounds at different temperatures and pressures. They used 4700 data belonging to experimental surface tension values of, approximately, 750 chemical compounds and 151 functional groups at different conditions. Although several methods have been proposed, none of them is able to reproduce the experimental data of cannabinoids or terpenes, additionally since the experimental information is scarce for these compounds it is certainly to propose a methodology to estimate them, in this particular a hybrid strategy of group contribution methods along with machine learning techniques and a big set of experimental data will be used. The manuscript will be presented as follows, first in the computational section, the database generation for and the group contribution method is described, second the neural network algorithms are presented, third the results are presented and finally the conclusion and future work is presented. The code and the best predicted model will be available after publication.

2. Methodology

2.1. Database features

In order to build the machine learning algorithms, a database with the property data was needed to train the method. To do so, an open database available on the U.S. Environmental Protection Agency (EPA) was used, this database contains the chemical name, the Simplified Molecular Input Line Entry Specification (SMILES) and different chemical and physical properties such as the melting point, the boiling point among many

others. This database has, approximately, 700,000 compounds with their respective melting point, as it is the property being studied. Additionally, the database created has the functional groups present in each compound, which are group of atoms that have their specific characteristics and contribute to the properties of the compound. The functional groups were obtained using RDKit, a library created for Python, which prints the functional groups of each compound based on its SMILES. Additionally, to evaluate the model a database of 50 cannabinoids and terpenes was generated, the information was obtained from the Cannabis Compound Database Version 1.0, a freely available electronic database containing detailed information about small molecules found in Cannabis sativa, Cannabis indica and Cannabis hybrids (Cannabis Database). The functional groups of the cannabinoids and terpenes were also found using RDKit and the compounds SMILES. Finally, it is important to mention that, as this was the database used for the evaluation of the model, the melting point was not included in it.

2.2. Machine learning algorithms

In this subsection, we introduce the concept of Artificial Neural Networks (ANN), and the process of its training and evaluation. ANN was inspired by the structure and functions of biological neurons, and it generally consists of an input layer where data is fed, the hidden layers standing for the internal structure of the model where information goes through until arrive at the output layer Schmidhuber (2015). Each layer comprises neurons that transmit the information downstream to connected neurons belonging to other layers. The previous layer acts as an input for the next one after applying nonlinear activation on it, and finally propagates the current result to the following layer. In a vector notation we have Wang et al. (2020):

$$h_{i+1} = W_{i+1}l_i + b_{i+1}; \quad l_{i+1} = f(h_{i+1}) \quad (1)$$

where l_i is the input row vector of the $i - th$ layer, W_{i+1} and b_{i+1} are the weights and biases trainable parameters, h_{i+1} is the intermediate vector after linear transformation, and f is the activation function. The goal of training an ANN is to minimize the difference between the predicted value given by the output layer $\hat{y} = f_{W,b}(h)$ and the ground truth y . The minimization is described through the loss function L , by optimizing the parameters W and b . In regression tasks Schmidhuber (2015), it is common to use L as the mean squared error (MSE), the mean absolute error (MAE) or the Huber loss (H):

$$MSE = \frac{1}{N} \sum (y - \hat{y})^2, \quad MAE = \frac{1}{N} \sum |y - \hat{y}|, \quad H_\delta(y, \hat{y}) = \begin{cases} 0.5(y - \hat{y}), & |y - \hat{y}| \leq \delta \\ \delta(|y - \hat{y}| - 0.5\delta), & \text{other } w \end{cases} \quad (2)$$

where N stands for the number of samples. The aforementioned functions will be adapted to this work. Following the chain rule, one could transmit gradients of the trainable parameters in the $i - th$ layer from the $(i + 1)th$ layer. During backward propagation the parameters are updated, then forward propagation is performed again. This is repeated until the loss function achieves the desired precision and thus, the NN is trained and ready to make predictions. Two types of artificial neural networks are implemented in this work: deep Feedforward Neural Networks (FNN) and Convolutional (Conv1D). The FNN described above does not have cycles contained within it and the connections between layers and the information flow are straightforward, while the Conv1D is a structured network with multiple sets of weights (filters) that "slide" or convolve across the input-space to analyze distance-pixel relationship opposed to individual node activation in the FNN Schmidhuber (2015). On the other hand, the dataset is split in training 70%, validation 15%, and test 15% sets. The first set is used to train the ANN, while the

validation and test set contain unseen values, relevant for testing the performance of the model. The latter can be achieved using a metric that in most of cases is similar to the loss function shown in Eq.(2) or thought the coefficient of determination (R^2) Hortúa et al. (2020):

$$R^2 = 1 - \frac{\sum_i(\hat{y} - y_i)^2}{\sum_i(y_i - \bar{y})^2}, \quad (3)$$

where \bar{y} is the average of the true parameters and the summations are performed over the entire test set. R^2 ranges from 0 to 1, where 1 represents perfect inference. R^2 will be used in this work as metric for measuring the performance of the NN models. Finally, a downside behind of the standard NNs is that they are sensitive to the training process and the hyper-parameter choices, finding different set of weights each time they are trained, which in turn produce different predictions and high variance. An attempt approach to reduce this variance is to train multiple models instead of a single one and to combine the predictions from these models. This is called ensemble learning and it helps to reduce the variance of predictions and can also result in better performance Tao (2019).

3. Analysis

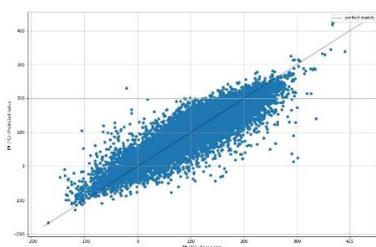
In order to predict the melting point in terms of the establish chemical groups, we started with single NN regression models. First, we built in TensorFlow a FNN consists in four dense layers with 256, 128, 64 and 1 neurons, followed by an ELU activation function, a Batch Normalization in order to optimize and accelerate the convergence, and a Dropout layer (with a rate of 0.1), for regularization purposes. Here, we adopt Adam as the optimizer, and a learning rate decay with 0.01 in its initial value, with a factor of 0.8 during each 10 epochs. On the other hand, the conv1D architecture comprises of four convolutional layers of 64, 32, 32, 16 filters, then a MaxPool layer is used to reduce the dimensions of the feature maps, followed by four dense layers of 256,126,64 and 1 neurons. As before, each layer is followed by an ELU activation function, Batch Normalization and Dropout (rate 0.1).

Table 1. Performance for different models in terms of R^2

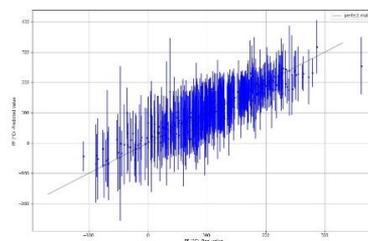
Metric/Model	Ensemble			Single		
	MSE	MAE	Huber	MSE	MAE	Huber
R^2 -FNN	0.88	0.84	0.84	0.83	0.82	0.82
R^2 -Conv1D	0.87	0.87	0.87	0.81	0.80	0.83

The results of the performance for both, FNN and Conv1D for different loss functions are reported in the right hand side of the Table 1. As we observed, both models seem to work well, reaching a decent coefficient of determination value. FNN model trained with MSE is the best model that we obtained through all single NN experiments, and the behavior for predicting unseen instances can be seen also in Fig.1a. Finally, we also worked with ensembles of Neural Networks (BNN) in order for including variances during the inference process, and improve the estimates for predictions. The schema for these ensembles with either FNN or Conv1D are displayed in Fig.(2)-(3) respectively. For these ensembles we added at the top of the models, an Average-layer which provides of an unique neuron for predicting the melting point. Furthermore, we could extract the information from the top-before layer to get the outputs for all the models in the ensemble and thus, to be able to compute the standard deviation for the predicted values. As we can

see in the left part of the Table 1, the ensemble models outperform the results that we have obtained in the single scenario. Indeed, we achieved coefficient of determination around 0.8 for the FNN ensemble trained with MSE loss function. The prediction for the test dataset can be also seen in Fig.1b. Here, we observed not only a better accuracy in the predictions, but also, we acquired new information about the uncertainty of those estimations. One of the main analysis that we can get from the ensemble models, is that there is a huge variance for chemical groups with melting points around 10-100 Celsius degrees, basically because of the few dataset in this range. Also, we can argue that five models in the ensemble is not good enough, and it requires more single NN for yield better results in terms of variability. These additional remarks along with the use of models which provide uncertainties such as Bayesian Neural Networks Hortúa et al. (2020) or Gaussian Processes are tools that we are proposing for a future work.



(a) Predicted values against true values for a single FNN model.



(b) Predicted values for an ensemble consisting of 5 FNN models.

Figure 1. Predicted vs true values for the best experiments found for a single and ensemble NN

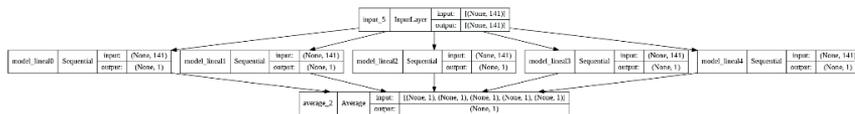


Figure 2. NN architecture used in the work based on ensembles of several NN Fully-connected models.

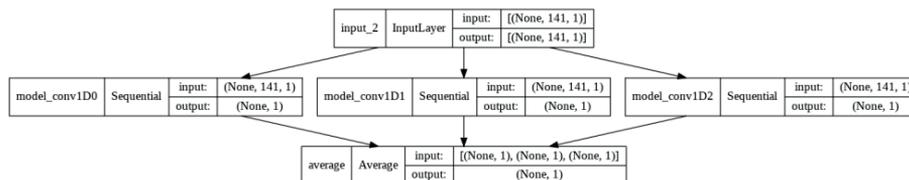


Figure 2. NN architecture used in the work based on ensembles of several NN convolutional models.

Table 2. Prediction provided by the ensemble FNN model for different cannabinoids

Compound Name	Predicted Value (°C)	Reported value (°C)	Standard Deviation
Mircene	-70.0	<-10.0	48.4
Apha-pinene	-30.8	-62.0	26.8
Menthol	87.0	36.0-38.0	36.0
Cannabidiol	152.4	66.0	39.2
Cannabigerol	152.3	49.0-52.0	36.4
Beta-pinene	-53.1	-61.5	35.4
Limonene	-45.8	-74.0	35.8

Table 2 summarizes the predictions for some of the studied cannabinoids, not previously used by the algorithm. Standard deviations are also reported. As shown: i) qualitatively speaking the model is able to predict positive and negative melting points, however, it still needs a lot of improvement in compounds such as cannabidiol, one of the most important in the current processes; ii) the obtained standard deviations are very high, indicating that a pre-treatment process is certainly necessary to improve the model.

4. Discussion and Conclusions

In this paper, we have explored several machine learning techniques for estimating the melting point property in chemical compounds. We contrasted two main approaches: the use of trained single NN models based on FNN and Convolutional layers, along with ensembles comprised of those single NN. We have found that these ensembles outperform single NN, and additionally provide uncertainties in its predictions. The best model found in this paper comes from the ensemble of FNN trained with MSE loss function, reaching a $R^2 = 0.88$. Even if the results are good enough and promising, we observed that the variance for the models built so far is huge. This effect can be associated with either a noisy dataset or because of few instances in the dataset for some temperature ranges. In order to overcome this issue, we suggest for future work, a strong reprocessing method before feeding the data into the NN model, and also the use of more robust techniques such as Bayesian Neural Network in order to get the aleatory and epistemic uncertainties which give a clearer picture about the huge variability in the predictions. Finally, we are currently working on improving the accuracy of this model as well as determining more thermodynamic properties of interest in the cannabinoid extraction process.

References

- M. Abyadeh, V. Gupta, J. A. Paulo, V. Gupta, N. Chitranshi, A. Godinez, M. H. D. Saks, M. M. A. Amirkhani, G. H. Salekdeh, P. A. Haynes, S. L. Graham, M. Mirzaei., 2021. A proteomic view of cellular and molecular effects of cannabis. *Biomolecules* 11, 1411.
- L. Constantinou, R. Gani, 1994. New group contribution method for estimating properties of pure compounds. *AIChE J.* 40, 1697–1710.
- F. Gharagheizi, A. Eslamimanesh, A. H. Mohammadi, D. Richon, 2011. Use of artificial neural network-group contribution method to determine surface tension of pure compounds. *J. Chem. Eng. Data* 56, 2587–2601.
- H. J. Hortúa, L. Malagò, R. Volpi, Sep 2020. Constraining the reionization history using bayesian normalizing flows. *Machine Learning: Science and Technology* 1 (3), 035014. URL <http://dx.doi.org/10.1088/2632-2153/aba6f1>
- K. G. Joback, R. C. Reid, 1987. Estimation of pure-component properties from group-contributions. *Chem. Eng. Commun.* 57, 233–243.
- R. Paduch, M. Kandefor-Szerszen, M. and Trytek, J. Fiedurek, 2007. Terpenes: substances useful in human healthcare. *Archivum immunologiae et therapeuticae experimentalis* 55(5), 315–327.
- J. Schmidhuber, Jan 2015. Deep learning in neural networks: An overview. *Neural Networks* 61, 85–117. URL <http://dx.doi.org/10.1016/j.neunet.2014.09.003>
- S. Tao, 2019. Deep neural network ensembles.
- J. O. Valderrama, L. A. Forero, R. E. Rojas, 2015. Extension of a group contribution method to estimate the critical properties of ionic liquids of high molecular mass. *Ind. Eng. Chem. Res.* 41, 3480–3487.
- G.-J. Wang, S.-Y. Li, J.-Q. Xia, Aug 2020. Ecopann: A framework for estimating cosmological parameters using artificial neural networks. *The Astrophysical Journal Supplement Series* 249 (2), 25. URL <http://dx.doi.org/10.3847/1538-4365/aba190>

Kinetic modelling of γ -linolenic acid production by *Cunninghamella echinulata*

Ziqi Song,^a Alexander Rogers,^a Bovinille Anye Cho,^a Keju Jing,^{b,c} Dongda Zhang^{a,d}

^aCentre for Process Integration, University of Manchester, Manchester, UK

^bDepartment of Chemical and Biochemical Engineering, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

^cThe Key Lab for Synthetic Biotechnology of Xiamen City, Xiamen University, Xiamen, China

^dCentre for Process Systems Engineering, Imperial College London, South Kensington Campus, London, UK

dongda.zhang@manchester.ac.uk

Abstract

Microbiological production of γ -linolenic acid (GLA) via a temperature-shift strategy has been found to shorten the batch cultivation period, improve GLA yields and reduce the operational cost. However, the underlining biochemical mechanistic phenomena are highly complex and challenging to model, thus hindering commercial upscaling applications of this fermentation technology. To bridge this gap, a dynamic model capable of simulating biomass growth, substrate consumption, and GLA biosynthesis of *Cunninghamella echinulata* for a wide temperature range was proposed for the first time. Once the model parameters were identified, the model's high simulation accuracy was demonstrated against data from a small scale 1L bioreactor. It was found that the optimal temperatures for biomass growth and GLA production were 37 °C and 14 °C, respectively. Model aided upscaling to a 5L bioreactor with a two-stage temperature-shift strategy showed a 69.6% increase in GLA production, which was verified experimentally. Therefore, this presents a significant advance for the upscaling of GLA production biotechnology from laboratory to pilot scale.

Keywords: fermentation; kinetic modelling; γ -linolenic acid; temperature-shift; process upscaling.

1. Introduction

The polyunsaturated fatty acid (PUFA), γ -linolenic acid (C18:3n-6, GLA), is widely utilised within the pharmaceutical and nutraceutical industries due to its outstanding value for treating diseases. As an essential precursor for the biosynthesis of several prostaglandins, GLA is an essential fatty acid with proven anti-inflammatory and anti-cancer effects (Wan, 2009). However, with the body unable to manufacture its own (Somashekar et al., 2003), GLA must be assimilated through the consumption of poultry, beef, pork or egg yolk, or else as a dietary supplement.

Plant seeds such as borage, black currant, evening primrose, and hemp have been used as commercial sources of GLA oil (Tanticharoen et al., 1994). However, cultivating these plants requires large swathes of arable land owing to their seeds low intracellular GLA

content. In contrast, a promising alternative source: the oleaginous fungus *Cunninghamella echinulata* (*C. echinulata*), possesses a much higher GLA content, and is simpler to cultivate (i.e., traditional fermentation). Taken together, this marks the fungus as an economically competitive source of GLA for large-scale production.

Temperature-shift: an approach where the operating temperature is switched during cultivation to increase the accumulation of the targeted metabolite has been reported to influence the intracellular PUFAs content of fungal species (Jang et al., 2005).

However, whilst *C. echinulata* is oleaginous fungi, the optimal temperature for biomass growth, lipid production, and product accumulation might differ significantly between strains. Hence, it is valuable to investigate the optimal temperature-shift strategy for industrial GLA production with *C. echinulata*. Moreover, model-based design of experiments is considered an effective tool to accomplish bioprocess scale-up from laboratory to pilot and industrial scale (Zhang et al., 2015).

Therefore, this work aims to: (i) investigate the effect of temperature on the *C. echinulata* biomass growth and GLA production through the construction of a rigorous kinetic model; (ii) evaluate the performance of the temperature-shift strategy when up-scaling the fermentation processes from a 1L to 5L bioreactor; (iii) evaluate the predictive accuracy and sensitivity of the kinetic model over different bioreactor scales.

2. Methodology

2.1. Experimental setup

In our lab, *C. echinulata* X-15 was a screened high-yield strain, which was maintained on potato dextrose agar (PDA) plates at 4 °C and transferred every 3 weeks to PDA plates. The culture was grown at 28 °C for 2 days and then stored at 4 °C until fermentation. GLA fermentation was carried out in 1L bioreactors (Infors-2015 Bioprocess controller, Netherland) containing 0.7L of medium and cultivated at different temperatures (14 °C, 28 °C, and 37 °C) with 10% (v/v) of the seed culture without pH controlled. In the two temperature-shift experiments, GLA fermentation was carried out in a 5L bioreactor (Infors-2015 Bioprocess controller, Netherland) containing 3.5L of medium, ultimately switching the temperature from 37 °C to 14 °C at either 168 hours or 96 hours. All the control conditions were the same as those in the single temperature experiments, and all cultivation experiments were performed in triplicate and analysed individually.

2.2. Kinetic model construction

2.2.1. Model structure identification

The Contois model accounts for the effect of cell flocculation and diffusional barriers that arise in high-density cell cultures. Thus, the Contois model was adopted in this study to describe biomass growth. Equation (1) simulates the total biomass growth rate, and Equation (2) simulates the fat-free biomass growth rate, assuming that the fat-free biomass growth rate is proportional to the total biomass growth rate.

$$\frac{dX_T}{dt} = \mu_m \cdot \frac{C}{K_{co} \cdot X_T + C} \cdot X_T - \mu_d \cdot X_T \quad (1)$$

$$\frac{dX_B}{dt} = k_0 \cdot \frac{dX_T}{dt} = k_0 \cdot \left(\mu_m \cdot \frac{C}{C + X_T \cdot K_{co}} \cdot X_T - \mu_d \cdot X_T \right) \quad (2)$$

Where X_T is total biomass concentration (g L^{-1}), C is glucose concentration (g L^{-1}), μ_m is the maximum specific growth rate (h^{-1}), μ_d is the specific cell death rate (h^{-1}), K_{co} is the half-saturation constant and k_0 is the ratio of fat-free biomass to total biomass (g g^{-1}).

$$\frac{dC}{dt} = -Y_{C0} \cdot \left(\mu_m \cdot \frac{C}{K_{co} \cdot X_T + C} \cdot X_T - \mu_d \cdot X_T \right) - m \cdot X_T \quad (3)$$

Equation (3) simulates glucose consumption rate, where Y_{C0} is the yield coefficient from glucose to biomass (g g^{-1}) and m is the biomass specific maintenance coefficient ($\text{g g}^{-1}\text{h}^{-1}$).

Finally, GLA accumulation was simulated by adapting the Luedeking-Piret equation to include a novel GLA consumption term which is used to describe the decay stage of GLA in experiment shown in Equation (4).

$$\frac{dX_G}{dt} = k_m \cdot \left(\mu_m \cdot \frac{C}{C + K_{C0} \cdot X_T} \cdot X_T - \mu_d \cdot X_T \right) + k_n \cdot X_T - k_d \cdot \frac{1}{C + K_p} \cdot X_T \quad (4)$$

Where k_m is the growth-dependent synthesis constant (g g^{-1}), k_n is the growth-independent synthesis constant ($\text{g g}^{-1}\text{h}^{-1}$), k_d is the specific GLA decay rate ($\text{g g}^{-1}\text{h}^{-1}$) and K_p is the saturation product constant (g L^{-1}) for GLA decay.

2.2.2. Simulating temperature effects

In order to investigate the temperature dependence of each kinetic parameter, seven parameters were modelled as a function of temperature (μ_m , K_{C0} , Y_{C0} , K_0 , k_m , k_d and K_p), whilst the remaining three parameters (m , μ_d , k_n) were considered temperature independent. The positive or negative temperature dependence was captured by either Equation (5), or Equation (6) (Laidler, 1984), respectively, where the former is simply the standard Arrhenius equation.

$$\theta_i = A_i \cdot \exp\left(\frac{B_i}{T}\right) \quad (5)$$

$$\theta_i = C_i - A_i \cdot \exp\left(\frac{B_i}{T}\right) \quad (6)$$

Where A_i , B_i and C_i in the above equations are specific parameters to be fitted for each kinetic parameter θ_i and T is the temperature in Kelvins K. Thus, two or three constants describe each of the seven temperature-dependent parameters.

2.3. Parameter estimation method

The 24 parameters required by the model were estimated by formulating the nonlinear least-squares optimisation problem defined by Equation 7, identifying parameter vector parameters θ by minimising the objective function (Del Rio-Chanona et al., 2015):

$$\min_{\theta} E(\theta) = \sum_{n=1}^{np} \left[\frac{(X_{Tn} - X_{Ten})^2}{X_{T_{emax}}^2} + \frac{(C_n - C_{en})^2}{C_{emax}^2} + \frac{(X_{Bn} - X_{Ben})^2}{X_{B_{emax}}^2} + \frac{(X_{Gn} - X_{Gen})^2}{X_{G_{emax}}^2} \right] \quad (7)$$

Where X_{Ten} , C_{en} , X_{Ben} and X_{Gen} are the measured concentrations of biomass, substrate, fat-free biomass and GLA respectively, at each sampling time n . Whilst, X_{Tn} , C_n , X_{Bn} and X_{Gn} are the respective concentrations computed by the model at each sampling time n , and $X_{T_{emax}}$, C_{emax} , $X_{B_{emax}}$ and $X_{G_{emax}}$ are the maximum measured concentrations and np is the number of sampling points.

This derivative-based nonlinear programming problem (NLP) was solved by adopting the parameter estimation framework widely employed. Given the high nonlinearity and stiffness of the system, the differential system of equations was discretised by direct transcription by orthogonal collocation into a series of nonlinear algebraic equations. The NLP was then solved using the interior point nonlinear optimisation solver IPOPT (Wächter & Lorenz T. Biegler, 2006).

2.4. Sensitivity analysis

Although the solution to a parameter estimation problem may provide point estimates that agree with the data instantaneously, questions about model stability remain unaddressed. Therefore, model sensitivity to parameter uncertainty was investigated by resampling each parameter in turn with all others fixed, simulating the state variable trajectories each

time. These trajectories were then aggregated into the propagated uncertainty bounds shown later. All the model parameters were assumed to be distributed uniformly between a lower bound (P_{min}) and an upper bound (P_{max}) centred on the mean (P_{mean}), defined as follows:

$$P_{min} = (1 - \%Variation) * P_{mean} \quad (8)$$

$$P_{max} = (1 + \%Variation) * P_{mean} \quad (9)$$

In total, Latin Hypercube Sampling (LHS) of the input space selected 200 parameter samples. All code was executed in Python version 3.7 using the *SMT 1.0.0*, *SciPy* libraries and *NumPy*.

3. Results and discussion

3.1. Results of model construction

Table 1: Parameters values with different operation temperatures

T (°C)	μ_m (h ⁻¹)	K_{CO} (g L ⁻¹)	Y_{CO} (g g ⁻¹)	K_0 (g g ⁻¹)	k_m (g g ⁻¹)
14	0.115	41.345	1.420	0.700	0.01160
28	0.164	49.586	1.307	0.780	0.003085
37	0.218	55.000	0.591	0.901	0.001173
	k_d (g g ⁻¹ h ⁻¹)	K_p (g L ⁻¹)	μ_d (h ⁻¹)	m (g g ⁻¹ h ⁻¹)	k_n (g g ⁻¹ h ⁻¹)
14	0.00352	25.013	0.0017	0.00498	0.00013814
28	0.00480	34.277	0.0017	0.00498	0.00013814
37	0.00727	37.888	0.0017	0.00498	0.00013814

The values of the parameter estimates are shown in Table 1.

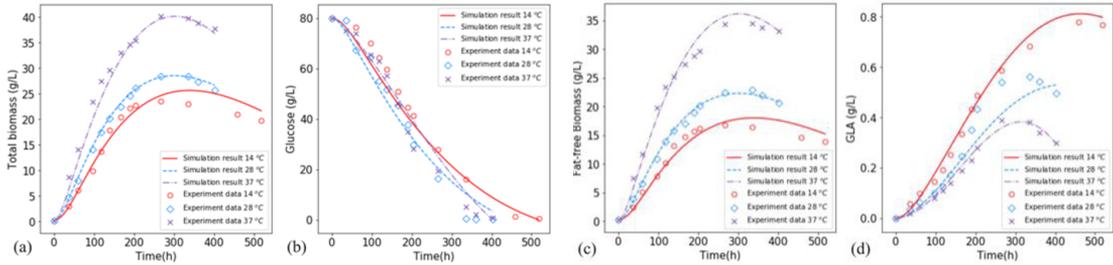


Figure 1. Experimental results and simulation fitting results for (a) total biomass, (b) glucose, (c) fat-free biomass and (d) GLA.

From Figure 1, it is evident that the model accurately captures the bioprocess dynamics over the different operating temperatures, with the state variables' mean relative percentage error falling within 20%. Furthermore, features observed experimentally are successfully described by the proposed model structure. However, whilst the model struggles to fit the GLA decay stage at 28 °C, the error remains relatively small.

3.2. Model sensitivity analysis

For a more comprehensive comparison of the model's sensitivity to each parameter, the mean relative percentage deviation (MRPD) was measured between the lower and upper bound for each state variable and operating temperature.

The result of MRPD indicated that the model is more sensitive to B_i than either A_i or the remaining temperature-independent parameters, an expected result given that B_i modifies the exponent of the temperature-dependent parameters. Of these, the MRPD is largest for

B_{μ_m} , $B_{K_{CO}}$ and $B_{Y_{CO}}$ in decreasing order, and therefore contribute most to model uncertainty.

3.3. Effects of temperature on bioprocess kinetics

The effect of temperature on *C. echinulate* fermentation was studied by inspecting the temperature correlations captured by fitting Equations (5) and (6).

Table 2. The coefficient of determination (R^2) of the optimal parameters fitting as functions of temperature and the obtained constant (A , B and C).

Parameter	R^2	A	B	C
μ_m	0.99	590.58	-2457.29	N/A
K_{CO}	0.98	1239.43	-966.81	N/A
K_0	0.96	18.41	-942.62	N/A
k_m	0.99	5.58E-16	8813.01	N/A
k_d	0.93	62.24	-2808.45	N/A
K_p	0.98	7670.33	-1640.48	N/A
Y_{CO}	0.99	5.60E+10	-7655.47	1.67

From Table 2, it can be seen that the majority of the kinetic parameters follow the standard Arrhenius relationship between 14 °C and 37 °C. However, whilst k_m and Y_{CO} exhibit the same trend overall, k_m decreases exponentially from the offset, unlike Y_{CO} which remains relatively static from 14 °C to 28 °C before plummeting over 28 °C to 37 °C. Since Y_{CO} is associated with the efficiency with which glucose is utilised to synthesis biomass constitutes, the sudden drop in Y_{CO} indicates that high temperatures deactivated the overflow metabolism that previously produced by-product.

3.4. Design of a temperature-shift strategy

A two-stage temperature-shift strategy promises to maximise GLA yield, given the different optimum temperatures for *C. echinulata* biomass growth and GLA accumulation. The bioreactors were operated at 37 °C to maximise biomass density in the first stage before dropping to 14 °C to maximise GLA accumulation in the second stage. The scale-up factor is not taken account in the work since the dynamic influence of scaling up from 1L to 5L can be neglected.

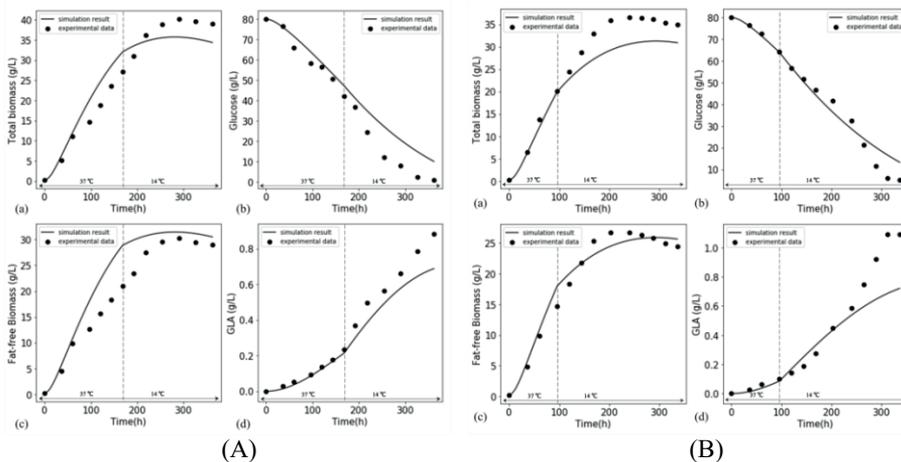


Figure 2. The experimental results and model prediction results for (a) total biomass, (b) glucose, (c) fat-free biomass and (d) GLA of temperature-shift strategy at 168hr (A) and 96 hr (B) in 5L fermenter.

Figure 2 compares the model predictions and corresponding experimental results when employing the chosen temperature-shift policy. The state variables were predicted accurately within 30% error in both cases. Of particular success was the upscaled temperature-shift experiment, which despite being a first for *C. echinulata*, achieved a GLA concentration of 1323 mg L⁻¹, a 69.6% increase over 780 mg L⁻¹ attained by the fixed temperature culture at 14 °C.

4. Conclusion

In this work, a temperature-dependent biokinetic model capable of simulating the fermentative biomass growth and GLA biosynthesis of *C. echinulata* was proposed for the first time. Using experimental data from a 1L bioreactor, the biokinetic parameters were identified and the prediction accuracy verified over a wide temperature range from 14 °C to 37 °C. Higher cultivation temperatures around 37 °C were found to benefit cell biomass growth, whilst GLA accumulation favoured lower temperatures around 14 °C. Thus, a two-stage temperature-shift strategy was designed and tested by optimising biomass growth and GLA biosynthesis of *C. echinulata* for the first time. Compared to fixed temperature cultivation at 14 °C, the optimised two-stage temperature shift strategy increased GLA production by 69.6% when verified experimentally. The proposed biokinetic model's high predictive accuracy when up-scaling the bioreactor from 1L to 5L demonstrates the model's reliability for continued scale-up of the biotechnology. However, further studies on the impact of scale-dependent transport phenomena such as mixing induce shear rate, aeration and eddy size is recommended to improve the upscaling predictions.

References

- Del Rio-Chanona, E. A., Dechatiwongse, P., Zhang, D., Maitland, G. C., Hellgardt, K., Arellano-Garcia, H., & Vassiliadis, V. S. (2015). Optimal Operation Strategy for Biohydrogen Production. *Industrial and Engineering Chemistry Research*, 54(24), 6334–6343.
- Jang, H.-D., Lin, Y.-Y., & Yang, S.-S. (2005). Effect of culture media and conditions on polyunsaturated fatty acids production by *Mortierella alpina*. *Bioresource Technology*, 96(15), 1633–1644.
- Laidler, K. J. (1984). The development of the arrhenius equation. *Journal of Chemical Education*, 61(6), 494–498.
- Somashekar, D., Venkateshwaran, G., Sambaiah, K., & Lokesh, B. R. (2003). Effect of culture conditions on lipid and gamma-linolenic acid production by mucoraceous fungi. *Process Biochemistry*, 38(12), 1719–1724.
- Tantichareon, M., Reungjitchachawali, M., Boonag, B., Vonkaveesuk, P., Vonshak, A., & Cohen, Z. (1994). Optimization of γ -linolenic acid (GLA) production in *Spirulina platensis*. *Journal of Applied Phycology*, 6(3), 295–300.
- Wächter, A., & Lorenz T. Biegler. (2006). Catalogue of an exhibition of the works of Dante Alighieri March to October 1909. In *Math. Program.* (Vol. 57).
- Wan, X. (2009). Production of Gamma-Linolenic Acid in *Pichia pastoris* by Expression of aDelta-6 Desaturase Gene from *Cunninghamella echinulata*. *Journal of Microbiology and Biotechnology*, 19(10), 1098–1102.
- Zhang, D., Chanona, E. A. D. R., Vassiliadis, V. S., & Tamburic, B. (2015). Analysis of green algal growth via dynamic model simulation and process optimization. *Biotechnology and Bioengineering*, 112(10), 2025–2039.

Techno-economic and environmental analysis of pyrolysis process simulation for plastic (PET) waste

Muhammad Shahbaz, Ahmed AlNouss, Gordon Mckay, Hamish Mackey,
Tareq-Al Ansari*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation,
Doha, Qatar*

*talansari@hbku.edu.qa

Abstract

The intensive use of plastic in modern society has accumulated a significant and damaging quantity of plastic waste. The management of plastic waste is challenging due to its un-degradable nature and increasing polluting impact on land, water, and air systems. Polyethylene terephthalate (PET) is one of the most important plastic types used and constitutes a large fraction of plastic waste. The pyrolysis thermochemical conversion process is an important technique used to convert the PET into value-added products such as char, oil and gas. This study develops a process simulation model for the pyrolysis of PET plastic to convert into char, bio-oil and gas. The process flow sheet model is developed using the Aspen Plus V11[®] and the impact of pyrolysis temperature and pressure on the production of char, pyrolysis oil and gas are investigated, where a techno-economic-environmental feasibility is also conducted using Aspen Plus built-in features. Furthermore, an optimisation is applied, where three sets of optimum operating parameters other than base case are generated through maximising the generation of each pyrolysis product. The base case demonstrates pyrolysis gas, char and oil production approximately at 330 kg/h, 490 kg/hr and 180 kg/hr, respectively at a temperature of 450 °C and 1 bar. The gas production is favourable at a high temperature of more than 450 °C, in contrast to oil and char. The techno-economic evaluation demonstrates the optimised capital and operating costs are obtained at a lower temperature with maximum yields of char and oil at 53% and 28%, respectively. Moreover, the increase in temperature to optimise gas production demonstrates enhancement in the gas quality and reduction in capital cost compared to the base case in addition to a reduction in environmental emissions (86 kg/h CO₂ emissions). In conclusion, this study provides a baseline for the utilisation of the pyrolysis process to convert PET into value-added products.

Keywords: PET, Aspen Plus, Char, Pyrolysis, Techno-economic-environmental analysis.

1. Introduction

As the global population increases, more waste streams are generated, for which effective waste management remains an issue. The most critical waste in municipal solid waste (MSW) is plastic waste, produced 99% from fossil fuel sources. Since the introduction of plastic in 1950, approximately 8.3 billion tons of plastic litter have been accumulated, and only 567 million tons have been recycled. More than 5 billion are disposed through landfill and only about 756 million tons recycled respectively (hub, 2019). The plastic production was approximately 393Mt/year in 2016, and it is expected to double in the

next 20 years (Lebreton and Andrady, 2019). The disposal of plastic waste is a salient issue as the degradation of plastic takes many years if it is disposed through landfilling and emits toxins upon incineration. The state of Qatar generates approximately 1.5 kg/capita/day, which is higher than the average global waste generation approximated at 3kg/capita/day (Miandad et al., 2017). It has been reported that the current 2.5 million population of Qatar produces approximately MSW of 4000 ton/day. Furthermore, plastic waste was approximately 8.8% of the total MSW in Qatar in 2006 (Al-Maaded et al., 2012), and later increasing to 13% (Hahladakis and Aljabri, 2019). The consumption of plastic is approximately 240000 ton, consisting of various plastics such as HDPE, LDPE, PP, PET, PS, and others, where PET is one of the most important types of plastic, and is used for packaging, especially for beverages and water bottles (Hahladakis and Aljabri, 2019).

The PET has a 14% share in total plastic waste in Qatar. Although PET is one of the plastic types that can be recycled, it is generally landfilled or incinerated (Hahladakis and Aljabri, 2019), noting that incineration of PET results in greenhouse gas emissions. In contrast, its conversion into value added products such as syngas, char and bio-oil through gasification and pyrolysis represent sustainable pathways (Anuar Sharuddin et al., 2016). There has been much work conducted in the pyrolysis of plastics domain (Anuar Sharuddin et al., 2016). For instance, the pyrolysis of PET in a fixed bed reactor yielded liquid oil and gaseous fuel in a ratio of 23.1% and 76.1%, respectively (Çepelioğullar and Pütün, 2013). Most studies demonstrate that the pyrolysis of PET produces yields of oil and gaseous products in the range of 24-40 wt% and 52-77 wt%, respectively, which renders PET as an attractive plastic for pyrolysis (Anuar Sharuddin et al., 2016). Although the pyrolysis of PET is well-investigated, very few studies have reported a techno-economic and environmental analysis and feasibility, especially for a Qatar case study. As such, this study aims to develop a process simulation model for the pyrolysis of PET for the generation of bio-oil, char and gas products, integrated with a sensitivity and optimisation of process parameters to maximise the yield. Furthermore, the study details a techno-economic and environmental analysis to investigate the investment and sustainable potential of PET pyrolysis.

2. Methodology

The approach followed for the process development begins with the pyrolysis model of plastic feedstock using Aspen Plus software. The proximate and ultimate analyses of PET presented in Table 1 are used to define the raw plastic feedstock (Oh et al., 2018). The base pyrolysis model is then evaluated in terms of economic and environmental performance. The base case is further optimised by means of sensitivity analyses to identify the optimum operating conditions that maximise the generation of bio-oil, char and gas products. The optimum operating conditions are then utilised to benchmark the different simulation models.

Table 1: Proximate and ultimate analyses of biomass feedstock as received basis (Oh et al., 2018).

Proximate Analysis	MC	VM	FC	Ash	
	0.2	87.2	12.6	0	
Ultimate analysis	C	H	N	S	O
	61.2	5.1	0	0	33.7

The current model is configured by considering the assumptions of; steady state operation, uniformity in heat and mass transfer, zero tar formation, and kinetic-free equilibrium as depicted in the study of AlNouss et al. (2021). The property package and equation of state PR-BM is selected for flowsheet development due to its applicability for hydrocarbon and refinery process (Shahbaz et al., 2020). The pyrolysis model illustrated in Figure 1 begins with an RYIELD reactor linked with calculator block to convert the feedstock attributes into conventional components based on the ultimate analysis as given in Table 1 (Shahbaz et al., 2021). The effluent stream enters the pyrolysis reactor simulated as RGIBBS to produce the pyrolysis main products. The process propagates based on Gibbs free energy minimisation by selecting the Gibbs equilibrium reactor (AlNouss et al., 2020). The hot product is then cooled prior entering a separator to flash out the pyrolysis gases. The remaining stream enters a solid separator to remove the char from the bio-oil liquid product.

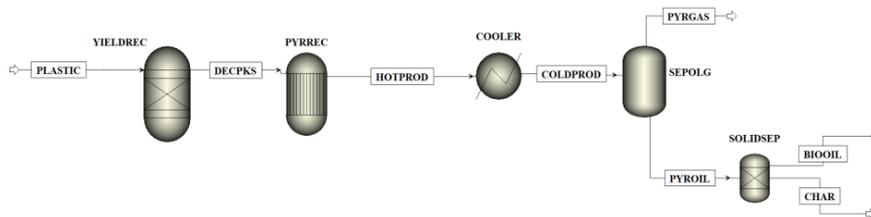


Figure 1: Aspen Plus model for plastic (PET) pyrolysis.

The techno-economic and environmental evaluation of the model is conducted using Aspen's built-in economic and environmental impact assessment features to estimate the capital, operating, and raw material costs for all cases in addition to CO₂ emissions. Moreover, the sensitivity analyses for the pyrolysis temperature and pressure are conducted to study their effect on the product yields. The pyrolysis temperature is varied between 150 and 850 °C at a constant pressure of 1 bar to cover the whole range of conventional pyrolysis process and find the optimum parameters for all three types of products including char, oil and gas. Whereas, the pressure is varied between 1 and 10 bar to observe the effect of different pressure models on the products yield by maintaining the temperature constant at 450 °C. The optimum values and static figures from the sensitivity analyses are then evaluated to benchmark their economic and environmental impact performance relative to the base case.

3. Results and discussion

The results of different PET (plastic) pyrolysis analysis are discussed in this section. The sensitivity analyses of pyrolysis reactor operating conditions illustrated in Figure 2 demonstrate the variation in the different product yields. From Figure 2, char production decreases from 527 kg/hr to 366 kg/hr with the increase in pyrolysis temperature from 150 to 850 °C. A similar trend is deduced for pyrolysis oil with increase in pyrolysis temperature. Whereas the gas generation is in direct relation with the elevation of pyrolysis temperature hence its production increases from 181 kg/hr to 633 kg/hr. The increase in pyrolysis temperature yields an increase in the gaseous product from 181 to 633 kg/h. The higher generation of gas relative to oil and char products from the pyrolysis of PET at a higher temperature of 500 °C is also noticed in other studies (Çepelioğullar and Pütün, 2013). In another study, the gas yield was about 53.13 % for PET pyrolysis (FakhrHoseini and Dastanian, 2013). Whereas, the increase in pressure demonstrates an

increase in the bio-oil production from 192 to 213, and a decrease in the gaseous and char yields from 317 and 489 to 303 and 483, respectively. The domination of gas and char yields as compared to oil at 1 bar of pressure is due to the higher fixed carbon and lower volatile content in PET plastic (Anuar Sharuddin et al., 2016). Based on the trend depicted in Figure 2, three optimum parameters sets are found for each type of product to provide flexibility for making decision in upscaling and applications listed in table 2.

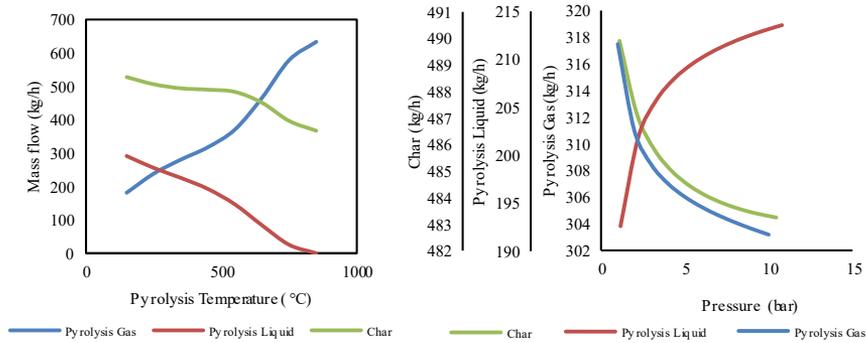


Figure 2: Sensitivity analyses trends for temperature and pressure of pyrolysis reactor.

The results from the sensitivity analysis are then utilised to evaluate the economic and environmental performance of three cases as summarised in Table 2. Optimised case 1 is expected to have higher bio-oil and char yields. Whereas, optimised case 2 is expected to demonstrate higher gaseous product yield while optimised case 3 is expected to demonstrate higher bio-oil yield.

Table 2: Optimum parameter cases for Pyrolysis of PET.

Case	Base	Optimised 1	Optimised 2	Optimised 3
Pyrolysis Temperature (°C)	450	150	750	150
Pyrolysis Pressure (bar)	1	1	1	10

The results of the economic benchmark for the different optimised cases are illustrated in Figure 3. The process is analysed based on the capital cost, operational cost, and total annualised cost in millions USD (M\$). The base case illustrates the highest capital cost (3.421 M\$), while the optimised case 2 illustrates the highest operating cost (2.255 M\$) and overall annualised cost. The higher operating cost in optimised case 2 is due to the higher heating load to operate the reactor at 750 °C. The total annualised cost is slightly higher for optimised case 2 because of the extra heating load. The optimised case 2 also demonstrates the highest in terms of gaseous product yield approximated at 59%. The higher heating yield of gas at high temperature is due to higher carbon conversion and endothermic reactions as evident from the lower volatile matter and higher carbon content (Anuar Sharuddin et al., 2016). Whereas, the highest bio-oil and char yields are achieved in optimised case 3 with values of 28% and 53%, respectively. Although pyrolysis is usually performed at a lower pressure of 1 bar, investigation of higher pressure pyrolysis is to determine the maximum pyrolysis oil.

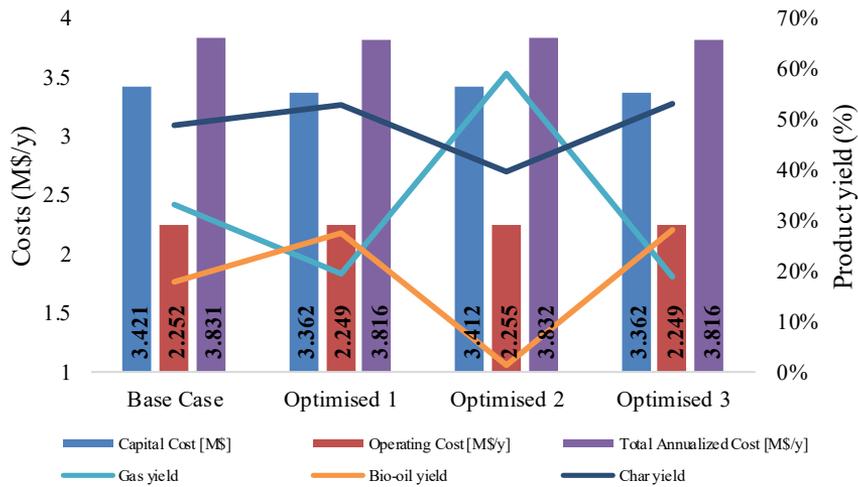


Figure 3: Techno-economic results of PET pyrolysis process.

The results of the environmental impact benchmark and gas quality for the different optimised cases are illustrated in Figure 4. The techno-environmental evaluation demonstrates optimised environmental emissions at a lower temperature of around 150 °C for the optimised cases 1 and 3 compared to the base case with 106 and 101 kg/h CO₂ emissions, respectively. However, the quality of pyrolysis gas is decreased with almost no hydrogen content. Moreover, the increase in temperature to 750 °C of optimum gas production (optimised case 2) demonstrates the best enhancement in the gas quality and the lowest environmental impacts approximated at 86 kg/h CO₂ emissions.

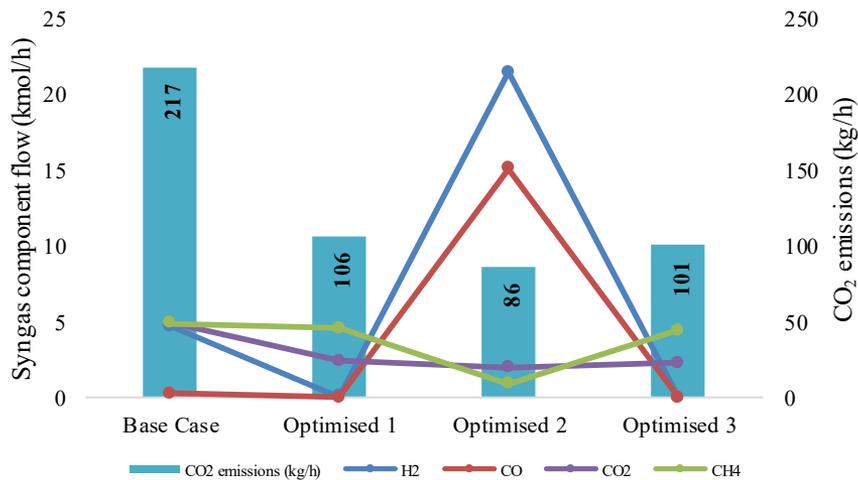


Figure 4: Techno-environmental results of PET pyrolysis process.

4. Conclusion:

The generation of value-added products such as char, oil and gas from the waste plastic (PET) don't contribute to effective PET plastic management only, however it can result

in environment and economic benefits. For this purpose, a process simulation model for the pyrolysis of PET is developed through Aspen Plus. The sensitivity and optimisation of process parameters such as temperature and pressure are conducted to maximise production. The outcomes indicate that the increase in pyrolysis temperature results in an increase in the gaseous product yield from 181 to 633 kg/h. The other two products; bio-oil and char, decrease from 291 and 527 to approximately 0 and 366 kg/h, respectively. The three optimised sets of temperature and pressure are found to maximise the yield of each pyrolysis product. The higher temperature favours the gas production, while oil generation is found to be maximum for higher pressure. The base case illustrates the highest capital cost (3.421 M\$), while the optimised case 2 illustrates the highest operating cost (2.255 M\$), and overall annualised cost with lowest environmental impact (86 kg/h CO₂ emissions). In conclusion, this study provides a basis for techno-economic-environmental feasibility which can support policy makers in making decisions as related to the conversion of PET to value added products.

References

- M. Al-Maaded, N.K. Madi, R. Kahraman, A. Hodzic, and N.G. Ozerkan, 2012, An overview of solid waste management and plastic recycling in Qatar, *Journal of Polymers and the Environment*, 20, 1, 186-94.
- A. AlNouss, P. Parthasarathy, H.R. Mackey, T. Al-Ansari, and G. McKay, 2021, Pyrolysis study of different fruit wastes using an Aspen Plus Model, *Frontiers in Sustainable Food Systems*, 5, 4.
- A. AlNouss, P. Parthasarathy, M. Shahbaz, T. Al-Ansari, H. Mackey, and G. McKay, 2020, Techno-economic and sensitivity analysis of coconut coir pith-biomass gasification using ASPEN PLUS, *Applied Energy*, 261, 114350.
- S.D. Anuar Sharuddin, F. Abnisa, W.M.A. Wan Daud, and M.K. Aroua, 2016, A review on pyrolysis of plastic wastes, *Energy Conversion and Management*, 115, 308-26.
- Ö. Çepelioğullar, and A.E. Pütün, 2013, Utilization of two different types of plastic wastes from daily and industrial life, *Journal of Selcuk University Natural and Applied Science*, 2, 2, 694-706.
- S.M. FakhrHoseini, and M. Dastanian, 2013, Predicting pyrolysis products of PE, PP, and PET using NRTL activity coefficient model, *Journal of Chemistry*, 2013.
- J.N. Hahladakis, and H.M.S.J. Aljabri, 2019, Delineating the plastic waste status in the State of Qatar: Potential opportunities, recovery and recycling routes, *Science of the Total Environment*, 653, 294-99.
- I. hub, 2019, Plastic: The Problem And Its Impact, Impact Hub GmbH, Accessed 10-14, <https://impacthub.net/plastic-the-problem-and-its-impact/>
- L. Lebreton, and A. Andrady, 2019, Future scenarios of global plastic waste generation and disposal, *Palgrave Communications*, 5, 1, 6.
- R. Miandad, M.A. Barakat, M. Rehan, A.S. Aburiazaiza, I.M.I. Ismail, and A.S. Nizami, 2017, Plastic waste to liquid oil through catalytic pyrolysis using natural and synthetic zeolite catalysts, *Waste Management*, 69, 66-78.
- D. Oh, H.W. Lee, Y.-M. Kim, and Y.-K. Park, 2018, Catalytic pyrolysis of polystyrene and polyethylene terephthalate over Al-MSU-F, *Energy Procedia*, 144, 111-17.
- M. Shahbaz, T. Al-Ansari, A. AlNouss, and G. McKay, 2021, Flowsheet development for the steam gasification of animal manure with subsequent CO₂ capturing using CaO: an Aspen Plus® modelling study, *International Journal of Global Warming*, 25, 3-4, 259-70.
- M. Shahbaz, A. AlNouss, P. Parthasarathy, A.H. Abdelaal, H. Mackey, G. McKay, and T. Al-Ansari, 2020, Investigation of biomass components on the slow pyrolysis products yield using Aspen Plus for techno-economic analysis, *Biomass Conversion and Biorefinery*.

Monte Carlo Simulation of the Mechanical Processing of Bulk Materials with Fluctuating Compositions – Compositional Probability Density

Karim Khodier^a, Tobias Krenn^a, Lisa Kandlbauer^a, Lisa Tatschl^a, Renato Sarc^a

^a*Chair of Waste Processing Technology and Waste Management, Montanuniversitaet Leoben, Franz-Josef-Strasse 18, 8700 Leoben, Austria*

Abstract

Compositions are omnipresent in process engineering, and often directly interact with process performance. In this work, a Monte Carlo based approach for considering their impact in the case of fluctuating compositions is presented. Based on experimental data, the first component – a model for the compositional probability density is targeted. A variety of descriptive approaches that involve parametric and non-parametric density estimations is discussed. Finally, for a D -dimensional composition, a multivariate Gaussian kernel density estimation for a bijective projection on $D-1$ linearly independent statically constrained coordinates in combination with boundary reflection is presented as a suitable approach.

Keywords: Monte Carlo simulation, kernel density estimation, compositional data, mixed solid waste, reflection

1. Introduction

Process engineers are dealing with compositions all the time. Examples are the chemical compositions of petroleum or combustion gases, but also compositions in terms of particle size fractions' shares. These compositions do not only describe product properties, but also affect the performance of individual processing steps. For example, the share of hydrogen in the reacting gas affects the kinetics of CO-methanation (Kopyscinski, 2010).

Some of these compositions are highly variable. This is especially the case when the input material stream to a process is (untreated) mixed solid waste (cf. Khodier et al., 2021).

Despite the waste's variability, its compositions are often only reported in terms of average values. Such information is not sufficient for calculating process optima, as soon as their dependence on the composition is non-linear. At the same time, optimizing waste processing is becoming more important, considering increasing legally required recycling rates, e.g., due to the Circular Economy Package of the European Union (European Union, 2018).

Based on these considerations, a concept was elaborated for calculating stationary optima for the mechanical processing of mixed solid waste (and in general for composition-dependent processes, with variable compositions): a material model and a process model are coupled through a Monte Carlo simulation. The material composition is described through the probability density distribution of compositions. Virtual model batches of material are created randomly, based on the probability density distribution. The condition of the processing product is then calculated, using a model for the individual process. Finally, all virtual processing products are joint, calculating the resulting overall

product condition. Based on this method, a static optimum of the process can be determined, and moreover the theoretical potential of dynamic material-adaptive processing can be evaluated.

Furthermore, a test scheme was designed, for evaluating the accuracy of the method, based on a lab-scale setup with a circular vibratory screen. The scheme consists of three experiments: first, a compositionally variable mixture of three model materials is sampled, taking at least 100 samples, for testing different descriptive approaches for the probability density distribution. Second, a quadratic regression model for the dependence of the screening efficiency on the composition is derived from a Design of Experiments-based test. Finally, five evaluation runs are performed, where compositionally variable mixtures of the model materials are sampled and screened. The expected compositions and amounts of the output streams are calculated using the suggested Monte Carlo approach and validated by a complete manual analysis of the processing products. The first experiment is covered in this paper, discussing a variety of descriptive approaches for the compositional probability density distribution, their advantages and limitations.

2. Experimental Data Generation

For evaluating different methods for describing the compositional probability density distribution, the variability of waste was experimentally simulated by mixing differently composed mixtures of three granular materials: undersize (with respect to the later used screen) gravel, undersize plastic particles, and oversize quadratic paper pieces. The choice reflects the relevance of an undersize fraction's share on one hand, and the negative influence of two-dimensional oversize particles on screening efficiency, due to the partial covering of the screen's perforation on the other hand (cf. Kaufeld et al., 2017).

The mixed material was then placed on a conveyor belt while ensuring a constant height of the flowing bulk. Finally, every fifth second, a sample, with a sampling duration of one second was taken. For designing the sampling process, Pierre Gy's Theory of Sampling, and its application on mixed solid waste (Khodier et al., 2020) were considered. In total 148 samples were taken. Their compositions are shown in Figure 1.

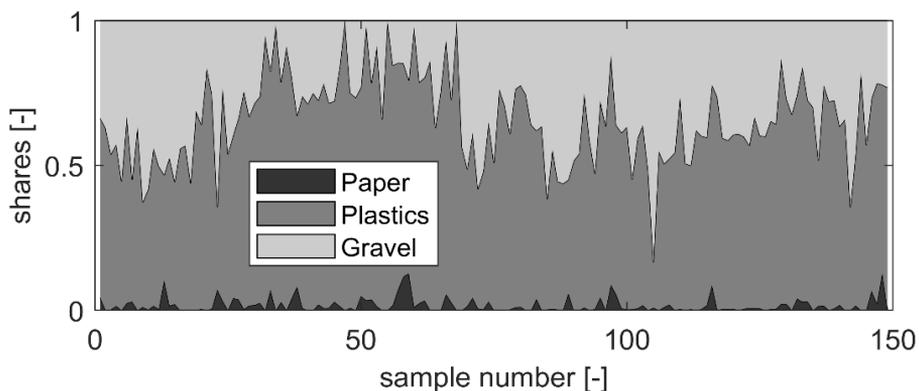


Figure 1: Compositions of the samples.

3. Probability Density Distribution

Compositions are naturally multivariate and constrained: a composition consists of at least two parts, of which each must be zero or positive, and the sum of all parts is constant

(usually 100%). Hence, a D -dimensional composition belongs to a $(D-1)$ -dimensional cartesian subspace of the D -dimensional real space. That subspace is called the D -dimensional simplex (Pawlowsky-Glahn et al., 2015). For three parts, the visualization of this subspace is well known in process engineering: the ternary diagram.

The mathematical nature of compositions must always be kept in mind when working with them. This nature, and also the relevance of correlations in the occurrence of different compositional parts, requires multivariate models for the probability density distribution.

3.1. Parametric Distributions

Parametric distributions describe the probability density through an analytical distribution with a limited number of parameters to be determined. They are hence favourable in terms of compactness and often require fewer individual samples for a reliable model of a population's distribution.

3.1.1. Multivariate Normal Distribution

The most commonly applied parametric distribution is the normal distribution. Its multivariate extension is the multivariate normal distribution. Strictly speaking, it always violates the constraints of the simplex, being positive from $-\infty$ to $+\infty$. Nonetheless, it can be a good approximation, if, e.g., at least 95% of the probability fall within the simplex. For the data at hand, the Henze-Zirkler test from the R-package “MVN” (Korkmaz et al., 2021) was applied, which neglects its multivariate normality. Considering, that multivariate normality requires univariate normality of each individual dimension (Wang, 2015), the univariate non-normality of the Paper fraction according to the Anderson-Darling test (MVN package) is at least one reason for that (Plastics and Gravel are univariate normal each, according to the test).

3.1.2. The Dirichlet Distribution

The Dirichlet distribution is the multivariate extension of the Beta distribution. It is simplicial by nature (i.e., fulfils the constraints of the simplex) and is generated by closing (normalizing to the summation constant) a set of independent, gamma-distributed random variables, with equal scale parameters. Consequently, it is quite inflexible in terms of fitting data, and hence only seldomly applicable for this purpose (Pawlowsky-Glahn et al., 2015).

Due to its inflexibility, in combination with the complexity of finding and applying packages that cover all of parameter estimation, distribution tests, and random number generation, for such a non-mainstream multivariate probability density function, it is not further investigated in this work, which targets finding widely applicable approaches.

3.2. Logarithmic and Log-ratio transformations

Many distributions, like the multivariate normal distribution, do not conform with the constraints of the simplex. For data with a one-sided boundary at zero, e.g., particle sizes, a log-transformation of the data is an often suitable approach. It can also facilitate dealing with compositional data, when the majority of compositional parts are much closer to zero than to one.

The more common approach in the community of compositional data mathematics is applying a so-called log-ratio transformation. For a D -dimensional composition, usually $(D-1)$ such log-ratios are calculated, corresponding to the number of degrees of freedom of the composition, due to the summation constraint. The log-ratios are logarithms of ratios of compositional parts, or products or sums of compositional parts (Greenacre, 2019), with so-called isometric log-ratios being the state of the art (Weise et al., 2020).

In the field of waste management, they have already been applied to waste compositions by Edjabou et al. (2017) and particle size class distributions by Khodier and Sarc (2021). For the data at hand, though, the most significant limitation of log-ratios is obstructive: they are not defined for zeros in the numerator or denominator of the ratio. For multiplicative log-ratios, they are hence not defined, as soon as one compositional part is zero. For the experimental data in this work, log-ratios involving amalgamations still do not solve the issue, since two out of three compositional parts include zeros. While there are zero-replacement approaches to treat that issue (cf. Pawlowsky-Glahn, 2015), their impacts have proven to be problematic, considering that around one third of the samples do not contain paper.

3.3. Non-parametric density estimation

While the compactness of parametric distributions is advantageous, due to the identified issues for the present data and the desired robustness of the modeling approach, non-parametric density estimation was finally targeted, in particular: kernel density estimation, with a multivariate Gaussian kernel.

The general kernel density estimator for the univariate case is shown in Equation (1), where q is the probability density, n is the number of samples, h is the bandwidth, K_i is the kernel function for the sample i , x is the point for which the probability density is estimated and x_i is the value of an individual sample. The kernel function is usually a non-negative, ideally continuous function, which is symmetrical to the origin. A widely applied kernel is the standard normal distribution. The more important choice, though, is the bandwidth, which reflects the amount of smoothing. It is often chosen based on a least-squares evaluation of the kernel density estimation in comparison to the sample data (Nedden, 2012).

$$q(x) = \frac{1}{nh} \sum_{i=1}^n K_i \left(\frac{1}{h} (x - x_i) \right) \quad (1)$$

The multivariate case, using a multivariate Gaussian kernel is shown in Equation (2), where the bandwidth matrix \mathbf{H} corresponds to the covariance matrix of the multivariate normal distribution (Wand and Jones, 1993).

$$q(\vec{x}) = \frac{1}{nh^D} \sum_{i=1}^n \frac{1}{\sqrt{(2\pi)^D |\mathbf{H}|}} \exp\left(-\frac{1}{2} (\vec{x} - \vec{x}_i)^T \mathbf{H} (\vec{x} - \vec{x}_i)\right) \quad (2)$$

3.3.1. Summation constraint

For compositional data, again the constraints of the simplex must be respected. A reliable approach for the summation constraint is a bijective projection of the D -dimensional composition onto $(D-1)$ linearly independent coordinates, as is usually done with the log-ratio transformations. The authors finally decided to follow a pragmatic approach, that does not induce issues with zeros, while the individual projected coordinates have static constraints, being defined from 0 to 1: the first coordinate is the share of one component in the overall composition. The second is the share of another component in the sub-composition that contains all components but the first. Continuing this principle, a defined

representation can be found for all cases, where at least one component is never zero. In this work, the first coordinate is the percentage of paper in the total composition, and the second is the percentage of gravel in the non-paper sub-composition.

3.3.2. Boundaries

While the chosen projection guarantees compliance with the summation constraint, each of the new coordinates is constrained with a minimum of zero and a maximum of 1. An approach for dealing with this in kernel density estimation is so-called reflection, where data is reflected in the boundary (cf. Jones, 1993).

3.3.3. Application on the experimental compositional data

Figure 2 shows the non-normalized histogram of the experimental compositional data and the corresponding bivariate kernel estimate, applying the projection from section 3.3.1 and a lower boundary of -0.00001 and upper boundary of 1.00001 for each dimension. As the Figure shows, the results are promising, while a detailed evaluation is subject to further research.

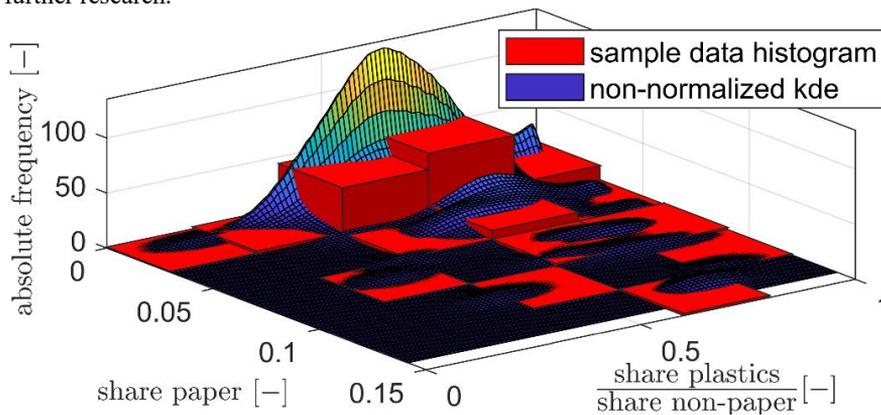


Figure 2: Non-normalized sample data histogram and kernel estimate of the compositional distribution of the experimental data

4. Conclusion and Outlook

There is a variety of approaches for estimating the density distribution of compositional data, including analytical simplicial distributions, non-simplicial distributions of log-ratio transformations and non-parametric density estimations which are adaptable to simplex data. For the investigated data, kernel density estimation, using a multivariate Gaussian kernel, a transformation to shares in sub-compositions and reflection boundaries appears to be a promising approach, and hence may contribute an essential component to the targeted Monte Carlo simulation of the processing of compositionally fluctuating bulks. The in-detail evaluation of the suggested method is still subject to further research. Furthermore, kernel density estimation, using a Dirichlet kernel (Aitchison and Lauder, 1985) is another interesting approach to be considered. Finally, the Monte Carlo approach as a whole will soon be evaluated.

Acknowledgements

Partial funding for this work was provided by: The Center of Competence for Recycling and Recovery of Waste for Future (acronym ReWaste F) (contract number 882 512) under the scope of the COMET – Competence Centers for Excellent Technologies – financially supported by BMK, BMDW and the federal state of Styria, managed by the FFG.

References

- J. Aitchison, I.J. Lauder, Kernel density estimation for compositional data. *Applied Statistics* 34 (2), pp. 129–137. doi: 10.2307/2347365
- M.E. Edjabou, J.A. Martín-Fernández, C. Scheutz, T.F. Astrup, Statistical analysis of solid waste composition data: arithmetic mean, standard deviation and correlation coefficients. *Waste Management* 69, pp. 13–23, 2017. doi: 10.1016/j.wasman.2017.08.036
- European Union, Directive (EU) 2018/851 of the European parliament and of the council of 30 May 2018 amending Directive 2008/98/EC on waste. 2018.
- M. Greenacre, *Compositional data analysis in practice*. CRC Press, Boca Raton, FL, USA, 2019. ISBN: 978-1-138-31643-0
- M.C. Jones, Simple boundary correction for kernel density estimation. *Statistics and computing* 3, pp. 135–146, 1993. doi: 10.1007/BF00147776
- S. Kaufeld, E. Coskun, M. Bosling, T. Pretz, Siebwirkungsgrad von Trommelsieben in Abfallbehandlungsanlagen: Untersuchung zum Einfluss zweidimensionaler Materialien [Screening efficiency of drum screens in waste processing plants: investigation on the influence of two-dimensional materials]. In: K.J. Thomé-Kozmiensky, D. Goldman (Eds.), *Recycling und Rohstoffe*. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 439–452, 2017.
- K. Khodier, S.A. Viczek, A. Curtis, A. Aldrian, P. O’Leary, M. Lehner, R. Sarc, Sampling and analysis of coarsely shredded mixed commercial waste: Part I: Procedure, particle size and sorting analysis. *International Journal of Environmental Science and Technology* 17, pp. 959–972, 2020. doi: 10.1007/s13762-019-02526-w
- K. Khodier, C. Feyerer, S. Möllnitz, A. Curtis, R. Sarc, Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: Coarse-shredding of commercial waste. *Waste Management* 121, pp. 164–174, 2021. doi: 10.1016/j.wasman.2020.12.015
- K. Khodier, R. Sarc, Distribution-independent empirical modeling of particle size distributions – coarse-shredding of mixed commercial waste. *Processes* 9 (3), 2021. doi: 10.3390/pr9030414
- J. Kopyscinski, Production of synthetic natural gas in a fluidized bed reactor: understanding the hydrodynamic, mass transfer, and kinetic effects. Doctoral Thesis at ETH Zürich, 2010. doi: 10.3929/ethz-a-006031831
- S. Korkmaz, D. Goksuluk, G. Zararsiz, Package ‘MVN’ (Version 5.9). 29/06/2021. Available online: <https://cran.r-project.org/web/packages/MVN/MVN.pdf>
- M. Nedden, Neue Methoden zur Charakterisierung der QSAR-Anwendungsdomäne: Modifizierte Kerndichteschätzung der Vorhersagegüte empirisch abgeleiteter Modelle in der Chemie [New methods for characterizing the QSAR application domain: modified kernel density estimation of the predictive performance of empirically derived models in chemistry]. Dissertation at the University of Bayreuth, 2012. doi: 10.1007/978-3-8348-2489-9
- V. Pawlowsky-Glahn, J.J. Egozcue, R. Tolosana-Delgado, *Modeling and analysis of compositional data*. Chichester, West Sussex: John Wiley & Sons Inc., 2015. ISBN: 978-1-118-44306-4.
- M.P. Wand, M.C. Jones, Comparison of smoothing parametrizations in bivariate kernel density estimation. *Journal of the American Statistical Association*, 88:422, pp. 520 – 528. doi: 10.1080/01621459.1993.10476303
- C.-C. Wang, A MATLAB package for multivariate normality test. *Journal of Statistical Computation and Simulation* 85 (1), pp. 166 – 188, 2015. doi: 10.1080/00949655.2013.808638
- D.R. Weise, J. Palarea-Albaladejo, T.J. Johnson, H. Jung, Analyzing wildland fire smoke emissions data using compositional data techniques. *Journal of Geophysical Research Atmospheres* 125 (6), 2020. doi: 10.1029/2019JD032128

Determination of the burst pressure of pillow plates using finite element methods

Alexander Zibart^a, Bernhard Spang^b, E.Y. Kenig^a

^a*Paderborn University, Chair of Fluid Process Engineering, Pohlweg 55, 33098 Paderborn, Germany*

^b*BUCO Wärmeaustauscher International GmbH, 21502 Geesthacht, Germany
eugeny.kenig@upb.de*

Abstract

Pillow plates are characterised by a high degree of geometrical and manufacturing flexibility, excellent structural strength, hermetic tightness and high thermo-hydraulic efficiency. These advantages open up a wide range of possible applications, and in the last two decades, numerous research studies have been dedicated to the thermo-hydraulic characteristics of pillow plates. In contrast, operational safety of pillow-plate-based equipment has gained limited attention. Due to the lack of design equations, the permissible operating pressure of such equipment is currently determined on the basis of the burst pressure obtained in time-consuming and cost-intensive experiments. In this work, we examine whether the burst pressure can be determined with the aid of finite element simulations. Furthermore, since preliminary studies have shown that the thermal resistance of pillow-plate heat exchangers can be significantly increased if they are fabricated from aluminium instead of stainless steel, we checked whether technically relevant pillow plates can be made of the aluminium alloy EN AW-5083.

Keywords: Pillow plates, Heat exchanger, Finite Element Analysis, Burst pressure

1. Introduction

Along with satisfactory thermal performance, heat exchangers must ensure sufficient operational safety, while both these criteria influence the equipment design. For conventional equipment, such as shell-and-tube heat exchangers, a calculation-based and hence non-destructive verification of the permissible operating pressure is possible by following established regulatory standards, e.g. AD2000 or ASME code. In contrast, pillow-plate-based heat exchangers represent a comparatively novel type of heat transfer equipment that has not yet been fully investigated. Therefore, currently, no appropriate design equations for the determination of the permissible operating pressure exist, mainly because of the high geometrical complexity and variability of pillow plates resulting from the simple fabrication process free of forming tools.

As a first manufacturing step, two superimposed metal sheets are spot-welded by means of a CNC-controlled laser welding machine, allowing the shape and spacing of the welding spots to be arbitrary selected. After welding of the edges and joining the nozzles, the sheets are inflated in a hydroforming process and achieve their characteristic pillow-like shape. The latter ensures high structural strength, which, along with the fully welded, hermetically sealed design, represents one of the main advantages of pillow plates. By arranging several plates in parallel, so-called pillow-plate heat exchangers (PPHX) can be assembled (cf. Fig. 1). These exchangers offer a flow path through the inner channels

of the pillow plates and another flow path through the wavy channels between adjacent plates. The welding spots and the complex, wavy-channel structure cause a periodic disturbance of the boundary layers, resulting in a high thermo-hydraulic performance of PPHX. Along with the planar design, pillow plates can also be shaped cylindrically to be applied as jackets in tempered pipelines and vessels.

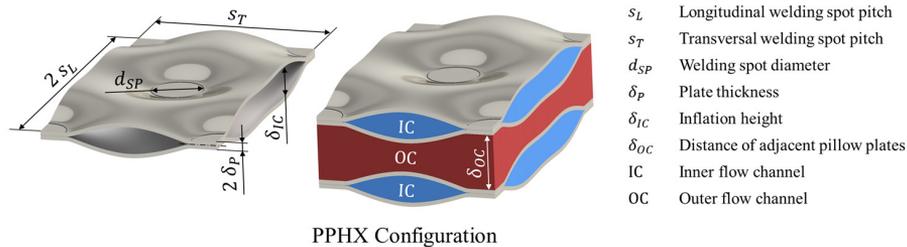


Figure 1. Periodic sections of a pillow plate (left) and of a PPHX composed of two pillow plates (right), adapted from Zibart et al., (2021).

Due to the lack of calculation methods, permissible operating pressure of pillow-plate-based equipment is currently evaluated in dependence on the burst pressure, which has to be determined experimentally by means of expensive and time-consuming burst tests. Previous studies were primarily concerned with the investigation and optimisation of the thermo-hydraulic behaviour of PPHX. Structural mechanics investigations were only carried out by Piper et al. (2015), with the focus on the correct reproduction of the pillow-plate geometry by means of FEM simulations. The burst pressure was determined for only one geometry, while no validation was performed. A comparison with experimental data obtained by the pillow plate manufacturer BUCO Wärmeaustauscher International GmbH showed that the simulations by Piper et al. (2015) significantly underestimate the burst pressure.

This work was aimed at establishing a correct prediction of the burst pressure of pillow plates using finite element methods which can replace experimental burst pressure determination. Furthermore, investigations by Zibart et al. (2021) showed that the thermal resistance of PPHX can be reduced by up to 25% by using aluminium instead of stainless steel. Therefore, in a second step, it was investigated whether technically relevant pillow plates can be manufactured from aluminium with regard to permissible operating pressures and achievable internal channel inflation height.

2. FEM Simulation

The finite element solver ABAQUS (version 2017) by Dassault Systèmes was used, which is well established in both academia and industry. The simulations were carried out in a transient manner with an explicit temporal discretisation. Since large displacements are encountered in the course of simulations, the stiffness matrix depends on geometry. To capture the arising geometric non-linearities, the stiffness matrix is updated in each time increment. In contrast to our approach, Piper et al. (2015) determined the stiffness matrix only once in the initialisation step of the simulation. Consequently, the increase in stiffness of the sheets associated with the formation of the shell-like pillow plate structure could not be captured in the simulations, and this was probably the main reason for the significant underestimation of the burst pressure in that work.

Further non-linearities result from the elasto-plastic material behaviour. In this work, along with the stainless steel AISI304, the most commonly used material for the manufacturing of pillow plates, the aluminium alloy EN AW-5083 was also considered.

Compared to other aluminium alloys, it is characterised by comparatively high strength combined with good forming capabilities. Furthermore, it offers good corrosion resistance and weldability. Thus, no filler material is required for welding, making EN AW-5083 ideally suited for laser welding. The non-linear plastic material behaviour exemplified in Fig. 2 was taken into account in the simulations.

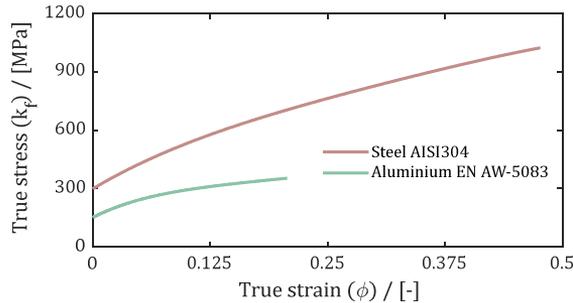


Figure 2. Flow curves of AISI304 and EN AW-5083 (Ostermann, 2007; Barthelmie, 2017).

For both steel and aluminium alloy, the Young's modulus, Poisson's ratio and yield strength, which characterise the linear elastic material behaviour, are summarised in Table 1. The so-called 'Ductile Damage Model' from ABAQUS was used to capture the bursting process in the simulations. The model parameters were fitted in such a way that computational elements completely lose their stiffness when they reach the material-specific elongation at break and are consequently eliminated from the computational mesh. Thus, cracks arise at the most highly stressed locations, which finally leads to the bursting of the pillow plate due to further crack growth. In order to decrease the computational effort, the simulation domain was reduced to the smallest possible characteristic section of a pillow plate, taking into account symmetries with regard to geometry and load. The chosen simulation domain has the dimensions s_L and $0.5s_T$. A mid-cut was performed in the thickness direction of the pillow plate, which separates the two metal sheets along their contact surfaces in the welded areas. The initial state represents a welded but still flat sheet as it is before the hydroforming process. The boundary conditions used are shown in Fig. 3.

Table 1. Elastic material parameters.

Material	Young's modulus / [GPa]	Poisson's ratio / [-]	Yield strength / [MPa]
AISI304	200	0.3	637
EN AW-5083	70.3	0.33	285

Symmetry boundary conditions in x - and y -direction were set at the lateral end faces, while symmetry boundary conditions in z -direction were specified at the free-cut contact surfaces of the plates in the welding spot regions. This causes a blockage of the translation in the respective coordinate direction as well as of the rotations around the other two coordinate axes. The free surface of the plate between the welding spots, which forms the inner channel wall of the pillow plate, was subjected to pressure. For the burst pressure determination, a physical duration time of $t_{max} = 1000$ s with a time step of $\Delta t = 0.001$ s was simulated. Here, the pressure was increased linearly from $p = 0$ bar at $t = 0$ s to p_{max} , which corresponds to t_{max} . At the beginning of each simulation, the burst pressure is not known. Therefore, p_{max} must be determined iteratively, until the condition $p_{max} > p_{Burst}$ is fulfilled and mechanical failure of the pillow plate is encountered. Care

was taken to ensure that p_{max} is only slightly higher than p_{Burst} , so that the strain rates and thus inertia effects are kept as low as possible. To determine the maximum inflation height, the simulation time was increased to $t_{max} = 2000$ s, while keeping the time step at $\Delta t = 0.001$ s. Up to $t = 1000$ s, pressure increases linearly to p_{max} . This is followed by a static phase until $t = 1500$ s, in which pressure is kept constant in order to eliminate inertial effects. In this case, p_{max} was iteratively determined in such a way that bursting during the static phase was avoided. Subsequently, pressure was linearly reduced to $p = 0$ bar at t_{max} in order to remove elastic strains.

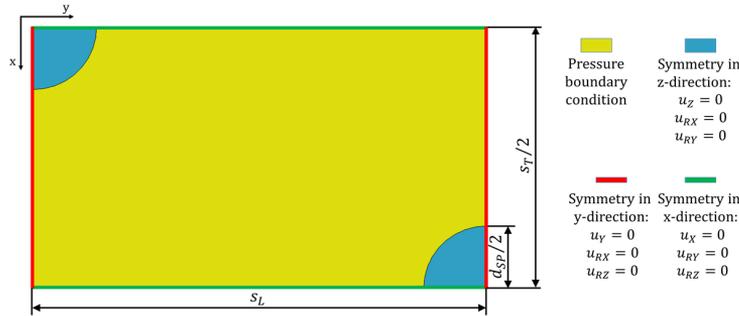


Figure 3. Applied boundary conditions shown on the inner surface of the initially undeformed plate.

Structured hexahedral computational grids were used to achieve a high computational accuracy combined with high computational efficiency. The ABAQUS specific computational element of type C3D8R was chosen, as this element type was proven in the work of Dancette et al. (2012) for the investigation of welding spot failures.

3. Validation

For the validation of the simulations, the pillow plate manufacturer BUCO provided protocols of burst tests carried out for two geometrically strongly different pillow plates. The geometry parameters of these plates are summarised in Table 2.

Table 2. Geometries used for validation.

Case	$2s_L$ / [mm]	s_T / [mm]	d_{SP} / [mm]	δ_p / [mm]	Reference
VAL1	95	55	10	1	BUCO
VAL2	52	30	11	1.5	BUCO

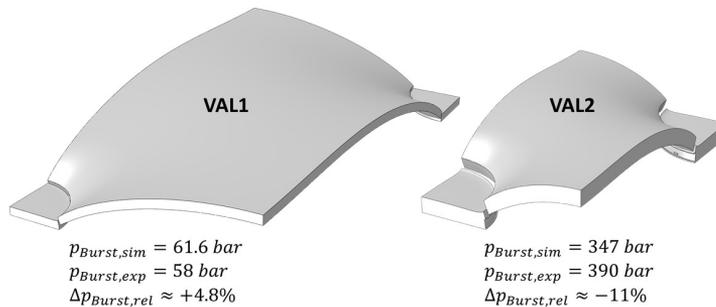


Figure 4. Results of the FEM simulations carried out for VAL1 and VAL2: the deformed geometries for the first time step after bursting of the pillow plates.

Both pillow plates were made of AISI304. The simulation results are illustrated in Fig. 4, showing the deformed plate geometries in the first time step after bursting. For the VAL1 geometry, the simulation with $p_{Burst,sim} = 61.6 \text{ bar}$ shows an overestimation of the experimentally determined burst pressure ($p_{Burst,exp} = 58 \text{ bar}$) of approx. 4.8%. The experimental burst pressure ($p_{Burst,exp} = 390 \text{ bar}$) for the VAL2 geometry is approx. 11% underestimated by the simulation with $p_{Burst,sim} = 347 \text{ bar}$. The agreement between simulated and measured values can be considered satisfactory for both geometries, and hence, the performed FEM simulations are successfully validated.

4. Results

After the successful validation of the FEM simulations, the next step was to investigate whether the aluminium alloy EN AW-5083 would be a suitable material for pillow plates. In this study, welding spot pitches $2s_L$ and s_T matching VAL1 and VAL2 geometries were chosen, while the welding spot diameter ($d_{SP} \in \{10; 11; 12\} \text{ mm}$) and the plate thickness ($\delta_p \in \{1; 1.5; 2\} \text{ mm}$) were varied. The burst pressure was found to increase almost proportionally with increasing welding spot diameter and plate thickness (cf. Fig. 5). This is due to the fact that in the region of the welding spots, the load-bearing material cross-section depends linearly on d_{SP} and δ_p . Furthermore, it can be seen that the burst pressures are significantly higher for smaller welding spot pitches. This stems from the fact that the number of welding spots per unit area increases. Thus, at equal acting pressure, the forces affecting each welding spot become smaller and consequently higher pressures can be withstood. Comparing the results with those of VAL1 and VAL2 geometries, it becomes apparent that the burst pressures are approx. 80% lower than for the counterpart made of AISI304. Considering the significantly lower tensile strength and elongation at break of EN AW-5083, this result is not surprising.

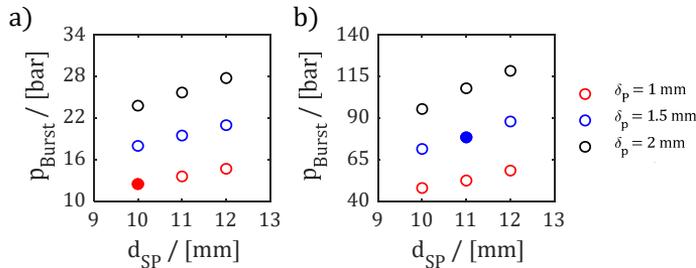


Figure 5. Burst pressure in dependence on the welding spot diameter for $2s_L = 95 \text{ mm}$, $s_T = 55 \text{ mm}$, $\delta_p \in \{1; 1.5; 2\} \text{ mm}$ (a), $2s_L = 52 \text{ mm}$, $s_T = 30 \text{ mm}$, $\delta_p \in \{1; 1.5; 2\} \text{ mm}$ (b); filled circles denote geometries matching VAL1 resp. VAL2.

Fig. 6 shows the maximum achievable inflation height ($\delta_{IC,max}$) for the investigated pillow-plate geometries plotted against welding spot diameter. Comparing Fig. 6a with 6b, it can be seen that $\delta_{IC,max}$ decreases with reducing welding spot pitch, which results from the decreasing free bending length. Furthermore, $\delta_{IC,max}$ decreases with increasing plate thickness due to the growing bending stiffness of the plates. The dependence of the maximum inflation height on the welding spot diameter is much less pronounced than for the burst pressure. It is also visible that the trends are different, namely, for $2s_L = 52 \text{ mm}$ & $s_T = 30 \text{ mm}$, $\delta_{IC,max}$ decreases with increasing d_{SP} , whereas for $2s_L = 95 \text{ mm}$ & $s_T = 55 \text{ mm}$, $\delta_{IC,max}$ increases with increasing d_{SP} . When considering the thermo-hydraulic performance of pillow plates, Piper et al. (2016) showed that the thermo-hydraulic efficiency (heat flowrate divided by required pumping power) for inner channel

flow depends almost quadratically on δ_{IC} . Small values of δ_{IC} mean small effective flow cross-sections, resulting in high flow velocities and thus high pressure drop. For this reason, expansions of $\delta_{IC} < 3 \text{ mm}$ should be regarded as technically irrelevant. Fig. 6a shows that with a welding spot pitch of $2s_L = 52 \text{ mm}$ & $s_T = 30 \text{ mm}$, technically relevant pillow plates can only be realised with a plate thickness of $\delta_p = 1 \text{ mm}$ and a further pitch reduction is not useful. For the remaining pillow plate geometries of this study, which are considered to be technically relevant, a maximum burst pressure of approx. 60 bar was obtained (cf. Fig. 5, for the geometry with $2s_L = 52 \text{ mm}$ & $s_T = 30 \text{ mm}$, $\delta_p = 1 \text{ mm}$, $d_{SP} = 13 \text{ mm}$). According to AD2000, a safety factor of 5 must be taken into account, so that a maximum permissible operating pressure of approx. 12 bar can be achieved with pillow plates made of EN AW-5083. Thus, pillow plates made of aluminium are mainly suitable for low-pressure applications, but can offer significant advantages due to their lower thermal resistance and much lower weight.

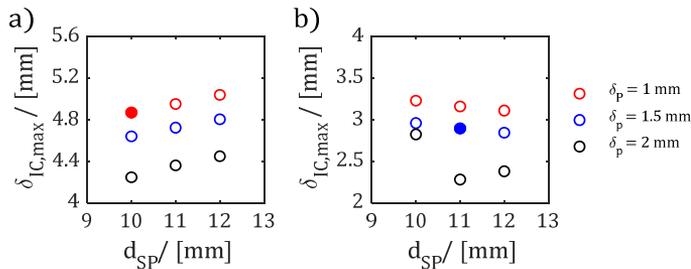


Figure 6. Maximum inflation height in dependence on the welding spot diameter for $2s_L = 95 \text{ mm}$, $s_T = 55 \text{ mm}$, $\delta_p \in \{1; 1.5; 2\} \text{ mm}$ (a), $2s_L = 52 \text{ mm}$, $s_T = 30 \text{ mm}$, $\delta_p \in \{1; 1.5; 2\} \text{ mm}$ (b); filled circles denote geometries matching VAL1 resp. VAL2.

5. Conclusions

It was demonstrated that FEM simulations are suitable for predicting the burst pressure and hence the operating pressure of pillow plates. A validation with experimental data showed a maximum deviation of 11%. Furthermore, it was found that technically relevant pillow plates with a burst pressure of up to 60 bar can be manufactured from the aluminium alloy EN AW-5083, which corresponds to a permissible operating pressure of approx. 12 bar . Thus, pillow plates made of aluminium alloys can be judged as a good alternative for low-pressure applications, providing lower thermal resistance and lower mass compared to steel-made pillow plates.

References

- J. Barthelmie, 2017, Untersuchung zur Flüssigmetallversprödung beim Widerstandspunktschweißen von hochmanganhaltigen Stählen. PhD thesis, TU Clausthal, Germany
- S. Dancette, D. Fabregue, R. Estevez, V. Massardier, T. Dupuy, M. Bouzekri, 2012, A finite element model for the prediction of Advanced High Strength Steel spot welds fracture, Eng. Fract. Mech. 87, 48-61.
- F. Ostermann, 2007, Anwendungstechnologie Aluminium, 2nd ed. Germany: Springer-Verlag Berlin Heidelberg
- M. Piper, A. Olenberg, J.M. Tran, E.Y. Kenig, 2015, Determination of the geometric design parameters of pillow-plate heat exchangers, Appl. Therm. Eng. 91, 1168-1175.
- M. Piper, A. Zibart, J.M. Tran, E.Y. Kenig, 2016, Numerical investigation of turbulent forced convection heat transfer in pillow plates, Int. J. Heat Mass Transf. 94, 516-527.
- A. Zibart, E.Y. Kenig, 2021, Numerical investigation of conjugate heat transfer in a pillow-plate heat exchanger, Int. J. Heat Mass Transf. 165, 120567.

A hybrid multi effect distillation and double reverse osmosis system for most economical brackish water desalination

O.M.A. Al-hotmani ^a, Mudhar A. Al-Obaidi ^b, Y.M. John ^a, Raj Patel ^a, and Iqbal M. Mujtaba ^{a,*}

^a *Department of Chemical of Engineering, Faculty of Engineering and Informatics, University of Bradford, Bradford, West Yorkshire, BD7 1DP, UK*

^b *Middle Technical University, Technical Institute of Baquba, Dayala – Iraq*
I.M.Mujtaba@bradford.ac.uk

Abstract

Brackish water desalination is one of the most promising methods to generate freshwater for the community in water-scarce regions. The current research proposes a hybrid Multi Effect Distillation and Thermal Vapor Compression (MED-TVC) and double Reverse Osmosis (RO) system for brackish water desalination for the Jordanian arid and semi-arid regions. In this regard, ten effects MED system was coupled with two permeate and retentate reprocessing designs of RO processes to desalinate brackish water. For this purpose, the model for the hybrid system developed by the same authors in the past has been used for simulation. For a given set of brackish water properties, the indicators such as the freshwater productivity, freshwater salinity, specific energy consumption, and disposed brine flow rate are used to evaluate the performance of the process in terms of economics and environment. The results show that freshwater can be produced from brackish water with high productivity and reduced specific energy consumption and with reduced brine flow rate into the environment compared to seawater desalination.

Keywords: Brackish Water Desalination; Multi Effect Distillation; Reverse Osmosis; Productivity; Specific Energy Consumption; Disposed Brine Flow Rate.

1. Introduction

Jordan is basically suffering from water shortage as it is located in a transitional position between arid and semi-arid climatological zones of mild rainy winter and hot dry summer. Menzel et al. (2007) stated that Jordan has approximately 90,000 km² of the semi-arid region stretching from the upper north of the Jordan basin to the south in the Gulf of Aqaba, and from the Mediterranean coast to the Jordanian Highland / Jordanian Plateau. Furthermore, the growth of population and industrialization besides the climate change and uneven spatial distribution of water resources with over-exploitation of aquifers have increased water demands that possibly would cause future water conflict. In such a complicated situation, the ground and surface water treatments are a long-term and vital solution to the issue of water scarcity specially for those coastal regions such as Red Sea region (Afonso et al., 2004). Furthermore, the wastewater reclamation and reuse is another alternative source of freshwater in arid and semi-arid regions in Jordan (Saidan et al., 2020).

The techno-economic feasibility of instilling RO system to desalinate brackish water in the Zarqa basin, Jordan was investigated by Afonso et al. (2004). This study demonstrated

the success of RO system for providing freshwater. However, the membrane fouling was one of many challenges in the water desalination. In this regard, Walschot et al. (2020) discussed in detail the challenges faced by RO desalination in Jordan despite the increasing utilisation of the process. They stated that RO desalination plants require higher maintenance than thermal desalination plants (due to fouling propensity) besides other environmental concerns. On the other hand, the thermal desalination system such as multi effect distillation is another commercially viable water desalination method which has been extensively used to generate a vast amount of freshwater, but with higher energy consumption and higher investment and operational cost compared to RO system (Filippini et al., 2018). This explains the expansion of water desalination using membrane technology and thermal processes. Jones et al. (2019) confirmed the existence of 16000 desalination plants around the world which use RO process, Multistage Flash (MSF), and MED that share 69%, 18%, and 7%, respectively, to produce 95 million m³/day of freshwater.

The potential of integrating both membrane and thermal technologies in a hybrid system was assessed by several colleagues and demonstrated its robustness to mitigate the drawbacks of individual processes and enhance the overall operation (Filippini et al., 2019). However, the high energy consumption of seawater desalination using the hybrid system still remains a challenge (Al-hotmani et al., 2021). Fera-Diaz et al. (2021) confirmed an intensive specific energy consumption between 14 to 21 kWh/m³ of MED-TVC system for seawater desalination. However, this is not the case for brackish water desalination using MED and RO hybrid system. Thus, it is imperative to analyse the potential of investigating the viability of a hybrid system of MED and double RO processes for brackish water desalination in arid and semi-arid regions in Jordan due to lower salinity of feed water compared to seawater. In other words, this research will introduce a feasible option of water desalination in Jordan compared to seawater desalination. For the first time, this study attempts to assess the feasibility of constructing a hybrid system of MED and permeate reprocessing and retentate reprocessing RO processes to desalinate brackish water in the coastal area of Red Sea in Jordan. The simulation results of this system including the performance indicators will be compared against the results of seawater desalination to evaluate the operational, economic, and environmental perspectives of brackish water desalination.

2. Description of MED-TVC and double RO processes

Fig. 1 shows a schematic diagram of the hybrid system of permeate reprocessing and retentate reprocessing RO processes (PRRO and RRRO) and MED-TVC system to desalinate brackish water. The PRRO process is designed as 20, 15, and 8 of pressure vessels (PVs) in a series where each PV contains eight spiral wound membranes synthesised by Toray, USA (brand: TM820M-400/SWRO of 37.2 m²). The water is fed into PRRO process using a high pressure pump of 85% efficiency. The forward MED-TVC system is designed of ten effects connected to thermal vapor compression (TVC). The combined brine streams of PRRO and MED-TVC are fed for further refining into the third process of RRRO. The temperature of the inlet stream of RRRO process is moderated to a specified temperature using a heat exchanger. The RRRO process is designed to process high flowrate of brackish water in 40, 30, and 16 PVs configuration. The fresh water of RRRO process is combined to the product water of PRRO and MED-TVC to form the final product stream of freshwater. However, the brine stream of RRRO process represents the brine disposal stream back to the environment.

10	Pressure at steam temperature	$P_s = P_{crit} e^{\left(\frac{T_{crit} + 273.15}{T_s}\right)^{-1}} \cdot \sum_{j=1}^8 f_j$	bar
----	-------------------------------	---	-----

Table 2. Mathematical modelling of an individual RO process (Filippini et al., 2018)

Eq. No.	Description	Equation	Unit
1	Freshwater Flux	$Q_p = A_{w(T)} \left(P_f - \frac{\Delta P_{drop,E}}{2} - P_p - \pi_w - \pi_p \right) A_m$	m ³ /s
2	Solute flux	$Q_s = B_{s(T)} (C_w - C_p)$	m ³ /s
3	Osmotic pressure in high-concentration and permeate sides	$\pi_w = 0.76881 C_w, \pi_p = 0.7994 C_p$	atm
4	Pressure droplet for each membrane and Reynolds number	$\Delta P_{drop,E} = \frac{9.8692 \times 10^{-6} A^* \rho_b Q_b^2 L}{2 d_h Re_b^3 (W \tau_f \epsilon)^2} Re_b = \frac{\rho_b d_h Q_b}{\tau_f W \mu_b}$	atm, -
5	Permeate concentration	$C_p = \frac{B_s C_f e^{\frac{J_w}{k}}}{J_w + B_s e^{\frac{J_w}{k}}}$	ppm
6	Rejection and water recovery rate	$Rej = \frac{C_f - C_p}{C_f}, Rec = \frac{Q_p}{Q_f}$	-
7	Specific energy consumption of PRRO	$E_{s,RO} = \left\{ \frac{[(P_{f(plant)} \times 101325) Q_{f(plant)}]}{\frac{\eta_{pump} Q_{p(plant)}}{3600000}} \right\} - \frac{(P_{r(block2)} \times 101325) Q_{f(block3)} \eta_{ERD}}{Q_{p(plant)} \frac{3600000}}{3600000}}$	kWh/m ³
8	Specific energy consumption of RRRO	$E_{s,RO} = \left\{ \frac{[(P_{f(plant)} \times 101325) Q_{f(plant)}]}{\frac{\eta_{pump} Q_{p(plant)}}{3600000}} \right\}$	kWh/m ³

4. Simulation and performance evaluation of the brackish water desalination system using MED-TVC and double RO processes

This section utilises the simulation of MED and double RO process (presented in Figure 1). The brackish water is simultaneously fed to the PRRO and MED-TVC processes at 3000 ppm and 25 °C of salinity and temperature, respectively. The operating pressure and feed flow rate of PRRO process are 50 atm and 5011.2 m³/day, respectively. For both cases of seawater desalination and brackish water desalination, the feed flow rate and brine temperature of MED-TVC system are same (16867.24 m³/day and 40 °C). The brine salinity for seawater is 60000 ppm for inlet salinity 39000 ppm while for brackish water is 4615 ppm for inlet salinity of 3000 ppm. Also, the motive steam of TVC is designed at 8 kg/s, 1300 kPa, and 70 °C of steam flow rate, pressure, and temperature, respectively. The combined brine of MED and PRRO is fed at 50 atm into the RRRO process. However, the inlet brine temperature of RRRO process is moderated to 25 °C using a heat exchanger (Fig. 1).

Table 3 shows the simulation results of desalinating brackish water including the most important performance indicators of the hybrid system of MED-TVC and double RO process. To justify the potential of this hybrid system, the simulation results of treating seawater of 39000 ppm at 25 °C are included in Table 3 for the purpose of comparison. This indicates that brackish water desalination using the hybrid system has a considerable increase of freshwater productivity of 20996.41 m³/day compared to 8516.66 m³/day of seawater desalination. This is an approximate growth of 146% leading to the reduction of

total specific energy consumption by 58.8%. The brackish water desalination of 3000 ppm produces a high quality water of less than 100 ppm compared to 277.87 ppm of seawater desalination of inlet salinity 39000 ppm besides attaining a high water recovery of 95.96%. Moreover, the concern of disposing brine into the environment is significantly reduced by 93.4%, which highlights a merit of brackish water desalination. The simulation results indicate the potential of constructing the proposed design of MED and double RO processes for arid and semi-arid regions in Jordan due to its superiority of producing freshwater of a high productivity compared to seawater desalination. In turn, this would be a cost-effective solution to tackle the issue of water shortage in Jordan.

Table 3. Simulation results of the hybrid system for two types of treated water

Water type	Salinity of feed water (ppm)	Productivity (m ³ /day)	Specific energy consumption (kWh/m ³)	Salinity of fresh water (ppm)	Water recovery%	Disposed brine flow rate (m ³ /day)	Salinity of disposed brine (ppm)
Brackish water	3000	20996.41	6.50	75.06	95.96	882.01	72762
Seawater	39000	8516.66	15.78	277.87	38.92	13361.77	63689.93

5. Conclusions

Water desalination from brackish water source was recognised as one of the most promising methods to produce freshwater in arid and semi-arid regions of water shortage. This research attempted to develop a new option of water desalination in Jordan based on brackish water compared to seawater desalination. Thus, a hybrid system of MED and double RO processes to desalinate brackish water has been introduced and analysed. The associated results confirmed the superiority of the proposed system and it looks very promising option in Jordan. The potential of the proposed hybrid system was assessed via comparing the simulation results with those obtained for seawater desalination. In turn, the proposed hybrid system has water recovery of 95.96% with lower energy consumption of 6.5 kWh/m³ compared to seawater desalination option.

References

- M.D. Afonso, J.O. Jaber, M.S. Mohsen, 2004, Brackish groundwater treatment by reverse osmosis in Jordan, *Desalination*, 164(2), 157-171.
- O.M.A. Al-hotmani, M.A. Al-Obaidi, R. Patel, I.M. Mujtaba, 2019, Performance analysis of a hybrid system of multi effect distillation and permeate reprocessing reverse osmosis processes for seawater desalination, *Desalination*, 470, 114066.
- O.M.A. Al-hotmani M.A. Al-Obaidi, Y.M. John, R. Patel, F. Manenti, I.M. Mujtaba, 2021, Minimisation of energy consumption via optimisation of a simple hybrid system of multi effect distillation and permeate reprocessing reverse osmosis processes for seawater desalination, *Computers and Chemical Engineering*, 148, 107261.

- J.J. Feria-Díaz, M.C. López-Méndez, J.P. Rodríguez-Miranda, L.C. Sandoval-Herazo, F. Correa-Mahecha, 2021, Commercial Thermal Technologies for Desalination of Water from Renewable Energies: A State of the Art Review, *Processes*, 9(2), 262.
- G. Filippini, M. Al-Obaidi, F. Manenti, I. Mujtaba, 2018, Performance analysis of hybrid system of multi effect distillation and reverse osmosis for seawater desalination via modelling and simulation, *Desalination*, 448, 21–35.
- G. Filippini, M.A. Al-Obaidi, F. Manenti, I.M. Mujtaba, 2019, Design and economic evaluation of solar-powered hybrid multi effect and reverse osmosis system for seawater desalination, *Desalination*, 465, 114-125.
- E. Jones, M. Qadir, M.T. van Vliet, V. Smakhtin, S.M. Kang, 2019, The state of desalination and brine production: A global outlook, *Science of the Total Environment*, 657, 1343-1356.
- L. Menzel, E. Teichert, M. Weiss, 2007, Climate change impact on the water resources of the semi-arid Jordan region, In *Proceedings of the 3rd International Conference on Climate and Water*, Helsinki (320-325).
- M.N. Saidan, M. Al-Addous, R.A. Al-Weshah, I. Obada, M. Alkasrawi, N. Barbana, 2020, Wastewater reclamation in Major Jordanian industries: a viable component of a circular economy, *Water*, 12(5), 1276.
- M. Walschot, P. Luis, M. Liégeois, 2020, The challenges of reverse osmosis desalination: solutions in Jordan, *Water International*, 45(2), 112-124.

Computational intelligence applied to the mathematical modeling of enzymatic syntheses of biosurfactants

Alice de C. L. Torres,^a Rafael A. Akisue,^a Lionete N. de Lima,^a Paulo W. Tardioli,^a Ruy de Sousa Júnior^a

^a*Federal University of São Carlos - UFSCar, Rod. Washington Luís, km 235 - São Carlos/SP - 13565-905 - Brazil*
ruy@ufscar.br

Abstract

Enzymatic esterification reactions of fatty acids with sugars generate non-ionic biosurfactants widely used in food, pharmaceutical and cosmetic industries, because of their high capability to reduce surface and interfacial tensions. In turn, mathematical modeling can be a useful tool, in its different approaches, for the simulation and optimization of enzymatic processes. Particularly, neural and fuzzy approaches are still scarcely evaluated for data of producing sugar fatty acid esters. Thus, this study aimed at using these approaches to model data of enzymatic esterification of fatty acids (oleic and lauric acids) with xylose, catalyzed by immobilized lipase B from *Candida antarctica* (CALB-IM-T2-350) and CALB immobilized on silica magnetic microparticles (SMMPs) modified with octyl groups (CALB-SMMP-octyl) or octyl+glutaraldehyde moieties (CALB-SMMP-octyl-glu). Using Matlab Neural Network Toolbox, five artificial neural networks (ANNs) were trained to predict the reaction rate, one for each type of biocatalyst and acid, obtaining R-squared values greater than 0.97. Furthermore, as an additional effort in neural modeling, two new ANNs were fitted (for two of the biocatalysts), each one of them incorporating, in its inputs, an option referring to the type of acid. R-squared values above 0.98 indicated good predictive capability. To carry out the modeling study by fuzzy inference systems, the Neuro Fuzzy Designer tool from ANFIS (Adaptive Network-Based Fuzzy Inference System) of Matlab was used. Fuzzy models were built for each of the three biocatalysts under study (CALB-IM-T2-350, CALB-SMMP-octyl and CALB-SMMP-octyl-glu), considering as input linguistic variables the type of acid, the temperature, the reaction time and the substrates molar ratio, to predict the conversion of the esterification process. Gaussian membership functions and linear output functions were used, in a Takagi-Sugeno's fuzzy approach. The fuzzy systems parameters were fitted by a hybrid parametric optimization method. The results showed that the fuzzy model outputs were very close to the targets, with RMSE (root mean squared error) values below 0.006. Finally, to demonstrate the potential of fuzzy modeling to optimize processes, response surfaces were built for the conversion of xylose as function of different operating conditions. The fuzzy surfaces indicated that higher values of xylose conversion are reached after 45 h of reaction, temperatures above 50°C, and at substrates molar ratio of 1:0.2 (acid:sugar). Thus, the present work presents, in a broad way, the potential of computational intelligence tools in the study of enzymatic production of biosurfactants.

Keywords: biosurfactants, artificial neural networks, fuzzy logic.

1. Introduction

Sugar fatty acid esters (SFAEs) are surfactants mainly applied in food, pharmaceutical and cosmetic industries (Colla *et al.*, 2010; Khan; Rathod, 2015). From the process point of view, different conditions can be applied to produce biosurfactants by heterogeneous biocatalysts using immobilized lipases. Therefore, the mathematical modeling of enzymatic reactions can be a useful tool for process simulation and optimization. This technique allows the prediction of the effects on the process due to changes in operating conditions, pointing towards factors that most influence the conversion of substrates. In addition, mathematical modeling can be applied in optimization projects, equipment sizing, economic feasibility analysis and scale-up (Nelles, 2001). However, in some situations, when complex phenomenological (white-box) models are not able to fully describe a certain reaction process, the use of black-box or gray-box models can be an interesting choice. Besides, rational exploration of the parametric space of a white-box model can impose a huge load of work (which is usually very time consuming), to guarantee that the set of mechanistic parameters satisfy chemical/biochemical and physical meanings (even more when considering possible correlations between them). Among black-box and gray-box models, stands out Artificial Neural Networks and fuzzy systems. Thus, the present work deals with the mathematical modeling of enzymatic syntheses to produce biosurfactants under two distinct approaches: use of artificial neural networks and the development of a fuzzy model.

2. Methodology

The experimental data for esterification modeling was provided by Lima *et al.* (2016). The dataset was acquired based on the syntheses of xylose esters obtained by esterification with oleic or lauric acid in tert-butyl alcohol. The experiments were conducted under the following conditions: stirring speed of 300 rpm, 72h of reaction, 1g of molecular sieve, acid-xylose molar ratio (substrates molar ratio) of 1:0.2 (base case), load of activity of $37.5 U_E/g_{acid}$ and reaction temperatures of 46°C and 55°C. The biocatalysts employed were immobilized lipase B from *Candida antarctica* (CALB-IM-T2-350) and CALB immobilized on silica magnetic microparticles (SMMPs) modified with octyl groups (CALB-SMMP-octyl) or octyl+glutaraldehyde moieties (CALB-SMMP-octyl-glu).

2.1. Modeling by artificial neural networks (ANNs)

Initially, five ANNs were trained based on xylose concentration data (mM) along time (h), one for each type of biocatalyst and acid. Training targets consisted of the experimental esterification data of sugar consumption rate ($mmol.L^{-1}.h^{-1}$) for both temperatures (46°C and 55°C). Construction and training of neural networks were performed using Matlab Neural Network Toolbox. A two-layer feed-forward network was designed for training, while classical backpropagation learning algorithm was used to adjust the weights and bias. After several tests, the architecture chosen for networks 1, 2, 4 and 5 was one hidden layer of 2 sigmoid neurons, followed by an output layer of a single linear neuron. Network 3, in turn, contained 5 sigmoid neurons followed by an output layer of a single linear neuron. Finally, to further study the potential of neural network modeling, two new ANNs were fitted (for two of the biocatalysts), each one of them incorporating, in its inputs, an option referring to the type of acid. Thus, the data initially used in the fit of networks 1, 2, 4 and 5 were grouped according to the type of biocatalyst, and two new (final) neural networks were obtained (6 and 7, “replacing” networks 1, 2, 4 and 5).

2.2. Modeling by fuzzy systems

To model the process using a fuzzy inference system (FIS), the linguistic input variables chosen were: temperature (T), substrates (acid-sugar) molar ratio (SMR), reaction time (t) and fatty acid type (FA). Xylose conversion was chosen as the output variable. Aiming to improve the quality of the FIS adjustment, further data was added from an interpolation made at 36h.

Matlab fuzzy logic toolbox was used to build the inference systems according to the type of biocatalyst. For each input variable, the respective linguistic values were identified: fatty acid (oleic and lauric), temperature (low and high), reaction time (low, medium and high) and SMR (low, medium and high). All membership functions chosen were Gaussian type. As the ANFIS (Adaptive Network-Based Fuzzy Inference System) tool used to generate the final FIS only supports Takagi-Sugeno fuzzy systems, the output was a linear function of the inputs as shown in Equation 1:

$$f(FA, T, t, SMR)_i = a * FA + b * T + c * t + d * SMR + e \quad (1)$$

Where “a”, “b”, “c”, “d”, and “e” represent fitting parameters.

For each model (or biocatalyst) 36 fuzzy If-Then rules were established based on the combination of the 10 linguistic values (e.g., Equation 2), where f_i is the linear combination described in Equation 1.

$$IF \text{ FA is oleic and T is low and t is low and SMR is low THEN conversion is } f_i \quad (2)$$

For the FIS training phase ANFIS tool was employed. This technique adjusts fuzzy parameters by mapping inputs and outputs based on process data and a feedforward neural network with a hybrid learning technique. In the forward pass, the consequent parameters are calculated by a least square method. In the backward pass, the error is back propagated, and the premises parameters are updated based on the gradient descent method.

3. Results and discussion

3.1. Modeling by ANNs

During the neural networks training to predict the rate of xylose consumption in the biosurfactant synthesis processes, R-quadratic values higher than 0.97 were obtained (from the comparison between the network output and its targets values), indicating a very good performance of the networks. Graphics of enzymatic kinetics were drawn from the output data of the networks. It was verified a very consistent fit of the five ANNs to the experimental data provided.

Figure 1 shows one of the enzymatic kinetics graphs, drawn from the output data of network 3, which refers to the syntheses of xylose laurate catalyzed by CALB-SMMP-octyl at two temperatures. The graph shows the reaction rates calculated from the experimental data and the rates estimated by the neural network modeling (dotted lines).

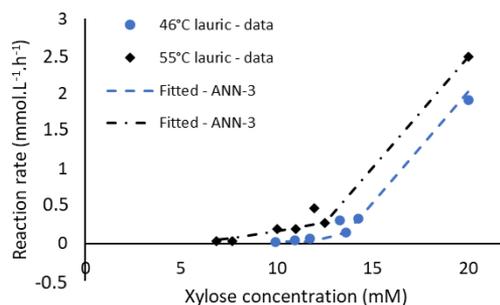


Figure 1: Kinetic modeling of xylose laurate syntheses catalyzed by CALB-SMMP-octyl at two different temperatures (ANN-3).

It can be observed that reaction rate curve at 55°C presented higher reaction rates than for the esterification at 46°C, throughout the profile. Therefore, it is indicated the possibility of acquiring xylose laurate synthesis behavior, reaction rates and sugar concentration by simple interpolation between the curves.

To further study the potential of neural network modeling, the data used in the fit of networks 1, 2, 4 and 5 were grouped according to the type of biocatalyst. Therefore, two new networks were created and trained: ANN-6 (with 5 neurons in the hidden layer, biocatalyst CALB-IM-T2-350, for oleic and lauric acids at 46 and 55°C) and ANN-7 (with 3 neurons in the hidden layer, biocatalyst CALB-SMMP-octyl-glu for oleic and lauric acids at 46 and 55°C). Regression analysis for networks 6 and 7 presented R-squared above 0.98, which suggests very good predictive capability. The kinetic graphs created after training the networks are shown in Figure 2.

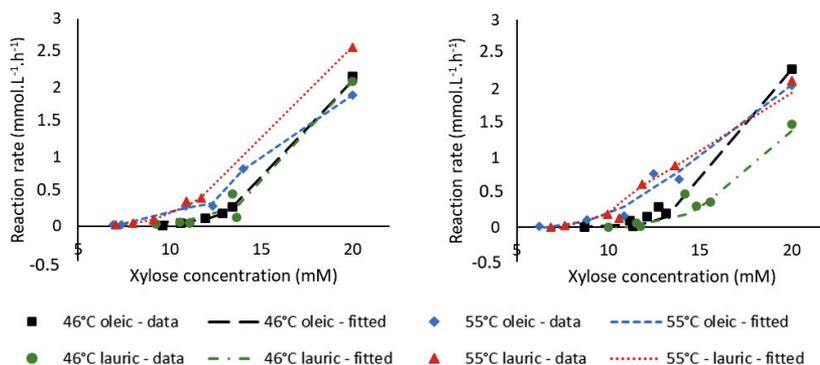


Figure 2: Kinetic modeling of xylose ester syntheses catalyzed by (A) CALB-IM-T2-350 (ANN-6) and (B) CALB-SMMP-octyl-glu (ANN-7) at different temperatures and with two fatty acids.

Once more, the modeled neural networks ANN-6 and ANN-7 presented a good fit to the experimental data. In Figure 2-A, an overlapping of the reaction curves occurred at 46°C for both acids, indicating that there is no influence regarding the type of fatty acid employed when the reaction occurs at this temperature. In general, it was possible to notice higher reaction rates for reactions at 55°C. For the CALB-SMMP-octyl-glu

biocatalyst, indicated in Figure 2-B, the curves at 55°C showed that there was no hard influence regarding the type of fatty acid.

3.2. Modeling by fuzzy systems

The adjustments of the input membership functions and output linear equation parameters of FIS were performed by ANFIS. During the testing phase, root mean squared error (RMSE) values below 0.006 were obtained for all inference systems, proving that the model outputs were similar to the experimental ones (indicating almost perfect fits). Proper graphs of xylose conversion along time were plotted (by fixing the type of fatty acid and the substrates molar ratio in 1:0.2). The graphs in Figure 3 were elaborated considering experimental esterification data and the predicted values by the fuzzy models at 46°C, 50°C and 55°C.

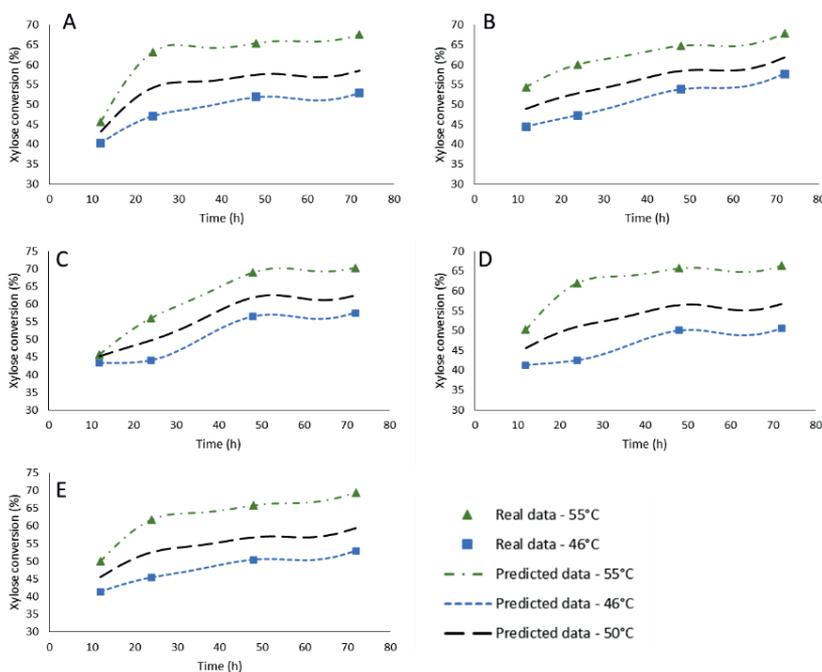


Figure 3: Fuzzy modeling of xylose ester syntheses with (A) oleic acid and CALB-IM-T2-350, (B) lauric acid and CALB-IM-T2-350, (C) oleic acid and CALB-SMMP-octyl-glu, (D) lauric acid and CALB-SMMP-octyl-glu, (E) lauric acid and CALB-SMMP-octyl.

Promising results regarding the modeling of the esterification process, by the fuzzy model, were obtained. It is worth noticing that the intermediate reaction conversion at 50°C was predicted by the fuzzy systems, at which no experimental data was available.

To further demonstrate the potential of fuzzy modeling for optimization of the operating conditions, response surface graphs were built combining the three input variables (temperature, reaction time and the substrates molar ratio) in pairs, for each biocatalyst, separated by fatty acid type. Figure 4 shows the response surfaces from the fuzzy model referred to the CALB-IM-T2-350 biocatalyst.

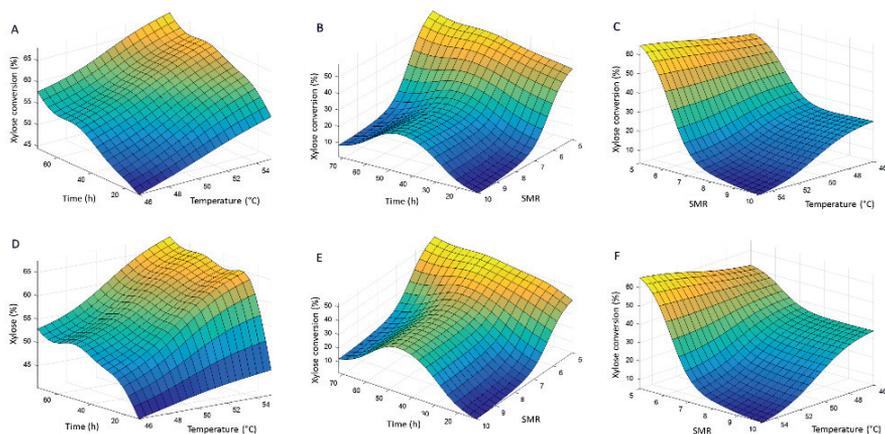


Figure 4: Response surfaces from the fuzzy model referred to xylose esterification with oleic (A, B, C) and lauric (D, E, F) acids by CALB-IM-T2-350.

It is possible to notice similarities between the response surfaces results with those from Lima *et al.* (2016). In general, the highest values of xylose conversion are reached after 45h of reaction, at high temperatures (55°C) and in the lowest condition of SMR (1:0.2 = 5).

4. Conclusions

ANN modeling successfully predicted the behavior of the xylose esterification process in biosurfactant syntheses for three different biocatalysts. All ANNs presented R-squared values higher than 0.97, which was translated to reliable mathematical models for the esterification processes. To expand the computation analysis, fuzzy logic was used to build fuzzy models of the esterification processes. Results point towards an excellent prediction of xylose conversion under experimental conditions, with RMSE values below 0.006, as well as under intermediate temperatures. In addition, it was showed the potential of fuzzy models for optimizing operational conditions as an alternative to the use of regression models in response surface methodologies.

References

- L. M. Colla, J. Rizzardi, M. H. Pinto, C. O. Reinehr, T. E. Bertolin, J. A. V. Costa, 2010. Simultaneous production of lipases and biosurfactants by submerged and solid-state bioprocesses, *Bioresource Technology*, v. 101, Issue 21, Pages 8308-8314.
- N. Khan, V. K. Rathod, 2015. Enzyme catalyzed synthesis of cosmetic esters and its intensification: A review. *Process Biochemistry*, v. 50, issue 11, p. 1793–1806.
- L. N. de Lima, G. N. A. Vieira, W. Kopp, P. W. Tardioli, R. L. C. Giordano, 2016. Mono- and heterofunctionalized silica magnetic microparticles (SMMPs) as new carriers for immobilization of lipases. *Journal of Molecular Catalysis B: Enzymatic*, v. 133, supplement 1, pages S491-S499.
- O. Nelles, 2001. *Fuzzy and Neuro Fuzzy Models. Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. 1^o ed. Springer, p. 299-340.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, by the São Paulo Research Foundation (FAPESP), grant #2016/10636-8, and by CNPq, grant #145244/2017-2.

A Practical Guide to Coffee Roaster Modelling

Cameron E. Bolt^a and Philip L. de Vaal^b

^a*University of Pretoria, Department of Chemical Engineering, Lynnwood Rd, Hatfield, Pretoria, 0002, South Africa, u17031096@tuks.co.za*

^b*University of Pretoria, Department of Chemical Engineering Lynnwood Rd, Hatfield, Pretoria, 0002, South Africa, philip.devaal@up.ac.za*

Abstract

It is desired to model the effect of three input features, the LPG flow rate, drum rotation speed and blower speed on the measured temperature profile in a rotating drum coffee roaster of a 30 kg capacity for operator prediction purposes. The performance of a first-principles model of this rotating drum coffee roaster (Schwartzberg, 2002) is compared to the performance of a composite/hybrid empirical model. The hybrid empirical model is composed of two gradient boosted regression trees and an artificial feedforward neural network which predicts two characteristic portions of the measured temperature profile by accurate prediction of the minimum in measured temperature. The first-principles model under-predicts the heat transfer to the coffee beans and as a result over-predicts the minimum temperature point when compared to true data. The relevant root mean squared error (RMSE) between the Schwartzberg model and the true measured temperature of an unseen test set of data is calculated as 23.05 as compared to the RMSE between the hybrid empirical model and the true measured temperature of 9.153. It is recommended that a larger set of data be used to train the empirical model to improve the generalisation of the model to new sets of input features.

Keywords: Coffee roaster, empirical modelling, machine learning, applied modelling

1. Introduction

1.1. Background

The process of converting the green *arabica* coffee bean into the cup of coffee one enjoys involves the roasting process, which brings to light an abundance of coffee flavours. The construction, design, and development of roasters and roaster technology alike is something a company located in South Africa, takes pride specialising in. There is a need for the development of a predictive model which will accurately predict changes in the temperature profile of the coffee beans due to changes in the inputs to the system. The inputs are the LPG (liquefied petroleum gas) flow rate to the burner, the drum rotation speed and the air blower speed. The model will be utilised in manually operated runs to preemptively adjust the inputs in order to shape the desired bean temperature profile, which is strongly correlated to the flavours produced. This paper will focus on the development of a model which generalises well to new sets of input features.

This paper investigates the modelling techniques applicable to the coffee roasting system based on theoretical first-principles (Schwartzberg, 2002) as well empirical (data-driven) methods such as *linear regression*, *decision trees*, *random forests*, *support-vector machines* and *neural networks* implemented using the *scikit-learn* and *tensorflow* python libraries (Géron, 2019). Comparison is

made between models based on performance metrics such as root mean squared error (RMSE) on a validation set of input features.

1.2. The coffee roaster

The coffee roaster of interest consists of a double walled ceramic coated drum, orientated horizontally on a central axis. The drum rotates above an open flame LPG burner. The flame heats the inlet air as well as the drum. Air is drawn through the roaster by an air blower located at the outlet of the roaster. As the air is drawn through the drum, smoke, steam, and chaff is removed from the drum and separated by a cyclone located at the outlet of the roaster. The blower is located at the top of the cyclone in order to separate out the coffee bean outer-skins (otherwise known as chaff) as well as additional solids. The gas separated from the emissions in the outlet is then passed to the stack and released to the environment. It is desired to model the measured bean temperature as a function of the roaster inputs, namely the LPG flow rate to the burner, the drum rotation speed and the blower speed.

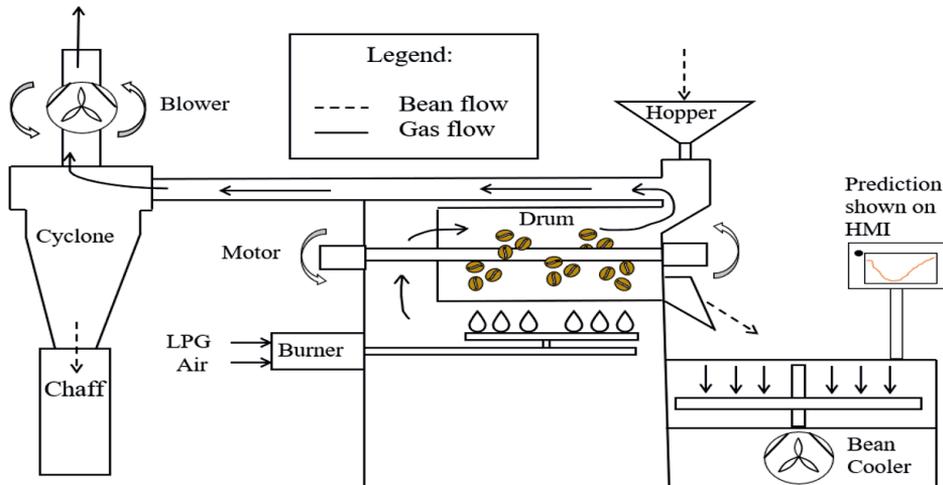


Figure 1: Traditional direct heated drum coffee roaster, adapted from Schwartzberg (2002).

The temperature profile shown in Figure 3 is obtained by measurement of the temperature of the coffee beans within the roaster. One will very quickly observe a local minimum in the curve followed by a steady rise in temperature. This inverse response is as a result of how the roaster is operated. Before the beans are released into the roaster drum through a hopper, the roaster operating temperature is primed to a set-point temperature, measured by a thermocouple located within the drum. At the point in time when the coffee beans are released into the drum (at the point $t = 0$ in Figure 3), the thermocouple within the roaster has reached the steady state priming temperature. As the room temperature beans begin making contact with the thermocouple, heat is transferred from the thermocouple to the beans. It is this dynamic lag of measured temperature which produces the measured temperature profile shown in Figure 3.

1.3. The adapted Schwartzberg model

The following semi-empirical model of the coffee roaster proposes a scalable model for a drum roaster. This model suggested by Palma et al. (2021) builds upon the model suggested by Schwartzberg (2002) by allowing for adjustments as a function of the size of the roaster i.e. the roaster mass and volume. The model can be summarised as follows:

$$\dot{T}_b = \frac{Q_{gb} - Q_{gm} + Q_{bm} + M_{bd}(Q_r + \lambda\dot{X})}{M_{bd}(1+X)c_b} \quad (1)$$

$$\dot{T}_m = \frac{Q_{gm} - Q_{bm}}{M_m c_m} \quad (2)$$

$$\dot{X} = -\frac{k_1}{D_b^2} \exp\left(-\frac{k_2}{T_b + 273.15}\right) \quad (3)$$

$$\dot{H}_e = A \frac{H_{et} - H_e}{H_{et}} \exp\left(-\frac{H_a}{R(T_b + 273.15)}\right) \quad (4)$$

$$T_{go} = \left(T_{gi} - \frac{T_b + FT_m}{1+F}\right) \left(1 - \exp\left(-\frac{h_e A_{gb}(1+F)}{G_g c_{pg}}\right)\right) \quad (5)$$

Nomenclature

A	Arrhenius constant
A_{gb}	air to bean heat transfer surface area
c_b	coffee bean specific heat capacity
c_{pg}	air specific heat capacity
c_m	metal specific heat capacity
D_b	coffee bean diameter
F	ratio of air-metal and air-beans thermal resistance
G_g	air (gas) flow rate
h_e	air to beans heat transfer coefficient
H_a	activation energy
H_e	cumulative heat of reaction
H_{et}	total reaction heat
k_1, k_2	semi-empirical constants
m_b	mass of a single bean
M_{bd}	mass of dry bean batch
M_m	mass of roaster metal
Q_{gb}	heat transfer from the air to the coffee beans
Q_{gm}	heat transfer from the air to the roaster metal
Q_{bm}	heat transfer from the beans to the roaster metal
Q_r	heat production by exothermic reaction
R	universal gas constant
T_b	measured coffee bean temperature
T_{gi}	inlet air temperature
T_{go}	outlet air temperature
T_m	roaster metal temperature
X	coffee bean moisture content
λ	latent heat of vaporisation of water

A notable advantage of the adapted Schwartzberg model is the scalability of the model to different sized coffee roasters. The adapted Schwartzberg model does however not take into account the effect of the drum rotation speed on the measured temperature. This is a considerable disadvantage if the proposed model is to be used for the purposes of process control such as in the case of a model predictive controller. The semi-empirical constants k_1 and k_2 will be chosen as proposed by Schwartzberg (Schwartzberg, 2002).

2. Methodology

Developing a realisable model of the coffee roaster which is multiple input – multiple output in nature can be achieved in one of two distinct approaches. Firstly, a model may be developed which makes use of first principle mass and energy balances in order to determine the effect of the three inputs, namely the LPG flow rate to the burner, the drum rotation speed and the blower speed on the measured temperature within the system. Available literature (Schwartzberg, 2002) will be utilised to simulate the “first-principles” model. Alternatively, measured temperature and input feature data sets may be used to develop empirical regression models which may accurately predict the measured temperature in the coffee roaster. Examples of such modelling techniques include *linear regression*, *random forests*, *decision trees* and *neural networks*. The theory behind the inner workings of these empirical methods will not be covered in this paper. The reader is referred to the excellent introduction to such topics provided by Géron (2019). These two approaches will be assessed and compared in order to determine inaccuracies that exist in either approach.

Data from a 30 kg (capacity) drum roaster is available which contains information about 42 runs (containing a total of 22 149 measurements). The current data acquisition software does not allow for the recording of total bean mass for each roast. As mentioned, the measured bean temperature exhibits a local minimum in temperature. The laws of heat transfer would lead one to believe that this local minimum’s location in time and temperature is strongly correlated to the initial starting temperature of the roast and the total mass of beans being roasted (assuming a constant specific heat capacity of beans for different temperatures and coffee bean species). Accurate prediction of this point in time will allow one to address the non-linear nature of the temperature profile by splitting the empirical model into two characteristic portions.

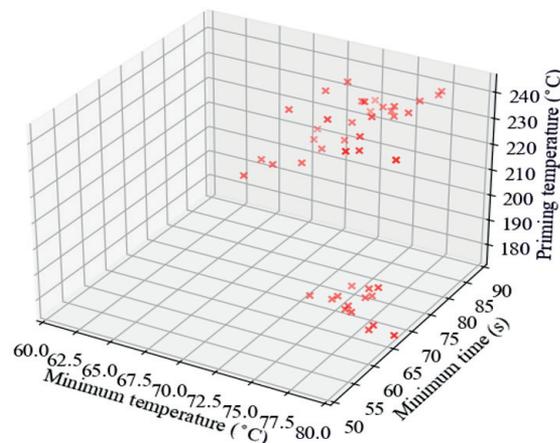


Figure 2: Relationship between the observed minimum temperature of the coffee beans and the associated priming temperature.

Available data from the roaster is plotted in Figure 2 which displays two distinct clusters of data. The first cluster of minimum times occur between 64–73 seconds from the beginning of each respective roast. This cluster is strongly correlated to a set of colder priming temperatures (in the region of 180–190 °C). The second cluster of minimum times occur between 73–90 seconds from the beginning of respective roast and is strongly correlated to a hotter set of priming temperature (in the region of 215–240 °C). A *gradient boosted regression tree* is used to model the relationship

between the priming temperature and the time at which the local minimum in temperature occurs. All the test data will be split into two portions, the “downward” portion (the region with a negative gradient in temperature) and the “upward portion”. A model will be developed for each portion applying the aforementioned regression techniques. The “downward” and “upward” data-sets will be split into training (data used to train the regression models) and test sets (data used to validate the regression models). The performance of each model will be assessed by calculating the RMSE on the relevant validation sets.

3. Results and Discussion

Upon splitting of the available data into the aforementioned “downward” and “upward” data sets, the following regression techniques are made use of: *gradient-boosted regression trees*, *linear-regression*, *support-vector machines*, *random-forests* and *neural networks*. The performance of each model is assessed against a validation data set using the root mean squared error (RMSE), the calculated performance norms are summarised in Table 1.

Table 1: RMSE of each model evaluated for the each characteristic portion.

Model	Downward portion	Upward portion
Linear	18.46	8.428
Gradient boosted regression tree	2.612	5.724
Support-vector machine	12.49	4.386
Random forest	2.659	1.161
Feedforward neural network	6.746	1.150

The results obtained indicate that a composite/hybrid modelling approach can be used to model the coffee roaster. The proposed hybrid model will be selected on the basis of the RMSE performance norms of the relevant models on the validation data set. A *gradient boosted regression tree* will be used to predict the downward portion of the measured temperature up to the predicted minimum time (where the local minimum in temperature is predicted to occur in time by use of an additional *gradient boosted regression tree*), subsequently the upward portion in measured temperature will be predicted by use of a *feedforward neural network*.

A visualisation of the results obtained by predicting the measured coffee bean temperature during roasting as a function of the roaster input features is shown in Figure 3. The supplied input features as shown in Figure 3 did not occur in the training or validation set of data and is typical of a roasting operation, the LPG input is initially maintained at 100% (an artefact of the priming process) before stepping down at a later stage. The empirical model predicts the measured bean temperature exceptionally well on the downward portion by use of a *gradient boosted regression tree*. The *neural network* deviates from the true measured temperature between 80–600 seconds before coinciding with the true temperature, interestingly a considerable reduction in LPG input occurs at 600 seconds. The Schwartzberg model was implemented using the semi-empirical constants and heat transfer coefficient approximations provided by Schwartzberg. The Schwartzberg model seems to model a system in which there is less heat transfer to the beans by the hot air and/or drum. This can be noted by the local minimum in the Schwartzberg model which occurs at a later point in time and at a hotter temperature than the true minimum. The relevant RMSE between the Schwartzberg model and the true measured temperature is 23.05 as compared to the RMSE between the hybrid model and the true measured temperature of 9.153. The RMSE for the hybrid model is surprisingly large when one considers the errors calculated in Table 1. This can be attributed to the size of the data set used (22 149 measurements). Use of a larger data set will allow the hybrid model to generalise with increased accuracy to unseen input features.

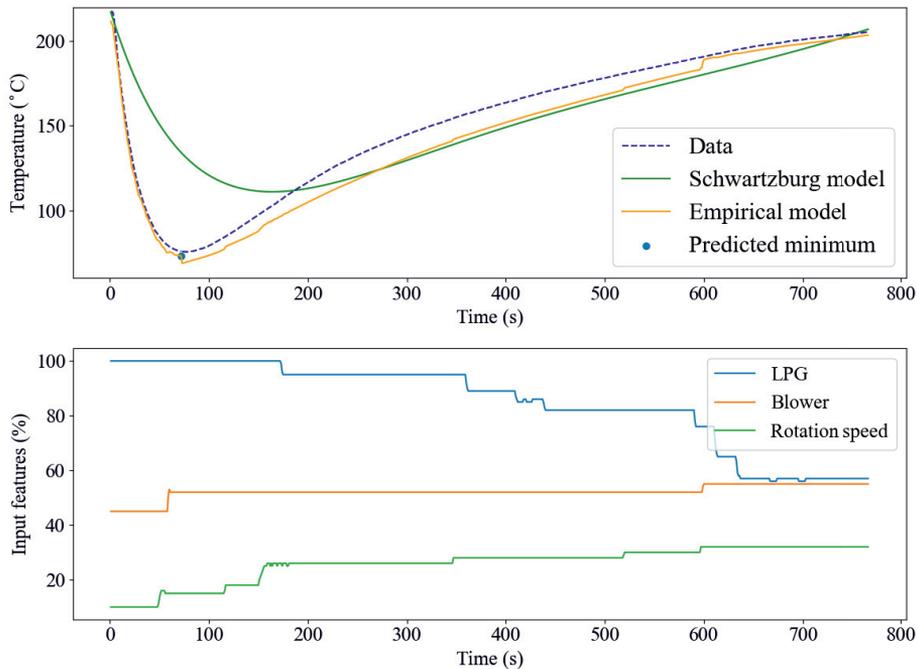


Figure 3: Comparison of the composite empirical model to the first-principles model.

4. Conclusions and Recommendations

The empirical hybrid model has been successfully developed. The empirical model is composed of two *gradient boosted regression trees* and an artificial *feedforward neural network* which predicts two characteristic portions of the measured temperature profile by accurate prediction of the minimum measured temperature. The relevant RMSE between the Schwartzberg model and the true measured temperature of an unseen test set of data is calculated as 23.05 as compared to the RMSE between the hybrid empirical model and the true measured temperature of 9.153. It is recommended that a larger set of data be used to train the empirical model to improve the generalisation of the model to new sets of input features. The methods shown can easily be applied to unrelated physical system modelling problems. The Schwartzberg model does not incorporate the effects of drum rotation speed on the measured temperature. It is suggested that with the development of an accurate model of the coffee roaster, the use of model-based control algorithms such as model-predictive control (MPC) be investigated.

References

- A. Géron, 2019. Hands-on Machine Learning with Scikit-learn, Keras and Tensorflow, 2nd Edition. O'Reilly, United States of America.
- F. D. Palma, F. Iancono, C. Toffanin, A. Ziccardi, L. Magni, 2021. Scalable model for industrial coffee roasting chamber. *Procedia Computer Science* 180, 122–131.
- H. G. Schwartzberg, 2002. *Engineering and Food for the 21st Century*, 1st Edition. McGraw-Hill, United States of America.

Modeling and simulation of anoxic-aerobic algal-bacterial photobioreactor for nutrients removal

Irina Bausa,^{a,b} Raúl Muñoz,^{a,c} Smaranda Podar,^{a,b} César de Prada^{a,b}

^a*Institute of Sustainable Processes, Dr. Mergelina s/n, Valladolid 47011, Spain*

^b*Department of Systems Engineering and Automatic Control, University of Valladolid, Real de Burgos s/n, Valladolid 47011, Spain*

^c*Department of Chemical Engineering and Environmental Technology, University of Valladolid, Real de Burgos s/n, Valladolid 47011, Spain*

irina.bausa@uva.es

Abstract

The increasing number of microalgae-based applications has contributed for development of new models to describe the complex interactions taking place in mixed algal-bacterial wastewater treatment systems. Mathematical models contribute to system optimization in terms of operation and control. The goal of this work is modeling and simulation of an anoxic-aerobic algal-bacterial photobioreactor with biomass recycling for wastewater treatment. Process model and simulation have developed in dynamic simulation software PROOSIS®. The model was set-up and calibrated with data from a pilot plant treating synthetic wastewater, located in facilities of University of Valladolid. Simulations have shown the capability of mathematical model to predict the removal efficiency of nutrients from wastewater. Removal efficiencies simulated are closely with experimental results ones.

Keywords: Modeling, Optimization, Simulation, Wastewater treatment.

1. Introduction

In recent years, microalgae-bacteria based technologies for wastewater treatment has generated a growing interest in scientific community. Microalgae-based technologies for wastewater treatment were proposed in the 1960s, but till now, remains certain limitations related with their exploitation at industrial scale. The European Directives concerning wastewater treatment processes; efficient gestion of nutrients; and transit to low-carbon economy, have reactivated the interest in microalgae-based technologies and have motivated the development of large number of improvements and applications (Muñoz & Guieysse, 2006).

The costs associated with mechanical aeration represent 45–75% of the total operational costs in conventional wastewater treatment plants (WWTPs) (Chae & Kang, 2013). These costs could be reduced using wastewater treatment systems based in microalgae-bacteria consortia. In addition, the capacity of microalgae to simultaneously remove carbon (C), nitrogen (N) and phosphorus (P) via mixotrophic assimilation represents an important advantage in comparison with aerobic activated sludge or anaerobic digestion technologies in terms of enhanced nutrient recovery.

The increasing number of microalgae-based applications has contributed for development of new models to studying of main processes, factors, and variables affecting microalgae growth in different cultures media, including wastewater (Casagli, et al., 2021), (Sánchez-Zurano, et al., 2021), (Solimeno, et al., 2019). Recently, design and improvement of

facilities for optimization of algae biomass yield and adequate depuration of wastewater has generated a growing interest. In this regard, in (De Godos, et al., 2014), an innovative anoxic–aerobic algal–bacterial photobioreactor configuration with biomass recycling was proposed. The configuration described in (De Godos, et al., 2014) was optimized in (Alcántara, et al., 2015) in order to promote N removal via denitrification and the development of a rapidly settling algal–bacterial population. Within this framework, the goal of this work is modeling and simulation for first time the anoxic-aerobic algal-bacteria photobioreactor for wastewater treatment proposed in (Alcántara, et al., 2015). In this work, simulation results for the concentration of Total Suspended Solids (TSS) and ammonium are presented. These values are used to estimate removal efficiencies in the anoxic-aerobic reactor configuration. Values of TSS and ammonium in effluents are key indicators for an adequate depuration of wastewaters. Process model and simulation have developed in dynamic simulation software PROOSIS®.

2. Materials and Methods

2.1. Plant Description

Experimental data were collected from anoxic–aerobic algal–bacterial photobioreactor configuration with biomass recycling located in facilities of University of Valladolid (Fig. 1). In (Alcántara, et al., 2015), the influence of the Hydraulic Retention Time (HRT), intensity and regime of light supply, and dissolved O₂ concentration (DOC) in the photobioreactor were analyzed in five-stage experimentation.

The aerobic tank (photobioreactor) was an enclosed jacketed 3.5 L glass tank (AFORA, Spain) with a total working volume of 2.7 L. The photobioreactor was continuously illuminated by LED lamps. The anoxic reactor consisted of a gas-tight 1 L polyvinyl chloride tank with a total working volume of 0.9 L maintained in the dark. The synthetic wastewater (SWW) was fed to the anoxic tank and continuously overflowed by gravity into the aerobic photobioreactor.

The algal–bacterial broth was continuously recycled at 3 L/d from the photobioreactor to the anoxic tank. An Imhoff cone with a volume of 1 L and interconnected to the outlet of the photobioreactor was used as a settler. The algal–bacterial biomass settled was recycled from the bottom of the settler into the anoxic tank at 0.5 L/d and wasted 3 days a week to control the algal–bacterial sludge retention time (SRT). A detailed description of the system, microorganisms and culture conditions, experimental design, and analytic procedures is provided in (Alcántara, et al., 2015).

2.2. Experiment Design

The design of the experimentation was conducted based on the hypothesis that algal–bacterial photobioreactors for wastewater treatment can support the oxidation of ammonium (N-NH₄⁺) into NO₂⁻/NO₃⁻, which can then be easily removed through denitrification (using the organic matter present in SWW) under anoxic conditions via internal recycling of the photobioreactor broth (De Godos, et al., 2014). Liquid samples of 100 mL were drawn three times a week from the SWW storage tank, anoxic tank, aerobic tank, wastage, and clarified effluent to monitor the concentration of dissolved N species (total nitrogen (TN), N-NH₄⁺, N-NO₂⁻, and N-NO₃⁻) and biomass concentration, expressed as TSS. The data used for simulations were obtained from previous studies in the Institute of Sustainable Process (University of Valladolid). Data used in this study were collected from May 2014 to July 2014 (corresponding with two experimentation stages with different HRT). Data from Stages I and II were selected, because same conditions of illumination and no oxygen supply are considered in both stages.

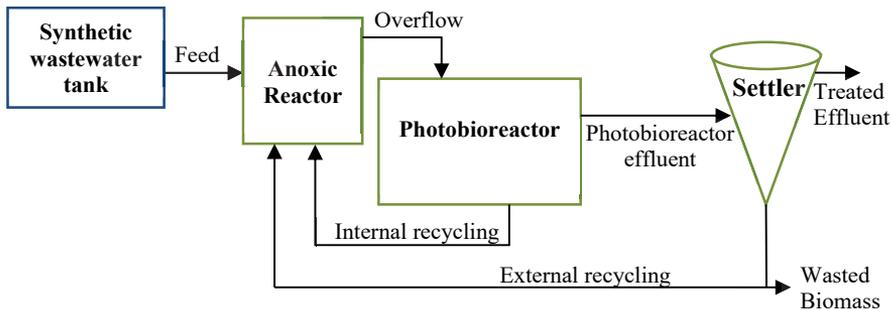


Fig. 1. Schematic of the anoxic-aerobic algal-bacterial photobioreactor configuration

2.3. Modeling

In the last two decades, several microalgae-bacteria models were developed (Reichert, et al., 2001), (Solimeno, et al., 2017), (Solimeno, et al., 2019), (Casagli, et al., 2021), (Sánchez-Zurano, et al., 2021). The review of these models has allowed to choose the model BIO_ALGAE 2 for the present work, since it includes the interactions between microalgae and bacteria, without presenting excessive complexity. Model BIO_ALGAE 2 (Solimeno, et al., 2019) has been used to represent the biochemical reactions and processes that take place in both anoxic and aerobic reactor described in (Alcántara, et al., 2015). Model BIO_ALGAE 2 uses the common nomenclature of the International Water Association models and considers 19 components – 6 particulate and 13 dissolved – implicated as variables in the physical, chemical and biokinetic processes. In (Solimeno, et al., 2017) those components are described, as well as their main role in the processes and their interactions with other components.

Settler model is described using the mass-balance expressions of Takács model (Takács, et al., 1991). Takács model is a multi-layer dynamic model for the clarification and thickening processes. In this work, a 5-layer settler is considered. A description of settler model used in this work can be found in (Bausa, et al., 2021).

Reactors and settler model are coded in dynamic simulation software PROOSIS® (Empresarios Agrupados Internacional, 2021).

2.4. Parameter Estimation

Previous to parameter estimation, a sensitivity analysis is realized with the aim of identify the parameters that have the greatest impact on the model. To this purpose, a subset of the most influential parameters on model outputs was analyzed.

The approach to solve a parameter estimation problem in terms of optimization considers that for each value of the vector of parameters θ (decision variables) the model provides a prediction of the response of the system in each experiment. For this purpose, a set of data samples from inputs $u(t)$ and outputs $y(t)$ of process is needed. The same sequence of process inputs is applied to model. For each time sample t , the prediction error is an indicator of model goodness. The dynamic optimization problem for the start-up optimization can be converted into a nonlinear programming (NLP) problem by means of a control vector parameterization technique and a proper procedure for computing the cost function. In this work, the SNOPT nonlinear programming algorithm has been used in the PROOSIS® dynamic simulation environment to solve the optimization problem. The selected integration method was IDAS. The fair function estimator (Huber, 2014) is used here as a robust objective function J against measurement outliers and gross errors. Dynamic optimization problem reads:

$$\min_{\hat{\theta}} J(\hat{\theta}, \theta) = \sum_{j \in M} c^2 \left[\frac{|\varepsilon_j|}{c} - \log \left(1 + \frac{|\varepsilon_j|}{c} \right) \right] \quad (1)$$

Subject to restrictions (2) and (3)

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) \quad (2)$$

$$\hat{\mathbf{y}}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) \quad (3)$$

$$\underline{\boldsymbol{\theta}} \leq \hat{\boldsymbol{\theta}} \leq \bar{\boldsymbol{\theta}}$$

Where $\varepsilon_j = (\hat{\theta}_j - \theta_j) / \theta_m$ represents the error between available process measurements (θ) and their estimated values $\hat{\theta}$ limited between user-defined minimum and maximum values, $c \in \mathbb{R}^+$ is an user defined fitting parameter to tune the slope for large residues, and θ_m is the media of process measurements.

3. Results

The results of sensitivity analysis in both reactors indicated that model outputs are especially sensitive to the maximum specific growth rate of microalgae (μ_{ALG}) and heterotrophic bacteria (μ_{H}), the decay-rate of microalgae ($k_{\text{death,ALG}}$) and heterotrophic bacteria ($k_{\text{death,H}}$), and the mass transfer coefficient for ammonia ($K_{\text{la, NH}_3}$). Limits for decision variables and initial values for optimization were established from similar studies reported in the literature. Values of decision variables resulting from optimization in both reactors are shown in Table 1. The optimized values of μ_{H} and $k_{\text{death,H}}$, are closely related with calibrated values reported in (Solimeno, et al., 2017) and (Casagli, et al., 2021), respectively. Value of $k_{\text{death,ALG}}$ coincide with value used in (Reichert, et al., 2001) and (Solimeno, et al., 2017). In sum, all parameter values are within values ranges adopted in literature for similar facilities. Results of sensitivity analysis and parameter estimation in settler is described in (Bausa, et al., 2021). Data from first 30 days of stage I were used for parameter estimation (corresponding with 14 samples). Validation was performed using data from stage I (days 33 to 47) and stage II (days 47 to 63). Both, parameter estimation, and model validation were performed using 14 samples of data.

Fig. 2 show simulation results for the concentration of TSS. Fig. 2a) and 2b) represent data set used for parameter estimation and for validation in anoxic reactor, respectively. Results for parameter estimation and validation in the aerobic reactor are presented in Fig. 2c) and 2d), respectively. Simulation results show model capability to reproduce dynamic behavior of the system. Fig. 3 presents simulation results for the concentration of ammonium. Fig. 3a) and 3b) represent data set used for parameter estimation and for validation in anoxic reactor, respectively. Results for parameter estimation and validation in the aerobic reactor are presented in Fig. 3c) and 3d), respectively. Although some discrepancies are observed in transient behavior prediction for ammonium concentration, average values (and, consequently, removal efficiencies) are closely with experimental results. Average values for simulated data are compared with experimental values reported in (Alcántara, et al., 2015) at steady state during stage I with the aim of evaluate model capability for prediction. This comparison is presented in Table 2.

Experimental removal efficiency of TSS for anoxic-aerobic configuration during stage I was roughly 98 %, the estimated value for removal efficiency of TSS was 97,7 %. Both experimental and estimated data are consistent with the percentage of reduction established by the European Union (CEE, 1991). Concentrations of ammonium in the effluent are above the maximum concentration permissible for wastewater discharge into the environment according to European Directive 91/271/CEE (CEE, 1991) on discharge of domestic waters. In this context, further research is needed to improve the N-NH_4^+ removal efficiency of the configuration.

Table 1. Values of estimated parameters in anoxic and aerobic reactor

Parameter	Value
μ_{ALG}	0.4 d ⁻¹
μ_H	1.4 d ⁻¹
$k_{death,ALG}$	0.1 d ⁻¹
$k_{death,H}$	0.9 d ⁻¹
$K_{la,NH3}$	22.3 d ⁻¹

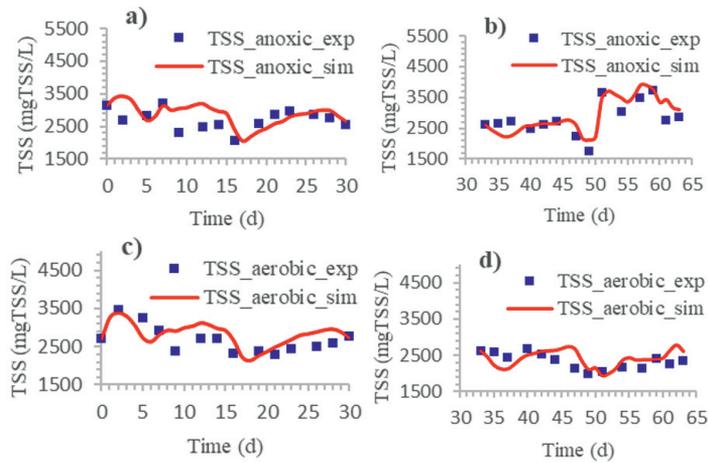


Fig. 2. Experimental and simulated TSS concentration in anoxic and aerobic reactor

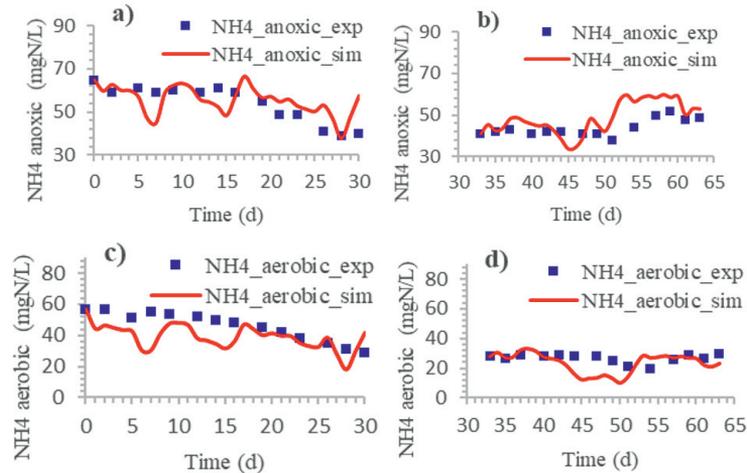


Fig. 3. Experimental and simulated N-NH₄⁺ concentration in anoxic and aerobic reactor

Table 2. Average values ± standard deviations at steady state during stage I

Parameter	Experimental	Simulated
TSS anoxic (mgTSS/L)	2575 ± 160	2627 ± 148
TSS aerobic (mgTSS/L)	2531 ± 191	2392 ± 111
N-NH ₄ ⁺ anoxic (mgTSS/L)	42±1	41.1±4.5
N-NH ₄ ⁺ aerobic (mgTSS/L)	28±1	22.9±7.4

4. Conclusions

In this paper, modeling of anoxic-aerobic algal-bacterial photobioreactor for wastewater treatment is presented. Parameter estimation via optimization is realized to fit experimental and simulated data. The model proved to be effective in reproducing dynamic behavior of different measured variables. Removal efficiencies simulated are closely with experimental results ones. Removal efficiency for total suspended solids is over 95% for experimental and simulation results.

5. Acknowledgments

This research was supported by Regional Government of Castilla y León and EU-FEDER program (CLU 2017-09 and UIC 233) and by Regional Government of Castilla y León and the European Social Found (Order EDU/601/2020).

References

- Alcántara, C. y otros, 2015. Evaluation of wastewater treatment in a novel anoxic–aerobic algal–bacterial photobioreactor with biomass recycling through carbon and nitrogen mass balances. *Bioresource Technology*, Volumen 191, pp. 173-186.
- Bausa, I., Muñoz, R., Podar, S. & de Prada, C., 2021. *Modelo para la estimación de la concentración de biomasa en una instalación reactor anóxico-fotobiorreactor aerobio de algas y bacterias para el tratamiento de aguas residuales domésticas*. Castellón, 1 a 3 de septiembre de 2021, Comité Español de Automática.
- Casagli, F. y otros, 2021. ALBA: A comprehensive growth model to optimize algae-bacteria wastewater treatment in raceway ponds. *Water Res.*, Volumen 190, p. 116734.
- CEE, 1991. *Directiva 91/271/CEE SOBRE EL TRATAMIENTO DE LAS AGUAS RESIDUALES URBANAS*, s.l.: s.n.
- Chae, K. & Kang, J., 2013. Estimating the energy independence of a municipal wastewater treatment plant incorporating green energy resources. *Energy Convers. Manage.*, Volumen 72, pp. 664-672.
- De Godos, I. y otros, 2014. Assessing carbon and nitrogen removal in a novel anoxic–aerobic cyanobacterial–bacterial photobioreactor configuration with enhanced biomass sedimentation. *Water Res.*, Volumen 61, pp. 77-85.
- Empresarios Agrupados Internacional, 2021. www.ecosimpro.com. [En línea].
- Huber, P., 2014. Robust statistics. En: M. Lovric, ed. *International Encyclopedia of Statistical Science*. s.l.:Springer Berlin, pp. 1248-1251.
- Muñoz, R. & Guieysse, B., 2006. Algal–bacterial processes for the treatment of hazardous contaminants: a review. *Water Res.*, Volumen 40, p. 2799–2815.
- Reichert, P. y otros, 2001. River water quality model no. 1 (RWQM1): II. Biochemical process equations. *Water Sci Technol : J. Int. Assoc. Water Pollut. Res.*, 43(5), pp. 11-30.
- Sánchez-Zurano, A. y otros, 2021. ABACO: A New Model of Microalgae-Bacteria Consortia for Biological Treatment of Wastewaters. *Appl. Sci.*, 11(998).
- Solimeno, A., Gómez-Serrano, C. & Ación, F., 2019. BIO_ALGAE 2: improved model of microalgae and bacteria consortia for wastewater treatment. *Environmental Science and Pollution Research*, Volumen 26, pp. 25855-25868.
- Solimeno, A., Parker, L., Lundquist, T. & García, J., 2017. Integral microalgae-bacteria model (BIO_ALGAE): Application to wastewater high rate algal ponds. *Science of the Total Environment*, Volumen 601-602, pp. 646-657.
- Takács, I., Nolasco, G. & Patryioand, D., 1991. A Dynamic model of the clarification-thickening process. *Water Research*, 10(25), pp. 1263-1271.

Gaussian-Process based inference of electrolyte decomposition reaction networks in Li-ion battery failure

Dr Peter J. Bugryniec^a, Aaron Yeardley^a, Aarjav Jain^b, Nicholas Price^c, Dr Sergio Vernuccio^a and Dr Solomon F. Brown^{a,*}

^a*Department of Chemical & Biological Engineering, The University of Sheffield, Sheffield, S1 3JD, UK*

^b*Cavendish Laboratory, Cambridge University, Cambridge, CB3 0HE, UK*

^c*Department of Chemistry, Imperial College London, London, W12 0BZ, UK*
s.f.brown@sheffield.ac.uk

Abstract

Li-ion batteries (LIBs) are widely adopted in EVs and stationary battery energy storage due to their superior performance over other battery chemistries. But LIBs come with the risk of thermal runaway (TR) which can lead to fire and explosion of the LIB. Hence, improving our understanding of TR is key to improving LIB safety. To achieve this, we aim to develop a detailed model of LIB TR, as existing models are oversimplified and often lead to inaccuracies when compared to experiments. To build a realistic representation of the reaction network (RN) for LIB TR, we present a case study on the ethylene carbonate (EC) solvent component of the LIB electrolyte. We use a RN for EC identified from literature to build a micro-kinetic model and optimize it against experimental data. Parameters optimisation and sensitivity analysis for a complex RN is made possible by using Gaussian Processes (GPs). It is found that the only four of the 14 parameters influence the simulation output significantly. Also, this work highlights areas of GP development for improved surrogate modelling of this type of problem. From this the methodology can be scaled to larger networks and can be applied LIB TR models to improve their accuracy, which in turn will help the development of safer LIBs.

Keywords: Thermal runaway, Gaussian Process, Li-ion battery, Reaction network analysis, Robust optimization

1. Introduction

Li-ion batteries (LIBs) have become the favoured electrochemical energy storage device in EVs and stationary applications as they benefit from high energy density and low cost (Keshan et al., 2016). However, they can undergo the rare but hazardous phenomenon of thermal runaway (TR), which through exothermic decomposition can lead to fire and explosion (Wang et al., 2012). Improved battery safety is paramount as the use of LIBs is increasing enormously.

Computational modelling is a proven method to aid the design of safer LIBs (Abada et al., 2016; Bugryniec et al., 2020b,a). However, models used in previous works can be considered oversimplified as they do not consider (1) the multi-pathway and interdependent reaction network (RN) (Wang et al., 2012), or (2) the generation of decomposition products which could be used for chemical hazard assessment. Also, existing models can be limited to the experimental conditions on which they are developed. Hence, there is a desire to build a more detailed model of LIB TR

based on a realistic representation of the RN. In this paper a new approach is developed to meet this goal.

RNs can be developed by inference of reactants and products from experimental observations, or from molecular simulations studies (Campion et al., 2005; Hirai and Jinnouchi, 2021). Micro-kinetic modelling techniques can be applied, which provide a detailed description of the reaction pathways, including each possible elementary step and reaction intermediate. In this approach, the kinetic parameters are typically estimated based on a solid theoretical basis and no a priori assumptions on the rate determining steps are needed (Hermes et al., 2019).

We have previously shown (Milton et al., 2019; Yeardley et al., 2020) that applying Gaussian Processes (GPs) surrogate modelling to LIB modelling is beneficial due to the efficiency of the GP prediction technique (Sacks et al., 1989). Further, as the GP is mathematically tractable it allows sensitivity analysis (SA) and optimisation of the complex RN. The implementation of the GP can be conducted through two main methods. The first, is by fitting the GP to a heuristic fit of the computer simulation, providing information about the error between simulation and experimental data. The second, is to fit multiple GPs to create a direct surrogate model that uses the same input variables as the computer simulation to predict the same outputs as computer simulation.

The aim of this work is to develop a methodology for the efficient generation of realistic RNs of LIB TR, from which the optimal and most influential reaction parameters can be determined. To assess the validity of the aforementioned methodology, an exploratory study has been conducted to analyse the thermal decomposition of ethylene carbonate (EC) as a model component of the electrolyte.

2. Methodology

As stated previously, this work focuses on the RN of EC oxidation decomposition. For this work, the energy diagram from Hirai and Jinnouchi (2021) (containing 3 reaction pathways for the EC decomposition) is used to define the RN. This size of network provides a suitable problem to test and validate our proposed methodology. From the energy diagram a micro-kinetic model can be created. The forward and backward reaction steps are listed in Table 1, for a total of 14 reactions and 8 species, along with the corresponding activation energies of each step. The micro-kinetic model (a.k.a. the full order model or FOM) consists of a system of 8 ordinary differential equations (ODEs) describing the change in concentration of the 8 species. Each ODE is expressed as the algebraic sum of all relevant reaction rates (R_i) corresponding to a given reaction step, where R_i is governed by an Arrhenius equation of the form $R_i = A_i \exp(-E_{a,i}/RT) \prod_j C_j$.

To analyse the RN, and to obtain the optimised frequency factors, GPs are applied. To generate training data for the GP, the desired parameter space of the 14 frequency factors is sampled using Latin Hypercube sampling, which through the FOM, is used to generate 10 000 outputs that are used for training.

The log base 10 of each input variable is taken before being standardised as the bounds of the variables extend over several orders of magnitude. In this research both a GP behaving as a direct surrogate model and a GP predicting a heuristic fit are chosen to analyse the data. A GP is a non-parametric machine learning technique which takes a $1 \times d$ row vector of inputs \mathbf{x} and returns a Gaussian random variable through calculations using the predictive equations shown in Yeardley et al. (2020). The difference between the direct surrogate model and the heuristic fit is the output that the GP predicts. For the surrogate model, numerous GPs are trained to predict the mole ratio every step in temperature. Whereas, the heuristic fit maps the log of the 14 parameters to a single output of interest. This output is the root mean squared error (RMSE) between the experimental and simulated mole ratio over the entire temperature range. In essence, the heuristic fit GP behaves as a blackbox function using training data to predict the RMSE from the log of the

Table 1: Calculated activation energies for the forward and backward reactions of EC decomposition according to the energy diagram from Hirai and Jinnouchi (2021). The notation R_i in brackets identifies each reaction step.

Reaction	Activation Energy (eV)	
	Forward	Reverse
$\text{EC} \longleftrightarrow \text{C}_2\text{H}_4\text{O} + \text{CO}_2$	1.12 (R_1)	1.86 (R_2)
$\text{EC} \longleftrightarrow \text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3$	0.06 (R_3)	1.31 (R_4)
$\text{EC} \longleftrightarrow * \text{C}_2\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3$	0.20 (R_5)	0.37 (R_6)
$\text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3 \longleftrightarrow \text{C}_2\text{H}_3\text{O} + \text{CO}_2 (+ \text{C}_3\text{H}_5\text{O}_3)$	0.84 (R_7)	1.55 (R_8)
$\text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3 \longleftrightarrow * \text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3$	1.08 (R_9)	0.00 (R_{10})
$* \text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3 \longleftrightarrow \text{C}_2\text{H}_5\text{O} + \text{CO}_2 (+ \text{C}_3\text{H}_3\text{O}_3)$	0.21 (R_{11})	1.58 (R_{12})
$* \text{C}_3\text{H}_5\text{O}_3 + \text{C}_3\text{H}_3\text{O}_3 \longleftrightarrow \text{C}_2\text{H}_3\text{O} + \text{CO}_2 (+ * \text{C}_3\text{H}_5\text{O}_3)$	0.84 (R_{13})	1.56 (R_{14})

model parameters

Both methods require GP learning which uses training data to optimise $d + 2$ hyperparameters found in the predictive equations, consisting of Λ , σ_f , and σ_e . This optimisation is completed by maximising the marginal likelihood $p[\mathbf{y}|\mathbf{X}]$ using the ROMCOMMA software library (Milton and Brown, 2019).

The GPs are used for a global sensitivity analysis (GSA) and to optimise the model simulation with respect to the experimental data from Lamb et al. (2015). For this work, we opted to implement a GSA using the variance based Sobol' indices technique (Sobol, 1993, 2001). As previously mentioned, GPs are mathematically tractable, therefore, we can compute the semi-analytic evaluation of complex multi-dimensional integrals resulting in Sobol' indices which measure the contribution to the outputs variance attributable to each input parameter. Hence, to understand how the input parameters impact the model simulations closeness to the experimental data, the GSA required an output that measures said closeness, resulting in a GP predicting the heuristic fit. Additionally, the direct surrogate model is required to be able to optimise the parameters throughout the model simulation as a function of temperature. In this way, the GP is trained to accurately copy the model simulation and is used to make fast, efficient predictions. Thus, the optimisation of model parameters is achieved by minimising the error between the GP predictions and the experimental data. GPs predict a probability distribution presenting the uncertainty in the predictions. Therefore, the optimisation should include constraints where the predicted standard deviation (STD) is limited to be within 2 standardised units from the mean. In this research, we compare both a constrained and an unconstrained GP for parameter optimisation.

3. Results and Discussion

3.1. GP Validation

Both techniques are tested using the 5-fold cross validation technique. The results of which are scrutinised by calculating the error metrics between the GP predictions and the model simulation data. In this work, we conduct a robust validation by calculating three popular error metrics and by plotting figures to provide a visual understanding of the error. Table 2 shows the coefficient of determination (R^2) and the root mean squared error (RMSE) values for each output. Both diagnostics measure a skill score, corresponding to the accuracy of the predicted mean (Al-Taweel, 2018), but with different scales. Further, the predictive distribution is analysed to ensure the GPs are not predicting with over confidence. This is shown by counting the outliers for any prediction where it's true standardised value is outside of the predictions 95% uncertainty distribution. Table 2 presents the outliers of both around the 5% as expected for a normal distribution. However, the two skill scores show the heuristic fit to have a worse accuracy than the direct surrogate

Table 2: Error metrics used to validate the GP models.

GP	R^2	RMSE	Percentage of outliers
Heuristic Fit	0.486	0.486	5.60 %
Direct Surrogate Model	0.931	0.279	5.25 %

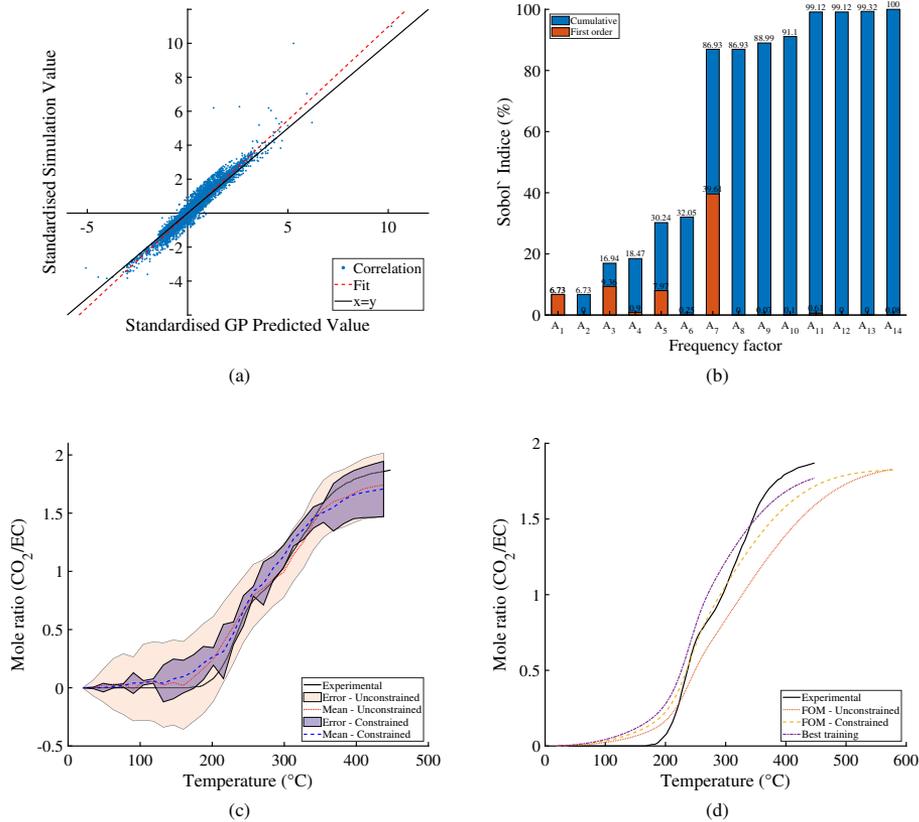


Figure 1: a) Validation, standardised value predicted by GP vs standardised value from FOM simulation, also shown is the linear fit (dashed line, $y = 1.102x - 0.041$ with $R^2 = 0.931$) of the scatter data. b) The cumulative Sobol' indices' overlaid with first order Sobol' indices' for each of the 14 frequency factors, first order totals 65.7 %. c) Optimisation output, GP predicted mean with error bounds (one STD) for scenarios where the STD is constrained and unconstrained. d) FOM output using optimised parameters determined by GP. (Note experimental data is from Lamb et al. (2015))

model as the R^2 is small and the RMSE is large. Further, Fig. 1(a) shows the correlation of GP predictions against the outputs of the FOM in a standardised form. The quality of prediction is shown by the closeness of the fit line to the $y = x$ line, with a coefficient of determination equal to 0.931. However, there are some large residuals, as well as instances of residuals far away from the straight line.

From this, it can be seen that the GP predicting the heuristic fit is satisfactory and appropriate for the calculation of Sobol' indices due to 95% of the true observed values fitting into the predicted distribution. However, the point estimations for the heuristic fit require further work to increase accuracy. Hence, the direct surrogate model is used for a parameter optimisation.

Table 3: Error (RMSE) from predicted output and FOM output (calculated against experimental) for optimised parameters under un/constrained conditions.

Type	GP predicted	FOM
Best training data	-	0.1190
Constrained GP	0.0980	0.1197
Unconstrained GP	0.0821	0.2304

3.2. Sensitivity Analysis

SA of the model output dependent on the 14 frequency factors can be realised by calculation of the Sobol' indices. The first order and cumulative Sobol' indices are presented in Fig. 1(b). From the first order Sobol' indices, it can be seen that the frequency factors relating to the initial decomposition of EC (A_1, A_3, A_5) by the forward reactions R_1, R_3, R_5 are of the most dominant parameters. This is besides parameter A_7 , which is the most dominant and relates to the second stage (forward) reaction from the products of R_2 and leads to the generation of CO_2 . Further, the first order Sobol' indices show that the frequency factors for all the backwards reactions and forward reactions R_9 and above have negligible influence of CO_2 generation.

From the cumulative Sobol' indices in Fig. 1(b) it can be seen that the increase in value is driven by the interactions from the frequency factors related to the first three forward reactions. The backwards reactions, with A_2, A_4, A_6 , have little effect on the increase in indices value. A_7 has a great effect on increasing the cumulative Sobol' indices value, equating to almost 87 % of the total Sobol' indices. Where over 50 % of the value is due to interactions with the first 6 frequency factors. After this, little change is seen in Sobol' indices until A_{11} at which there is an increase to 99 %. From Table 1, the influence of A_{11} can be rationalised by the fact that it is a second stage reaction that leads to CO_2 generation by a relatively low activation energy. Overall, first order Sobol' indices total 65.7 %, thus interaction account for 34.3 %. Hence, it can be seen that the frequency factors relating to R_1, R_3, R_5 and R_{11} , which relate to the forward reaction of the first stage of EC decomposition and the forward reaction of the decomposition of the products from R_5 , have the most influence on predicted output.

3.3. Parameter Optimisation

The best training data, with the lowest RMSE relative to experimental, was chosen from FOM simulations, see Table 3. So the aim is to optimise the GP against experimental data to determine parameters that lead to a smaller RMSE than 0.1190. The results of the direct surrogate model for parameter optimisation are presented in Figs. 1(c) and 1(d). Fig. 1(c) shows the GP predicted mean with STD for two scenarios, where the deviation is (1) unconstrained and (2) constrained. The constraint was applied to minimise the size of the STD around the predicted mean. It can be seen from Table 3 that the unconstrained GP predicts a smaller RMSE than the constrained GP. However, when the optimised parameters from the GP are implemented in the FOM the resulting output (see Fig. 1(d)) leads to RMSEs larger than the GP predicted in both scenarios. Further, the parameters from the unconstrained GP lead to a much larger RMSE than the best training data, while the parameters from constrained GP lead to a similar RMSE to the training data.

It is shown that constraining the GP increases the RMSE of the prediction but leads to a GP that is better at capturing the FOM behaviour. However, as the GPs both predict better than the FOM output for the same parameters it indicates that our surrogate model is not accurately capturing the behaviour of the FOM. This may be due to the slight deviation of the correlation plot from the $y=x$ line (see Fig. 1(a)). Or it may be possible that the FOM is already as close to optimal, due to the large number of starting point considered, and the GP cannot improve on it. To further refine the GP to improve the correlation plot, then the follow can be applied: 1) restricting the training data

to sets with lower RMSE, 2) using both cross-validation and a test set to validate the GP so that there is a data set purely for testing, 3) optimise against experimental data using the continuous ranked probability score to account for uncertainty rather than using the RMSE.

4. Conclusion

The modelling of chemical reactions is usefully for calculating reaction barriers, the effects of catalysts, product yields and analysing theoretical reactions schemes. Here, decomposition of EC is studied using micro-kinetic modelling. A GP approach is applied for SA and optimisation of the frequency factors. The SA shows that most of the error between the simulation output and experimental data comes from 4 of the 14 parameters. The GP, while predicting a mean close to the simulation, currently does not capture the FOM to a standard required to calculate optimised model parameters. This work provides lessons for GP development applicable to RN analysis. It can be used for safer LIB design by applying it to LIB TR models so that the most likely reaction pathway and production of hazardous species can be predicted.

5. Conflicts of Interest & Acknowledgements

There are no conflicts to declare. This work was supported by the Faraday Institution [grant number FIRG028].

References

- S. Abada, G. Marlair, A. Lecocq, M. Petit, V. Sauvant-Moynot, F. Huet, 2016. Safety focused modeling of lithium-ion batteries: A review. *Journal of Power Sources* 306, 178–192.
- Y. Al-Taweel, 2018. Diagnostics and Simulation-Based Methods for Validating Gaussian Process Emulators. Phd thesis, University of Sheffield.
- P. J. Bugryniec, J. N. Davidson, S. F. Brown, 2020a. Advanced abuse modelling of Li-ion cells – A novel description of cell pressurisation and simmering reactions. *Journal of Power Sources* 474, 228396.
- P. J. Bugryniec, J. N. Davidson, S. F. Brown, 2020b. Computational modelling of thermal runaway propagation potential in lithium iron phosphate battery packs. *Energy Reports* 6, 189 – 197, 4th Annual CDT Conference in Energy Storage & Its Applications.
- C. L. Campion, W. Li, B. L. Lucht, 2005. Thermal Decomposition of LiPF₆-Based Electrolytes for Lithium-Ion Batteries. *Journal of The Electrochemical Society* 152 (12), A2327.
- E. D. Hermes, A. N. Janes, J. R. Schmidt, 2019. Micki: A python-based object-oriented microkinetic modeling code. *The Journal of Chemical Physics* 151 (1), 014112.
- H. Hirai, R. Jinnouchi, 2021. Discovering chemical reaction pathways using accelerated molecular dynamics simulations and network analysis tools – application to oxidation induced decomposition of ethylene carbonate. *Chemical Physics Letters* 770, 138439.
- H. Keshan, J. Thornburg, T. Ustun, January 2016. Comparison of lead-acid and lithium ion batteries for stationary storage in off-grid energy systems. *IET Conference Proceedings*, 30 (7 .)–30 (7 .)(1).
- J. Lamb, C. J. Orendorff, E. P. Roth, J. Langendorf, 2015. Studies on the Thermal Breakdown of Common Li-Ion Battery Electrolyte Components. *Journal of The Electrochemical Society* 162 (10), A2131–A2135.
- R. Milton, P. Bugryniec, S. Brown, 2019. Parameter estimation for thermal runaway of li-ion cells: a gaussian process approach. In: A. A. Kiss, E. Zondervan, R. Lakerveld, L. Özkan (Eds.), 29th European Symposium on Computer Aided Process Engineering. Vol. 46 of *Computer Aided Chemical Engineering*. Elsevier, pp. 775 – 780.
- R. A. Milton, S. F. Brown, 2019. ROMCOMMA.
URL <https://github.com/C-0-M-M-A/rom-comma>
- J. Sacks, W. Welch, T. Mitchell, H. Wynn, 1989. Design and analysis of computer experiments. *Journal of Statistical Science* 4 (4), 409–435.
- I. M. Sobol, 1993. Sensitivity analysis for nonlinear mathematical models. *Mathematical Modelling Computational Experiments* 1 (4), 407–414.
- I. M. Sobol, 2001. Global sensitivity indices for nonlinear mathematical models. Review. *Mathematics and Computers in Simulation* (55), 271–280.
- Q. Wang, P. Ping, X. Zhao, G. Chu, J. Sun, C. Chen, 2012. Thermal runaway caused fire and explosion of lithium ion battery. *Journal of Power Sources* 208, 210–224.
- A. S. Yeardley, P. J. Bugryniec, R. A. Milton, S. F. Brown, 2020. A study of the thermal runaway of lithium-ion batteries: A gaussian process based global sensitivity analysis. *Journal of Power Sources* 456, 228001.

Energy evaluation of processes for the production of hydrogen from biomass biodigestion under Aspen Plus

Lokmane Abdelouahed, Fatma Kourdourli, Bechara Taouk, Lionel Estel

*Normandie Univ, INSA Rouen Normandie, UNIROUEN, LSPC-Laboratoire de Sécurité des Procédés Chimiques, 76800, Saint Etienne Du Rouvray, France
lokmane.abdelouahed@insa-rouen.fr*

Abstract

Hydrogen is considered as a clean energy carrier that can be produced from several renewable resources. Biomass as an alternative energy resource can be used to produce hydrogen to substitute fossil fuels. The objective of this work is to propose a coupled process modelling under Aspen for H₂ production based on anaerobic digestion for syngas production and methane transformation to hydrogen based on steam reforming (process 1) and partial oxidation (process 2). The modelling was validated with experimental results. The modelling has allowed to establish an evaluation of the energetic performance of each installation with a quantification of the CO₂ production which is considered as greenhouse gas. The feed of the process was set equal to 0.33 l/day of cow manure biomass for both modelling. The results show that the process 1 produced 5.71 l/day of H₂ with an energy efficiency of 82.72% and the emission of the CO₂ is about 12.83 kg per kg of H₂ produced. These values are highly advantageous compared to the second process.

Keywords: biomass, anaerobic digestion, hydrogen, methane, modelling.

1. Introduction

The current production of hydrogen is mainly based on fossil fuels as feedstocks (coal, natural gas, oil, etc.), and 4% is produced by electrolysis of water [1]. Hydrogen could be the fuel of the future because of its high energy content, as well as it presents environmental, economic and social issues, it is considered a clean energy that does not cause air pollution. The production of hydrogen from renewable sources, in particular biomass, instead of fossil fuels, is a green way and shows many advantages due to its availability, abundance and its carbon neutrality.

Anaerobic digestion, also called "methanization", is the transformation of organic materials into a gas rich in methane and carbon dioxide called biogas containing about 20 to 40% of the energy value of the original biomass. It is carried out in the absence of oxygen by a complex microbial community in a closed chamber (digester, fermenter or reactor), and is widely used for the treatment of wet organic waste (80 to 90% moisture) [2,3].

The different hydrogen production techniques are mainly based on syngas and methane reforming techniques, such as (i) steam reforming (methane and steam converted to syngas), (ii) partial oxidation reforming (methane and oxygen converted to syngas), (iii) autothermal reforming (methane, oxygen and steam converted to syngas), (iv) dry reforming (methane and carbon dioxide converted to syngas), (v) bi-reforming (methane,

carbon dioxide and steam converted to syngas), (vi) tri reforming (methane carbon dioxide, steam and oxygen to syngas).

In the light of these considerations, this work is aimed at studying two different processes of hydrogen production from biogas. The two main steps involved in these processes are the production of biogas from biomass by anaerobic digestion and the transformation of the methane produced to produce pure hydrogen.

2. Process modeling in Aspen Plus

2.1. Process description

The main steps in the production of hydrogen from biogas or methane are shown in the Figure 1. The bio-digester model was established and validated by [4]. Steam reforming, partial oxidation reforming and the autothermal reforming models were validated by Hajjaji et al. [5] while the bi-reforming, dry reforming and tri-reforming models were validated by Phan et al. [6].

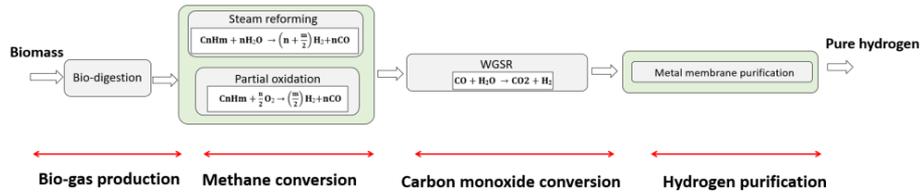
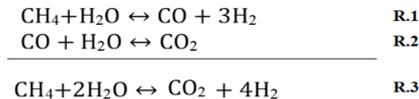


Figure 1: Main steps in the production of hydrogen from biogas or methane

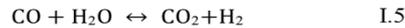
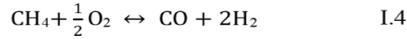
The thermodynamic method selected was NRTL (Non-Random Two Liquid) which is recommended for anaerobic digester simulation [7]. The cow manure is considered as the feedstock with a loading rate of 0.33 l/days. The first step of anaerobic digestion which is the hydrolysis of biomass was not modeled due to complexities such as the difficulty associated with specifying the composition of biomass in terms of its constituents, therefore, the flow for this process is assumed to be pretreated (hydrolyzed) [6,7]. Two streams were produced at the outlet of the digester (i) biogas and (ii) undigested slurry (mixture of solid and liquid phases) which can be used for agricultural purposes.

Steam Reformation of Methane (SRM) is the most industrially applied and commonly used to produce hydrogen in large quantities. This technique consists in reacting methane with steam to produce a synthesis gas, in the presence of a nickel-based catalyst, at a high temperature of 800 to 1000 °C, and at a moderate pressure of about 5 to 25 bars. This transformation is followed by different operations which lead to the production of hydrogen and carbon monoxide, carbon dioxide, methane and water [1,8], as shown in the following reaction system:



The first reaction (R.1) corresponds to the steam reforming, it is endothermic and is characterized by a ratio H_2/CO of about 3 where the second reaction (R.2) corresponds to the conversion of CO, known as the water gas shift reaction, and it is slightly exothermic. A desulfurization pre-treatment is usually applied to the feedstock, to avoid poisoning the catalysts.

The partial oxidation of methane consists in converting hydrocarbons into hydrogen and carbon monoxide in the presence of oxygen, carried out at high temperature from 1200 to 1500°C and at high pressure from 20 to 90 bars. The reaction is exothermic and takes place with or without catalyst depending on the load and the reactor used, this technique remains however much less common than steam reforming because of its high cost. Shell and Texaco processes are the most known, the operating conditions for these two processes involve a temperature of around 1000°C and a pressure of 35 bar [9].



The CO conversion reaction (Water Gas Shift reaction) carries out the conversion of the CO existing in the syngas by reaction with steam, this reaction is also the same in all reforming processes. The WGSR is balanced as follows:



This conversion is carried out in two successive stages, the first the conversion at high temperature which is carried out in HTS reactor (High Temperature Shift) at temperatures between 350°C and 450°C, then, the conversion at low temperature in a LTS reactor (Low Temperature Shift) at temperatures between 200°C and 250°C [10].

Metal membranes operate at high temperature and convert catalytically hydrogen gas into hydrogen atoms that can pass through solids. The atoms recombine on the product side into molecular hydrogen [11]. These membranes operate at temperatures ranging from 300°C to 600°C and the purified hydrogen exits the separator at an atmospheric pressure [8][12].

2.2. Process flowsheet and operating conditions

Several blocks were used to achieve this modelling for both processes. The process 1, for example, is essentially composed of 4 blocks:

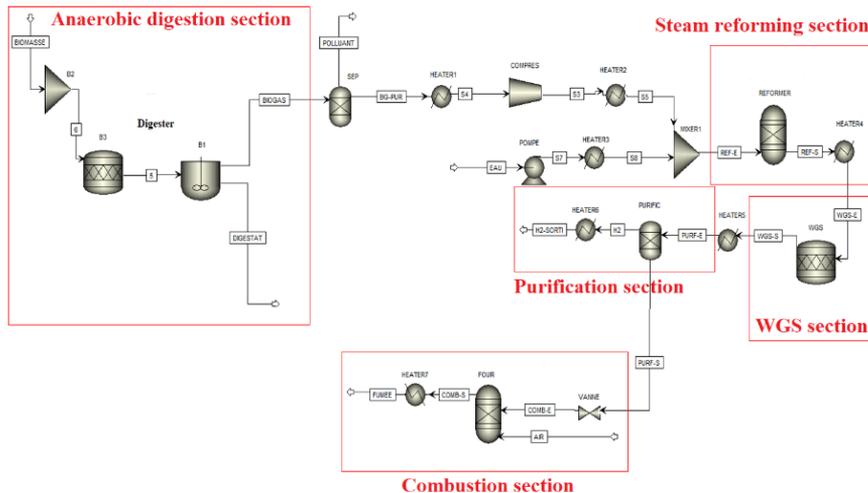


Figure 2: Flowsheet of hydrogen production from biomass via steam reforming
 (i) **Reformer block:** The process feeds the reformer modeled by the RGIBBS model at 973K and 10 atm by methane and steam at high pressure (10 atm). The syngas at the exit of the reformer block (REF-S stream) is cooled to 573K to be returned to the

WGS block. **(ii) The Water Gas Shift block:** The syngas at 573 K and 10 atm is converted by steam into a H₂ and CO₂ rich gas. The WGS reactor is modeled by the RSTOIC with a conversion rate of CO equals 0.98. The (WGS-S stream) from this block is heated to 723 K before being fed into the purification block. **(iii) The purification block:** The purification block is modeled by the Sep model under Aspen plus. The gas leaving the purifier still containing CH₄, CO, CO₂, H₂O and residual H₂ (PURF-S stream) is expanded to atmospheric pressure, to be returned to the combustion block. **(iv) The combustion block:** The combustion furnace is fed with the gas from the purification block to provide heat to the reforming reactor. The global flowsheet in the case of process 1 (coupling with steam reforming) is shown in figure 2.

2.3. Process energy analysis

The energy performance of coupled “anaerobic digestion-reforming” processes has been examined based on several [8,13]:

$$\eta_{biomass} = \frac{D_{H_2} \times LHV_{H_2}}{D_{biomass} \times LHV_{biomass} + W_{compres} + W_{pump}} \quad \text{Eq.1}$$

$$\eta_{biogas} = \frac{D_{H_2} \times LHV_{H_2}}{D_{biogas} \times LHV_{biogas} + W_{compres} + W_{pump}} \quad \text{Eq.2}$$

$$\eta_{methane} = \frac{D_{H_2} \times LHV_{H_2}}{D_{CH_4} \times LHV_{CH_4} + W_{compres} + W_{pump}} \quad \text{Eq.3}$$

With :

D_i : The mass flow rate of i component (hydrogen, biomass, biogas, methane) (kg/s), processes (kg/s), LHV_i : Lower Heating Value of i component (hydrogen, biomass, biogas, methane) (kJ/kg), W_i : Mechanical work of the compressor and pump (kW)

The biomass considered as hydrolyzed cow manure. The LHV of this last was set equal to 12.47 MJ/kg [14].

3. Results and discussion

The simulations of the "DA-Reforming" processes have been modeled on Aspen Plus, these models allow to compare the production of hydrogen from biomass via the coupling of the anaerobic digestion process with the different reforming models.

Table 1: comparison between the two processes

Process	Process 1	Process 2
H ₂ (L/day)	5.57	3.58
$\eta_{biomass}$ (%)	30.58	19.67
η_{biogas} (%)	82.72	47.21
$\eta_{Methane}$ (%)	83.24	53.51
Exhaust gas (L/day)	30.06	33.18
CO ₂ (L/day)	5.2	5.2
kg CO ₂ per kg H ₂	12.83	19.95

The results given by the simulations carried out show that the process 1 "DA-steam reforming" is the most productive of hydrogen, this process has the greatest volume flow in H₂ produced, for a feeding of 0.33 l/day of biomass, we obtained 5.57 l/day of hydrogen produced. The purified biogas containing CH₄ (51.97%), CO₂ (32.07%), H₂O (15.77%), H₂ (0.19%) where the presence of CO₂ and moisture in the biogas is an advantage or the biogas can be used directly, without the removal of steam because it promotes the WGS reaction [7], this reduces energy consumption and improves the ratio

H₂/CO, the biogas is reformed with an excess of steam. Figure 3 shows the mass and energy balance for the two simulations.

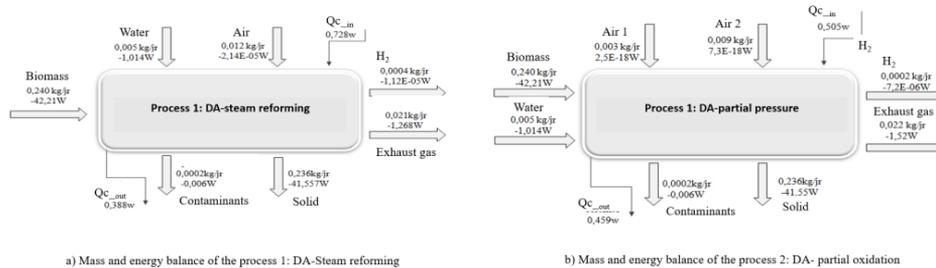


Figure 3: Mass and energy balance for the two process models

based on the energy balance and the different efficiencies shown above, the process1 (DA-steam reforming) has shown more interesting efficiencies compared to process 2 (DA- partial pressure (table 1).

It is necessary to consider the related pollutant emissions for a comparison of the two coupling processes for hydrogen production from biogas. Table 1 presents a comparison between the processes in terms of CO₂ emissions and exhaust gas released into the air from each process at identical temperatures and pressures. The CO₂ emissions per kg of H₂ produced in both processes show that process 1 emits less CO₂ than the process for producing the same amount of H₂. Overall, the amount of CO₂ released is almost the same for both process and equal to 5.2 L/day.

Conclusion

This study focuses on hydrogen production processes from biogas obtained from anaerobic digestion of biomass. The overall process consists of two main steps: (1) biogas production and (2) biogas reforming to generate pure hydrogen adapted from the literature. The important results of this section are summarized in the following points:

- The process 1 coupling DA and a steam reforming section is the most productive of H₂, with a flow of H₂ of 5,57 l/day, for a feed of 0.33 l/day of biomass.
- An energy efficiency η_{biogas} (hydrogen to biogas) of the process1 is about 82.72% which is the highest.
- The process 1 remains significantly interesting compared to process 2 in terms of quantity of CO₂ and gas exhaust released for 1 kg of H₂ produced

This global model can be improved by replacing the hydrogen separators by a purification section based on more efficient models for a better estimation of the energy consumed in the installation. Also, the energy optimization based on the pinch method seems indispensable followed by an exergetic study for each coupling process to calculate the exergetic efficiency and to determine the thermodynamic irreversibility.

References

- [1] N. Hajjaji, Marie-Noëlle Pons, Viviane Renaudin, 2011, Ammar Houas, Analyse de cycle de vie exergetique de systèmes de production d'hydrogène, Institut National Polytechnique de Lorraine.
- [2] B. Benabdallah, Y. Schenkel, A. Riedacker, P. Girard, 2005, Organisation internationale de la francophonie, Agence de l'environnement et de la maîtrise de l'énergie (France), B. Centre de recherches agronomiques (Gembloux, *Guide Biomasse-Energie*.
- [3] G. De Gioannis, L. F. Diaz, A. Muntoni, A. Pisanu, 2008, *Waste Management*, 28, 1801.

- [4] K. Rajendran, H. R. Kankanala, M. Lundin, M. J. Taherzadeh, 2014, *Bioresource Technology* 168, 7.
- [5] N. Hajjaji, 2011, Analyse de cycle de vie exergétique de systèmes de production d'hydrogène, Institut National Polytechnique de Lorraine.
- [6] T. S. Phan, Élaboration, caractérisation et mise en œuvre d'un catalyseur dans le reformage du biogaz en vue de la production d'hydrogène vert, 2020, Ecole des Mines d'Albi-Carmaux.
- [7] R. K. Balaji, K. P. Rajan, U. B. R. Ragula, 2020, *International Journal of Hydrogen Energy* 45, 18226.
- [8] A. P. Simpson, A. E. Lutz, 2007, *International Journal of Hydrogen Energy*, 32, 4811.
- [9] M. Roby, 1997, "Gazéification des résidus pétroliers par le procédé Shell : Dossier complet | Techniques de l'Ingénieur," can be found under <https://www.techniques-ingenieur.fr/base-documentaire/procedes-chimie-bio-agro-th2/procedes-industriels-de-base-en-chimie-et-petrochimie-42329210/gazeification-des-residus-petroliers-par-le-procede-shell-j5440/>.
- [10] B. Smith, L. Muruganandam, L. Murthy, S. Shantha, 2010, *International Journal of Chemical Reactor Engineering - INT J CHEM REACT ENG* 8, DOI 10.2202/1542-6580.2238.
- [11] SDTC Sustainable Development Technology Canada, 2006, *Rapport d'investissement Du Développement Durable Carburant Renouvelable – Hydrogène Analyse d'investissement Du DDMC*.
- [12] S. Adhikari, S. Fernando, 2006, *Industrial & Engineering Chemistry Research - IND ENG CHEM RES* 45, DOI 10.1021/ie050644l.
- [13] N. Hajjaji, V. Renaudin, A. Houas, M. Pons, 2010, *Chemical Engineering and Processing: Process Intensification* 49, 500.
- [14] A. Szymajda, G. Łaska, M. Joka, 2021, *Energies* 14, 1192.

Life cycle optimization of energy systems integrated with carbon capture and utilization

Iasonas Ioannou,^a Alex Marcandetti,^a Gonzalo Guillén Gosálbez^{a*}

^a*ETH Zürich, Vladimir-Prelog-Weg 1, 8093 Zürich, Switzerland*

**gonzalo.guillen.gosalbez@chem.ethz.ch*

Abstract

Carbon capture and utilization (CCU) based chemicals' have become an appealing strategy towards more sustainable production patterns. Several previous CCU studies often overlook the variability of renewable power technologies (by assuming average capacity factors) and focus on single technologies. To enlarge the scope of these studies, we optimize location-specific renewable energy systems (ES) clustered with CCU, while accounting for the yearly variability, to evaluate their environmental and economic performance more accurately. Our approach, based on a mixed-integer linear programming (MILP) model, coupled with life cycle assessment (LCA) principles, was applied to evaluate nine locations in the European Union (EU) and considering solar photovoltaics (PV), wind turbines, pumped hydro storage (PHS), and the interactions with the national grid. Overall, our tool could complement future assessments, expanding their scope and boosting the successful industrial implementation of CCU.

Keywords: Life cycle optimization; Energy system; Carbon capture and utilization.

1. Introduction

The fossil-based chemical industry should transition towards a low-carbon future aligned with the Paris climate agreement target (United Nations, 2016). Notably, even though the EU has adopted several policies, the envisioned carbon neutrality goal by 2050 is still beyond reach (European Commission, 2019). Thus, reducing the chemical sector's footprint could aid in achieving the climate mitigation target, which could be attained by replacing fossil-based resources with renewable carbon. Among the proposed alternatives, CCU, which requires large amounts of energy (often via an energy carrier, e.g., electrolytic H₂) to activate the carbon dioxide (CO₂), has received substantial attention (Ioannou et al., 2021). Moreover, by using CO₂ feedstock captured directly from the air –while using low-carbon energy (Deutz and Bardow, 2021)–, CCU could significantly aid the direct atmospheric carbon removal.

Within this general context, the design, location, and operation of the ES, covering the power demand, strongly affect the performance of the low-carbon chemicals (Ioannou et al., 2020). At present, LCA assessments of CCU assume a single power technology, e.g., wind or solar, for H₂ production while overlooking energy storage and links with the power sector. Furthermore, CCU chemicals are often economically inferior relative to their fossil-based counterparts due to their high energy consumption and the low capacity factor of the electrolyzer, which is powered by renewable technologies (Parkinson et al., 2019). Therefore, the variability and availability of renewable power could act as barriers to the successful implementation of CCU.

Here we develop a MILP model to design an ES-CCU cluster, considering nine EU locations, delivering an energy demand at a minimum levelized cost of electricity (LCOE). We integrate the ES with electrolytic H₂, CO₂ capture, and their transformation to methanol (MeOH), benefiting the sector's downstream applications. Our analysis

shows that since the ES's performance can vary significantly within and among countries, an integrated design of the ES-CCU clusters is key to assessing the chemicals' production feasibility and viability accurately.

2. Methods

To carry out our analysis, we developed a network model to design location-specific renewable ES, with energy storage, to minimize the *LCOE*. The latter aim could aid in coping with the variability of renewable technologies while attaining reductions in the cost of the CCU chemicals. The temporal representation is 2015 –on an hourly discretization–, which is adequate for the scope of the current design study. Hence, our model's unique feature is that it explicitly considers the hourly synergetic effects of hybrid ES –and storage– for CCU applications. The proposed methodology is divided into five steps, described briefly in the following sections: (i) selection of location(s), (ii) life cycle inventory generation (LCI) –based on assumptions, process simulation data or other sources–, (iii) model execution to obtain a tailored ES for the location defined in (i), while delivering the hourly power consumption defined in (ii), (iv) techno-economic and environmental assessment, and (v) interpretation of results.

2.1. Locations

We focus on three EU countries and three regions within each of them –Germany: Schwerin, Essen, and Haundorg; Austria: Parndorf, Tulln, and Zwettl; and Spain: Tarragona, Huelva, and Cartagena. Thus, we create a representative sampling to interpret the potential behaviour of the location-specific ES-CCU systems.

2.2. LCI of the CO₂-based MeOH production cluster

We design an ES that supplies the power requirements of a chemical facility, based on CO₂ capture and electrolytic H₂ generation, that produces 11.08 kt y⁻¹ of MeOH, acting as the functional unit. Within this general context, we assume an electrolytic efficiency of 80 % based on the LHV of H₂ and consider the co-product oxygen as a burden-free byproduct. Furthermore, we retrieved from the literature the LCI of the CO₂ hydrogenation to MeOH (González-Garay et al., 2019) and of the direct air capture (DAC) process (Fasihi et al., 2019). Finally, the LCI of raw materials and utility inputs are retrieved from the Ecoinvent V3.5 database (Wernet et al., 2016). Based on the latter, the hourly delivered power should be 10.84MW (amounting to 94.95 GWh y⁻¹), where 10.00MW is consumed for the generation of electrolytic H₂, and the remaining is used to synthesize MeOH and DAC (0.38 and 0.46MW, respectively). Furthermore, the facility will require 9.13 GJ of steam to capture 16.06 kt y⁻¹ of CO₂. The remaining inputs and direct emissions for MeOH production can be found in the original source. A cradle-to-gate LCA is carried out using the global warming (GW) indicator of the ReCiPe 2016 Midpoint (H) methodology (Huijbregts et al., 2017).

2.3. Mathematical model

We briefly describe the ES model, expressing parameters with regular letters and variables with italics. Our model integrates a set of power technologies (solar PV and wind turbines, $i \in I$), exploiting the complementary strengths to design the hybrid ES with storage at minimum *LCOE* –Eq.(1). The *LCOE* is based on the total annualized cost (*TAC*), and the predefined power demand of electrolysis, DAC, and MeOH production ($H_t + D_t$).

$$\min \left(LCOE = \frac{TAC}{\sum_{t \in T} (H_t + D_t)} \right) \quad (1)$$

The TAC considers the costs of purchasing power –from the national grid–, and the capital and operational expenses for installing (i) a PHS facility, with installed capacity S_{max} , (ii) solar PV, with installed capacity $CAP_{i=pv}$ and 3 kW nominal capacity per unit (NCAP), and (iii) wind turbines, with installed capacity $CAP_{i=wind}$ and 3 MW NCAP.

$$TAC = \sum_{t \in T} Grid_t \cdot LCOE_{gr} + Storage_{CAPEX} \cdot crf_s \cdot S_{max} + \sum_{i \in I} ((CAPEX_i \cdot crf_i + OPEX_i) \cdot CAP_i) \quad (2)$$

, where crf is the annual capital charge. The grid interactions and the PHS are necessary to satisfy the ES hourly availability, and thus, cope with the intermittent character of the installed renewables. CAP_i is calculated based on the $NCAP_i$ of the technology i , and the respective number of units –integer variable N_i .

$$CAP_i = NCAP_i \cdot N_i, \quad \forall i \in I \quad (3)$$

We assume that, within the ES lifespan, the annual local solar irradiation and wind speed patterns (obtained by "European Commission: PVGIS tool") will remain constant during the time horizon, and thus, the power generated from a unit i ($NP_{i,t}$) –Eq.(4)–, for 2015 is representative for the upcoming years.

$$NP_{i,t} = f(\text{wind speed, solar irradiation}), \quad \forall i \in I, t \in T \quad (4)$$

The number of renewable units is then connected with the hourly power generation from the respective renewable source – $NP_{i,t}$.

$$P_{i,t} = NP_{i,t} \cdot N_i, \quad \forall i \in I, t \in T \quad (5)$$

An energy balance is defined –Eq.(6)– and expressed in power units since the hourly representation allows us to omit the $\Delta t = 1h$.

$$H_t + D_t + PHS_{ch,t} + \sum_{i \in I} Cur_{i,t} = \sum_{i \in I} P_{i,t} + Grid_t + PHS_{di,t}, \quad \forall t \in T \quad (6)$$

The power demand ($H_t + D_t$) is covered by the renewable technologies ($P_{i,t}$), by purchasing power from the grid ($Grid_t$), and by discharging the PHS ($PHS_{di,t}$). Notably, the PHS capacity may not be sufficient to store the excess renewable power ($PHS_{ch,t}$) at all times. Thus, we consider power curtailment ($Cur_{i,t}$) to the national grid with zero cost. We further constrain the curtailment since it cannot physically exceed the amount delivered from the respective renewable source.

$$Cur_{i,t} \leq P_{i,t}, \quad \forall i \in I, t \in T \quad (7)$$

The PHS stored energy (S_t) at period t is determined from a second balance –Eq.(8)–, including the charging and discharging of power, with an efficiency η_s , and a self-discharge, with a coefficient λ_{di} . Therefore, due to losses by (i) charging-discharging efficiency or (ii) curtailment, the model prioritizes the direct use of renewable power.

$$S_t = (1 - \lambda_{di}) \cdot S_{t-1} + PHS_{ch,t-1} \cdot \eta_s - PHS_{di,t-1} / \eta_s, \quad \forall t \in T > 1 \quad (8)$$

Furthermore, since we consider a yearly base to design the ES –8760 periods–, a periodic condition is necessary to model the storage facility dimensions appropriately.

$$S_{t=1} = S_{t=8760} \quad (9)$$

Subsequently, the highest value of S_t is defined as the maximal amount of stored energy (S_{max}) –Eq.(10)–, which is needed for the cost calculations in Eq.(2).

$$S_t \leq S_{max}, \quad \forall t \in \mathbf{T} \quad (10)$$

We further define the binary variable y_t which restricts the PHS discharging when charging, and vice versa, while an upper bound, C_{ub} , limits both energy flows.

$$PHS_{ch} \leq C_S, \quad \forall t \in \mathbf{T} \quad (11)$$

$$PHS_{ch,t} \leq C_{ub} \cdot y_t, \quad \forall t \in \mathbf{T} \quad (12)$$

$$PHS_{di,t} \leq C_S, \quad \forall t \in \mathbf{T} \quad (13)$$

$$PHS_{di,t} \leq C_{ub} \cdot (1 - y_t), \quad \forall t \in \mathbf{T} \quad (14)$$

Subsequently, we assume that the designed PHS facility must have stored power able to satisfy at least a day of operation $-t_1=24 \text{ h}-$ at all times, whereas we further ensure that in the first period of the year it can deliver power for at least two days of operation $-t_2=48 \text{ h}$. Both of the latter limits affect the PHS design.

$$S_t \geq (H_t + D_t) \cdot t_1, \quad \forall t \in \mathbf{T} \quad (15)$$

$$S_{t=1} \geq (H_{t=1} + D_{t=1}) \cdot t_2 \quad (16)$$

Finally, since the solution for minimum *LCOE* might be degenerate, we apply a post-process step to minimize the GW of the ES (CW_{ES})—Eq.(17)—, subject to not increasing the minimum cost identified by solving Eq.(1). Please note that the GW of the renewable technologies $GW_{i \in I}$ should be adjusted to reflect the location-specific capacity factor.

$$CW_{ES} = (GW_{grid} \cdot \sum_{t \in \mathbf{T}} Grid_t + \sum_{i \in I} \sum_{t \in \mathbf{T}} GW_i \cdot P_{i,t}) / \sum_{t \in \mathbf{T}} (H_t + D_t) \quad (17)$$

3. Results and discussion

Focusing on the *LCOE* (Figure 1, left), we observe that the designed ESs show a substantial economic enchantment in Spain (reduction by 46.2–59.8 %), followed by more moderate improvements in Germany (22.7–42.8 %) when compared to the respective national grid. In contrast, the selected locations in Austria are less appealing (reduction by 2.9–17.3 %). The ESs' benefits emerge mainly due to the high *LCOE* of the national grid—significant in Spain for 2015 and less prominent in Germany and Austria ("eurostat 2016"). Furthermore, local synergies of wind and solar generation, and their storage, significantly influence their lower *LCOE* (see Table 1 and Figure 1, e.g., Cartagena and Schwerin), as also discussed in other studies (Demirhan et al., 2021; Fasihi and Breyer, 2020). For Huelva and Haundorg, the hourly wind speed and solar irradiation characteristics fail to complement each other. Thus, their ES operates without harvesting wind power, while the national grid supplies a considerable share to the ES (12.5 and 19.5 %, respectively). Even though both solar and wind power are being harvested, significant synergies are absent in Essen and the Austrian selected locations. For the latter regions, we observe a significantly lower supply of power from the PHS facility and substantial shares from the national grid—32.4–56.4 %—to attain the required high availability (Table 1).

Compared to the national grids, the designed ESs provide significant GW benefits in all locations (Figure 1, middle). Putting these reductions into perspective, the designed ESs in Spain reduce the GW by 79.3-90.6 %, 64.4-86.5 % in Germany, and 37.8-59.6 % in Austria. The latter improvements are attained via high renewables share in the hybrid ES, i.e., 43.6 to 94.7 % in Zwettl and Cartagena, respectively (see Table 1, accounting for both the direct and indirect –via PHS– use of renewable power). Notably, the high shares of grid power at Essen and in the three locations in Austria lead to a more moderate GW improvement, which influences the CCU performance –as discussed next. Finally, we observe high power losses for most regions due to the charging-discharging and curtailment (Table 1). The model's scope could be expanded to investigate alternatives with a lower curtailment, which we will leave as future work.

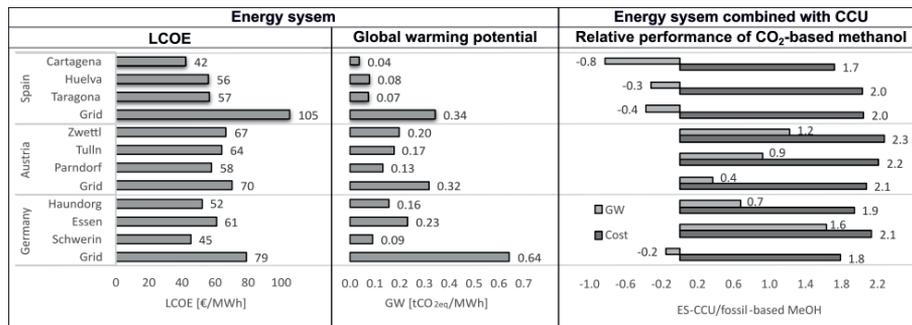


Figure 1. *LCOE* (left) and *GW* (middle) of the designed ESs compared to the national grid –for 2015. Notably, in 2019 the Spanish grid showed a lower *LCOE* compared to 2015 (17.1 % less), whereas the reported *LCOE* for Germany's and Austria's grid increased by 5.7 and 10.12 %, respectively, within the same period ("eurostat 2016"). On the right, we provide the relative cost and *GW* of the CO₂-based MeOH compared to the conventional counterpart.

Table 1. ES delivered power shares as a percentage of the total. The percentage of stored renewable energy is the same as in the ES, e.g., 13.6/86.4 % of solar/wind for Cartagena.

Country	Region	ES delivered power shares in %				ES losses as an additional %	
		Solar	Wind	PHS	Grid	Curtailed	PHS losses
Spain	Cartagena	9.7	61.9	23.1	5.3	17.7	6.1
	Huelva	43.6	0.0	43.9	12.5	8.2	11.5
	Tarragona	40.9	14.6	32.9	11.5	9.4	8.7
Austria	Zwettl	20.1	20.7	2.8	56.4	1.5	0.8
	Tulln	18.6	27.9	4.8	48.7	3.0	1.3
	Parndorf	15.9	38.2	11.3	34.6	4.6	3.0
Germany	Haundorg	42.8	0.0	37.7	19.5	3.3	9.9
	Essen	17.0	40.5	10.1	32.4	4.8	2.7
	Schwerin	20.9	47.3	21.2	10.6	9.6	5.6

Finally, green MeOH production based on the designed ESs (Figure 1, right) is economically unappealing in all selected locations, due to the vast power consumption for H₂ production. Furthermore, we show that in two areas, Essen and Zwettl, the *GW* of the MeOH is even higher –by 1.6- and 1.2-fold, respectively– compared to the fossil-based counterpart, and thus, the investigated CCU application is unfavourable. The latter burdens emerge due to the high shares of grid power in the ES. Compared to the benchmark, a lower *GW* for the green MeOH is observed in the remaining regions, indicating a successful application due to the effective utilization of atmospheric CO₂.

Finally, we observe negative GW values for MeOH (on a cradle-to-gate) in Schwerin, Cartagena, Huelva, and Tarragona, due to the better performance of their ESs.

4. Conclusions

Here we addressed the integrated design of power mix and CCU plants within a single cluster. In seven out of nine locations, we found that carbon footprint reductions (compared to fossil methanol) can be attained via CO₂-based methanol when location-specific hybrid ES powers the production. Our results highlight the benefits of hourly synergetic effects of renewable ES, while providing a roadmap for assessing CCU chemicals more accurately. An optimal and cost-effective integrated design of ES-CCU clusters, predominantly based on renewables, could help overcome the renewable's inherent variability. Notably, our assessment highlighted that even though a renewable-based ES can provide significant benefits compared to the national grid, these might be, in cases, insufficient to make CCU environmentally appealing.

Acknowledgements

This publication was created as part of NCCR Catalysis, a National Centre of Competence in Research funded by the Swiss National Science Foundation.

References

- Demirhan, C.D., Tso, W.W., Powell, J.B., Pistikopoulos, E.N., 2021. A multi-scale energy systems engineering approach towards integrated multi-product network optimization. *Appl. Energy* 281, 116020.
- Deutz, S., Bardow, A., 2021. Life-cycle assessment of an industrial direct air capture process based on temperature–vacuum swing adsorption. *Nat. Energy* 6, 203–213.
- European Commission. Photovoltaic Geographical Information System (PVGIS).
- European Commission, 2019. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions :The European Green Deal. European Commission, Brussels.
- eurostat 2016., URL <https://ec.europa.eu/eurostat/databrowser/view/ten00117/default/table?lang=en>
- Fasihi, M., Breyer, C., 2020. Baseload electricity and hydrogen supply based on hybrid PV-wind power plants. *J. Clean. Prod.* 243, 118466.
- Fasihi, M., Efimova, O., Breyer, C., 2019. Techno-economic assessment of CO₂ direct air capture plants. *J. Clean. Prod.* 224, 957–980.
- González-Garay, A., Frei, M.S., Al-Qahtani, A., Mondelli, C., Guillén-Gosálbez, G., Pérez-Ramírez, J., 2019. Plant-to-planet analysis of CO₂-based methanol processes. *Energy Environ. Sci.* 12, 3425–3436.
- Huijbregts, M.A.J., Steinmann, Z.J.N., Elshout, P.M.F., Stam, G., Verones, F., Vieira, M., Zijp, M., Hollander, A., van Zelm, R., 2017. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *Int. J. Life Cycle Assess.* 22, 138–147.
- Ioannou, I., D'Angelo, S.C., Galán-Martín, Á., Pozo, C., Pérez-Ramírez, J., Guillén-Gosálbez, G., 2021. Process modelling and life cycle assessment coupled with experimental work to shape the future sustainable production of chemicals and fuels. *React. Chem. Eng.* 6, 1179–1194.
- Ioannou, I., D'Angelo, S.C., Martín, A.J., Pérez-Ramírez, J., Guillén-Gosálbez, G., 2020. Hybridization of Fossil-and CO₂-Based Routes for Ethylene Production using Renewable Energy. *ChemSusChem* 13, 6370–6380.
- Parkinson, B., Balcombe, P., Speirs, J.F., Hawkes, A.D., Hellgardt, K., 2019. Levelized cost of CO₂ mitigation from hydrogen production routes. *Energy Environ. Sci.* 12, 19–40.
- United Nations, 2016. Framework Convention on Climate Change (UNFCCC), Paris Agreement.
- Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-ruiz, E., Weidema, B., 2016. Theecoinvent database version 3 (part I): overview and methodology. *Int. J. Life Cycle Assess.* 3.

A comparative study of swarm intelligence and artificial neural networks applications in modeling complex reaction processes

Min Wu^a, Ulderico Di Caprio^a, Furkan Elmaz^b, Bert Metten^c, Dries De Clercq^c, Olivier Van Der Ha^c, Siegfried Mercelis^b, Peter Hellinckx^b, Leen Braeken^a and M. Enis Leblebici^{a,*}

^a*Center for Industrial Process Technology, Department of Chemical Engineering, KU Leuven, Agoralaan Building B, 3590 Diepenbeek, Belgium*

^b*IDLab, Faculty of Applied Engineering, University of Antwerp—imec, Sint-Pietersvliet 7, 2000 Antwerp, Belgium*

^c*Ajinomoto Bio Pharma Services, Coopallaan 91, 9230 Wetteren, Belgium*

**muminenis.leblebici@kuleuven.be*

Abstract

This study investigated two artificial intelligence techniques, Swarm Intelligence(SI) and Artificial Neural Networks(ANN), aiming to overcome the difficulties of simulating complex processes with unknown reactions and intermediates. These techniques are incorporated in reaction modeling via mass balances and reaction kinetic models. The accuracy and the applicability of the resulting models from ANN and SI were compared in the trained semi-batch reactors and the new continuous flow reactors. The ANN-based model is recommended when the extrapolation is unnecessary, and the data is high in volume and variety at the applied space. In this case, no profound reaction knowledge is required. Otherwise, the SI-based model should be employed, which provides detailed information of the target process and is constrained by physical meaning parameters.

Keywords: swarm intelligence, artificial neural networks, hybrid models, semi-batch reactors, continuous flow reactors

1. Introduction

The modeling and simulation of a chemical process are vital for digitization in the modern chemical industry, where the reaction process plays a significant role. The duty of downstream processes can be reduced dramatically if a substantial improvement in reaction controls is achieved. An optimal process design can be achieved by using accurate models, resulting in reducing working time and human errors. However, the modeling of the reaction process is complicated to perform, especially when it includes complex reactions that have not yet been studied earlier. The white-box (first-principle), black-box (data-driven), and grey-box (hybrid) approaches all have got decent performances in chemical reaction process modeling (Rojnuckarin et al., 1993; Xie et al., 2018; Babanezhad et al., 2020). Depending on the knowledge of the target process and the available data, the optimal modeling approach differs. Nowadays, people spend much time searching for the appropriate methods, and the inappropriate use of Artificial Intelligence(AI) techniques is one of the greatest threats because most chemical engineers have limited training in computer science and data analysis (Dobbelaere et al., 2021). This paper gives insights about which method is suitable according to the limited data at

hand by modeling unknown reaction processes via two AI techniques from different data scenarios. The data scenarios were analyzed by 4V characteristics - Volume, Velocity, Variety, and Veracity (Anagnostopoulos et al., 2016). The data volume refers to the involved data quantity, and the low data volume can lead to unreliable derived information. The data velocity shows the speed of getting and feeding new data. The data variety is the diversity of data, and the data veracity defines the data accuracy and precision.

The data coming from two different reactors are employed in this study, namely the Continuous Flow Reactor(CFR) and the Semi-Batch Reactor(SBR). The model aims to predict the time profiles of the process outputs (concentrations, temperatures). Because of a limited amount of CFR data, the kinetic model is trained using only the information coming from SBR. It contains a higher amount of data with a greater variety. This approach is possible since the model focuses only on the common chemical reactions for both CFRs and SBRs. The SBR and CFR models are developed taking mass and heat balances into account. The model of SBR is divided into fed-batch and batch subsections, and the CFR modeling is through a series of continuous stirred tank reactors.

Swarm Intelligence(SI) and Artificial Neural Networks(ANN) are the two AI methods applied in this paper. SI is a class of heuristic optimizers utilized to tune the parameters of a system. They are a sub-field of AI inspired by the social behavior of animals and other living beings. The collaboration of simple individuals performs the optimization task. Each individual has the freedom to move within the parameter search space, and its movement is affected both by its own experience and the experiences of the other group components (Chakraborty and Kar, 2017). The applications of SI algorithms have been popular in recent years. Schlueter et al. (2009) applied the ant colony optimization successfully in the design and control of wastewater treatment plants. ANNs are inspired by the biological neurons and their structures in the human brain. They have been widely applied in chemical engineering. Schweidtmann et al. (2019) proposed a hybrid modeling approach with ANN for the deterministic global process optimization. The available applications of SI and ANN in the open literature have proven their strong abilities in developing or improving the models. This paper compared these two techniques in modeling complex reaction processes from different data scenarios, aiming to offer insights about which method is suitable according to the limited data at hand for chemical engineers.

2. Methodology

The data employed in this paper for the training of the models was generated from experiments performed on SBRs. The available data were sorted into five scenarios representing the problems or situations in reality. The causes of these problems are usually safety requirements, device restrictions, or time limitations when collecting data in the chemical industry, resulting in challenges in data volume, velocity, variety, and veracity (Chiang et al., 2017). The data volume is represented by n , and the data velocity (s) is the number of measured points in one minute in this case. Commonly, the measurements from plants are restricted to a specific range or even a certain point, which can lead to low data variety. The mean chi-squares (χ^2) were calculated to assess the variety quantitatively, and a lower χ^2 indicates a wider variety. For the utilized data, the data veracity of all scenarios is considered equal. Among the available five data scenarios, Scenario 1 was considered the baseline to achieve a vision for the lower or higher volume, velocity, and variety. The two modeling approaches by ANN and SI integrated with Arrhenius equations were performed with different data scenarios as shown in Figure 1.

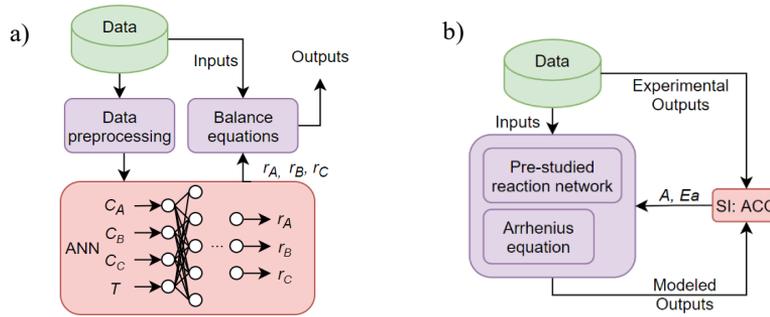


Figure 1: Modeling approaches: a) Artificial Neural Networks(ANN): The preprocessing of the available data is necessary. Each model is trained with an optimal ANN architecture for each data scenario; b) Swarm Intelligence(SI): Based on the possible reaction networks and Arrhenius equations, the related parameters for the model are determined by minimizing the error between modeled and experimental results through SI.

In the ANN approach (Figure 1a), it is necessary to preprocess the data because of the existing high noise and the indirect inputs/outputs. A Savitzky-Golay filter was applied firstly to the generated data from Raman spectra. Its inputs are molar concentrations (C_A, C_B, C_C) and temperatures (T), and the outputs are the reaction rates (r_A, r_B, r_C). The molar concentrations of the components were calculated from the mass weights obtained from Raman spectra. For each data scenario, an optimal ANN architecture was obtained by searching the number of layers and the neuron number of each layer. 70% of the data is utilized to form the training set, while the remaining 30% are used for the validation set. The inputs/outputs were both scaled to the range from 0.00 to 1.00 before feeding into the model. The searching regions of the hidden layer number and the neuron number of each hidden layer are up to 5 and 50, respectively, and the activation function is “tanh”. A $\lambda_1 = 10^{-5}$ L1 regularization term is also utilized to avoid overfitting. The developed reaction model was then combined with the material balance equations to simulate both SBRs and CFRs. There is no need for reaction knowledge in this approach, while the knowledge is compulsory for the following SI approach (Figure 1b).

Table 1: Possible reaction pathways for the SI approach

No.	Reaction	Reaction rates	Remarks
1	$A + B \leftrightarrow C$	$r_1 = A_1 e^{-\frac{Ea_1}{RT}} C_A C_B - A'_1 e^{-\frac{Ea'_1}{RT}} C_C$	r : reaction rate
2	$B + C \leftrightarrow D$	$r_2 = A_2 e^{-\frac{Ea_2}{RT}} C_B C_C - A'_2 e^{-\frac{Ea'_2}{RT}} C_D$	A : frequency factor
3	$A + E \leftrightarrow F$	$r_3 = A_3 e^{-\frac{Ea_3}{RT}} C_A C_E - A'_3 e^{-\frac{Ea'_3}{RT}} C_F$	Ea : activation energy
4	$A + C \leftrightarrow G$	$r_4 = A_4 e^{-\frac{Ea_4}{RT}} C_A C_C - A'_4 e^{-\frac{Ea'_4}{RT}} C_G$	R : gas constant
5	$A + B \leftrightarrow H$	$r_5 = A_5 e^{-\frac{Ea_5}{RT}} C_A C_B - A'_5 e^{-\frac{Ea'_5}{RT}} C_H$	T : temperature
6	$I + B \leftrightarrow J$	$r_6 = A_6 e^{-\frac{Ea_6}{RT}} C_I C_B - A'_6 e^{-\frac{Ea'_6}{RT}} C_J$	C_i : i concentration

Table 1 gives the pre-studied possible reactions from the open literature. The parameters (A_s, Eas) are determined by minimizing the error between modeled and experimental results through SI. R is the gas constant, and T is the temperature. Only three components (A, B, C) were tracked among the ten components listed in the table because of the limited reliable information derived from Raman spectra. The applied objective function is the

normalized mean squared error adding the $\lambda_1 = 10^{-5}$ L1 regularization term. The SI algorithm used in this paper is Ant Colony Optimization(ACO) (Riadi, 2014).

3. Results

The five data scenarios are shown in Figure 2, and the corresponding values indicating data volume, velocity and variety are listed in the following:

- 1) Baseline ($n=1083$, $s=1 \text{ min}^{-1}$, $\text{mean}(\chi^2)=227$);
- 2) Less volume, same velocity, more variety ($n=361$, $s=1 \text{ min}^{-1}$, $\text{mean}(\chi^2)=37$);
- 3) More volume, same velocity, less variety ($n=1805$, $s=1 \text{ min}^{-1}$, $\text{mean}(\chi^2)=288$);
- 4) More volume, more velocity, less variety ($n=2163$, $s=2 \text{ min}^{-1}$, $\text{mean}(\chi^2)=441$);
- 5) Less volume, less velocity, more variety ($n=365$, $s=0.2 \text{ min}^{-1}$, $\text{mean}(\chi^2)=64$).

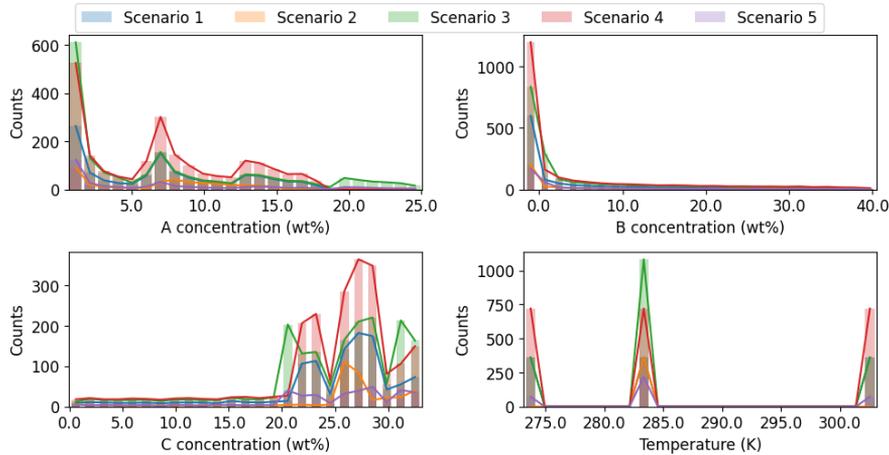


Figure 2: Data distribution of different scenarios. The x-axes are the data ranges, and y-axes are the data amount in each bin. The total number of bins is 25. Scenario 1): Baseline; 2): Less volume, same velocity, more variety; 3): More volume, same velocity, less variety; 4): More volume, more velocity, less variety; 5): Less volume, less velocity, more variety.

Data volume, velocity, and variety are correlated with each other in this case, while their importance for different applications is diverse. The highest amount in this paper is contained in Scenario 4 with 2163 samples, but its data variety is lower, meaning the data majority is in a certain range. It can be seen from Figure 2 that for A concentrations, almost no data in the range from 20.0 to 25.0 in Data scenario 4. Data scenarios 2 and 5 have similar data volumes, while their velocities differ, leading to different data varieties. The imbalanced data is common for all data scenarios, especially for B concentrations and temperatures.

In the ANN approach, a unique architecture of MultiLayer Perceptron(MLP) was employed for each scenario. The optimal architecture was searched using the “keras-tuner Hyperband” algorithm (Li et al., 2018). For all data scenarios, the input neurons are 4 and the output neurons are 3. The resulting optimal structures for scenario 1-5 are: 1) 19-5-36; 2) 8-22-13-5-41; 3) 6-33-12-23; 4) 28; 5) 41-5-35, where the k^{th} position from left is the neuron number of the k^{th} hidden layer. From the optimal trained structures, it can be concluded that an ANN trained with a higher data amount results in a lower amount of hidden layers, because the training of ANN parameters with more data is more efficient. For those with more than one hidden layer, the number of neurons usually increases first and then decreases, increases again, overcoming the difficulties from the imbalance data.

The final models were trained with 1,000 epochs with optimal ANN structures. The training processes are shown in Figure 3a, and all of the final losses were in the magnitude of 10^{-3} .

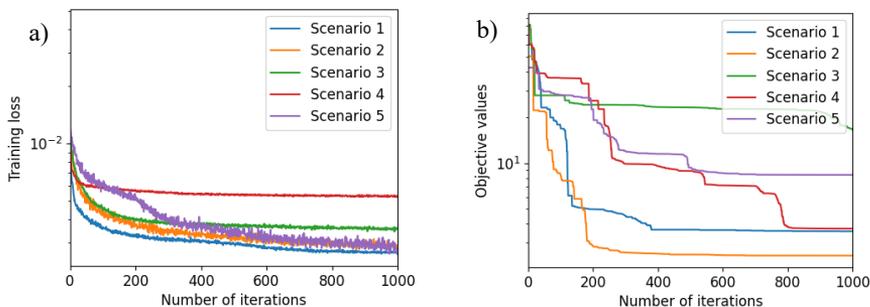


Figure 3: The training process: a) Artificial Neural Networks(ANN): The loss function is the mean squared error with the L1 regularization term, and the used optimizer is “Adam”; b) Swarm Intelligence(SI): The objective values are the sum of the normalized mean square errors and the L1 regularization term.

Scenarios 1, 2, 5 with the greater variety surpass the others. However, the data set with the highest amount (Scenario 4) returns a high loss function value without any significant improvement after the 200th training iteration. This number for the other scenarios is over 600. It can be discussed that a large amount of data but with a low variety does not help the ANN training and can even lead to unfavorable overfitting.

Figure 3b shows the progress of the parameter search executed using the ACO algorithm. The total number of parameters (A_i, Ea_i) is 24 for 6 possible reversible reactions, and the literature study of similar reactions determined their search boundaries. For each scenario, 1,000 iterations were performed and 30 ants were employed to obtain the optimal A s and Ea s. The decreasing of the objective values is stepped because ACO used the same particle until it found another outperforming one, requiring a large number of iterations to mitigate this phenomenon. Within the limited iterations, the training of Scenario 3 ended the earliest with the highest objective value and the least accurate model. Scenario 2 is superior to the others due to the greatest variety. Compared to the ANN approach, fewer efforts were put on the data processing, but more efforts were given to the chemistry reaction study. If the target reactions are already well-studied, this approach is suggested because of less pre-stage work and more delivered information. However, inappropriate assumed reactions can lead to totally different pathways from reality for the unknown complex reactions. The ANN approach can give more insights into unconsidered reaction routes.

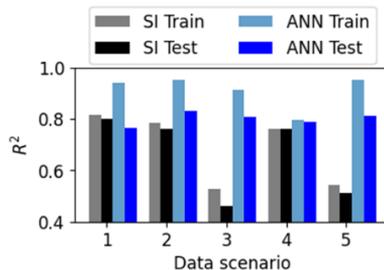


Figure 4: R^2 score of different data scenarios with SI and ANN approaches in semi-batch reactors. The training results were based on mentioned data scenarios. The SI/ANN test results were calculated from a brand-new test set.

The determination coefficients (R^2) of ANN and SI models are shown in Figure 4. The training results were from the data scenarios mentioned above. A brand-new data set of SBR gave the testing results. The ANN approach outperforms the SI approach for most data scenarios. In Scenario 1, the SI obtains a slightly higher R^2 for the testing set, which

shows the strong approximation ability of ANN for the same reactors. In the further stages, the obtained models from different scenarios were applied as the reaction model of a flow reactor. Unexpectedly, ANN models for CFRs gave $R^2 < 0.5$ because their temperatures are out of the ranges of the training data. The SI-based models with $R^2 > 0.8$ showed their advantages for extrapolating since the involved parameters are related to the physical meanings and are temperature independent.

4. Conclusions

This paper applied and compared two artificial intelligence techniques, the SI approach with ACO integrated with Arrhenius equations and the ANN approach with MLP to model complex reaction processes. The ANN model is recommended when the data is high in volume and variety at the applied space, where no profound reaction knowledge is required. The data variety is more important than other data characteristics, especially when the data amount reaches a certain level, such as 300 in this paper. However, the ANN approach is sensitive to the applied region and may be inappropriate for extrapolations. The SI model can provide more detailed information for straightforward extrapolations to other reactors.

Acknowledgments: The authors acknowledge the funding from VLAIO, DAP²CHEM: Real-time data-assisted process development and production in chemical applications (HBC.2020.2455).

References

- I. Anagnostopoulos, S. Zeadally, E. Exposito, 2016. Handling big data: research challenges and future directions. *Journal of Supercomputing* 72 (4), 1494–1516.
- M. Babanezhad, I. Behroyan, A. T. Nakhjiri, A. Marjani, M. Rezakazemi, S. Shirazian, 2020. High-performance hybrid modeling chemical reactors using differential evolution based fuzzy inference system. *Scientific Reports* 10 (1), 1–11.
- A. Chakraborty, A. K. Kar, 2017. Swarm Intelligence: A Review of Algorithms. In: *Nature-Inspired Computing and Optimization*. pp. 475–494.
- L. Chiang, B. Lu, I. Castillo, 2017. Big data analytics in chemical engineering. *Annual Review of Chemical and Biomolecular Engineering* 8, 63–85.
- M. R. Dobbelaere, P. P. Plehiers, R. Van de Vijver, C. V. Stevens, K. M. Van Geem, 2021. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering* 7 (9), 1201–1211.
- L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, A. Talwalkar, 2018. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research* 18, 1–52.
- I. Riadi, 2014. Cognitive Ant colony optimization: A new framework in swarm intelligence. Ph.D. thesis, University of Salford.
- A. Rojnuckarin, C. A. Floudas, H. Rabitz, R. A. Yetter, 1993. Optimal control of a plug flow reactor with a complex reaction mechanism. *Journal of Physical Chemistry* 97 (45), 11689–11695.
- M. Schlueter, J. A. Egea, L. T. Antelo, A. A. Alonso, J. R. Banga, 2009. An Extended Ant Colony Optimization Algorithm for Integrated Process and Control System Design. *Industrial and Engineering Chemistry Research* 48, 6723–6738.
- A. M. Schweidtmann, D. Bongartz, W. R. Huster, A. Mitsos, 2019. Deterministic Global Process Optimization: Flash Calculations via Artificial Neural Networks. *Computer Aided Chemical Engineering* 46 (2000), 937–942.
- Q. Xie, H. Liu, D. Bo, C. He, M. Pan, 2018. Data-driven Modeling and Optimization of Complex Chemical Processes Using a Novel HDMR Methodology. *Computer Aided Chemical Engineering* 44, 835–840.

Combined optimization of start-up shutdown and grade transition of a multistage continuous crystallization process

Jiaxu Liu, Brahim Benyahia*

*Department of Chemical Engineering, Loughborough University, Loughborough, LE11 3TU United Kingdom
B.Benyahia@lboro.ac.uk*

Abstract

Continuous pharmaceutical campaigns may exhibit extremely short operating windows. As a result, the impact of the start-up and shutdown on both costs of production and environmental footprint is extremely important. In this paper, a systematic model-based dynamic optimization strategy was developed to address the combined start-up, shutdown and grade transition of a multistage combined cooling and antisolvent continuous crystallizer. Firstly, the optimal shutdown process was addressed, and several scenarios were considered. Subsequently, a combined start-up and shutdown optimization aimed at the maximization of on-spec production based on the maximization of the steady-state operating window for a fixed total manufacturing schedule. It was shown that 5510.2 g of on-spec products can be obtained during an operating window of 800 minutes, which, compared to the non-optimized scenario, corresponds to a 10% increase in production along with a significant reduction of wastes. In addition, the optimization of grade transition was also considered to improve the dynamic performance and flexibility of the process and helped reduce the transition time from 38.5 minutes to 2.7 minutes.

Keywords: Dynamic Optimization, MSMPR, Start-Up, Shutdown, Grade Transition, combined start-up and shutdown optimization

1. Introduction

Continuous manufacturing is increasingly considered as the most flexible option for pharmaceutical manufacturing (Mascia et al., 2013). Despite its great advantages, such as lower costs and increased productivity, the achievement of the full potential of continuous pharmaceutical manufacturing is still hindered by technical challenges such as the lack of optimal strategies to operate the plant during dynamic transitions, particularly under short operation windows (Benyahia et al., 2018). Hence, the development of systematic and rigorous strategies to optimize start-up, shutdown, and grade transition of single processes and integrated continuous pharmaceutical processes is critical to minimize wastes and further enhance resilience and cost efficiency.

Most upstream pharmaceutical processes require at least one crystallization unit, which is adopted as the key purification technology. Various model-based and model-free techniques have been developed to optimize and control crystallization processes (Nagy et al., 2020; Parekh et al., 2018). Typically, there are three types of widely used continuous crystallization technologies namely mixed suspension mixed product removal (MSMPR) crystallizers, plug flow crystallizers, and continuous oscillatory baffled crystallizers. A systematic optimization of the start-up, to minimize the time required to

reach steady state, has already been investigated and multiple decision variables and optimization strategies have been already implemented, including the optimization of discrete profiles of the jacket temperatures, antisolvent flow rates, seeding policies, initial state of the crystallization vessels, and several discretization methods (Liu & Benyahia, 2022). However, despite the importance of shutdown optimization and its significant impact on the cost and environmental performance of continuous pharmaceutical processes, given the large amounts of wastes generated at the end of unoptimized operation, the problem has been seldom reported in the literature.

In this work, a systematic optimization of the shutdown of a multistage continuous crystallization process was developed to maximize on-spec production by manipulating and adopting different discretization methods for the jacket temperatures and antisolvent flow rates. With the experience gained from previous start-up optimization (Liu & Benyahia, 2022), an effective optimization approach was proposed to optimize a scenario that combines start-up, steady-state, and shutdown. The optimal operating profiles of the control variables were identified in the case of the cooling and antisolvent crystallization of Aspirin in a three-stage MSMPR crystallizer. To enhance process flexibility and resilience, the problem of optimal grade transition of the crystal product is also considered to pave the way for the optimal operation of future continuous pharmaceutical processes focused on multiproduct manufacture. Optimal grade transition which has been widely investigated in the polymer industry to produce different product grades inherent to dynamic market demand (Wang et al., 2000, Benyahia et al., 2011). However, it has never been addressed in the context of continuous pharmaceutical manufacturing. In the current scenario, the product grade is associated with different targeted mean crystal sizes. The optimal grade transition aims at improving the flexibility of the multistage continuous crystallization process by minimizing the transition time based on a dynamic optimization approach which manipulates a set of decision vectors in a similar way as in the start-up and shutdown optimization.

2. Material and Method

2.1. Process model

As a case study, the crystallization of Aspirin (acetylsalicylic acid, ASA) in ethanol (solvent) and water (antisolvent) was used to demonstrate the dynamic optimization strategies. The dynamic mathematical model of the continuous crystallization process, which consists of the population balance model based on the standard method of moment, as well as the energy and mass balances, was developed under a set of assumptions as clearly discussed elsewhere (Liu & Benyahia, 2022). The model was modified to capture the shutdown procedure and allow additional decision options for the optimization problem. The full set of modified equations will not be presented here for the sake of brevity.

2.2. Formulation of the optimization problem

A three-stage MSMPR crystalliser was used, and the operating profiles of the antisolvent flow rate, jacket temperature were used as decision vectors based on several discretization methods. Several shutdown scenarios were developed to maximize on-spec production under both fixed and optimized total shutdown time. The shutdown process starts when the fresh feed to the first MSMPR is stopped. The operation profile of antisolvent and jacket temperature at each stage are used as decision variables for the optimization. When

a vessel undergone a complete shutdown, the antisolvent flow rate is also switched off as well as the flowrate of the coolant to the jacket. The combined process involves the start-up, steady state and shutdown. The start-up part has already discussed in great details in the previous work (Liu & Benyahia, 2022). For sake of brevity, only the mathematical formulation of the optimal shutdown is introduced here which can be described by the generic optimization problem shown below:

$$\begin{aligned}
 & \max_{T_{J,i,j}, F_{AS,i,j}, \Delta t_{i,j}} M_{OS} \\
 \text{s.t.} \quad & \dot{x} = f(x, y, u, p, t) & x_{t=0} = x_0 \\
 & 0 = g(x, y, u, p, t) \\
 & C1: 0 \leq F_{AS,i,j} \leq 20 \\
 & C2: 25 \leq T_{J,i,j} \leq 40 \\
 & C3: 0.5 \leq \Delta t_{i,j} \leq 8 \\
 & C4: \sum t_{1,j} \leq 8 \\
 & C5: \sum t_{2,j} \leq 15 \\
 & C6: \sum t_{3,j} \leq 21 \\
 & C7: \omega_{as,i} \leq 70\% \\
 & C8: S_i(t) \geq 1 \\
 & i = 1, 2, 3, j = 1, 2, \dots, n_d
 \end{aligned}$$

M_{OS} is the overall on-spec production during shutdown, the Decision variables are the vectors of discrete jacket temperatures and antisolvent flow rates, and time intervals ($T_{J,i}, F_{AS,i}, \Delta t_i$). The $\omega_{as,i}$ is the antisolvent mass ratio and n_d is the number of discretization point considered the same for all vessels.

For the shutdown optimization the initial conditions were set at the steady-state values. Based on the optimization problem above, several optimization scenarios were developed. The jacket temperature of the i th stage and j th time intervals were regarded as the decision variables in Scenario 1 (Constraints: C2 and C8), and antisolvent flow rate were regarded as decision variables in Scenario 2 (Constraints: C1, C7 and C8). The combined optimization of jacket temperature and antisolvent flow rate are considered in Scenario 3 (Constraints: C1, C2, C7 and C8). In these three cases, the shutdown time was fixed as 30 minutes. The discretisation method of the time intervals is also important, and as such, Scenario 4 was developed based on the combination of several decision vectors including jacket temperatures, antisolvent flow rates and time intervals (C1-C8). In scenario 4, the shutdown time is not fixed. Constraints C1-C3 represent the upper and lower bounds of the decision variables. C4-C6 are the linear constraints for the shutdown of each stage. C7 is a nonlinear constraint which is used to keep the antisolvent mass ratio

in each stage below 70% which is dictated by the prediction capability of the polynomial model of the solubility. C8 is a critical nonlinear constraint used to help avoid dissolution at all stages.

To address the optimization problem with different product requirements, a on the problem of optimal grade transition was also developed. Here, the optimization is designed to shift from a mean particle size of $417 \mu\text{m}$ to $300 \mu\text{m}$. This is considered to present the case where 2 different pharmaceutical products are required to have different dissolution profiles which may be encountered in the context of multiproduct or on-demand continuous manufacture. The combined manipulation of jacket temperatures, antisolvent flow rates and discretisation method was adopted to address the optimal grade transition problem. The objective is to minimize the transition time which in turn minimizes wastes and enhances operation flexibility. The full description of the optimization problem will not be presented here for the sake of brevity, but overall, it can be formulated in a similar way as the optimization problem described above.

3. Result and Discussion

The optimization problem was solved using a combination of a genetic algorithm and a deterministic method (SQP algorithm form MATLAB). The optimization results for the shutdown optimization are shown in table 1. It can be clearly seen that the total on-spec product mass collected during shutdown increases with the increased number of decision variables (increased degree of freedom). The scenario which combines the jacket temperatures and antisolvent flow rates gave the maximum on-spec product mass within the 30-minute shutdown time. The fourth scenario, with the shutdown time extended to 42 minutes which very close to the considered upper bound of the shutdown time, demonstrated additional gain in production capacity but on the expenses of a large shutdown time.

Following the successful shutdown optimization, the dynamic optimization of the overall campaign was developed to optimize the combined start-up, steady-state and shutdown to maximize on-spec production over a fixed continuous manufacturing window. The jacket temperature, antisolvent flow rate, seeding policies and discretisation method were all used as decision variables. The optimal mean crystal size and on-spec product profiles obtained with the resulting optimized operation profiles (Figure 2) are shown in Figure 1.

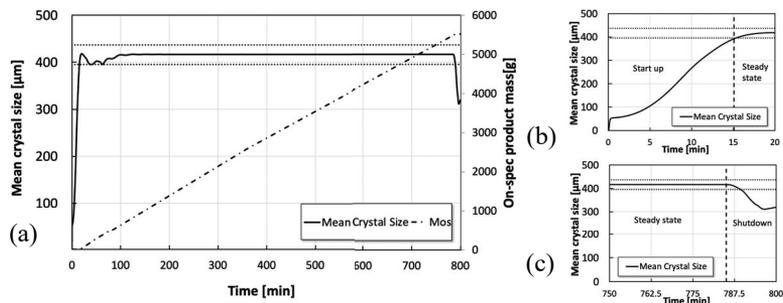


Figure 1. Dynamic profile of the mean crystal size, (a) On-spec production and mean crystal size over the whole process. (b) Mean crystal size during start-up and (c) Mean crystal size during shutdown.

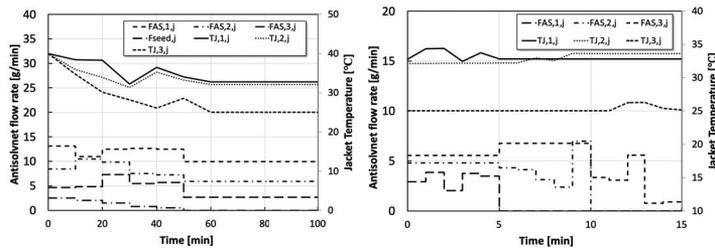


Figure 2. Optimal operating profiles during start-up and shutdown.

Compared to the shutdown process (scenario 4: 42 minutes shutdown and scenario 3: 30 minutes shutdown), in the integrated process optimization, the shutdown is shortened to 15 minutes and less on-spec product mass is collected during this period. When the total operating window is fixed as 800 minutes, shorter shutdown time will prioritise a production rate at the steady-state operating conditions which is higher than the one associated with the shutdown optimization. In other words, longer shutdown time will increase yield of the shutdown process, but shorter shutdown time will increase the production over the fixed total operating window. It is worth mentioning that the maximum on-spec production does not mean to rule out completely the shutdown optimization. It can be observed in figure 1 (c), that the shutdown starts at 785 minutes, and on spec products are still generated until the mean crystal size dropped below the 5% lower bound.

Table 1. Summary of the optimal shutdown results.

Scenario	1	2	3	4
M_{OS} (g)	177.6	205.6	208.0	216.7

Figure 2 shows that that the jacket temperature varies more significantly that in the case of the shutdown which tends to indicate that it is more sensitive to a variation in the antisolvent flowrates. In the shutdown, there are small variations in the jacket temperature which may be associated with the need to maintain the supersaturation at a lower level when a large amount of antisolvent is added. The addition of antisolvent has a conflicting effect as it increases supersaturation but also causes dilution of the system.

In the case of grade transition shown in Figure 3, the operation conditions associated with the second steady state (second product grade) were performed as a single step change for all manipulated variables to generate the base case scenario for the grade transition problem. With the optimal profile, the transition time was shortened from 38.49 minutes to 2.68 minutes, and a total of 290.3 g of the second grade on-spec products were collected.

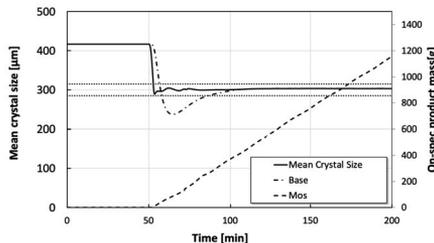


Figure 3. Performance of the optimized grade transition vs base case scenario.

4. Conclusion

The shutdown, combined start-up and shutdown and grade transition scenarios were optimized through a systematic model-based approach. To maximize on-spec production during the shutdown period, several strategies were implemented including discretized jacket temperature, antisolvent flow rate and discretization. Overall, antisolvent showed a more significant impact on the production, and additional on-spec products can be collected by slightly extending the shutdown time.

The dynamic optimization of the whole production process, which combines start-up, shutdown, and steady-state, were developed to maximize on-spec product mass by manipulating the jacket temperatures, antisolvent flowrates, seeding policies along with various discretization methods. With the optimized operation profile, 5510.2 g of on-spec products can be collected over a fixed operating window of 800 minutes. The problem of optimal grade transition was also addressed based on a similar dynamic optimization strategy which demonstrated a significant improvement of process flexibility by minimizing transition time by more than 90% which allows the minimization of wastes and maximization of on-spec production of different product grades.

References

- Benyahia, B. Applications of a plant-wide dynamic model of an integrated continuous pharmaceutical plant: Design of the recycle in the case of multiple impurities. In *Computer Aided Chemical Engineering* (1st ed., Vol. 41). Elsevier B.V, 2018.
- Benyahia, B., Latifi, M. A., Fonteix, C., Pla F. 2011. Multicriteria dynamic optimization of an emulsion copolymerization reactor. *Computers & Chemical Engineering*, 35, 12, 2886-2895.
- Liu, J., & Benyahia, B. 2022. Optimal start-up strategies of a combined cooling and antisolvent multistage continuous crystallization process. *Computers and Chemical Engineering*, 159, 107671.
- Mascia, S., Heider, P. L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P. I., Braatz, R. D., Cooney, C. L., Evans, J. M. B., Jamison, T. F., Jensen, K. F., Myerson, A. S., & Trout, B. L. 2013. End-to-end continuous manufacturing of pharmaceuticals: Integrated synthesis, purification, and final dosage formation. *Angewandte Chemie - International Edition*, 52(47), 12359–12363.
- Nagy, Z. K., Szilagyi, B., Pal, K., & Tabar, I. B. 2020. A novel robust digital design of a network of industrial continuous cooling crystallizers of dextrose monohydrate: From laboratory experiments to industrial application. *Industrial and Engineering Chemistry Research*, 59(51), 22231–22246.
- Parekh, R., Benyahia, B., & Rielly, C. D. 2018. A Global State Feedback Linearization and Decoupling MPC of a MIMO Continuous MSMPR Cooling Crystallization Process. *Computer Aided Chemical Engineering*, 43, 1607–1612.
- Wang, Y., Seki, H., Ohyama, S., Akamatsu, K., Ogawa, M., & Ohshima, M. 2000. Engineering Optimal grade transition control for polymerization reactors. *Computers & Chemical Computers and Chemical Engineering*, 24, 1555-1561.

Systematic dynamic modelling of heat exchanger network

Bertrand Zitte^a, Isabelle Pitault^{a,*}, Boussad Hamroun^a and Françoise Couenne^a

^a*Univ Lyon, Université Claude Bernard Lyon 1, CNRS, LAGEPP UMR 5007, 43 boulevard du 11 novembre 1918, F-69100, VILLEURBANNE, France*
isabelle.pitault@univ-lyon1.fr

Abstract

A systematic way to compute the finite dimensional dynamic model of the counter current heat exchanger network based on the graph theory is proposed. This contribution focuses to the serial interconnection of heat exchanger. The proposed models are build from the elementary block composed of two compartments corresponding to the hot and cold streams. The method is based on the graph based input-output representation of the convective heat flux between the heat exchangers.

Keywords: Heat exchanger network, Dynamic modelling, Graph theory

1. Introduction

Modelling and simulation are essential tools for the design of heat exchangers (HE) and Heat Exchangers Networks (HEN) (Roetzel et al. (2020)). Typical HE and HEN design approaches are primarily based on steady-state models. They are suitable for HENs retrofitting, upgrades and for estimating optimal steady-state operating points. However, dynamic performance of HEN has also to be explored in order to improve flexibility and controllability properties of HEN (Yang et al. (2021)) as well as their energy efficiency.

Roetzel et al. (Roetzel et al. (2020)) cited several mathematical models for transient analysis of heat exchangers: (i) the lumped parameter model, (ii) the distributed parameter model and (iii) the cell model. The first model is a systemic approach wherein the HE is considered as a single box; each fluid in the whole HE has a uniform temperature which is used to calculate the heat transfer through the partition wall and the energy balance involves only the inlet and outlet temperatures. The second model is an infinite dimensional model for which the HE space must be discretized (Michel and Kugi (2013)). The latter, also called lumped-distributed parameter model, consists in dividing the HE space into many elements along its length and applying the lumped method to each element (Correa and Marchetti (1987)). All the models were used for a better understanding of HE behaviour and the model building methods were discussed for small systems with simple topology like HE that are not nearly as large as district heating or heat networks. In addition, energy balances were written based on temperatures as state parameters and do not separate the convective and heat transfer parts, which makes it impractical to use them for scalable and extensible HENs.

For this purpose, we propose a systematic way to compute the finite dimensional dynamic model of a network of counter-current heat exchangers based on graph theory. This method gives the dynamic model in an iterative way and is well suited for modification of the network topology. The obtained model is given in a structure form that can be also used for control purposes.

2. Model formulation

This contribution focuses to the serial interconnection of heat exchangers. The proposed models are built from the elementary block named i composed of two well stirred compartments corresponding to the hot and cold streams as shown in Figure 1. Notations h, \bar{h} stand for specific enthalpies and Q, \bar{Q} for mass flows of the two streams. The indexes in and out stand for inlet and outlet flows resp. In this paper, we considered these two compartments were separated by a wall without heat accumulation for pedagogical purpose. However, this assumption could be overcome using mean heat capacities including both the fluid and wall properties as described in Correa's paper annex (Correa and Marchetti (1987)).

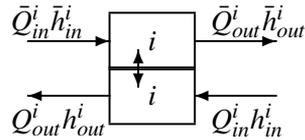


Figure 1: Elementary block i with enthalpy flows

The method is based on the graphic input-output representation of two heat exchangers in series with $[i]$ and $[j]$ elementary blocks, resp. as shown in Figure 2. The arrows are the convective inlet or outlet flows. They are also the edges of the graphs linked to the matter convection. The nodes of the matter and energy graphs are (i) the grey circles which represent the connections of the input and output convective fluxes (grey nodes without matter and energy accumulation) and (ii) each compartment where energy accumulation occurs. There are two graphs, since there are one hot and one cold streams by compartment, but they are treated simultaneously in order to implicitly take into account the heat transfer through the block partition walls (the heat transfer is represented in the block and the edge does not appear in the graph representation). The general construction of the model is based on the adjacency matrix of the convective graphs based on the material and energy balances. With this method, dynamic models of block series can be easily obtained. One of the main advantage of this approach is its modularity as the network can be extended as much boxes (representing a single or network heat exchangers) and streams as needed. So it can be used for network retrofitting or upgrade. Furthermore this modelling method gives structured matrices for which mathematical network analysis can be easily applied and used for control purposes such as dynamic energy assessment approach.

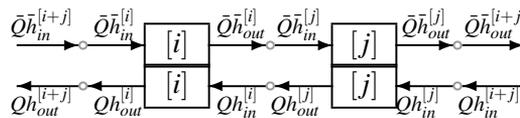


Figure 2: Two heat exchangers or two heat exchanger networks in series

The main assumptions for the modelling of blocks are the following:

- (a) The compartments are perfectly stirred.
- (b) The two streams remain in liquid phase.
- (c) The pressure P, \bar{P} in the compartments as well as the mass densities ($\rho, \bar{\rho}$), the heat capacities (c_p, \bar{c}_p) and the mass flow rates (Q, \bar{Q}) are constant.
- (d) There is no heat accumulation in the block partition wall and no heat exchange with the environment.
- (e) The global heat transfer coefficient λ and the volumes V, \bar{V} are constant and equal for all the compartments.

With the previous assumptions, the mass balances in each block are reduced to the equalities of the inlet and outlet mass flow rates, that for brevity we denote as \bar{Q} et Q :

$$0 = \bar{Q}_{in}^i - \bar{Q}_{out}^i \quad 0 = Q_{in}^i - Q_{out}^i$$

and the enthalpies are defined as $H = \rho V h(T)$, $\bar{H} = \bar{\rho} \bar{V} \bar{h}(\bar{T})$

Proposition 1 *With assumptions (a) to (e), the energy balances in an elementary block 1 write as:*

$$\begin{cases} \frac{d}{dt} \mathbb{H}^1 &= -\lambda (I_1 \otimes A_t) \mathbb{T}^1 - \frac{Q}{\rho V} \mathbb{A}_1 \mathbb{H}^1 - \frac{\bar{Q}}{\bar{\rho} \bar{V}} \bar{\mathbb{A}}_1 \mathbb{H}^1 + \mathbb{B}_1 \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^1 \\ \bar{H}_{in}^1 \end{bmatrix} \\ F_{out}^1 &= \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \mathbb{C}_1 \mathbb{H}^1 \end{cases} \quad (1)$$

where $\mathbb{H}^1 = \begin{bmatrix} H^1 \\ \bar{H}^1 \end{bmatrix}$, $\mathbb{T}^1 = \begin{bmatrix} T^1 \\ \bar{T}^1 \end{bmatrix}$ and $\begin{bmatrix} H_{in}^1 \\ \bar{H}_{in}^1 \end{bmatrix}$ are the enthalpy state, temperature and input enthalpy vectors resp. F_{out}^1 is the output enthalpy flow vector. \otimes stands for the Kronecker product. Matrices are as follow: $A_t = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, $\mathbb{A}_1 = ((I_1 - S_1) \otimes A_c)$, $\bar{\mathbb{A}}_1 = ((I_1 - S_1^T) \otimes \bar{A}_c)$ with $I_1 = 1$, $S_1 = 0$, $A_c = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\bar{A}_c = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbb{V} = \begin{bmatrix} (\rho V)^{-1} & 0 \\ 0 & (\bar{\rho} \bar{V})^{-1} \end{bmatrix}$, $\mathbb{B}_1 = \begin{bmatrix} B & \bar{B} \end{bmatrix}$ with $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\bar{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbb{C}_1 = \begin{bmatrix} C \\ \bar{C} \end{bmatrix}$ with $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $\bar{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}$,

The indexes t and c stand for heat transfer through the partition wall and for convection, resp. The matrices B_c et \bar{B}_c are null in this case but not for several blocks in series. The proof is trivial in this case. It suffices to write the energy balances and identify the matrices.

3. Model of two interconnected blocks

The objective of this section is to show on a simple case how the systematic modelling is constructed. For this purpose, we consider two elementary HEs composed of one block each as shown in Figure 3.

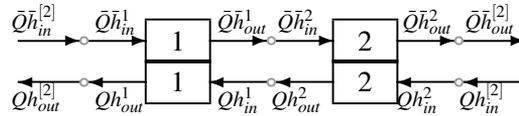


Figure 3: Two elementary HEs composed of one block each in series

Proposition 2 *With assumptions (a) to (e), the system of energy balances for 2 blocks in series writes as:*

$$\begin{cases} \frac{d}{dt} \mathbb{H}^{[2]} &= -\lambda (I_2 \otimes A_t) \mathbb{T}^{[2]} - \frac{Q}{\rho V} \mathbb{A}_2 \mathbb{H}^{[2]} - \frac{\bar{Q}}{\bar{\rho} \bar{V}} \bar{\mathbb{A}}_2 \mathbb{H}^{[2]} + \mathbb{B}_2 \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^{[2]} \\ \bar{H}_{in}^{[2]} \end{bmatrix} \\ F_{out}^{[2]} &= \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \mathbb{C}_2 \mathbb{H}^{[2]} = \begin{bmatrix} Q h_{out}^{[2]} \\ \bar{Q} \bar{h}_{out}^{[2]} \end{bmatrix} \end{cases} \quad (2)$$

with $\mathbb{H}^{[2]T} = [H^1 \quad \bar{H}^1 \quad H^2 \quad \bar{H}^2]$, $\mathbb{T}^{[2]T} = [T^1 \quad \bar{T}^1 \quad T^2 \quad \bar{T}^2]$. The matrices are given by: $\mathbb{A}_2 = ((I_2 - S_2) \otimes A_c)$, $\bar{\mathbb{A}}_2 = ((I_2 - S_2^T) \otimes \bar{A}_c)$ with I_2 the 2×2 identity matrix, $S_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbb{B}_2 = \begin{bmatrix} 0_2 & \bar{B} \\ B & 0_2 \end{bmatrix}$, $\mathbb{C}_2 = \begin{bmatrix} C & 0_2^T \\ 0_2^T & \bar{C} \end{bmatrix}$ with $0_2^T = \begin{bmatrix} 0 & 0 \end{bmatrix}$.

Proof First let us consider two systems of energy balances as written in Proposition 1, the first for the block 1 and the second for the block 2. After vector and matrix concatenations, considering the global state vector $\mathbb{H}^{[2]}$ and the temperature vector $\mathbb{T}^{[2]}$ as written in Proposition 2 and denoting by \mathbb{O} the zeros matrices when the size is not specified, we can immediately write:

$$\begin{cases} \frac{d\mathbb{H}^{[2]}}{dt} = -\lambda \begin{bmatrix} (I_1 \otimes A_r) & \mathbb{O} \\ \mathbb{O} & (I_1 \otimes A_r) \end{bmatrix} \mathbb{T}^{[2]} - \frac{Q}{\rho V} \begin{bmatrix} \bar{A}_1 & \mathbb{O} \\ \mathbb{O} & \bar{A}_1 \end{bmatrix} \mathbb{H}^{[2]} - \frac{\bar{Q}}{\bar{\rho} \bar{V}} \begin{bmatrix} \bar{A}_1 & \mathbb{O} \\ \mathbb{O} & \bar{A}_1 \end{bmatrix} \mathbb{H}^{[2]} \\ + \begin{bmatrix} \mathbb{B}_1 \\ \mathbb{O}_{2 \times 2} \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^1 \\ \bar{H}_{in}^1 \end{bmatrix} + \begin{bmatrix} \mathbb{O}_{2 \times 2} \\ \mathbb{B}_1 \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^2 \\ \bar{H}_{in}^2 \end{bmatrix} \\ F_{out}^1 = \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} [\mathbb{C}_1 \quad \mathbb{O}_{2 \times 2}] \mathbb{H}^{[2]}, \quad F_{out}^2 = \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} [\mathbb{O}_{2 \times 2} \quad \mathbb{C}_1] \mathbb{H}^{[2]} \end{cases} \quad (3)$$

Clearly the first right hand side term of (3) is equal to first right hand side term of (2). In order to write the other terms, let us consider the interconnection matrix A_d for the enthalpy flows. The matrix elements are 1 if a grey node links two enthalpy flows, 0 else. By premultiplying on both sides the flow equality equations by ρV , $\bar{\rho} \bar{V}$ for flow related to Q , \bar{Q} respectively, we obtain:

$$\begin{bmatrix} QH_{in}^1 \\ \bar{Q}\bar{H}_{in}^1 \\ QH_{in}^2 \\ \bar{Q}\bar{H}_{in}^2 \\ QH_{out}^{[2]} \\ \bar{Q}\bar{H}_{out}^{[2]} \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}}^{A_d} \begin{bmatrix} QH_{out}^1 \\ \bar{Q}\bar{H}_{out}^1 \\ QH_{out}^2 \\ \bar{Q}\bar{H}_{out}^2 \\ QH_{in}^{[2]} \\ \bar{Q}\bar{H}_{in}^{[2]} \end{bmatrix} = \begin{bmatrix} A_{dlu} & A_{dru} \\ A_{dl4} & \mathbb{O}_{2 \times 2} \end{bmatrix} \begin{bmatrix} QH_{out}^1 \\ \bar{Q}\bar{H}_{out}^1 \\ QH_{out}^2 \\ \bar{Q}\bar{H}_{out}^2 \\ QH_{in}^{[2]} \\ \bar{Q}\bar{H}_{in}^{[2]} \end{bmatrix} \quad (4)$$

A_d is intrinsic to the series of two HEs independently on their sizes.

- Computation of the inputs of the 2 subsystems that become global input variables. We have:

$$\begin{bmatrix} QH_{in}^{[2]} \\ \bar{Q}\bar{H}_{in}^{[2]} \end{bmatrix} = A_{dru}^T \begin{bmatrix} QH_{out}^1 \\ \bar{Q}\bar{H}_{out}^1 \\ QH_{out}^2 \\ \bar{Q}\bar{H}_{out}^2 \end{bmatrix} = \begin{bmatrix} QH_{in}^2 \\ \bar{Q}\bar{H}_{in}^2 \end{bmatrix} \quad (5)$$

- Computation of the internal inputs: They are given by:

$$QH_{in}^1 = A_{dlu_1} [QH_{out}^1 \quad \bar{Q}\bar{H}_{out}^1 \quad QH_{out}^2 \quad \bar{Q}\bar{H}_{out}^2]^T \quad (6)$$

$$\bar{Q}\bar{H}_{in}^2 = A_{dlu_4} [QH_{out}^1 \quad \bar{Q}\bar{H}_{out}^1 \quad QH_{out}^2 \quad \bar{Q}\bar{H}_{out}^2]^T \quad (7)$$

where A_{dlu_1} and A_{dlu_4} are the first and fourth line of A_{dlu} respectively.

Inserting the expressions of the internal outputs with respect to global state vector, we obtain the following relations for the internal inputs:

$$\begin{bmatrix} H_{out}^1 \\ \bar{H}_{out}^1 \\ H_{out}^2 \\ \bar{H}_{out}^2 \end{bmatrix} = \overbrace{\begin{bmatrix} \mathbb{C}_1 & \mathbb{O}_{2 \times 2} \\ \mathbb{O}_{2 \times 2} & \mathbb{C}_1 \end{bmatrix}}^{\mathcal{C}_2} \begin{bmatrix} \mathbb{H}^1 \\ \mathbb{H}^2 \end{bmatrix} \implies \begin{cases} QH_{in}^1 = QA_{dlu_1} \mathcal{C}_2 \mathbb{H}^{[2]} \\ \bar{Q}\bar{H}_{in}^2 = \bar{Q}A_{dlu_4} \mathcal{C}_2 \mathbb{H}^{[2]} \end{cases} \quad (8)$$

Let us rewrite the following terms in (3): $\begin{bmatrix} \mathbb{B}_1 \\ \mathbb{O}_{2 \times 2} \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^1 \\ \bar{H}_{in}^1 \end{bmatrix} + \begin{bmatrix} \mathbb{O}_{2 \times 2} \\ \mathbb{B}_1 \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^2 \\ \bar{H}_{in}^2 \end{bmatrix}$ in order to recompose with respect to internal and global inputs. We have :

$$\begin{bmatrix} \mathbb{B}_{11} & \mathbb{O}_2 \\ \mathbb{O}_2 & \mathbb{B}_{12} \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^1 \\ \bar{H}_{in}^1 \end{bmatrix} + \begin{bmatrix} \mathbb{O}_2 & \mathbb{B}_{12} \\ \mathbb{B}_{11} & \mathbb{O}_2 \end{bmatrix} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^2 \\ \bar{H}_{in}^2 \end{bmatrix}$$

where $\mathbb{B}_{11}, \mathbb{B}_{12}$ are the first and second column of \mathbb{B}_1 , that are respectively B and \bar{B} .

So for the global inputs, the last right hand side term of (2), we can identify that

$$\mathbb{B}_2 = \begin{bmatrix} \mathbb{O}_2 & \mathbb{B}_{12} \\ \mathbb{B}_{11} & \mathbb{O}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{O}_2 & \bar{B} \\ B & \mathbb{O}_2 \end{bmatrix}.$$

Finally, for stream Q , let us add the convective term in (3) and the term on internal inputs and replace H_{in}^1 by the formula given in (8), we obtain:

$$-\frac{Q}{\rho V} \begin{bmatrix} \mathbb{A}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{A}_1 \end{bmatrix} \mathbb{H}^{[2]} + \frac{Q}{\rho V} \begin{bmatrix} \mathbb{B}_{11} \\ \mathbb{O}_2 \end{bmatrix} H_{in}^1 = -\frac{Q}{\rho V} \begin{bmatrix} \mathbb{A}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{A}_1 \end{bmatrix} \mathbb{H}^{[2]} + \frac{Q}{\rho V} \begin{bmatrix} \mathbb{B}_{11} \\ \mathbb{O}_2 \end{bmatrix} A_{dlu_1} \mathcal{C}_2 \mathbb{H}^{[2]}.$$

It remains to show that $\begin{bmatrix} \mathbb{A}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{A}_1 \end{bmatrix} - \begin{bmatrix} \mathbb{B}_{11} \\ \mathbb{O}_2 \end{bmatrix} A_{dlu_1} \mathcal{C}_2 = \mathbb{A}_2$. It can be checked that $A_{dlu_1} \mathcal{C}_2$ is equal to the third line of \mathcal{C}_2 . So $\begin{bmatrix} \mathbb{B}_{11} \\ \mathbb{O}_2 \end{bmatrix} [0 \ 0 \ 1 \ 0] = \begin{bmatrix} \mathbb{O}_{2 \times 2} & A_c \\ \mathbb{O}_{2 \times 2} & \mathbb{O}_{2 \times 2} \end{bmatrix}$. So $\begin{bmatrix} \mathbb{A}_1 & \mathbb{O} \\ \mathbb{O} & \mathbb{A}_1 \end{bmatrix} - \begin{bmatrix} \mathbb{B}_{11} \\ \mathbb{O}_2 \end{bmatrix} A_{dlu_1} \mathcal{C}_2 = \left((I_2 - S_2) \otimes A_c \right)$.

The same computation can be made for stream \bar{Q} .

- Computation of the global outputs. From the equations (4) and (8), we have:

$$F_{out}^{[2]} = \begin{bmatrix} QH_{out}^{[2]} \\ \bar{Q}\bar{H}_{out}^{[2]} \end{bmatrix} = A_{dld} \begin{bmatrix} QH_{out}^1 \\ \bar{Q}\bar{H}_{out}^1 \\ QH_{out}^2 \\ \bar{Q}\bar{H}_{out}^2 \end{bmatrix} = \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \underbrace{A_{dld} \mathcal{C}_2}_{\mathbb{C}_2} \mathbb{H}^{[2]}. \text{ It is easy to check that } \mathbb{C}_2 = A_{dld} \mathcal{C}_2 = \begin{bmatrix} C & \mathbb{O}_2^T \\ \mathbb{O}_2^T & \bar{C} \end{bmatrix} \text{ This ends the proof. } \blacksquare$$

This method can be iteratively used in order to obtain models for HE with i blocks.

4. Model of Heat Exchanger Network interconnection in series

The objective of this section is to give the model of two series interconnected HENs as shown in Figure 2. The results are given without proof.

As in the previous section, we do not present the graphic representation of matter flows since the graph is trivial and leads to equality of matter flows with two flows only: Q and \bar{Q} . However, as soon as mixer or splitter units will be consider this matter graph representation will be important to analyse.

First the model of a HE composed of i blocks in series is given by:

Proposition 3 *With assumptions (a) to (e), the system of energy balances for a heat exchanger composed of i blocks in series writes as:*

$$\begin{cases} \frac{d}{dt} \mathbb{H}^{[i]} &= -\lambda (I_i \otimes A_T) \mathbb{T}^{[i]} - \frac{Q}{\rho V} \mathbb{A}_i \mathbb{H}^{[i]} - \frac{\bar{Q}}{\rho V} \bar{\mathbb{A}}_i \mathbb{H}^{[i]} + \mathbb{B}_i \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_{in}^{[i]} \\ \bar{H}_{in}^{[i]} \end{bmatrix} \\ F_{out}^{[i]} &= \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \mathbb{C}_i \mathbb{H}^{[i]} = \begin{bmatrix} Qh_{out}^{[i]} \\ \bar{Q}\bar{h}_{out}^{[i]} \end{bmatrix} \end{cases} \quad (9)$$

with $\mathbb{H}^{[i]}, \mathbb{T}^{[i]}$ are the enthalpy and temperature $2i$ vectors of the i -blocks HE resp. The matrices are given by: $\mathbb{A}_i = \left((I_i - S_i) \otimes A_c \right)$, $\bar{\mathbb{A}}_i = \left((I_i - S_i^T) \otimes \bar{A}_c \right)$ with S_i the upper shift $i \times i$ matrix and

I_i the $i \times i$ identity matrix, $\mathbb{B}_i = \begin{bmatrix} 0_{2i-2} & \bar{B} \\ B & 0_{2i-2} \end{bmatrix}$, $\mathbb{C}_i = \begin{bmatrix} C & 0_{2i-2}^T \\ 0_{2i-2}^T & \bar{C} \end{bmatrix}$ with 0_{2i-2} the $2i-2$ zero vector.

Now let us consider the series of the i -blocks HE with a j -blocks HE or HEN. We have for the representation of the system of energy balances:

Proposition 4 *With assumptions (a) to (e), the system of energy balances for the global system composed of two systems of the form (9) in series (the first of size i and the second of size j) writes:*

$$\begin{cases} \frac{d\mathbb{H}^{[i+j]}}{dt} = -\lambda(I_{i+j} \otimes A_i) \mathbb{T}^{[i+j]} - \frac{Q}{\rho V} \mathbb{A}_{i+j} \mathbb{H}^{[i+j]} - \frac{\bar{Q}}{\rho V} \bar{\mathbb{A}}_{i+j} \mathbb{H}^{[i+j]} + \mathbb{B}_{i+j} \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_m^{[i+j]} \\ \bar{H}_m^{[i+j]} \end{bmatrix} \\ F_{out}^{[i+j]} = \mathbb{V} \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \mathbb{C}_{i+j} \mathbb{H}^{[i+j]} \end{cases} \quad (10)$$

with $\mathbb{H}^{[i+j]T} = \begin{bmatrix} \mathbb{H}^{[i]T} & \mathbb{H}^{[j]T} \end{bmatrix}$, $\mathbb{T}^{[i+j]T} = \begin{bmatrix} \mathbb{T}^{[i]T} & \mathbb{T}^{[j]T} \end{bmatrix}$ and the notations are the same as previously using $i+j$ instead of i .

The generalization of this approach when heterogeneous blocks parameters are considered can be easily deduced. Consider the $i \times i$ heat transfer parameter diagonal matrices Λ_i with notations $\lambda_{i,k}$ for the k^{th} element. Since volume of compartment can be different for the blocks, the previous (2×2) matrix \mathbb{V} has to be indexed by the number of the block: \mathbb{V}_n for the n^{th} block. Let us consider the block diagonal matrix \mathbb{W}_i with \mathbb{V}_n on the diagonal for the n^{th} block and the matrix $W_i^T = [\mathbb{V}_1 \quad \mathbb{V}_2 \quad \dots \quad \mathbb{V}_i] = \mathbf{1}_{2i}^T \mathbb{W}_i$ with $\mathbf{1}_{2i}$; the $2i$ vector of 1. The matrix \mathbb{W}_i is a $(2i \times 2i)$ matrix while the matrices W_i and W_i^T have the same dimensions as \mathbb{B}_i and \mathbb{C}_i , resp. Finally let us denote the Hadamard product as \odot .

Corollary 4.1 *With assumptions (a) to (d), the system of energy balances for a heat exchanger composed of i blocks in series writes as:*

$$\begin{cases} \frac{d\mathbb{H}^{[i]}}{dt} = -(\Lambda_i \otimes A_i) \mathbb{T}^{[i]} - Q \mathbb{W}_i \mathbb{A}_i \mathbb{H}^{[i]} - \bar{Q} \mathbb{W}_i \bar{\mathbb{A}}_i \mathbb{H}^{[i]} + \left(\mathbb{B}_i \odot W_i \right) \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} H_m^{[i]} \\ \bar{H}_m^{[i]} \end{bmatrix} \\ F_{out}^{[i]} = \begin{bmatrix} Q & 0 \\ 0 & \bar{Q} \end{bmatrix} \left(W_i^T \odot C_i \right) \mathbb{H}^{[i]} = \begin{bmatrix} Q h_{out}^{[i]} \\ \bar{Q} \bar{h}_{out}^{[i]} \end{bmatrix} \end{cases} \quad (11)$$

5. Conclusion

The proposed systematic way of modelling series coupling of heat exchangers or heat exchanger networks has the advantage to have intrinsic matrices (with only 1 and 0 as elements) such \mathbb{A}_i , \mathbb{B}_i and \mathbb{C}_i . The perspective is to extend this method to parallel interconnections in order to model collectors or distributors. The genericity of the proposed structured dynamical model can be easily implemented as a computer aided process engineering.

References

- D. J. Correa, J. L. Marchetti, 1987. Dynamic simulation of shell-and-tube heat exchangers. *Heat Transfer Engineering* 8 (1), 50–59.
- A. Michel, A. Kugi, 2013. Accurate low-order dynamic model of a compact plate heat exchanger. *International Journal of Heat and Mass Transfer* 61, 323–331.
- W. Roetzel, X. Luo, D. Chen, 2020. *Design and Operation of Heat Exchangers and their Networks*. Academic Press.
- R. Yang, C. T. Tran, A. Zoughaib, 2021. An analytical model to study the dynamic response of heat exchanger network. *International Journal of Heat and Mass Transfer* 176, 121461.

Technical and economic assessment of a castor bean biorefinery to produce renewable aviation fuel: a computer-aided design

Araceli Guadalupe Romero-Izquierdo ^a, Claudia Gutiérrez-Antonio ^a, Fernando Israel Gómez-Castro ^b, Salvador Hernández ^{b,*}

^a *Facultad de Ingeniería, Universidad Autónoma de Querétaro, Campus Amazcala, Carretera a Chichimequillas s/n km. 1, El Marqués, Querétaro, 76225, México*

^b *Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato, Noria Alta S/N, Guanajuato, Gto. 36050, México. hernasa@ugto.mx*

Abstract

2020 has been identified as the worst year in the history of the aviation sector. Now, this sector faces its sustainable economic recovery; one of the alternatives to reach this is the use of renewable aviation fuel. Aviation biofuel is produced from several biomasses through different processing routes, which are technically feasible, but still not economically competitive with its fossil counterpart. In this context, biorefineries are an interesting processing scheme to produce aviation biofuel. Therefore, in this work the computer-aided modelling of a biorefinery for the conversion of castor bean plant (stems, leaves, husk and kernel) to produce biojet fuel is presented. The biorefinery is simulated in Aspen Plus, including mechanical pressing of kernel, pyrolysis of press cake, gasification, pyrolysis, fermentation of stems and leaves, transesterification and hydrotreating of vegetable oil. The fermentation stage produces bioethanol, which is fed to the alcohol-to-jet process; also, the gasification produces biohydrogen, used into the hydroprocessing. In this biorefinery, products as biochar, biogases, bio-oil, biohydrogen, bioethanol, biodiesel, light-gases, naphtha, green diesel and biojet fuel are obtained. The assessment of the biorefinery considers the calculation of economic (total annual cost and net gross profit of products), environmental (CO₂ emissions from electricity and steam), and energetic indicators (total energy invested in the processing, and total energy delivered by products). Based on results, the main factors affecting the TAC are the steam requirements (48.6 %) and the raw material cost (22.5 %); biojet fuel represents 35.4 % of the net gross profit. Also, in this scheme 4.97 ton CO₂ per kg of products are generated, and 5.49 kW are invested per kW of energy delivered by products. This biorefinery can be further improved by process intensification to reduce energy consumption.

Keywords: biorefinery scheme, castor bean plant, biojet fuel, computer-aided design.

1. Introduction

The pandemic caused by the accelerated spread of SARS-CoV2 virus has affected all economic sectors. In particular, the aviation sector has suffered a dramatic decreasing in its operation, mainly due to the social isolation measures and border closure. According to the International Air Transport Association, in 2019 the net profit of the sector was 26.4 billion USD, while in 2020 this indicator drops to -137.7 billion USD (IATA, 2021). In spite of a slight improvement during 2021, this sector is facing its sustainable economic

recovery, for which one of the alternatives is the use of renewable aviation fuel. The aviation biofuel can be produced from different biomasses through several processing routes; these processes are technically feasible, but still not economically competitive with its fossil counterpart (Gutiérrez-Antonio et al, 2021). In this context, the conversion of complete biomasses through a biorefinery scheme is a promising alternative to produce renewable aviation fuel, bioenergy and value-added products. According to the literature, the production of biojet fuel through a biorefinery scheme has been proposed for microalgae, *Jatropha curcas* fruit, sugarcane, soybean, waste cooking oil, rice straw (Romero-Izquierdo et al, 2021). Nevertheless, the conversion of castor bean plant is still missing in the literature; this study is relevant in México, since it is the energetic crop with major productive potential (Gutiérrez-Antonio et al, 2021). Thus, in this work the computer-aided modelling of a biorefinery for the conversion of castor bean plant to produce biojet fuel is presented. The complete castor bean plant is processed, which includes stems, leaves, husk and kernel. The biorefinery is simulated in Aspen Plus V.10, considering mechanical pressing of kernel, pyrolysis of press cake, gasification, pyrolysis and fermentation of stems and leaves, as well as transesterification and hydrotreating of vegetable oil.

2. Modelling and simulation of castor bean plant biorefinery

The biorefinery feedstock is castor bean plant, which consist of stems (38.8 wt %), leaves (14.4 wt %), husk and kernel (seeds, 46.8 wt %). The composition of stems and leaves includes cellulose (47 wt %), hemicellulose (28 wt %), and lignin (25 wt %); the husks consist of cellulose (4 wt %), hemicellulose (5 wt %), lignin (36 wt %), proteins (24 wt %), and other compounds (31 wt %). The feed stream is assumed as 3,531,954.98 kg/h of castor bean plant; this flowrate is obtained from an average yield for castor bean seed (1.41 ton/ha/year) and assuming 10.286 million ha available for its cultivation (Romero-Izquierdo, 2020). To generate biojet fuel and value-added products, the biorefinery consist of 4 internal processing zones: oil extraction (Z1), pyrolysis of stem, leaves, husk and press cake (Z2), lignocellulosic residues treatment (Z3), oil processing and purification of products (Z4). The structure of the biorefinery is shown in Figure 1. Zone Z1 considers the modelling of mechanical pressing of the kernel (Belaid et al., 2011) with module Crusher, and NRTL thermodynamic model; 90% of oil is obtained in this step. Zone Z2 includes the modelling of the pyrolysis of stem and leaves (Kan et al., 2016; Kaur et al., 2018), husk (Mohammed et al., 2014), and press cake (Santos et al., 2015) to generate bio-gases, biochar and bio-oil; these thermochemical processes were modelled with the modules Rbatch and the NRTL model. On the other hand, the lignocellulosic residues are processed in zone Z3 through acid hydrolysis plus co-fermentation (Conde-Mejía et al., 2013) and gasification (Inayat et al., 2010a), using RStoic and Rbatch modules with the NRTL model, respectively; also, in Z3 the alcohol is converted to biojet fuel through the Alcohol-to-Jet (ATJ) process based on the models reported by Ristic and Pacolli (2017), Heveling et al. (1988), Gounder and Iglesia (2011) and Tshabalala and Ojwach (2018), using RStoic modules and NRTL as thermodynamic model. In Zone Z4, the hydrotreating of the oil is described (Liu et al., 2015), using a Rstoic module and the Peng-Robinson equation. Transesterification of the oil is also modelled (Lima Da Silva et al., 2006), using a RStoic module and the UNIFAC model. Also, in Z4 the purification of the hydrocarbons produced in the ATJ and the hydrotreating processes is carried out in a distillation train, first designed with DSTW module assuming key components' recoveries of 99%, and later simulated using the Radfrac module with equilibrium stage model and BK-10 as thermodynamic model. The simulation of the

biorefinery is performed in Aspen Plus; indicators are calculated based on the results of the simulation. The economic and environmental indicators are the total annual cost (TAC) and the counting of released CO₂ emissions. The TAC involves capital cost, calculated by Aspen Economics, adding 18% and 61% due to contingencies, installation fees and equipment maintenance (Turton et al., 2012); and the operating cost, which considers utilities cost (heating and cooling), raw material cost (castor bean, hydrogen, ethanol, methanol) and additional reagents cost (glycerol, urea, H₂SO₄, etc), excluding catalyst cost and filters. The CO₂ emissions due to the production of steam and electricity are considered, and they are calculated through the methodology presented by Gutiérrez-Antonio et al. (2016). The gross profit is calculated with the product volume estimated from simulation results, and the market price of each product. Moreover, two proposed indicators are calculated: IE, energetic indicator defined as invested energy used for heating, regarding to energy delivered by the products; and IA, environmental indicator defined as CO₂ emissions regarding to total product mass obtained from the biorefinery.

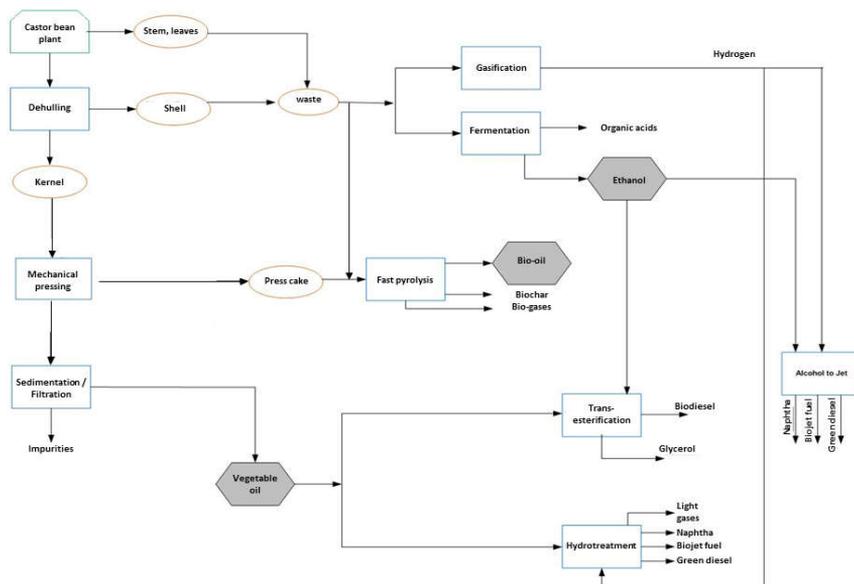


Figure 1. Block diagram of the proposed biorefinery for the conversion of castor bean plant.

3. Analysis of results

Table 1 presents the products obtained on the sections of the biorefinery and their mass flowrate. The kind of product is also indicated. RM means that the obtained product is used as raw material inside the process, while SP implies that the product is intended to be externally sold. It is observed that vegetable oil corresponds to approximately 11% of the initial biomass. From the mass of vegetable oil, approximately 50% is converted into hydrocarbons through hydrotreating, while 31% is transformed into biodiesel. It is important to notice that the hydrocarbons products from hydrotreating consists approximately on 62% of the biojet fuel fraction, which is the main product for this

biorefinery. From the gasification of biomass, an important quantity of hydrogen is obtained, which can be used to partially satisfy the H₂ requirements of the hydrotreating.

Table 1. Products obtained from the biorefining scheme.

Zone	Process	Products	kg/h	Kind of product
1	Mechanical pressing	Vegetable oil without impurities	398,320.81	RM
		Residual cake	678,428.38	RM
2	Fast pyrolysis (residue)	Bio-oil	190,461.98	SP
		Charcoal	15,665.26	SP
		Bio-gases	357,562.55	RM
	Fast pyrolysis (husk)	Bio-oil	56,843.35	SP
		Charcoal	55,991.02	SP
		Bio-gases	4971.70175	RM
	Gasification	Syngas (H ₂ , 98.75 mol%)	58,903.99	RM
	DA-AHCF	Glycerol	6,322.45	Non-recovered
		Succinic acid	7,349.43	
		Acetic acid	15,282.89	
		Furfural	114,555.53	
Ethanol		312,808.46	RM	
ATJ	Naphthas	72,119.12	SP	
	Biojet fuel	45,224.3	SP	
	Green diesel	11,149.31	SP	
3	Pyrolysis (residual cake)	Bio-oil	424,209.79	SP
		Charcoal	26,134.88	SP
		Bio-gases	240,266.56	RM
4	Hydrotreating	Naphthas	46,533.57	SP
		Biojet fuel	124,627.77	SP
		Green diesel	28,648.65	SP
	Transesterification	Biodiesel	125,457.45	SP
		Glycerol	11,839.08	RM

From the ATJ process, it can be observed that approximately 41% of the ethanol is transformed into hydrocarbons, where the biojet fuel fraction corresponds to approximately 35%.

Table 2 shows the results for the economic assessment of the biorefinery. It is observed that operational costs have the high contribution to the total annual cost, higher than 99.5%. Particularly, heating represents 48.7% of the operational costs, followed by others (which includes the cost for reactants) with 31.6%. This is a clear indicator of the need for reducing the energy requirements of the biorefinery, which could be achieved through the application of process intensification and energy integration. As reference, in January

2021 the price of fossil jet fuel was 1.51 USD/gal, while at September 2021 the price increased to 4.29 USD/gal; it is observed that the price of fossil jet fuel is increasing, so taking this aspect under consideration as well as a reduction in the processing costs could help to renewable aviation fuel to be economically competitive with its fossil counterpart.

Table 2. Results for the economic assessment of the biorefinery.

Operational cost (USDx10⁶/y)					
<i>Heating utilities</i>	<i>Cooling utilities</i>	<i>Others</i>	<i>Castor oil plant</i>	<i>Electricity selling</i>	Total operational cost
9,513.766	20.155	6,163.281	4,413.256	605.521	19,504.626
Capital cost (USDx10⁶)					
<i>Equipment cost</i>		<i>A1</i>	<i>A2</i>	Total equipment cost	
180.996		32.568	110.407	323.971	
TAC 19,569.732 USDx10 ⁶ /y					

From the environmental assessment, it has been determined that the total CO₂ emissions are 51,626.31 Mt CO₂/y. Those emissions of carbon dioxide are all associated with the production of steam for heating purposes. There is no production of CO₂ due to the generation of electricity, since the changes of pressure in the process are used to produce electricity, where this generation is even high than the electricity needs of the process. Thus, the excess electricity is sold. Then, reductions on heating requirements through advances technologies would also allow reducing the emissions of carbon dioxide.

4. Conclusions

The design and simulation of a biorefinery for the conversion of castor bean plant into biojet fuel, other biofuels and value-added products has been presented. The processing scheme considers four internal zones, where thermochemical, chemical, as well as biochemical processes are carried out. The proposed biorefinery allows obtaining biojet fuel as main product from two different routes: hydrotreating and ATJ, making use of the whole castor oil plant. It has been determined that the main contribution to the total annual cost and the environmental impact is due to the heating requirements, thus further enhancements are required to reduce those requirements and turn the production scheme economically feasible.

Acknowledgments

The authors gratefully acknowledge financial support from CONACyT for the realization of the doctoral and post-doctoral studies of Araceli Guadalupe Romero-Izquierdo.

References

- IATA, 2021, Industry Statistics – Fact Sheet October 2021, Available from: < <https://www.iata.org/en/iata-repository/pressroom/fact-sheets/industry-statistics/>>, accessed on 16.11.2021
- C. Gutiérrez-Antonio, A.G. Romero-Izquierdo, F.I. Gómez-Castro, S. Hernández, 2021, Production processes of Renewable Aviation Fuel: Present Technologies and Future Trends, Elsevier, ISBN: 978-0-12-819719-6.
- A.G. Romero-Izquierdo, C. Gutiérrez-Antonio, F.I. Gómez-Castro, S. Hernández, 2021, Advances biorefineries for the production of renewable aviation fuel: challenges and opportunities, In Sustainable Alternatives for Aviation Fuels, Elsevier, In press.
- A. G. Romero-Izquierdo, 2020, Diseño, Modelado y Simulación de un Esquema de Biorefinería para el aprovechamiento integral de mezclas de materias primas renovables, PhD Thesis, Universidad de Guanajuato.
- M. Belaid, E. Muzenda, G. Mtilene, M. Mollagee, 2011. Feasibility Study for a Castor oil Extraction Plant in South Africa. *World Acad. Sci. Eng. Technol.* 740–744.
- T. Kan, V. Strezov, T.J. Evans, 2016. Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters. *Renew. Sustain. Energy Rev.* 57, 1126–1140.
- R. Kaur, P. Gera, M.K. Jha, T. Bhaskar, 2018. Pyrolysis kinetics and thermodynamic parameters of castor (*Ricinus communis*) residue using thermogravimetric analysis. *Bioresour. Technol.* 250, 422–428.
- T.H. Mohammed, R. Lakhmiri, A. Azmani, I.I. Hassan, 2014. Bio-oil from Pyrolysis of Castor Shell. *Int. J. Basic Appl. Sci.* 1–5.
- N.A.V. Santos, Z.M. Magriotis, A.A. Saczk, G.T.A. Fássio, S.S. Vieira, 2015. Kinetic study of pyrolysis of castor beans (*Ricinus communis* L.) presscake: An alternative use for solid waste arising from the biodiesel production. *Energy and Fuels* 29, 2351–2357.
- C. Conde-Mejía, A. Jiménez-Gutiérrez, M.M. El-Halwagi, 2013. Assessment of combinations between pretreatment and conversion configurations for bioethanol production. *ACS Sustain. Chem. Eng.* 1, 956–965.
- A. Inayat, M.M. Ahmad, S. Yusup, M.I.A. Mutalib, 2010. Biomass steam gasification with in-situ CO₂ capture for enriched hydrogen gas production: A reaction kinetics modelling approach. *Energies* 3, 1472–1484.
- M. Ristic, M. Pacolli, 2017. Oligomerization of Ethylene and Ethanol into Fuel Through Heterogeneous Catalysis Marko Ristic & Mergim Pacolli.
- J. Heveling, C.P. Nicolaiades, M.S. Scurrell, 1998. Catalysts and conditions for the highly efficient, selective and stable heterogeneous oligomerisation of ethylene. *Appl. Catal. A Gen.* 173, 1–9.
- R. Gounder, E. Iglesia, 2011. Catalytic hydrogenation of alkenes on acidic zeolites: Mechanistic connections to monomolecular alkane dehydrogenation reactions. *J. Catal.* 277, 36–45.
- T.A. Tshabalala, S.O. Ojwach, 2018. Kinetics and chemoselectivity studies of hydrogenation reactions of alkenes and alkynes catalyzed by (benzoimidazol-2-ylmethyl)amine palladium(II) complexes. *Inorganica Chim. Acta* 483, 148–155.
- S. Liu, Q. Zhu, Q. Guan, L. He, W. Li, 2015b. Bio-aviation fuel production from hydroprocessing castor oil promoted by the nickel-based bifunctional catalysts. *Bioresour. Technol.* 183, 93–100.
- N. Lima Da Silva, M. Regina, W. Maciel, C. Batistella, R.M. Filho, 2006. Optimization of Biodiesel Production N IVEA DE L IMA DA S ILVA ,* M ARIA R EGINA W OLF M ACIEL ,.
- R. Turton, R.C. Bailie, W.B. Whiting, J.A. Shaeiwitz, D. Bhattacharyya, 2012. Analysis, Synthesis, and Design of Chemical Processes, Fourth Edition.
- C. Gutiérrez-Antonio, A.G. Romero-Izquierdo, F.I. Gómez-Castro, S. Hernández, A. Briones-Ramírez, 2016. Simultaneous energy integration and intensification of the hydrotreating process to produce biojet fuel from *Jatropha curcas*. *Chem. Eng. Process. Process Intensif.* 110, 134–145.

SiCN fibers as advanced materials for electromagnetic shielding in X-band: experiments and computational modelling and simulation

Heloisa Ramlow,^a Liangrid Lutiani da Silva,^b Braulio Haruo Kondo Lopes,^b Maurício Ribeiro Baldan,^b Ricardo Antonio Francisco Machado^a

^a*Graduate Program in Chemical Engineering, Federal University of Santa Catarina, Universitário Reitor João David Ferreira Lima Campus, Florianópolis, 88040-900, Brazil*

^b*Coordination of Applied Research and Technological Development, National Institute for Space Research, Avenida dos Astronautas 17518, São José dos Campos, 12227-010, Brazil*

Abstract

Materials prepared via the polymer-derived ceramic route have been increasingly studied for protection against electromagnetic energy to mitigate electromagnetic interference. Both experimental and computational evaluations of electrospun SiCN fibers applied to electromagnetic shielding are not yet reported in the literature. This work aims to evaluate the electromagnetic properties of SiCN electrospun fibers by experiments and computational modelling and simulation. Polysilazane and polyacrylonitrile were used respectively as the ceramic precursor and spinning aid. After electrospinning, the fibers were pyrolyzed at 1000 °C. The samples were named SiCN_0, SiCN_40, and SiCN_70 respectively for 0, 40, and 70 wt.% polyacrylonitrile. The scattering parameters, impedances, and reflection losses were collected under X-band (8.2-12.4 GHz) in a vector network analyzer employing the waveguide propagation setup. The experimental scattering parameters were converted through the Nicolson-Ross-Weir method together with the shielding effectiveness and numerical electromagnetic computational studies. Simulations of scattering parameters were performed, and introductory electromagnetic scattering calculations in free space were computed including the radar cross-section (RCS) study. The relative complex electrical permittivity was approximately 3, 4.5, and 4 (real part) and 0.05, 0.22, and 0.1 (imaginary part) respectively for SiCN_0, SiCN_40, and SiCN_70. The SiCN_40 could experimentally store and lose more electromagnetic energy in the material, exhibiting a minimum reflection coefficient of -1.4 dB at 12.4 GHz. The computational simulation corroborated the better performance of SiCN_40 in reflection loss as well as in other electromagnetic spectral responses. Additionally, correlations between electromagnetic properties extracted from experiments and computational results from the RCS study were observed. The free space electromagnetic scattering of SiCN_40 showed better features when compared to the other samples. Owing to the microstructure and product design, SiCN_40 fibers demonstrated satisfying electromagnetic shielding properties in X-band. The computational experiments showed to be a new modelling and simulation approach to evaluate the electromagnetic properties of electrospun SiCN fibers. Further research will focus on material optimization and computational evaluation.

Keywords: complex electrical permittivity, polymer-derived ceramic, radar cross-section, reflection loss, shielding effectiveness.

1. Introduction

Materials prepared via the polymer-derived ceramic route have been increasingly studied for protection against electromagnetic energy to mitigate electromagnetic interference (Chen et al., 2021). The lightweight ceramic provided by these advanced materials is an important advantage over metallic alloys classically used in the military and aircraft industries (Hou et al., 2021). The protection against electromagnetic waves in X-band (8.2-12.4 GHz) is especially important since this frequency range is used for radar, satellite communication, and wireless computer networks. Recent research on electrospun preceramic fibers applied to electromagnetic shielding in X-band has been boosted with the manufacturing of SiC, SiCN, and SiBCN systems. The first work was published in 2017 reporting the electrospinning of polycarbosilane and polycaprolactone followed by crosslinking and pyrolysis at 1300 °C to produce SiC fibers with a reflection loss of -25 dB at approx. 10 GHz (thickness of 3 mm) (Hou et al., 2017).

Various methods based on modelling and simulation approaches have been addressed to investigate the shielding effectiveness of various materials (Liang et al., 2021; Ud-Din Khan et al., 2020). Despite some efforts, the manufacturing of polymer-derived ceramic fibers for high-performance in X-band electromagnetic shielding remains a daunting technical challenge. This work aims to evaluate the electromagnetic properties of SiCN electrospun fibers by experiments and computational modelling and simulation. This work contributes to studying electrospun SiCN fibers for electromagnetic shielding applications by experimental-computational technique. The results obtained here open new grounds for developing these advanced materials by computational techniques, which will be useful to assess and develop these ceramics to final applications, thus, reducing the cost of the experimental procedure. A comprehensive electromagnetic simulation is urgently needed to improve the safety and reliability of different applications in complex electromagnetic environments.

2. Material and methods

Polysilazane synthesized from crosslinked Durazane 1800 (Merck KGaA, Germany) and polyacrylonitrile 200 kDa (Polysciences, Inc., USA) were used respectively as ceramic precursor and spinning aid. The solutions were prepared using dimethylformamide 99.5% (Fisher Chemical, USA) as solvent and dicumyl peroxide 98% (Sigma-Aldrich Corporation, Germany) as crosslinker. After electrospinning, the fibers were pyrolyzed at 1000 °C in nitrogen. The samples were named SiCN_0, SiCN_40, and SiCN_70 respectively for 0, 40, and 70 wt.% of polyacrylonitrile. Five samples were prepared in silicon matrices: 5, 10, or 20 wt.% (thickness of 2.35 ± 0.32 mm). The scattering parameters, impedances, and reflection losses were collected under X-band in a vector network analyzer (VNA) N5230C-PNA-L (Agilent Technologies, USA) employing the waveguide propagation setup. The experimental parameters were converted through the Nicolson-Ross-Weir (NRW) method. The complex permittivity (ϵ) in a dispersive (frequency-dependent), homogeneous, and isotropic material medium can be expressed by **Equation (1)** (Naito and Suetake, 1971),

$$\epsilon(\omega) = \epsilon_0 \epsilon_r(\omega) = \epsilon_0 (\epsilon_r'(\omega) - j \epsilon_r''(\omega)) = \epsilon_0 \epsilon_r'(\omega) (1 - j \tan[\delta_e(\omega)]), \quad (1)$$

where ϵ_0 is the electric permittivity in vacuum, ϵ_r is the relative complex electrical permittivity, ϵ_r' and ϵ_r'' are the real and imaginary parts of the relative complex electrical permittivity, respectively. The electrical tangent loss term, $\tan[\delta_e(\omega)] = \epsilon_r''(\omega)/\epsilon_r'(\omega)$, directly relates the imaginary and real components of the permittivity and measure the inherent dissipation energy due to electrical losses. The term $\omega = 2\pi f$ is the angular

frequency ($\text{rad}\cdot\text{s}^{-1}$) with f being the frequency in Hz. The reflection loss (RL) assuming the slab considerations ended by a conduction plane in decibel units (dB) is given by **Equation (2)**,

$$RL_{dB}(\omega) = 20 \log_{10}(S_{11}(\omega)) = 20 \log_{10} \left(\frac{z_{in}(\omega) - z_0}{z_{in}(\omega) + z_0} \right), \quad (2)$$

where S_{11} is the scattering parameter obtained from VNA measurements (observed from the port 1 of the waveguide), z_{in} is the input impedance, z_0 is the reference to the impedance load. The equation for shielding effectiveness (SE) related directly to the scattering parameters in dispersive and reciprocal media is given by **Equation (3)** (Al-Saleh et al., 2013),

$$SE_{TdB}(\omega) = 10 \log_{10} \left(\frac{1}{1 - |S_{11}(\omega)|^2} \right) + 10 \log_{10} \left(\frac{1 - |S_{11}(\omega)|^2}{|S_{12}(\omega)|^2} \right) = SE_{RdB}(\omega) + SE_{AdB}(\omega), \quad (3)$$

where $SE_T(\omega)$, $SE_R(\omega)$, and $SE_A(\omega)$ represent respectively total, reflected, and absorbed shielding effectiveness and the scattering term $S_{12} = S_{21}$ in reciprocal media. The radar cross-section (RCS) is a far-field parameter employed to characterize the scattering properties of a radar target (Balanis, 2012). For three-dimensional objects, the RCS can be calculated considering the spherical coordinate system by **Equation (4)** (Balanis, 2012),

$$\sigma_{RCS}(r, \theta, \phi) = \lim_{r \rightarrow \infty} \left[4\pi r^2 \frac{\bar{S}_e(r, \theta, \phi)}{\bar{S}_i(r, \theta, \phi)} \right] = \lim_{r \rightarrow \infty} \left[4\pi r^2 \frac{|\bar{E}_e(r, \theta, \phi)|^2}{|\bar{E}_i(r, \theta, \phi)|^2} \right], \quad (4)$$

where r is the distance between the target object and the observer, $\bar{S}_e(r, \theta, \phi)$ is the scattered power spectral density, $\bar{S}_i(r, \theta, \phi)$ is the incident power spectral density (or the incident Poynting vector), $\bar{E}_e(r, \theta, \phi)$ is the scattered electric field intensity vector, and $\bar{E}_i(r, \theta, \phi)$ is the incident electric field intensity vector. The unit of $\sigma_{RCS}(r, \theta, \phi)$ is m^2 , although the RCS is also commonly presented in dBsm (decibel squared milliwatt), as expressed by **Equation (5)** (Knott et al., 2004),

$$\sigma_{RCS_{dBsm}}(r, \theta, \phi) = 10 \log_{10}[\sigma_{RCS}(r, \theta, \phi)]. \quad (5)$$

Based on the experimental electromagnetic properties, numerical electromagnetic computational studies were implemented. Simulations of reflection losses were performed in a guided electromagnetic wave study, and introductory electromagnetic scattering calculations in free space including the RCS study were computed. The computational experiments were simulated with FEKO® software (Altair, USA), which allows solving Maxwell's equations in three dimensions for solving full electromagnetic waves. The method of moments (MoM) was used by both the waveguide electromagnetic propagation with one and two ports and the free space electromagnetic propagation simulations.

3. Results and discussion

According **Equation (1)**, the relative complex electrical permittivity was approximately 3, 4.5, and 4 (real part) and 0.05, 0.22, and 0.1 (imaginary part) respectively for SiCN_0, SiCN_40, and SiCN_70 (**Figure 1**). The studies showed that 5 wt.% fibers concentration in silicone was effective on electromagnetic shielding. The SiCN_40 could experimentally store and lose more electromagnetic energy in the material, exhibiting a minimum reflection coefficient of -1.4 dB at 12.4 GHz (**Figure 2-a**). Likewise, the

computational simulation corroborated the better performance of SiCN_40 in reflection loss (**Figure 2-b**), considering the incident electric intensity field vector with unitary amplitude value, $z_0 = 377 \Omega$ (approximately the free space impedance), 2 mm of sample thickness, and WR-90 rectangular waveguide section dimensions ($a = 22.86$ mm and $b = 10.16$ mm) for all cases according to **Equation (2)**.

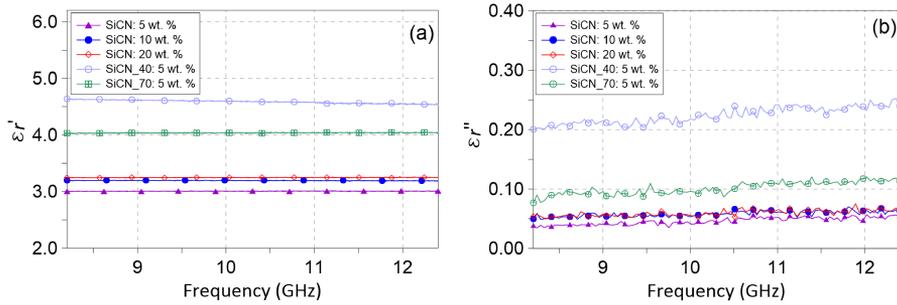


Figure 1. a) Real and b) Imaginary components of the complex relative electrical permittivity obtained experimentally by the NRW method in X-band.

There is a slight difference between experimental and computational results, which can be attributed to the alignment and agglomeration of fibers, matrix curing, and variation of sample thickness. The ability to adjust shielding properties through the addition of the carbon precursor (spinning aid), the nanostructure of SiCN, and fiber morphology opens new strategies for the development of electromagnetic shielding materials.

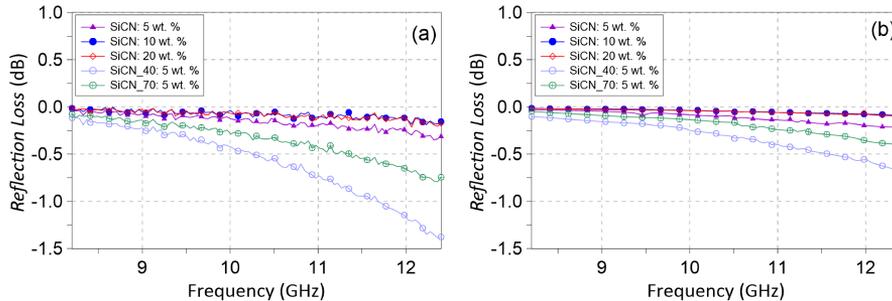


Figure 2. a) Experimental and b) Computational measurement of reflection loss in X-band.

As reported by **Equation (3)**, the total shielding effectiveness calculation in experimental electromagnetic guided conditions suggests the dominance of the waveguide experimental S_{11} scattering term contribution, i.e., the scattering reflection mechanism contribution is predominant (**Figure 3-a**). The correlations on the amplitudes are also pointed out (**Figure 3-b**).

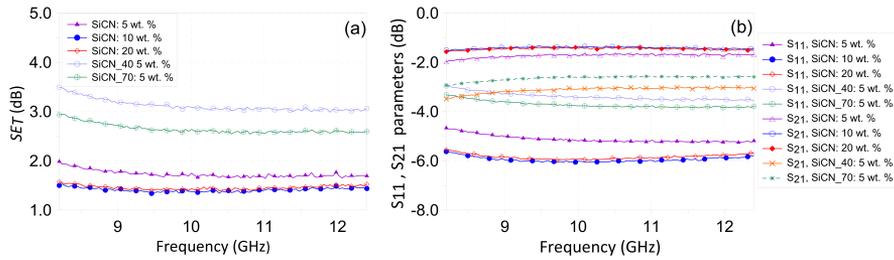


Figure 3. a) Total shielding effectiveness calculation and b) Experimental waveguide S_{11} and S_{21} scattering parameters in X-band.

The free space scattering radiation performance of an object describes how energy is scattered when an object is exposed to under a plane wave electromagnetic incident field (see the free space computational setup in **Figure 4-a**). The simulated target was defined with a thickness of 2 mm and 300 mm of side length. The blue arrow represents the direction of the electromagnetic wave propagation (i.e. the incident Poynting vector, \vec{S}_i) as normal orientation ($-z$ axis direction) assuming the x - y target plane, while the green arrow represents its polarization vector direction (i.e. the incident electric intensity field vector \vec{E}_i) defined in $-x$ axis direction, whereas its amplitude value was set as unitary (**Figure 4-a**). For clarification, the three-dimensional (3D half sphere) RCS simulated results (**Equations (4)** and **(5)**) using experimental electromagnetics properties from SiCN: 5 wt.% were calculated assuming the frequencies of $f = 8.2$ GHz (**Figure 4-b**), $f = 10.3$ GHz (**Figure 4-c**), and $f = 12.4$ GHz (**Figure 4-d**). The main RCS lobe amplitude is enhanced when the electromagnetic wave frequency increases as well the number of lateral RCS side lobes also increase due to the increments in the frequency of the wave excitation. The free space electromagnetic scattering of SiCN_40 defined with arbitrary frequency ($f = 10.3$ GHz) showed better features when compared to the other samples (**Figure 4-e**).

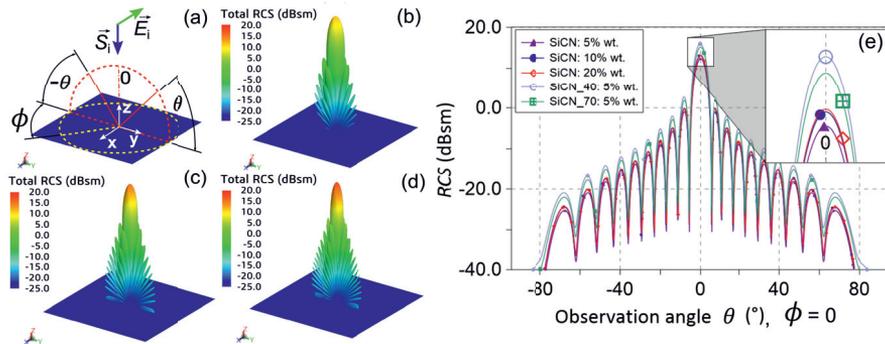


Figure 4. Computational results of RCS scattering study in free space. a) Computational setup illustration for the far-field spectral diagram of radiation scattering from SiCN: 5 wt.% at the frequency of b) 8.2 GHz, c) 10.3 GHz, and d) 12.4 GHz. e) Radiation diagram with theta, (θ), observation angle from -90° to 90° , for all sample materials at the frequency of 10.3 GHz which was chosen arbitrarily.

Direct correlations between the amplitudes from the electromagnetic properties extracted from experiments (**Figure 1**) and computational results from the amplitudes of the RCS

study were observed (**Figure 4-e**). This fact suggests the possibility of exciting opportunities owing to the advanced studies evaluations which can be performed over real and practical applications of shielding materials supported by simulation methods. Computer simulation is important for real applications that have specific targets since the RCS is a function of the effective permittivity of the material, its roughness, geometry, and angle. The simulation enables in analyzing the performance of a wide range of applications to answer RCS questions for military and aircraft industries. The simulation study showed the wave scattering in free space over the SiCN₄₀ material in electromagnetic shielding, showing that final applications can be simulated using the experimental parameters (**Figure 4**).

4. Conclusion

The experimental and computational results suggest that polymer-derived ceramics act as electromagnetic shielding materials, greatly attenuating the incoming radiation for certain frequency values. Owing to the microstructure and product design, SiCN₄₀ fibers demonstrated satisfying electromagnetic shielding properties in X-band. The computational experiments showed to be a new modelling and simulation approach to evaluate the electromagnetic shielding of electrospun SiCN fibers. The practical implications of this work are the scattering simulations with the use of the SiCN fibers for final applications including military and aircraft industries. Using electromagnetic simulations, various types of electromagnetic problems involving the use of polymer-derived ceramic fibers can be quickly and accurately solved. Further research will focus on material optimization with computational evaluation.

References

- Al-Saleh, M.H., Saadeh, W.H., Sundararaj, U., 2013. EMI shielding effectiveness of carbon based nanostructured polymeric materials: A comparative study. *Carbon N. Y.* 60, 146–156. <https://doi.org/10.1016/j.carbon.2013.04.008>
- Balanis, C.A., 2012. *Advanced Engineering Electromagnetics*, 2nd ed. John Wiley & Sons, Danvers.
- Chen, Q., Li, D., Liao, X., Yang, Z., Jia, D., Zhou, Y., Riedel, R., 2021. Polymer-Derived Lightweight SiBCN Ceramic Nanofibers with High Microwave Absorption Performance. *ACS Appl. Mater. Interfaces* 13, 34889–34898. <https://doi.org/10.1021/acsami.1c07912>
- Hou, Y., Cheng, L., Zhang, Y., Du, X., Zhao, Y., Yang, Z., 2021. High temperature electromagnetic interference shielding of lightweight and flexible ZrC/SiC nanofiber mats. *Chem. Eng. J.* 404, 126521. <https://doi.org/10.1016/j.cej.2020.126521>
- Hou, Y., Cheng, L., Zhang, Y., Yang, Y., Deng, C., Yang, Z., Chen, Q., Wang, P., Zheng, L., 2017. Electrospinning of Fe/SiC hybrid fibers for highly efficient microwave absorption. *ACS Appl. Mater. Interfaces* 9, 7265–7271. <https://doi.org/10.1021/acsami.6b15721>
- Knott, E.F., Schaeffer, J.F., Tulley, M.T., 2004. *Radar Cross Section*, 2nd ed. Scitech Publishing, Inc., Raleigh.
- Liang, J., Bai, M., Gu, Y., Wang, S., Li, M., Zhang, Z., 2021. Enhanced electromagnetic shielding property and anisotropic shielding behavior of corrugated carbon fiber felt composite and its sandwich structure. *Compos. Part A Appl. Sci. Manuf.* 149, 106481. <https://doi.org/10.1016/j.compositesa.2021.106481>
- Naito, Y., Suetake, K., 1971. Application of ferrite to electromagnetic wave absorber and its characteristics. *IEEE Trans. Microw. Theory Tech.* 19, 65–72. <https://doi.org/10.1109/TMTT.1971.1127446>
- Ud-Din Khan, Salah, Ud-Din Khan, Shahab, Almutairi, Z., Haider, S., Ali, S.M., 2020. Development of theoretical-computational model for radiation shielding. *J. Radiat. Res. Appl. Sci.* 13, 606–615. <https://doi.org/10.1080/16878507.2020.1812798>

Continuous-Time Surrogate Models for Data-Driven Dynamic Optimization

Burcu Beykal,^{a,b} Nikolaos A. Diangelakis,^{c,d} Efstratios N. Pistikopoulos^{c,d}

^a*Department of Chemical and Biomolecular Engineering, University of Connecticut, Storrs, CT 06269, USA*

^b*Center for Clean Energy Engineering, University of Connecticut, Storrs, CT 06269, USA*

^c*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA*

^d*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA*
stratos@tamu.edu

Abstract

This work addresses the control optimization of time-varying systems without the full discretization of the underlying high-fidelity models and derives optimal control trajectories using surrogate modeling and data-driven optimization. Time-varying systems are ubiquitous in the chemical process industry and their systematic control is essential for ensuring each system to be operated at the desired settings. To this end, we postulate nonlinear continuous-time control action trajectories using time-varying surrogate models and derive the parameters of these functional forms using data-driven optimization. Data-driven optimization allows us to collect data from the high-fidelity model without pursuing any discretization and fine-tune candidate control trajectories based on the retrieved input-output information from the nonlinear system. We test exponential and polynomial surrogate forms for the control trajectories and explore various data-driven optimization strategies (local vs. global and sample-based vs. model-based) to test the consistency of each approach for controlling dynamic systems. The applicability of our approach is demonstrated on a motivating example and a CSTR control case study with favorable results.

Keywords: Data-driven optimization, dynamic optimization, time-varying systems, optimal control, surrogate modeling.

1. Introduction

Time-varying processes are omnipresent in chemical engineering and postulating the correct optimal control laws are critical for a feasible and safe operation of any process. Commonly, such processes are studied and modeled with a set of differential or differential-algebraic equations that captures the overall dynamic balances and other algebraic relationships (e.g., rate laws) (Diangelakis et al., 2017).

A vast number of dynamic optimization techniques rely on approximations that utilize uniform and nonuniform discretization, leading to the solution of large-scale linear, nonlinear, or linearized problems (Biegler, 1984). Yet, such problems are solely addressed using calculus of variations due to the inherent complexity of the approach and its applicability is commonly limited to linear ODE systems. For highly nonlinear problems, linear control schemes are often insufficient to portray an appropriate nonlinear control strategy for a given nonlinear process. In such cases, the full discretization of the

time-varying problem is preferred to take a manipulated action at each discrete time point, which is computationally prohibitive.

Recent efforts focused on using machine learning and data-driven techniques to overcome computational challenges associated with optimizing dynamic systems. Sorek et al. (2017) used interpolation and polynomial approximations to reduce the computational burden of discretized controls for maximizing the profit in oil production optimization problems. Beykal et al. (2020) used Support Vector Machines (SVMs) to model implicit constraints and numerical infeasibilities of dynamic systems as a classification problem. The SVM model is later used within a deterministic optimization framework to find the best initial conditions of a steam cracking reactor that maximizes the profit of operation.

Although these studies have shown that data-driven modeling and optimization provides a feasible path for solving such a difficult class of optimization problems, their applications were limited to deterministic formulations and failed to address the direct derivation of optimal control trajectories within a dynamic optimization framework. In this work, our goal is to bridge this gap in dynamic optimization by avoiding the use of full discretization-based control schemes in linear and nonlinear time-varying models via surrogate modeling and data-driven optimization. To this end, we parametrize the control/input actions of a dynamic problem by postulating nonlinear surrogate models that are functions of time. The parameters of these control actions are then derived by collecting input-output information from the dynamic models and by using data-driven optimization techniques. The fine-tuned parameters are finally used to retrieve the optimal state trajectories across the time horizon. We test the applicability of our approach on a stable motivating example, as well as on a nonlinear unstable CSTR case study.

2. Continuous-Time Surrogate Models and Data-Driven Optimization

Our key idea is to represent the decision variables of a dynamic optimization problem (i.e., the control actions) with a continuous-time model rather than with discrete decisions taken at every time point. By representing the decision variables as a functional form, the decision variables of the dynamic optimization problem are parametrized and reduced to the number of parameters considered in the surrogate models. Previously, various forms of exponential and polynomial surrogate models have shown to be effective in deterministic optimization problems (Sorek et al., 2017; Beykal et al., 2018). Inspired by this, we heuristically construct one exponential (Eq. (1)) and one polynomial (Eq. (2)) continuous-time surrogate model as candidate control actions, $u(t)$.

$$u(t) = \alpha \cdot \exp(\beta t) + \gamma \quad (1)$$

$$u(t) = \alpha t^2 + \beta t + \gamma \quad (2)$$

The bounds on the unknown parameters are carefully identified by plotting these continuous actions across a wide range of parameter space. This ensures that the continuous-time action models capture all possible actions within the bounds of $u(t)$ while being flexible and generic such that they can be applied to various dynamic optimization problems.

Once the parametrization is completed and the dynamic model is constructed for a given problem, our goal is to find the optimal continuous control action that will minimize the control objective by fine-tuning the surrogate model parameters, α , β , and γ , using data-driven optimization. The data-driven optimization algorithm creates samples for the model parameters and these sampling points are used to construct $u(t)$. This continuous-

time control action is then fed to the dynamic model, which is simulated to collect the final value of the control objective as output information. If there are terminal constraints in the dynamic model, these can also be collected from the simulation and will be treated as additional output information (i.e., grey-box constraints). Using this input-output data, the solvers identify promising regions for exploration based on their specific data-driven optimization strategies and converge to a solution. We then use the fine-tuned parameters to generate the optimal state trajectories and evaluate the results. In this work, we explore two different data-driven optimization algorithms for fine-tuning: (1) ARGONAUT, a global model-based solver (Boukouvava & Floudas, 2017), and (2) NOMAD, a local sample-based algorithm (Le Digabel, 2011). The performance of these techniques is demonstrated in the following motivating example.

3. Motivating Example: A Stable Dynamic Model

We study the following dynamic formulation as our motivating example:

$$\begin{aligned} & \min_{u(t)} x^T(t_f) \cdot P \cdot x(t_f) + \int_0^{t_f} x(t)^T \cdot Q \cdot x(t) + u(t)^T \cdot R \cdot u(t) dt \\ & \text{s. t. } \frac{dx(t)}{dt} = A \cdot x(t) + B \cdot u(t), \quad x(0) = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} \\ & A = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, P = \begin{bmatrix} 0.082 & 0.082 \\ 0.082 & 1.082 \end{bmatrix}, Q = I, R = 0.01 \\ & -2 \leq u(t) \leq 2, u(t) \in \mathbb{R}, x(t) \in \mathbb{R}^2 \end{aligned} \quad (3)$$

where t_f is the final time, $u(t)$ is the control action that has the surrogate model form of either Equation 1 or 2 and is bounded between $[-2, 2]$, $x(t)$ are the two states in the problem with initial points $[-0.5, 0.5]$, and $dx(t)/dt$ is the rate of change in the states with respect to time. The unknown parameters of the surrogate control action (i.e., α , β , and γ) are the decision variables of this dynamic optimization problem and optimized using the recipe described above.

The results of the motivating example are summarized in Figures 1A and B which show that the exponential continuous-time control action can drive the states to zero and produce favorable results with a stable objective function trajectory across the time horizon. The trajectory of the optimal control action is also provided in Figure 1C. However, when polynomial continuous-time control action is employed, the states do not converge, and favorable results are not achieved (Figure 1D). This shows that not all surrogate functional forms are appropriate for pursuing data-driven dynamic optimization. Specifically for the motivating example, the exponential continuous-time surrogate model outperforms the polynomial control action by allowing states to converge at zero.

We also test the effect of changing initial conditions on the dynamic problem (Figure 2A). We observe that even with more “difficult” initial conditions, our data-driven approach will lead the states to convergence. In addition, we characterize the effect of using global versus local data-driven optimizers on the dynamic optimization performance. Over 5 random repeated runs, Figures 2A and B show that ARGONAUT, as a global solver, consistently leads to the convergence of states at zero whereas in some instances NOMAD, as a local solver, will return results that diverges the states. Hence, we observe that global optimization will consistently provide the best parameters necessary for the

control of the motivating example. Here, it is important to note that global optimality is not theoretically guaranteed due to the data-driven nature of the approach.

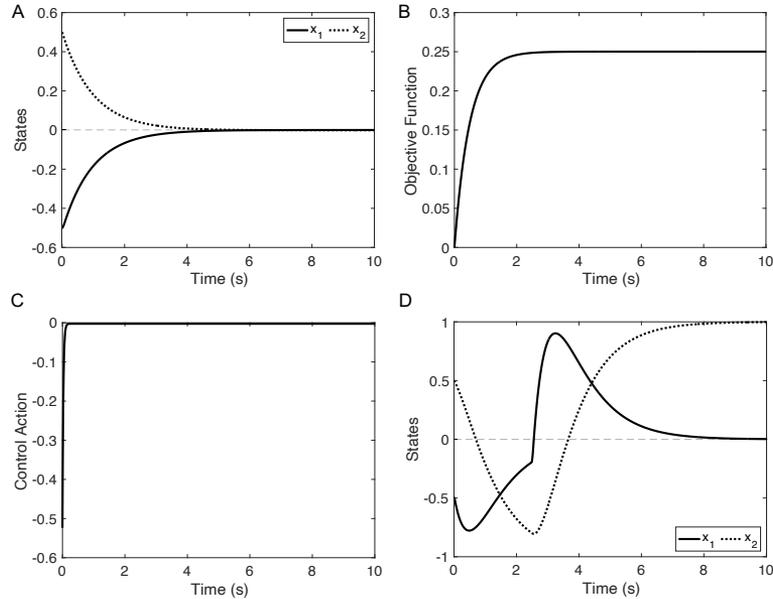


Figure 1 (A) State trajectories and (B) objective function trajectory with exponential continuous-time control action; (C) Optimal control action derived from the exponential surrogate model and data-driven optimization; (D) State trajectories with polynomial continuous-time control action.

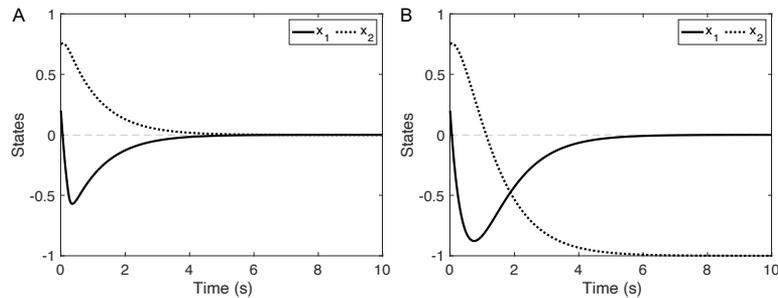


Figure 2 State trajectories for the new initial conditions, $x(0) = [0.2, 0.75]$. The parameters of the continuous-time model are fine-tuned with (A) ARGONAUT and (B) NOMAD algorithm.

We further compare the data-driven solution to the linearized Simulink solution to quantify the performance of our approach. The state profiles obtained from these two different approaches are provided in Figure 3 and the errors are quantified by calculating the area under the curve (AUC) of each state profile. The results show that the data-driven solution matches the linearized solution with high accuracy. For the case of $x(0) = [-0.5, 0.5]$ shown in Figure 3A, the relative and absolute errors between the data-driven and the linearized solution are very small (State 1: Abs. Err. = 0.0013; Rel. Err. = 0.0025; State 2: Abs. Err. = 0.0167; Rel. Err. = 0.0335). Likewise, for the case of $x(0) = [0.2, 0.75]$ shown in Figure 3B, the absolute and relative errors are very small (State 1:

Abs. Err. = 0.0005; Rel. Err. = 0.0007; State 2: Abs. Err. = 0.0070; Rel. Err. = 0.0077), which shows that continuous-time surrogate models can capture the correct dynamics and provide satisfactory results. Next, we test our approach on a nonlinear unstable system.

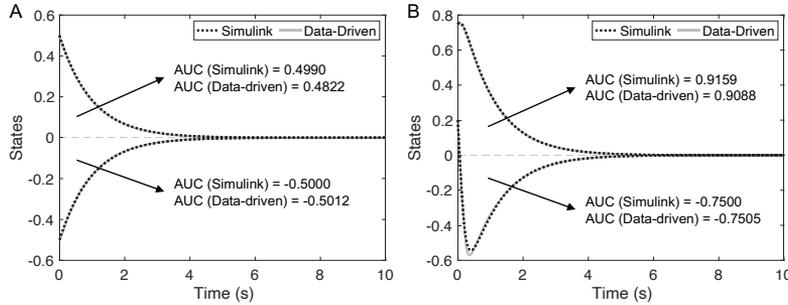


Figure 3 Comparison of state trajectories obtained from the linearized solution and the data-driven solution for: (A) $x(0) = [-0.5, 0.5]$, and (B) $x(0) = [0.2, 0.75]$.

4. Continuous-Stirred Tank Reactor Example: An Unstable Nonlinear Dynamic Model

To further test the limits of our data-driven approach, we formulate the following nonlinear setpoint tracking problem for an unstable CSTR system:

$$\begin{aligned}
 & \min_{u(t)} \int_0^{t_f} \begin{bmatrix} C_A(t) - 0.5 \\ T(t) - 400 \end{bmatrix}^T \cdot QR \cdot \begin{bmatrix} C_A(t) - 0.5 \\ T(t) - 400 \end{bmatrix} dt \\
 & \text{s. t. } \frac{dC_A}{dt} = \frac{\dot{m}}{\rho V} (C_{A0} - C_A) - k_0 C_A \exp\left(-\frac{E_a}{RT}\right) \\
 & \quad \frac{dT}{dt} = \dot{m} C_p (T_0 - T) + V \Delta H_{rxn} k_0 C_A \exp\left(-\frac{E_a}{RT}\right) + \frac{UA(T_c - T)}{V \rho C_p} \\
 & \quad QR = \begin{bmatrix} 50 & 0 \\ 0 & 1 \end{bmatrix}, C_A(0) = 9, T(0) = 400, 0 \leq \dot{m}(t) \leq 100, 10 \leq T_c(t) \leq 600
 \end{aligned} \tag{4}$$

where C_A is the concentration of reactants with the reaction mechanism $A \rightarrow B$, T is the reactor temperature, \dot{m} is the mass flowrate of the reactant into the CSTR, V is the reactor volume, T_c is the cooling jacket temperature, U is the overall heat transfer coefficient, A is the heat transfer area, C_p is the heat capacity at constant pressure, ρ is the density, ΔH_{rxn} is the heat of reaction, k_0 is the pre-exponential factor, E_a is the activation energy, and R is the gas constant. The values of the model parameters are available upon request.

As this problem is more challenging compared to the motivating example, we use the rolling horizon approach to perform data-driven dynamic optimization with continuous-time surrogate models with $t = 0.2s$ intervals. The results in Figure 4A show that the concentration of the reactants is depleting as the reaction is carried out over time. The temperature profile of the reactor presented in Figure 4B shows that the continuous-time control action can track the temperature at the required setpoint (400 K) by adjusting the cooling temperature as shown in Figure 4C. Overall, the data-driven dynamic optimization with continuous-time control actions shows favorable results in both linear stable and nonlinear unstable case studies. For a comparison with an MPC scheme based on a linearized version of the CSTR, please refer to Avraamidou and Pistikopoulos (2017).

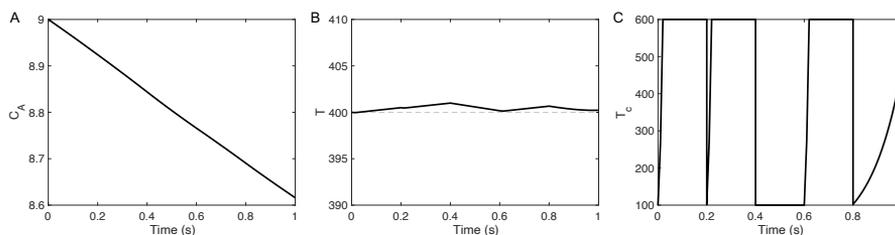


Figure 4 (A) Concentration profile for the reactants over time; (B) Reactor temperature profile over time; (C) Control action for the cooling temperature to keep the reactor temperature at the setpoint.

5. Conclusions

In this work, we present a dynamic optimization strategy based on continuous-time surrogate models and data-driven optimization. By parametrizing the input control actions on the dynamic system using surrogate models that are functions of time, we avoid the full discretization of the dynamic model. The parameters of the control actions are fine-tuned by data-driven optimization which is based on collecting input-output information from the dynamic system. We explore two different continuous-time surrogate models and rigorously test the effect of initialization and type of data-driven optimizer on the control performance. Results showed that exponential continuous-time control actions lead to stable convergence for the states and perform setpoint tracking in dynamic optimization problems. Global optimization of the surrogate model parameters is also found to play an important role in achieving stable solutions. Finally, when compared to the linearized solution, our data-driven approach provides results with minimal errors and matches the linearized state profiles perfectly. This research was funded by the U.S. National Institutes of Health (NIH) grant P42 ES027704.

References

- S. Avraamidou, E.N. Pistikopoulos, 2017, A multi-parametric bi-level optimization strategy for hierarchical model predictive control. *Computer Aided Chemical Engineering*, 40, 1591-1596.
- B. Beykal, F. Boukouvala, C.A. Floudas, N. Sorek, H. Zalavadia, E. Gildin, 2018, Global optimization of grey-box computational systems using surrogate functions and application to highly constrained oil-field operations, *Computers & Chemical Engineering*, 114, 99-110.
- B. Beykal, M. Onel, O. Onel, E.N. Pistikopoulos, 2020, A data-driven optimization algorithm for differential algebraic equations with numerical infeasibilities, *AIChE Journal*, 66, 10, e16657.
- L.T. Biegler, 1984, Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation, *Computers & Chemical Engineering*, 8, 3-4, 243-247.
- F. Boukouvala, C.A. Floudas, 2017, ARGONAUT: AlgoRithms for Global Optimization of coNstrAined grey-box compUTational problems, *Optimization Letters*, 11, 5, 895-913.
- N.A. Dangelakis, B. Burnak, J. Katz, E.N. Pistikopoulos, 2017, Process design and control optimization: A simultaneous approach by multi-parametric programming, *AIChE Journal*, 63, 11, 4827-4846.
- S. Le Digabel, 2011, Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm, *ACM Transactions on Mathematical Software*, 37, 4, 1-44.
- N. Sorek, E. Gildin, F. Boukouvala, B. Beykal, C.A. Floudas, 2017, Dimensionality reduction for production optimization using polynomial approximations. *Computational Geosciences*, 21, 247-266.

Parameter estimation in dynamic metabolic models applying a surrogate approximation

Rafael D. de Oliveira¹, Dielle P. Procópio¹, Thiago O. Basso¹ and Galo A.C. Le Roux¹

¹*Department of Chemical Engineering, Polytechnic School, University of São Paulo, Prof. Luciano Gualberto Ave., 380, São Paulo, 05508-900, Brazil*
galoroux@usp.br

Abstract

Dynamic Flux Balance Analysis (dFBA) models are widely applied in the system biology field. The model connects the cellular genome information to the cell's phenotype, it can therefore be applied to predict the effect of gene deletions or the insertion of new enzymes into the metabolic network. The dFBA model consists of a system of differential equations and an optimization problem that is performed to compute the internal flux distribution. Furthermore, kinetic equations (i.e. Michaelis-Menten) are used to model the uptake of substrates. The kinetic equations have some parameters that must be estimated from batch fermentation experimental data. When the dFBA model is inserted in a parameter estimation architecture, a bi-level optimization problem arises. As is well known, bi-level optimization problems are hard to solve and suffer from convergence problems. A recent method to replace the FBA optimization with a surrogate model was developed in our research group and applied to the simulation of a model predictive control of a bioreactor. Here, that recently developed surrogate dFBA model is applied to a parameter estimation problem. The approach was tested in a case study of *Saccharomyces cerevisiae* fermentation using glucose and xylose as substrates. *S. cerevisiae* is the main microorganism for industrial alcoholic fermentation and there is research to amplify the range of substrates that can be used, such as xylose. dFBA models can be applied to link possible genetic modifications strategies with the bioreactor performance. In order to achieve this aim, the kinetic parameters in the dFBA model must be well defined. First, the surrogate model was trained using Flux Balance Analysis simulations of the Yeast 8.3 genome-scale model. After that, the kinetic parameters on the surrogate dFBA were fitted to *in silico* data. The surrogate dFBA outperformed the sequential approach where the nested LP needs to be solved many times during the estimation. Furthermore, solving the parameter estimation with the surrogate model in a simultaneous approach can considerably reduce the computational time. The results indicated that the surrogate dFBA can be an important tool for the parameter estimation and optimal design of experiments of dynamic metabolic models.

Keywords: Parameter estimation, Surrogate model, dFBA, Systems Biology, Metabolic Engineering

1. Introduction

System biology models are widely applied to characterize microorganisms and aid strain optimization (Maranas, 2016). Kinetic models of cell metabolism are a very promising category of models because of their high prediction capabilities. However, the lack of knowledge of enzymatic reactions mechanisms and the need to estimate an unbearable number of parameters makes

the application of kinetic models restricted to the description of small enzymatic pathways so far. Dynamic Flux Balance Analysis (dFBA) models appear as an alternative approach where the internal flux distribution is described by a steady-state model and the solution computed by an optimization problem. Therefore, in dFBA models, only a small number of parameters must be estimated from experimental data making the problem solvable in practice. Despite the fact that the number of parameters to be estimated is reduced, the dFBA model consists of a system of differential equations and an embedded optimization problem, making the parameter estimation problem challenging to solve.

The sequential solution of the problem consists in solving the nested LP and the ODE system inside the optimization. The lack of gradient information and the non-smoothness of dFBA makes the problem hard to solve. Leppävuori et al. (2011) developed a sequential gradient-based solution with direct sensitivities equations. They estimated 8 parameters and used a metabolic network of 1266 enzymatic reactions and 1061 metabolites. Waldherr (2016) reformulated the bi-level problem as a Mixed Integer Quadratic Program, however, due to the computational burden they used a small-scale network of 10 reactions and 12 metabolites. Raghunathan et al. (2003, 2006) reformulated the problem as a Mathematical Program with Complementary constraints (MPCC). MPCC cannot be solved by standard NLP solvers, therefore they relaxed the complementary constraints using a barrier parameter. They applied the approach to a small-scale metabolic network of 39 reactions and 43 metabolites.

Recently, our group developed a methodology to insert dFBA models in Model Predictive Control (MPC) architecture (Oliveira et al., 2021a). The methodology consists in replacing the embedded optimization problem solution by a surrogate model making the solution of FBA faster and allowing gradient computation. Here, we investigate the suitability of this methodology in order to reduce the computational load of parameter estimation problems using dFBA models. As a case study, a dFBA model is formulated to describe batch cultivation of glucose and xylose mixtures by *Saccharomyces cerevisiae*. *S. cerevisiae* is the main microorganism for industrial alcoholic fermentation; however, the spectrum of substrates is almost restricted to sugars, such as glucose and fructose. *S. cerevisiae* does not naturally consume xylose and the development of strains of *S. cerevisiae* for the conversion of xylose into ethanol by *S. cerevisiae* has been implemented (Kuyper et al., 2004). However, studies are needed in order to make the xylose fermentation by *S. cerevisiae* more efficient. Metabolic models can be very useful to achieve this aim, because they allow the understanding of xylose fermentation in a multi-scale approach, from a genome-scale level to the bioreactor operation.

2. Methodology

Flux Balance Analysis (FBA) has become one of the most popular mathematical methods for simulating metabolism using genome-scale models in the past years (Maranas, 2016), the reason for that relies on the simplicity and applicability of the method. FBA is a method based on an optimization approach using the stoichiometric matrix, by exploiting the fact that all the possible flux distributions of the cell are present in the null space of the stoichiometric matrix. FBA uses an objective function, to try to select one phenotype between the many possibilities. Mathematically, the FBA can be formulated as a Linear Programming problem as follows:

$$\begin{aligned} \max \quad & c^T v \\ \text{subject to:} \quad & S \cdot v = 0 \\ & lb \leq v \leq ub \end{aligned} \tag{1}$$

where lb and ub are lower and upper bound vectors for metabolic fluxes, respectively. The fluxes constraints are typically measured external fluxes or derived from thermodynamics data. S is the

stoichiometric matrix and c is the vector of coefficients that multiplies the flux vector in order to express the objective function. The most applied objective function on FBA is to maximize the biomass formation (Maranas, 2016), where the vector c is a vector with zeros except for the biomass reaction. Many genome-scale reconstruction models for *S. cerevisiae* have been developed so far. Among them, the consensus yeast metabolic network model is the largest network (Lu et al., 2019). Yeast 8.30 has 2666 metabolites and 3928 enzymatic reactions. The genome-scale *S. cerevisiae* model Yeast version 8.30 (Lu et al., 2019) was downloaded from the project's website: <http://sysbiochalmers.github.io/yeast-GEM/>. The uptake rates of the substrates were fixed to solve the FBA problem (note, not dFBA), the uptake rate of oxygen was set to zero, and the objective function was the maximization of biomass yield.

The dynamic version of the FBA (dFBA) model can be built by incorporating mass balance equations for the external metabolites and Michaelis–Menten kinetics equations for describing the uptake of substrates. For the anaerobic fermentation of glucose and xylose by *S. cerevisiae*, the dFBA model can be formulated as follows:

$$\begin{aligned}
 \frac{dX}{dt} &= \mu X & \frac{dG}{dt} &= v_g X \\
 \frac{dE}{dt} &= v_e X & \frac{dZ}{dt} &= v_z X \\
 v_e, \mu &= \Xi(v_g, v_z) & v_g &= v_{g,max} \frac{G}{K_g + G} \\
 X, G, Z, E &\geq 0 & v_z &= v_{z,max} \frac{Z}{K_z + Z} \frac{1}{1 + (G/K_{ig})}
 \end{aligned} \tag{2}$$

where μ , v_g , v_z , and v_e are the growth rate, and the exchange fluxes of glucose, xylose, and ethanol, respectively. X , G , Z , and E represent the biomass, glucose, xylose, and ethanol concentrations, respectively. $v_{g,max}$ and $v_{z,max}$ are the maximum uptake rate for glucose and xylose respectively. K_g and K_z are the saturation constants, and K_{ig} is the glucose inhibition constant. The exchange flux of ethanol v_e and cellular growth μ are computed by the mapping Ξ , which can be solved either by the optimization FBA problem (Equation 1) or by the surrogate approximation.

The methodology to generate the surrogate FBA model was performed as described in Oliveira et al. (2021a). First, a series of optimization problems (FBA) were solved covering the whole flux input domain (v_g and v_z). After that, a polynomial model for each output (μ and v_e) was fitted to the data by Partial Least Square (PLS) to avoid over-fitting. A parameter estimation problem for estimating the five parameters in Equation 2 was implemented. The measurements data of the extracellular metabolites were generated by *in silico* experiments using the nominal parameter values presented in Table 1. The parameter estimation was solved as a nonlinear constrained least-squares problem as follow:

$$\min_{\theta} \sum_j (y_j^c(\theta) - y_j^m)^2 \quad \text{subject to: Equation 2} \tag{3}$$

where θ is the vector of the parameter to be estimated, and y is the vector of extracellular concentrations. Indexes c and m indicates calculated and measured quantities respectively. Three different methods were applied to solve the problem in Equation 3:

1. **dFBA + lsqnonlin**: Solved as a bi-level optimization problem. The outer parameter estimation problem was solved by *lsqnonlin* routine in MATLAB with the levenberg-marquardt

method. The ODE system was solved by ODE15s with the embedded LP (FBA) solved in GUROBI. No gradient information was supplied.

2. **dfBA surrogate + *lsqnonlin***: Solved as a single-level optimization problem by *lsqnonlin* routine in MATLAB with the levenberg-marquardt method. The ODE system was solved by ODE15s with the embedded LP (FBA) replaced by the surrogate model. No gradient information was supplied.
3. **dfBA surrogate + IPOPT**: Solved as a single-level optimization problem in Julia language with the interior point NLP solver IPOPT. The ODE system was solved by orthogonal collocation with the embedded LP (FBA) replaced by the surrogate model. Automatic differentiation package was used to compute the gradient.

3. Results

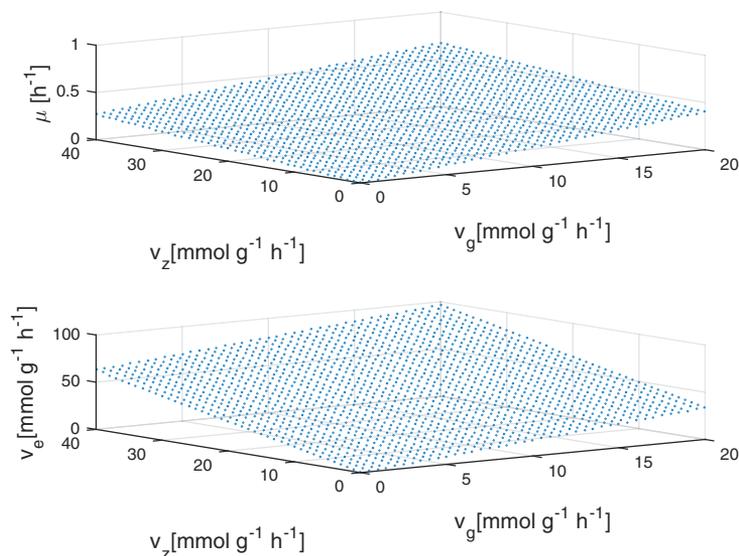


Figure 1: Profiles of the FBA (Equation 1) solutions for different values of the uptake rates of glucose v_g and xylose v_z . We solved the FBA problem for every value of the independent variables in an equidistant 40 by 40 grid.

First, the surrogate FBA model was trained using FBA simulations performed in COBRA Toolbox for MATLAB. The FBA was solved for every value of the v_g and v_z uptake fluxes in an equidistant 40 by 40 grid (Figure 1). Both response surfaces for μ and v_e are flat, which means that the amount of ethanol and biomass being produced are linearly proportional to the uptake of each substrate. Different from the non-linear response surface in Oliveira et al. (2021a) where a piecewise polynomial surrogate model was needed, here a single polynomial could fit the data. The relative Root Mean Square Error (RMSE) was $6.67e-8\%$ and $1.41e-4\%$ for μ and v_e respectively. While the relative Root Mean Square Error of Prediction (RMSEP) was $2.99e-10\%$ and $1.91e-6\%$ for μ and v_e respectively. The relative RMSE and RMSEP were computed by dividing the fluxes by the maximum value of the correspondent uptake flux.

After the FBA surrogate model was trained, the parameter estimation problem was solved using the three different methods described in the methodology section. Ten different initial guesses were supplied to solve the problem by each method and the solution with the lower objective function (OF) was selected. Figure 2 compares the predicted concentrations with the best-fitted set of parameters for each case. Visually, the methods that applied the surrogate FBA fit the *in silico* data adequately, on the other hand, the method that uses the nested LP to solve the dFBA model was unable to fit the data. In fact, all the attempts to solve the estimation problem using the nested LP resulted in a set of parameters close to the initial guess. Because of the embedded optimization problem, the *lsqnonlin* solver was not able to compute efficiently the gradient of the problem. Furthermore, the attempts of using derivative-free methods like *simplex* (i.e. *fminsearch*) have failed as well.

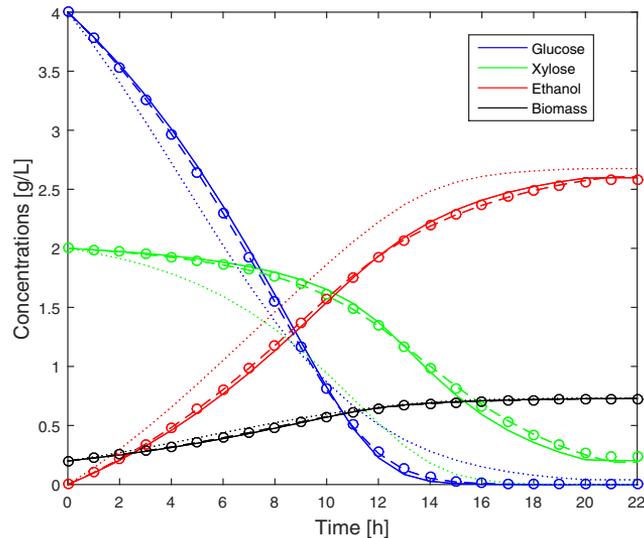


Figure 2: Simulated profiles for dFBA surrogate in Julia (solid), dFBA surrogate in MATLAB (dashed) and dFBA DA in MATLAB (dotted). *in silico* data points are presented as circles.

The performance of each method for solving the problem is presented in Table 1. The dFBA with the embedded optimization had a CPU time about 60 times higher than the methods that used the surrogate FBA. Despite the larger CPU time, the solver performed only 5 iterations and 36 function evaluations. The need of solving the nested LP at each step of the ODE solver makes this method computationally expensive and ineffective, as it can be seen by the poor fit as well (Figure 2). Comparing the methods that applied the surrogate approximation, the utilization of automatic differentiation can improve performance. However, the utilization of the surrogate FBA is enough to guarantee a good fitting. These results illustrate the advantage of the surrogate approximation FBA to solve parameter estimation problems. The time and effort to train the surrogate model must be taken into account in that analysis, but for a small number of input fluxes the task can be easily done. Moreover, Table 1 also presents the set of parameters estimated in each case, as well as the set of parameters used to generate the measurements. The set of estimated parameters was different from the one used to simulate the measurements data even when a good fit was achieved. In fact, the set of parameters in dFBA models are typically dependent and cannot be uniquely identified (Leppävuori et al., 2011). In order to make a complete analysis, the parameter uncertainty must be taken into consideration. This is a practical identifiability issue that should be discussed with more depth. A possibility is to apply a recent methodology developed in our group that used sparse Principal Component Analysis to access the identifiability of metabolic fluxes on carbon labeling experiments (Oliveira et al., 2021b).

Table 1: Comparison of the Computational Performance of each method and the parameter values used in model simulation to yield measurements.

	Model simulation	dFBA surrogate IPOPT	dFBA surrogate <i>lsqnonlin</i>	dFBA surrogate <i>lsqnonlin</i>
CPU	-	0.05 s	26.47 s	27.82 min
iterations	-	29	20	5
function evaluations	-	50	126	36
Objective function	-	0.049	0.009	7.17
v_g^{max}	7.30	6.44	7.13	30.01
K_g	1.03	0.64	0.94	12.02
v_z^{max}	32.00	26.07	4.69	7.99
K_z	14.85	10.48	1.60	0.80
K_{ie}	0.50	0.39	0.81	1.00

4. Conclusion

Dynamic metabolic models are a powerful tool that can be applied from metabolic engineering to bioprocess fields. The estimation of the model parameters poses a challenge due to the bi-level optimization architecture and the non-smooth behavior of the dFBA model. Here, the replacement of the embedded optimization problem by a surrogate model was evaluated. The results demonstrated that the surrogate model can be easily trained from FBA simulations and improve the performance of the estimation problem. In the future, the methodology should be applied to parameter estimation and uncertainty quantification problems using experimental data.

5. Acknowledgments

RDO and GACR would like to express their acknowledgement to CAPES (001), CNPq (312049/2018-8 fellowship) and RCGI/FAPESP (2020/15230-5) for their financial support. DPP and TOB acknowledged funding from FAPESP, grants #2015/50612-8, #2018/17172-2, #2018/01759-4.

References

- M. Kuyper, A. A. Winkler, J. P. van Dijken, J. T. Pronk, Mar. 2004. Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS yeast research* 4 (6), 655–664.
- J. T. Leppävuori, M. M. Domach, L. T. Biegler, Nov. 2011. Parameter Estimation in Batch Bioreactor Simulation Using Metabolic Models: Sequential Solution with Direct Sensitivities. *Industrial & Engineering Chemistry Research* 50 (21), 12080–12091.
- H. Lu, F. Li, B. J. Sánchez, Z. Zhu, G. Li, I. Domenzain, S. Marčišauskas, P. M. Anton, D. Lappa, C. Lieven, M. E. Beber, N. Sonnenschein, E. J. Kerkhoven, J. Nielsen, Aug. 2019. A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nature Communications* 10 (1), 3586, number: 1 Publisher: Nature Publishing Group.
- C. D. Maranas, 2016. *Optimization Methods in Metabolic Networks*, 281.
- R. D. Oliveira, M. N. Guedes, J. Matias, G. A. C. Le Roux, Oct. 2021a. Nonlinear Predictive Control of a Bioreactor by Surrogate Model Approximation of Flux Balance Analysis. *Industrial & Engineering Chemistry Research* Publisher: American Chemical Society.
- R. D. Oliveira, C. S. Nakama, V. Novello, J. G. Gomez, G. A. Le Roux, 2021b. Identifiability of metabolic flux ratios on carbon labeling experiments. In: *Computer Aided Chemical Engineering*. Vol. 50. Elsevier, pp. 1983–1989.
- A. U. Raghunathan, J. R. Pérez-Correa, E. Agosin, L. T. Biegler, Nov. 2006. Parameter estimation in metabolic flux balance models for batch fermentation—Formulation & Solution using Differential Variational Inequalities (DVI). *Annals of Operations Research* 148 (1), 251–270.
- A. U. Raghunathan, J. R. Pérez-Correa, L. T. Bieger, 2003. Data reconciliation and parameter estimation in flux-balance analysis. *Biotechnology and Bioengineering* 84 (6).
- S. Waldherr, Jul. 2016. State estimation in constraint based models of metabolic-genetic networks. In: *2016 American Control Conference (ACC)*. pp. 6683–6688, iSSN: 2378-5861.

A Benchmark Model to Generate Batch Process Data for Machine Learning Testing and Comparison

Margarida L.C. Vicente^{a,b}, José F.O. Granjo^a, Ruomu Tan^c and Franz D. Böhner^{b,*}

^a*CERENA, DEQ, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal.*

^b*Bayer AG, 51368 Leverkusen, Germany*

^c*ABB Corporate Research Germany, Automation & Analytics, Wallstadter Straße 59, 69117 Ladenburg, Germany*
franz.baehner@bayer.com

Abstract

This work presents a benchmark model of a batch process to generate data where active challenges, disturbances, and noise are fully controlled. The process is viewed at the unit level with filling, processing, draining and cleaning operations. The simulation is limited to mass balances. Chemical reactions and kinetics are omitted, as they are not relevant for the intended Machine Learning (ML) application. Several scenarios were generated where 21 disturbances of different types, causes, or fault origins are injected (isolated or combined) in the simulation to test and compare ML algorithms. The scenarios were grouped into six benchmark cycles with increasing levels of complexity in terms of intensity and duration, tackling some major challenges in ML.

Keywords: Batch Process, Machine Learning, Process Modeling, Data Generation, Hybrid Dynamic Simulation

1. Introduction

Batch processes are ubiquitous in pharmaceutical and fine-products industries (Böhner and Huusom, 2019). They are characterized by cyclic operations of one or several units whose sequence and operation parameters follow a specific recipe and a production schedule, typically based on heuristics, and regulatory constraints. Decisions in planning and scheduling are based on batch demonstrated practices (aggregate production rates) and quality constraints. Targeted tracking of batch cycle times enables improved scheduling, but this information is not always available. Time-series values of sensors and controllers are usually available in process historians, and they implicitly contain information about start, end, delays, and variability in the processes. Pre-processing, analyzing and modeling these data are of interest to the planner to support the decision-making.

The development and application of mechanistic models are often not economical in an industrial setting, especially if complex batch processes involve many units and operations. Building those models can be challenging, and extensive experiments for parameter estimation and model validation may be needed. In addition, the resulting nonlinear differential algebraic models are non-smooth, demanding specialized solvers for integration (Floudas and Lin, 2004). Due to the inherent repetitiveness of batch processes, one would expect the Machine Learning (ML) algorithms to be able to recognize patterns and generate meaningful predictions. However, their performance in practice is degraded by data noise (transient, stochastic, or discrete), scarcity of good quality training data, or disturbances in the processes. Incomplete monitoring and complex manual oper-

ations can create reasonable doubts about the quality of data acquired during production (Böhner and Huusom, 2019). Also, the time-varying and nonlinear profiles in batch systems complicate the predictions of ML algorithms. Consequently, it is hard to know which challenges constitute fundamental limitations to the performance of an ML algorithm when working with real process data. Hence, a simulated batch process with modeled data challenges that serve as a benchmark model for the development and study of data-driven techniques, in particular, ML algorithms, is valuable.

This contribution presents a benchmark model for generating simulated data that can be used to develop and test ML algorithms for batch phase-detection problems. By working with simulated data where the active challenges, disturbances, and noise are fully controlled, ML methods can be tested systematically. The dynamic simulations were conducted with a hybrid computational framework where the discrete-events are modeled as state machines in the Stateflow[®] environment, while the continuous components of the model were implemented in Simulink[®].

2. Benchmark Process Model

The benchmark process features *filling, processing, draining, and cleaning* operations on a single unit. Three liquid raw materials (Educt1, Educt2, Educt3) are valve-controlled; and solids are added through a hatch placed on top of the vessel.

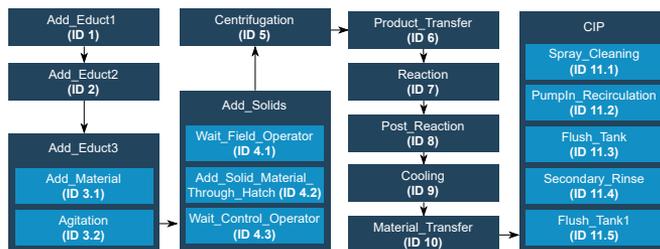


Figure 1: Benchmark Model Recipe.

Table 1: Characteristics of the simulation.

Property	Value	Unit
Vessel volume	12.5	m ³
Threshold volume †	6.25	m ³
Batch duration	19.5	h
Cleaning procedure duration	1.4	h
Simulation Time ‡	333	days
Inactivity Period † ‡	5 min-2 h	-

† The vessel is curved, and as a result, the height has a logarithmic profile until a threshold value, followed by a linear behavior. The height is given as a function of the vessel volume with the threshold mentioned above.

‡ Operation year has 333 days, already leaving a month for cleaning and maintenance.

† ‡ The inactivity period is the time between the end of a cleaning procedure and the beginning of a new batch. It is set as a random duration between 5 min and 2 h, differing on each new batch production.

Besides raw materials, it is considered that before and after the reaction, hot (steam) and cold (water) utilities are, respectively, needed to obey the process design parameters and batch recipe. Fig. 1 illustrates the path, operations (dark blue), and phases (light blue) for the reference process designed. A batch is produced from ID 1 to ID10 with filling, processing, and draining operations. The last operation (ID 11.1 to ID11.5) of each batch is the cleaning procedure.

Modeled variables are vessel contents and their sensors, valves, flows, and the hatch position. While the process just described is apparently simple, complexity is induced through mathematical expressions which emulate various disturbances. Besides linear profiles, the vessel level can

follow an exponential, step, stair function. The recipe is based on duration and volume conditions. For this purpose, the flow rate is adjusted as a degree of freedom to connect volume with time. Characteristics of the process, such as level profile and conditions, can be found in Tables 2 and 1.

Table 2: Transition trigger conditions, nominal durations, and level profiles of the benchmark process.

Event ID	Transition Trigger Condition †	Nominal Duration	Level Profile
1	L1_PV \geq 30%	3 min	Linear
2	L1_PV \geq 60%	3 min	Linear
3.1	L1_PV \geq 65% or $t >$ 5 min	5 min	Linear
3.2	L1_PV \geq 70%	5 min	Noise (\pm 5%)
4.1	$t >$ 15 min	15 min	Linear
4.2	L1_PV \geq 85%	5 min	$G = 16/(s^2 + 3.2s + 16)$
4.3	$t >$ 5 min	5 min	Linear
5	L1_PV \geq 90% Design: 5 steps	increase: 3 min inactivity: 7 min	Step
6	L1_PV \leq 60%	3 h	Linear
7	L1_PV \geq 85%	12 h	$G = 10e^{-0.05s}/(s^2 + 18s + 15)$
8	$t >$ 2 h	2 h	Linear
9	L1_PV \leq 70%	1 h	$G = 5e^{-0.05s}/(0.001s^2 + 0.2s + 3)$
10	L1_PV \leq 0%	5 min	Linear
11.1	L1_PV \geq 10%	30 min	Noise ‡
11.2	$t >$ 14 min	14 min	
11.3	L1_PV \leq 0%	5 min	
11.4	L1_PV \geq 5%	30 min	
11.5	$t >$ 2.5 min	2.5 min	

† L1_PV corresponds to filling percentage. ‡ The injected noise is part of the process nominal profile.

3. Process Model Implementation

The tank volume is calculated from a mass balance where incompressible flow and constant density at the vessel entry and exit are assumed:

$$\frac{d \text{ Volume}}{dt} = valve_{out} \cdot q_{out} - valve_{in} \cdot q_{in} \quad (1)$$

where $valve_{in}$ and $valve_{out}$ are binary variables indicating the opening of the inlet valve and the outlet valve. Changing $valve_{in}$ and $valve_{out}$ makes it possible to control the liquid flow direction. To model realistic instrumentation, the volume is transformed to a level signal representing the vessel filling level (L1_PV in Table 2). The flow rates of heating/cooling agents are assumed to increase/decrease the profile according to the transfer functions of ID7 and ID9 in Table 2. In specific states, noise is added to the level signal to obtain a specific level profile. During cleaning procedures, it is worth noting that heavy sensor noise is often masked due to the filling and draining of water, which induces agitation and oscillation behaviors.

3.1. Disturbances Mapping and Implementation

Disturbances are classified as a *step*: which mirrors time effects such as recipe changes; *seasonality*: oscillation emulating time effects such as season of the year; *stair*: for example, consecutive steps (changing supplier multiple times); and *drift*: such as a linear term overlay. Table 3 lists key disturbances scenarios studied. All disturbances are modeled stochastic wherein the user controls likelihood and severity. Twenty one simulation scenarios have been implemented in Stateflow® (ID1-ID10, ID13) or Simulink® (ID11, ID12, ID14-ID21).

4. Evaluation of ML algorithms

As a first step towards assessing how ML algorithms respond to the challenges mentioned above, Just (2021) compared the performance of Random Forests (RF) and the Hidden Markov Model (HMM). These algorithms are capable of classifying sequences in time-series data (Tsay and Chen, 2019). Several studies compare the two in terms of overall accuracy on labeling, classification speed, memory consumption, feature computation, and model complexity. A comparison between both algorithms is provided in Munther et al. (2016) regarding internet traffic data. Two criteria discussed in Just (2021) were used to evaluate the algorithms learning performance. Accuracy describes the number of correct labels divided by the total number of labels. The mean absolute error (MAE) measures the average error in identifying change points (CPs) using the distance between real (CP_{real}) and predicted change ($CP_{\text{predicted}}$) points in time units.

$$\text{MAE} = \frac{\sum_{i=1}^{\#\text{CP}} |CP_{\text{predicted}} - CP_{\text{real}}|}{\#\text{CP}} \quad (2)$$

Table 3: Disturbances mapping and description, classified by cause. ID identification for each benchmark cycle with multiple disturbances scenarios (second column).

ID	Cycles	Cause of Noise/Fault	Behavior Description
1	A-F	Delay of single-phase end	Phase end is delayed, no change in actuation occurs during the phase
2	A-F	Delay of multiple phases ends	e.g., if the yield was found to be insufficient after quality sample, an additional reaction step is provided
3	C-F	An irregular single-phase occurs	The CIP does not always occur at the batch end
4	C-F	An irregular series of phases occur	After a specific number of batches, the volume decreases between 30-70% of the batch size. Also, small variability of 2% to 5% can be added to final filling stages
5	A-F	Tank is filled to different final fill levels	End of phase 1 happens, but logged time of phase 1 is at the beginning of phase 2, the label of which is delayed
6	†	Start/End points of the label(s) are shifted	It appears wrongly labeled as phase 1 lasts for the time of the 2 phases, but the sensor measurements remain the same
7	F	Phase wrongly labeled	
8	†	Two consecutive phases have the name of the first phase	
9	A-F	A valve is opened and closed several times, and nothing happens before the material transfer starts	Instead of being opened and closed once, it happens more times. Valve starts open
10	A-B, D-F	Valve opens and closes several times, and nothing happens after material transfer ends	Instead of being opened and closed once, it happens more times. Valve starts closed
11	B-F	Phase name remains the same, but the actuation changes	A different valve is opened at several instances without that behavior having meaning to the reference process
12	A-F	A pump slowly supplies less throughput	The flow rate decreases in time. To reach the same level of liquid in the tank, a certain task takes more time to complete
13	A, D-F	The motor provides less agitation	Rotation number decreases randomly, resetting after a batch
14	C-F	Utility Flow masked with noise	Incremental changes in the flow with increasing spikes
15	C	Loss of signal	A series of data points are not written for the sensors
16	†	Value outside of sensor range	If the level measurement deviates from normal values (from 0 to 100), an error message appears
17	C, D	A sensor suffers from a gradual drift for a period of time	Drift lasting until the end of the simulation, resulting in volume shift, as much as the slope chosen
18	B, E-F	A sensor suffers from a gradual drift and is suddenly recalibrated	Similar to ID15 but the sensor has a sudden recalibration after a specific number of batches
19	B, D-F	A sensor suddenly has an offset (recalibration or fault)	At random points in time, with small durations, the sensor has offsets. The actuation remains unaltered
20	B-F	A wide variety of sensor noise	Statistical distributions: Gaussian, uniform, and waveform
21	A-F	White-Band Noise	Added to the signal introducing spikes in measurements

† Implemented but not included in the benchmark cycles reported here.

5. Results

The profile L1_PV at nominal behavior (i.e., disturbances-free) of the reference model for a single batch is represented in Fig. 2. By following a sequence with an increasing complexity level, disturbances scenarios were grouped by six benchmark cycles and listed in Table 3. Simulation

runs were conducted for the duration of a year (ca. 382 batches). Fig. 3 summarizes the accuracy and MAE obtained by RF and HMM for the benchmark cycles A-F.

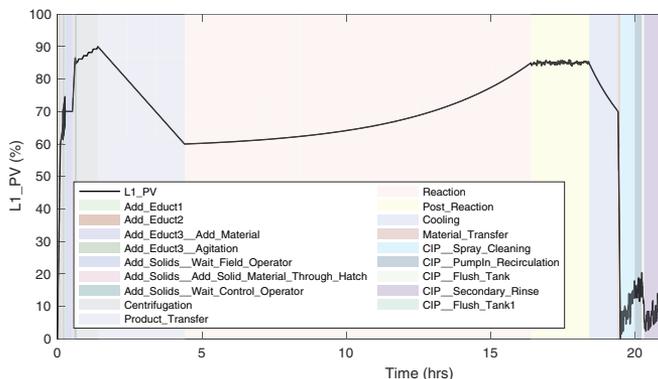


Figure 2: Nominal behavior of a batch and a cleaning procedure for the reference model with batch phase identification.

Cycles A and B yield the best results for accuracy, with the MAE decreasing slightly for the HMM on Cycle B. HMM correctly estimates in Cycle A the number of change points and the correct recipe sequence is recognized. Nonetheless, there are cases where the location of the change points is incorrectly estimated. Just (2021) show that the most problematic transitions are from *Centrifugation* to *Product_Transfer* and between all phases of the cleaning procedure. Between the HMM and RF, the latter confuses these state transitions more often than the HMM. This confusion between states might be due to heavy noise being added to the signal for the *CIP*, where the overlap of volumes for consecutive states is a limitation.

Prediction performance was worst for Cycle C, especially for the HMM. Because the implementation of ID15 is independent of the number of batches, fewer occurrences of the *CIP* affected the level profile for a longer time. While the RF can order input characteristics according to their influence on the classification decision, the HMM always outputs the same phase for these batches. The HMM could not predict Cycle F because individual observations could not be assigned to any state. By enabling a label disturbance, this algorithm failed in recognizing states.

The overall results indicate the RF to be less generalizable and tends to overfit. Also, the RF confuses more states (lower MAE), whereas the HMM can identify the correct sequences because it can learn the sequence of recipe steps. This is particularly relevant when ID4 is active for Cycles C to F. The HMM delivered accuracies above 85% in all but Cycles C and F. In these two cycles, the performance of RF was not affected, achieving accuracies of 79%–81%. The disturbances that caused the most degradation in performance were ID7 (purpose incorrectly labeled training data) and ID15 (loss of several signal points) combined with ID4 (phase occurring irregularly).

6. Conclusions

The focus of this work was to develop a benchmark model to support the testing and comparison of ML methods concerning their capability of recognizing and predicting batch recipes based on labeled time-series data. This is motivated by the fact that there are few comparable applications of ML and no formal methods for assessing the capability of the shown algorithms in the open literature. The entire process model is viewed at the unit level with filling, processing, draining,

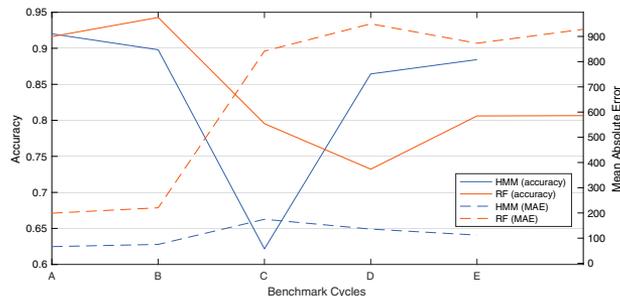


Figure 3: Evaluation of accuracy and MAE for each benchmark cycle with RF and HMM (Just, 2021).

and cleaning operations. Since active disturbances and data noise are fully controlled, the benchmark model offers capabilities that enable the user to design several simulation scenarios, building a deeper understanding of the ML algorithms.

Different scenarios were generated where up to 21 disturbances of several types, causes, or fault origins are injected into the simulated data to compare the RF and HMM algorithms. The scenarios were grouped into six benchmark cycles with increasing levels of complexity in terms of intensity, duration, and probability. The disturbances shown to be more challenging were the ones where likelihood and dependency between consecutive states are affected, which correspond to ID4, ID5, ID7, and ID11. The results show that the HMM deals with most of the implemented process disruptions and makes better predictions than the RF.

The process model here presented can be readily expanded to incorporate additional features inspired in real-life process data as those discussed in Stief et al. (2019), including: (i) valves that partially open, (ii) inclusion of equilibrium- and kinetically-controlled chemical reactions, (iii) temperature monitoring and incorporating heat transfer, and (iv) more nonlinear profiles can be considered in the batch recipe. A performant model implementation could be achieved by reducing the continuous part of the model to a simple design of experiments and emulating disturbance profiles based on process insight rather than building a detailed mechanistic model. Separating the discrete model part (Stateflow[®]) and continuous part (Simulink[®]) allowed building a modular model, facilitating its expansion and adaptation.

Acknowledgements

The authors acknowledge the financial support by CERENA (strategic project FCT-UIDB/UIBD/04028/2020) and by the Federal Ministry for Economic Affairs and Energy of Germany (BMWi) within the innovation platform “KEEN - Artificial Intelligence Incubator Laboratory in the Process Industry” (grant number 01MK20014T).



References

- F. D. Böhner, J. K. Huusom, 2019. A Bottlenecking Study of an Industrial Pharmaceutical Batch Plant. *Ind. Eng. Chem. Res.* 58 (43), 20003–20013.
- C. A. Floudas, X. Lin, 2004. Continuous-time versus discrete-time approaches for scheduling of chemical processes: A review. *Computers & Chemical Engineering* 28, 2109–2129.
- G. Just, 2021. *Zeitreihenanalyse zur Erkennung von Batchphasen in der Prozessindustrie*. Technische Universität Dresden.
- A. Munther, R. R. Othman, A. S. Alsaadi, M. Anbar, 2016. A performance study of Hidden Markov Model and Random Forest in internet traffic classification. In: K. J. Kim, N. Joukov (Eds.), *Information Science and Applications (ICISA) 2016*. Springer Singapore, Singapore, pp. 319–329.
- A. Stief, R. Tan, Y. Cao, J. R. Ottewill, N. F. Thornhill, J. Baranowski, 2019. A heterogeneous benchmark dataset for data analytics: Multiphase flow facility case study. *Journal of Process Control* 79, 41–55.
- R. S. Tsay, R. Chen, 2019. *Nonlinear time series analysis*. John Wiley & Sons, Inc.

Global warming impact of electric city buses in Chile: Critical stages of their fabrication and use

Guillermo Valenzuela-Venegas^a, Daniel Peña-Torres^{a,b}, Franco Lizama-Valenzuela^a
and Melanie Colet-Lagrille^{a,b,*}

^a*Department of Chemical Engineering, Biotechnology and Materials, Faculty of Physical and Mathematical Sciences, Universidad de Chile, Beauchef 851, Santiago 8370456, Región Metropolitana, Chile*

^b*Advanced Mining Technology Center (AMTC), Faculty of Physical and Mathematical Sciences, Universidad de Chile, Av. Tupper 2007 (AMTC Building), Santiago 8370451, Chile
mcolet@ing.uchile.cl*

Abstract

The government of Chile has pledged that 100% of the city buses fleet will be powered by electricity by 2040. Since the fabrication processes of electric vehicles and electricity generation are not emissions-free, process analysis appears as a valuable tool to quantify their implications towards the optimization of the industrial processes in terms of better use of resources and recycling. This work aims to quantify the global warming potential and identify critical stages in the fabrication and use of electric and diesel city buses in Chile by process simulation of: (i) metallurgical processes (aluminum, copper and steel), (ii) thermoelectric power plants, and (iii) diesel production plants. A total of 618 kg of aluminum, 187 kg of copper and 11,538 kg of steel are required for the fabrication of an electric bus, while for a conventional bus are required 61 kg of aluminum, 10,354 kg of steel and 268 kg of cast iron. The results indicate that 12.94 tons of CO_{2,eq} per ton of aluminum, 1.22 tons of CO_{2,eq} per ton of copper and 1.38 tons of CO_{2,eq} per ton of steel are generated, with the major contribution to emissions coming from the Hall-Héroult process, the electrorefining stage and the iron blast furnace, respectively. A natural gas power plant in Chile produces 1.0 kg of CO_{2,eq} per kWh generated, which corresponds to 0.51 kg of CO_{2,eq} per km driven for electric buses. The processing of crude oil in the United States (main supplier of diesel in Chile) generates 0.61 kg of CO_{2,eq} per litre of diesel produced, corresponding to 1.38 kg of CO_{2,eq} per km driven for conventional buses. These results are in good agreement with previous reports which suggest that the positive impact of switching to electricity-powered vehicles is strongly correlated with the carbonization-level of the electricity grid related to buses fabrication and charging.

Keywords: city buses, metallurgical raw materials, electricity generation, diesel production, greenhouse gas emissions

1. Introduction

Electricity-powered vehicles do not produce greenhouse gas emissions directly during their use. However, the processes associated with their fabrication (mainly mining and metallurgical processes) and electricity generation (still strongly based on combustion of fossil fuels) are not emissions-free. Therefore, numerous life cycle assessment (LCA) studies have been issued in the past two decades aiming to quantify the real impact of switching from a transport system

powered by fossil fuels to one powered by electricity (Bouter et al., 2020; Cooney et al., 2013; Nordelöf et al., 2014; Tagliaferri et al., 2016; Xylia et al., 2019).

Nordelöf et al. reported an LCA of city buses powered by electricity, hydrogenated vegetable oil, or diesel. The dependence of the environmental impacts on the buses degree of electrification and the charging electricity generation system were studied for Sweden, the European Union and the United States of America. Their results show that the impact decrease in climate change is strongly correlated with the carbonization-level of the electricity grid related to buses fabrication and charging (Nordelöf et al., 2019c). The importance of the electric grid makeup in the life cycle inventory assessment (LCIA) of electric public transportation buses is also suggested by other authors (Cooney et al., 2013; Ercan and Tatari, 2015). Moreover, Nordelöf et al. suggested that the abiotic resource use in the fabrication of the electrical automotive traction machines, automotive power electronic inverter units, and permanent magnet electric traction motors (which includes copper, cobalt, steel, aluminum and rare earth elements such as neodymium, dysprosium and samarium) can have significant impacts in climate change (Nordelöf and Tillman, 2018; Nordelöf et al., 2018, 2019a,b).

This work aims to quantify the global warming potential (GWP) associated with the higher-impact stages in the fabrication and use of city buses powered by electricity or diesel in Santiago, Chile, pursuing the industrial processes optimization via a circular economy approach. Its novelty is centred in the methodology employed, which consist in the direct simulation of processes for key raw materials production and energy generation to determine the stages with the higher GWP associated. This method comes to complement the conventional LCA method based on the use of commercial software databases to assess the environmental impacts of production processes, which can generate data for each stage in a life cycle but cannot go into the operational details (for example, kinetics, efficiencies, stream compositions, specific energy consumptions, and their interrelations).

2. Case Study

This study is centered in two different vehicles of the bus line no. 506 in Santiago, Chile: (i) an all-electric powertrain bus, model BYD K9 (hereafter electric bus); and (ii) a diesel engine powertrain bus, model Mercedes Benz O500U (hereafter conventional bus). The latter was used as a reference for the assessment of the GWP impact associated with the fabrication and use of an electric bus. For the electric bus the analysis considered the key raw materials production processes associated with the fabrication of the lithium iron phosphate (LFP) batteries module, chassis, frame and body, and the electricity generation according with the existing energy matrix in Chile. For the conventional bus the analysis considered the key raw materials production processes associated with the fabrication of the engine, chassis, frame and body, the production of diesel, and the emissions associated with the bus operation. Data related to buses parts and operation was provided by a local bus operator stakeholder (Reborn Electric SpA¹), while information related to electricity generation and diesel production (import) was obtained from public Chilean government databases (Comisión Nacional de Energía, 2020a).

Technical information indicates that a considerable weight fraction of the LFP batteries module is represented by aluminum (33 wt.%) and copper (10 wt.%) (Golubkov et al., 2014). The fabrication of a bus chassis, frame and body requires considerable amounts of steel, while the fabrication of a conventional bus engine requires steel, cast iron and aluminum (Harr et al., 1999; BYD, 2018). Because the production of aluminum, copper, steel and cast iron is associated with difficult-to-decarbonize processes, they were considered critical stages in the fabrication of electric and conventional city buses (Davis et al., 2018). Table 1 summarizes the amounts of aluminum, copper, steel and cast iron required for the fabrication of an electric bus and a conventional bus.

¹<https://www.rebornelectric.cl/>

Regarding electricity generation in Chile, fossil fuel power plants stand for almost a 49% of the total electricity generation capacity. Of these a 16% corresponds to pulverized coal combustion, integrated gasification combined cycle and fluidized bed combustion power stations, while another 16% corresponds to gas turbine and natural gas combined cycle technologies (Comisión Nacional de Energía, 2020a; Gaete-Morales et al., 2019). Concerning diesel consumption in Chile, an 82% of the imports comes from the United States of America (Comisión Nacional de Energía, 2021).

Table 1: Weight of aluminum, copper, steel and cast iron required for the fabrication of an electric bus and a conventional bus.

<i>Electric bus</i>		
Bus part	Metal required	Weight / kg
Cathode current collector	Aluminum	618
Anode current collector	Copper	187
Chassis, body and frame	Steel	11,538
<i>Conventional bus</i>		
Bus part	Metal required	Weight / kg
	Steel	205
Engine	Cast iron	268
	Aluminum	61
Chassis, body and frame	Steel	10,149

3. Methodology

The conventional Bayer process in conjunction with the Hall-Héroult process were simulated for aluminum production from a gibbsite rich mineral (Habashi, 1998). For copper the pyrometallurgical via for the metal refining was simulated since this is the method mostly applied for processing sulfide minerals, particularly chalcopyrite, which are the mostly abundant source of copper worldwide (Schlesinger et al., 2011). Similarly, for steel and cast iron the pyrometallurgical via was simulated considering that most of its production worldwide comes from processing hematite rich minerals (Seetharaman, 2014). The computational simulations consisted of the conceptualization of the three processes flowsheets, the resolution of steady state mass and energy balances for each stage in these flowsheets, and the determination of the GWP associated with each of these stages. For this purpose, data reported in scientific publications, metallurgy books and technical reports was used to characterize the input streams flow rates and compositions, the operational parameters (e.g., solid particle sizes, temperatures, conversions) and thermodynamic values (standard enthalpies and specific heat values) associated with each process flowsheet. Suitable software for processes simulation, such as HSC Sim 10, was used.

Regarding electricity generation, pulverized coal power plants (PCPPs) and natural gas combined cycles (NGCCs) were simulated considering they are the mostly used technologies in Chile (Gaete-Morales et al., 2019). Details on the stages and data about the operational parameters associated with both processes were obtained from environmental reports published by the Servicio de Evaluación Ambiental (2015) and the information provided by the Comisión Nacional de Energía (2020b). Averages of the gross power generated in Chile by each technology were considered to facilitate the simulation of the processes.

Concerning the production of diesel, a common process for petroleum refining in the United States of America was considered since this country is the main supplier of this fossil fuel in Chile.

Details on the stages and data about the operational parameters associated with this process were obtained from (Comisión Nacional de Energía, 2020a; Liu et al., 2018).

The computational approach followed for the simulation of electricity generation and production of diesel was the same as described for the metallurgical processes.

4. Results

The results of the simulation indicate that an electrical power consumption of 984.16 MW and a heating power consumption of 6.28 GW are associated with the Bayer process in the aluminum production. These values correspond to total emissions flows of 341.97 tons of CO_{2,eq} per hour and 1,253 tons of CO_{2,eq} per hour associated with electricity generation and natural gas combustion, respectively. An electrical power consumption of 4.96 GW was determined for the Hall-Héroult process, which corresponds to a total emissions flow of 1,720 tons of CO_{2,eq} per hour. Additionally, 346.08 tons of CO_{2,eq} per hour are generated as product of the graphite anodes oxidation in the electrolysis cells. Therefore, considering a production rate of 282.91 tons of aluminum per hour (matching with the alumina production capacity of the Alcoa Pinjarra plant in Australia), a global emissions factor of 12.94 tons of CO_{2,eq} per ton of aluminum produced was determined. The process stages identified with the higher GWP are the digestion stage (Bayer process): 34.2% of the total CO_{2,eq} emissions, the calcination stage (Bayer process): 9.3% of the total CO_{2,eq} emissions, and the electrolysis stage (Hall-Héroult process): 56.4% of the total CO_{2,eq} emissions.

Regarding copper production, an electrical power consumption of 27.77 MW and a heating power consumption of 130.35 MW were determined. These values correspond to total emissions flows of 17.22 tons of CO_{2,eq} per hour and 29.52 tons of CO_{2,eq} per hour associated with electricity generation and natural gas combustion, respectively. Therefore, considering a production rate of 38.2 tons of copper per hour (matching with the copper production capacity of the “El Teniente” Codelco plant in Chile) a global emissions factor of 1.22 tons of CO_{2,eq} per ton of copper produced was determined. The process stages identified with the higher GWP are the drying stage: 18.2% of the total CO_{2,eq} emissions, the fire refining stage: 16.5% of the total CO_{2,eq} emissions, and the electrorefining stage: 58.6% of the total CO_{2,eq} emissions.

Finally, the simulation of the steel production process indicates a heating power consumption of 627 MW. This value corresponds to a total emissions flow of 126.18 tons of CO_{2,eq} per hour. Additionally, 663.8 tons of CO_{2,eq} per hour are generated as product of the reactions taking place in the iron blast furnace and the basic oxygen furnace. Considering a production rate of ca. 570.78 tons of steel per hour (matching with the steel production capacity of the Arcelor Mittal Asturias plant in Spain) a global emissions factor of 1.38 tons of CO_{2,eq} per ton of steel produced was determined. Since cast iron is an intermediate product of the steel production process an emissions factor of 1.26 tons of CO_{2,eq} per ton of iron cast produced was determined, which indicates that in this case the iron blast furnace is the stage with the higher GWP (96.7% of the total CO_{2,eq} emissions).

Concerning electricity generation, a PCPP that generates a gross power of 249.1 MW and a NGCC that generates a gross power of 390.4 MW (with 2/3 of this power generated in the gas turbine and the remaining power generated in the steam turbine) were simulated. A total emissions flow of 576.8 tons of CO_{2,eq} per hour was determined for the PCPP technology, of which a 48.7% is associated with the pulverized coal boiler operation and a 37.8% is associated with the desulfurization stage. Correspondingly, a total emissions flow of 173.2 tons of CO_{2,eq} per hour was determined for the NGCC technology associated with the combustion chamber, turbine operation, and heat recovery steam generator (HRSG) stages. Finally, considering capacity factors of 0.81 and 0.53 for the PCPP and the NGCC, respectively (Gaete-Morales et al., 2019), emission factors of 3.1 and 1.0 kg of CO_{2,eq} per kWh generated were determined. Since an electric bus consumes 1.1 kWh per km driven and assuming that the electricity matrix in Chile includes a 48.6% of natural

gas based power plants, an emissions factor of 0.51 kg of CO_{2,eq} per km driven was determined for this type of vehicle.

To conclude, the simulation of a petroleum refining plant processing 100,000 barrels of crude oil per day indicates that a total amount of 1.32 million litres of fuel per day (fuel coming from the same plant) is required to satisfy the utilities requirements of the process. This is associated with an emissions factor of 0.61 kg of CO_{2,eq} per litre of diesel produced. Therefore, since a conventional bus has a performance of 2.4 km per litre of diesel and the combustion of this fuel produces 2.69 kg of CO_{2,eq} per litre (Dirección de Transporte Público Metropolitano, 2014; US Environmental Protection Agency, 2018), the emissions factor associated with its use is 1.38 kg of CO_{2,eq} per km driven.

5. Conclusions

The results obtained indicate that the electrolysis stage (Hall-Héroult process) in the aluminum production process is related to a significant GWP due to its high electricity consumption and the generation of CO₂ as a by-product. Similarly, the electrorefining stage in the copper production process is related to the higher GWP due to its elevated electricity and heat power requirements. This represents a big challenge for the experts in electrochemical engineering to develop alternatives to these conventional processes which offer similar kinetics and the absence of undesirable by-products operating under moderate conditions. In the case of the steel production process, the CO₂ generated as by-product in the iron blast furnace is another sign of the need of developing new technologies which do not depend on the use of fossil fuels or an energy matrix with high carbonization levels. Since the fabrication of an electric bus requires a higher amount of these raw materials than the fabrication of a conventional bus (having in mind that this study does not include other materials such as lithium and rare earths also related to high environmental impacts), it is required to consider new developments and a circular economy approach for the effective implementation of these vehicles.

The resulting emissions factor associated with the use of an electric bus (0.51 kg of CO_{2,eq} per km driven) is, as expected, lower than the emissions factor associated with the use of a conventional bus (1.38 kg of CO_{2,eq} per km driven). This is a positive result considering that the electricity matrix in Chile is still highly dependent on fossil fuels. However, the fabrication impact must not be disguised by this result since it is also associated with the urgent need of recycling strategies at the end of life considering, for example, that batteries modules have an optimal lifetime of approximately six years (a half of the optimal lifetime of a conventional bus).

It is thought that the proposed methodology successfully complements the LCA studies reported in the literature since it allows to identify the processes stages with the higher GWP associated and the operational aspects that explain it, thus offering guidelines for future research and optimization in the fabrication and use of electric vehicles.

References

- A. Bouter, E. Hache, C. Ternel, S. Beauchet, 2020. Comparative environmental life cycle assessment of several power-train types for cars and buses in France for two driving cycles: “worldwide harmonized light vehicle test procedure” cycle and urban cycle. *The International Journal of Life Cycle Assessment* 25, 1545–1565.
URL <https://doi.org/10.1007/s11367-020-01756-2>
- BYD, 2018. *Bus Urbano Eléctrico K9G*, 2018.
URL <https://bydelectrico.com>
- Comisión Nacional de Energía, 2020a. *Estadísticas para Electricidad e Hidrocarburos*.
URL <https://www.cne.cl>
- Comisión Nacional de Energía, 2020b. *Installed Capacity in the SIC and SING*.
URL <https://www.cne.cl>
- Comisión Nacional de Energía, 2021. *Importaciones y Exportaciones de Hidrocarburos*.
URL <http://energiaabierta.cl>

- G. Cooney, T. R. Hawkins, J. Marriott, 2013. Life Cycle Assessment of Diesel and Electric Public Transportation Buses. *Journal of Industrial Ecology* 17, 689–699.
URL <https://doi.org/10.1111/jiec.12024>
- S. J. Davis, N. S. Lewis, M. Shaner, S. Aggarwal, D. Arent, I. L. Azevedo, S. M. Benson, T. Bradley, J. Brouwer, Y.-M. Chiang, C. T. M. Clack, A. Cohen, S. Doig, J. Edmonds, P. Fennell, C. B. Field, B. Hannegan, B.-M. Hodge, M. I. Hoffert, E. Ingersoll, P. Jaramillo, K. S. Lackner, K. J. Mach, M. Mastrandrea, J. Ogden, P. F. Peterson, D. L. Sanchez, D. Sperling, J. Stagner, J. E. Trancik, C.-J. Yang, K. Caldeira, 2018. Net-zero emissions energy systems. *Science* 360 (6396), eaas9793.
URL <https://www.science.org/doi/10.1126/science.aas9793>
- Dirección de Transporte Público Metropolitano, 2014. Proyecto de Mejoramiento Tecnológico en Buses del Sistema de Transporte Público de Santiago.
URL <http://biblioteca.digital.gob.cl>
- T. Ercan, O. Tatari, 2015. A hybrid life cycle assessment of public transportation buses with alternative fuel options. *The International Journal of Life Cycle Assessment* 20, 1213–1231.
URL <https://doi.org/10.1007/s11367-015-0927-2>
- C. Gaete-Morales, A. Gallego-Schmid, L. Stamford, A. Azapagic, 2019. Life cycle environmental impacts of electricity from fossil fuels in Chile over a ten-year period. *Journal of Cleaner Production* 232, 1499–1512.
URL <https://doi.org/10.1016/j.jclepro.2019.05.374>
- A. W. Golubkov, D. Fuchs, J. Wagner, H. Wiltche, C. Stangl, G. Fauler, G. Voitc, A. Thaler, V. Hacker, 2014. Thermal-runaway experiments on consumer Li-ion batteries with metal-oxide and olivin-type cathodes. *RSC Advances* 4, 3633–3642.
URL <https://doi.org/10.1039/C3RA45748F>
- F. Habashi, 1998. *Handbook of Extractive Metallurgy*. Wiley-VCH.
- T. Harr, K.-G. Rolker, H. Bergmann, C. Argachoy, 1999. The New Mercedes-Benz Medium Duty Engines OM 924 LA and OM 926 LA of the 900 Engine Series for Trucks and Buses. *Tech. rep.*
URL <https://doi.org/10.4271/1999-01-3086>
- Y. A. Liu, A.-F. Chang, K. Pashikanti, 2018. *Petroleum Refinery Process Modeling: Integrated Optimization Tools and Applications*. John Wiley & Sons.
- A. Nordelöf, M. Alatalo, M. L. Söderman, 2019a. A scalable life cycle inventory of an automotive power electronic inverter unit—part I: design and composition. *The International Journal of Life Cycle Assessment* 24, 78–92.
URL <https://doi.org/10.1007/s11367-018-1503-3>
- A. Nordelöf, E. Grunditz, S. Lundmark, A.-M. Tillman, M. Alatalo, T. Thiringer, 2019b. Life cycle assessment of permanent magnet electric traction motors. *Transportation Research Part D: Transport and Environment* 67, 263–274.
URL <https://doi.org/10.1016/j.trd.2018.11.004>
- A. Nordelöf, E. Grunditz, A.-M. Tillman, T. Thiringer, M. Alatalo, 2018. A scalable life cycle inventory of an electrical automotive traction machine—Part I: design and composition. *The International Journal of Life Cycle Assessment* 23, 55–69.
URL <https://doi.org/10.1007/s11367-017-1308-9>
- A. Nordelöf, M. Messagie, A.-M. Tillman, M. Ljunggren Söderman, J. Van Mierlo, 2014. Environmental impacts of hybrid, plug-in hybrid, and battery electric vehicles—what can we learn from life cycle assessment? *The International Journal of Life Cycle Assessment* 19, 1866–1890.
URL <https://doi.org/10.1007/s11367-014-0788-0>
- A. Nordelöf, M. Romare, J. Tivander, 2019c. Life cycle assessment of city buses powered by electricity, hydrogenated vegetable oil or diesel. *Transportation Research Part D: Transport and Environment* 75, 211–222.
URL <https://doi.org/10.1016/j.trd.2019.08.019>
- A. Nordelöf, A.-M. Tillman, 2018. A scalable life cycle inventory of an electrical automotive traction machine—Part II: manufacturing processes. *The International Journal of Life Cycle Assessment* 23, 295–313.
URL <https://doi.org/10.1007/s11367-017-1309-8>
- M. E. Schlesinger, M. J. King, K. C. Sole, W. G. Davenport, 2011. *Extractive Metallurgy of Copper*. Elsevier.
- S. Seetharaman, 2014. *Treatise on Process Metallurgy, Volume 3: Industrial Processes*. Elsevier.
- Servicio de Evaluación Ambiental, 2015. *Sistema de Evaluación de Impacto Ambiental*.
URL <https://www.sea.gob.cl>
- C. Tagliaferri, S. Evangelisti, F. Acconcia, T. Domenech, P. Ekins, D. Barletta, P. Lettieri, 2016. Life cycle assessment of future electric and hybrid vehicles: A cradle-to-grave systems engineering approach. *Chemical Engineering Research and Design* 112, 298–309.
URL <https://doi.org/10.1016/j.cherd.2016.07.003>
- US Environmental Protection Agency, 2018. *Emission Factors for Greenhouse Gas Inventories*.
URL <https://www.epa.gov>
- M. Xylia, S. Leduc, A.-B. Laurent, P. Patrizio, Y. van der Meer, F. Kraxner, S. Silveira, 2019. Impact of bus electrification on carbon emissions: The case of Stockholm. *Journal of Cleaner Production* 209, 74–87.
URL <https://doi.org/10.1016/j.jclepro.2018.10.085>

Sustainable Ammonia Production via Electrolysis and Haber-Bosch Process

Gbemisola Ojo,^aKyle Camarda^a

^a*University of Kansas, Lawrence, KS 66047, United State*

Abstract

The production of ammonia accounts for nearly 2% of global carbon emissions (Palys et al. 2018) and therefore finding ways to make ammonia using renewable energy sources is paramount. The current mechanism for NH₃ production is dominated by the Haber-Bosch process, in which the hydrogen is produced from a water-gas shift reaction (Martín, 2016). Rather than obtaining this hydrogen from fossil fuels, this project employs electrolysis, driven by a solar plant coupled with battery storage. In this work, a dynamic simulation of a chemical process plant for the synthesis of the production of ammonia electrolysis is simulated using Aspen Hysys. Since renewable sources of energy are naturally variable, the plant's dynamic performance over varying electricity availability is also investigated. Case studies are used to demonstrate the feasibility of the plant as well as to explore the maximum and minimum battery capacities required.

Results show production profiles utilizing variable energy availability while avoiding shutdowns and minimizing battery charging cycles. The results also demonstrate the feasibility of completely sustainable ammonia production. This study produced 60.25 tons/day of ammonia using 8 tons/day of hydrogen at a current density range of 43 – 90 A/cm². The minimum battery capacity required is 56.25 MW.

Keywords: Electrolysis, Ammonia, Battery, Renewable, Solar Energy.

1. Introduction

Ammonia (NH₃) is the second most produced chemical in the world. Traditionally, hydrogen is produced from methane steam reforming, while nitrogen is separated from the air. This project utilizes Aspen Hysys to simulate a dynamic model of an ammonia production plant that utilizes sustainable hydrogen production. Nitrogen is obtained from an air separation unit using a pressure swing adsorption technique. Hydrogen is produced via water electrolysis at moderate pressure and temperature conditions to ensure overall energy usage is reduced and high purity of hydrogen is obtained.

While renewable energy is crucial to handling global warming, a major challenge is intermittency. In this work, rechargeable lithium-ion batteries are implemented to allow continuous production of ammonia. The specific objectives of this study are (1) to simulate a dynamic model for hydrogen production, nitrogen separation, and ammonia synthesis (2) to integrate the process with a solar energy and battery source, (3) to test and analyze the effects of varying process parameters (4) to perform an assessment of the system to determine battery charging and discharging schedules.

2. Background

2.1. Hydrogen Production

Several production methodologies for hydrogen production can be considered in the context of ammonia production, including Alkaline Water Electrolysis (AWE), Proton Exchange Membrane Electrolysis (PEM), and Solid-Oxide Water Electrolysis (SOWE). These methods are differentiated by two major parameters: pH, and temperature. The criteria used to decide the most effective technique is based on the temperature range, the electrolytes and durability of the system. AWE was chosen because of its low capital cost, low corrosivity, high durability, and thermodynamic stability. Based on these factors, the AWE process was chosen for this project. The simulation for the proposed AWE was completed using Aspen Custom modeler.

2.2. Nitrogen Separation

A renewable-resourced ammonia production facility should obtain the required nitrogen directly from the air. Three methods of air separation were considered for this project: cryogenic distillation, pressure swing adsorption, and membrane separation. Although cryogenic distillation is the most common process employed in industry, it requires large energy input. While membranes offer a low-energy solution for air separation, the current production of membranes does not provide sufficient selectivity for practical use. PSA, on the other hand, provides a trade-off between practicality and energy consumption. Furthermore, high-quality adsorbents for oxygen already are in industrial practice. PSA was not modelled in this work, it is recommended for future work.

3. Process Description

A dynamic process is simulated which integrates a Haber-Bosch reactor with water electrolysis for hydrogen production. It utilizes a solar plant as the main source of energy with lithium-ion batteries to buffer the power availability. The process model was developed using Aspen Hysys and Aspen Custom Modeler. An overall schematic of this process is shown in Figure 1.

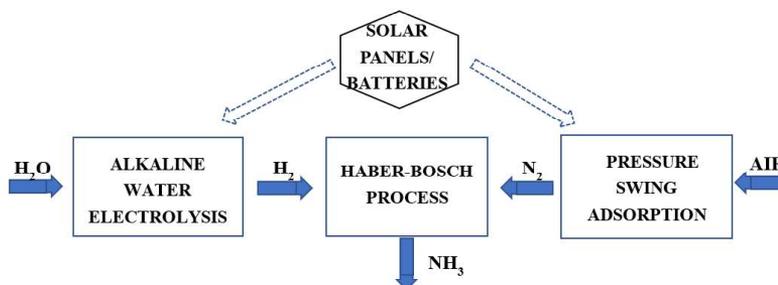


Figure 1: Overall schematic for ammonia production

3.1. Hydrogen Production Section

Alkaline water electrolysis is a process by which water is decomposed into hydrogen and oxygen in an electrolytic cell. Water is fed to the cell, which includes a porous nickel electrode and a diaphragm. Figure 2 shows the anode circulation loop,

which produces oxygen while the cathode circulation loop produces hydrogen for the Haber-Bosch process. This affects a direct separation between the hydrogen and oxygen produced. The AWE was simulated using a custom unit operation created in this work.

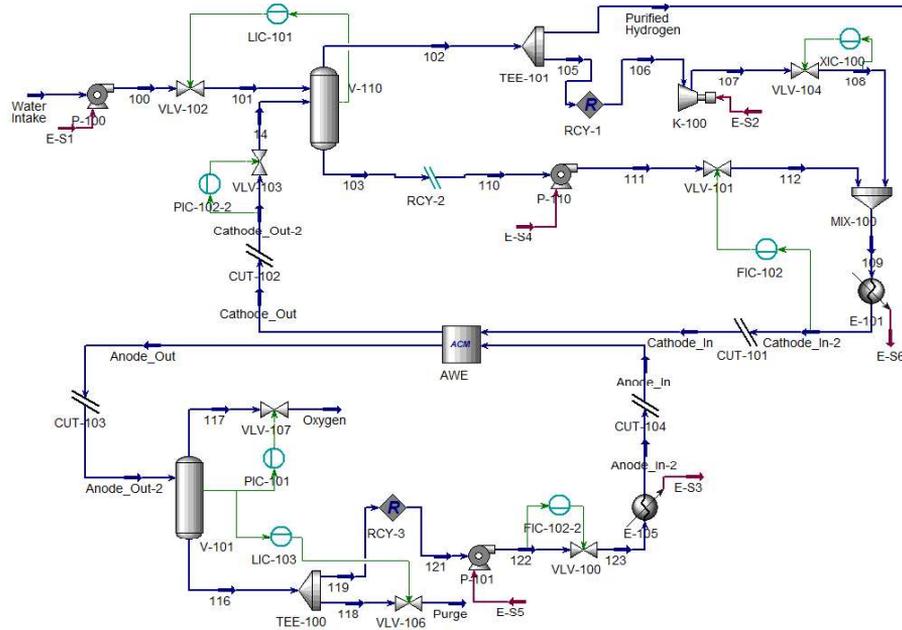


Figure 2: Alkaline Water Electrolysis showing anode and cathode circulation

3.2. Ammonia Production System

As discussed earlier, the Haber Bosch process is operated in a plug flow reactor temperature of 580 K and 150 bar. Heat integration is applied to recover excess energy from the product. Unreacted nitrogen and hydrogen are recycled to reduce energy usage.

3.3. Energy Specification

Upon the implementation of the above process model, a dynamic simulation was completed for three different case studies to demonstrate the energy requirements of the system. The objective of these case studies is to show the dynamic profile of the plant given various weather and location scenarios. The case studies are also used to understand the battery capacity required for the plant.

The storage energy system is an essential part of most off-grid renewable energy systems. Batteries are employed to store the surplus power produced from solar power allowing them to serve the load demand while balancing the hybrid system's fluctuations. The model used in this paper estimates the state of charge of the battery storage following the methods by Guezgouz et al. (2019).

4. Results and Discussions

The case studies shown describe the relationship between power availability and ammonia production level. Knowing that water electrolysis dominates the power

consumption operation of this plant, it is important to understand the power requirement of the alkaline water electrolysis based on the overall ammonia produced. The study uses a minimum and maximum current density specification to control the amount of hydrogen produced. This controls the overall ammonia production rate. The maximum and minimum current density is identified by using the specific requirements of AWE, at a limiting temperature of 90° C. Current density is directly properly proportional to temperature of the electrolytic cell. At current density ranging from 0.2 A/cm² to 0.3 A/cm², temperature ranges from 66 ° C – 90 ° C respectively.

The dynamic behavior of the integrated process was modeled using Aspen Hysys. In the system, controllers are placed in strategic places to stabilize the open loop response of the system as the current density of the process is changed. The controllers in the AWE will be tuned electrolysis to decrease the settling time and overshoot by tuning the control parameters. Dynamic studies will be continued to determine optimal operating conditions and charging/discharging schedules

For each case, the electrolyzer inlet temperature is set to 25 °C, while the inlet temperature of the for the Haber-Bosch PFR is set to 257 °C. This gives a 92% conversion rate of hydrogen to ammonia.

Case study one shows a startup for this process leading to full-capacity operation. The plant then runs at a constant density of 0.3 A/cm². Figure 3 describes the dynamic profile of the plant. It can be observed that at a constant density of 0.3 A/cm², the maximum amount of ammonia produced is 27.45 tons/day, which corresponds to a small industrial facility. The energy requirement is estimated to be 5.54e +04 kW per day. This case study serves as a base case to compare with real-word scenario.

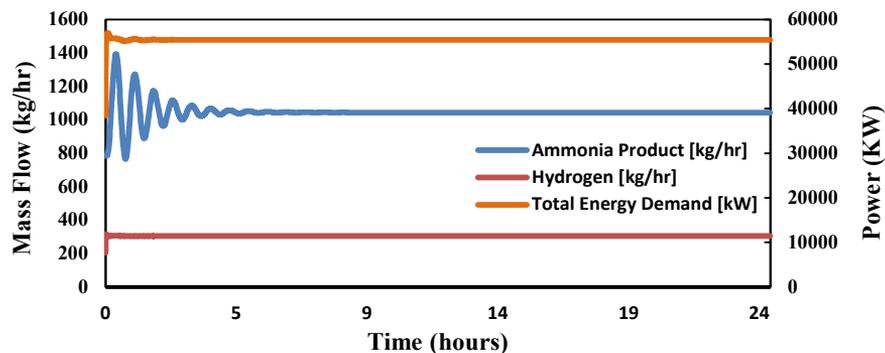


Figure 3: Ammonia production profile for case study one

Case study two assumes construction at a location closer to the equator, where there is approximately 12 hours of daylight. The current density at high solar hours is assumed to be 0.3 A/cm² which gives an ammonia production rate of 27.84 tons/day, as seen in the previous case study. This production rate decreases to 19 tons/day at a lower current density of 0.2 A/cm² when power is obtained from batteries. The excess energy produced from the solar panel at peak hours is transferred into batteries for use during the night. A total of 600 solar modules are utilized for functionality of the plant. Figure 4 shows the response to the change in current density of the overall ammonia production.

The dynamic profile also illustrates a second order response to the change in current density.

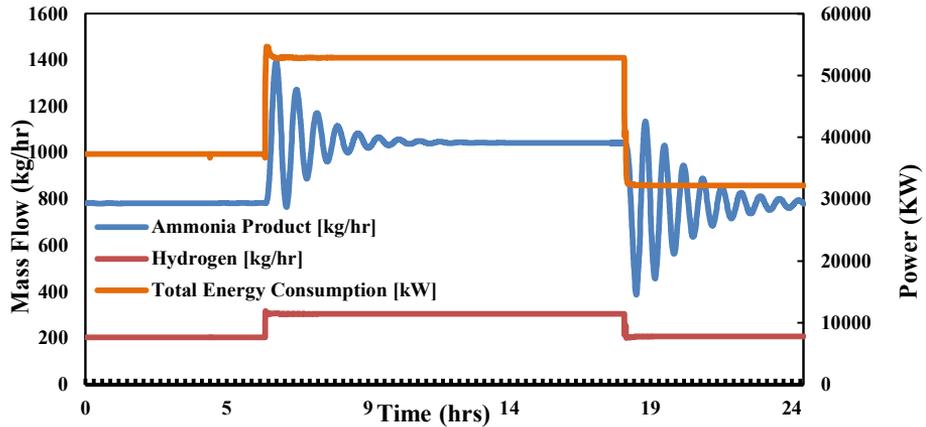


Figure 4: Ammonia production profile for case study two

Case study three involves a location further away from the equator, where there is lower amount of sunlight. The plant runs at 0.2 A/cm^2 for 16 hours at night while it runs at a higher current density 0.3 A/cm^2 for 8 hours. From the result shown in Figure 5, it can be inferred that there is an 60% increase in battery required to make the same amount of ammonia. excess energy supplied during the data is stored because it is needed to power the plant at night.

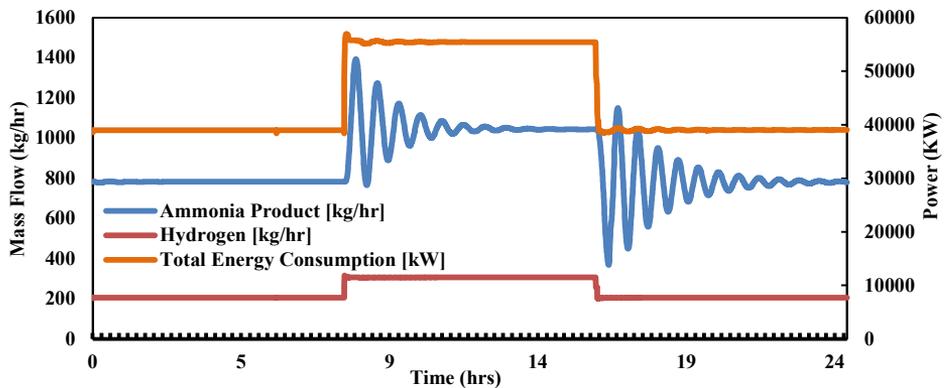


Figure 5: Ammonia production profile for case study two

5. Conclusions and Future Work

In this project, the feasibility of a 100% renewable energy chemical production plant to produce ammonia was considered. Due to the intermittency of renewable energy sources, the dynamic production profile was considered, to size the energy storage requirements. Water alkaline electrolysis and pressure swing absorption were applied to produce hydrogen and separate nitrogen respectively. The resulting profiles show the feasibility of a full-scale ammonia plant utilizing 100% renewable energy by integrating the entire Haber Bosch process with solar power. Analyzing these case studies at different current densities and time intervals gave insights on the effects of solar radiation and battery capacity on the ammonia production rate.

The battery system employed is large but not beyond currently used industrially sized systems. The key conclusion explained from the analysis shows that, the scale of the ammonia plant is necessary for the renewable energy to be economical. Several small-scale renewable plant is more cost effective than one large renewable plant. The results also show that case study 2 is more economical for the ammonia plant because it is much closer to the equator, therefore more energy can be harvested and stored in the battery for night-time. Lastly, the results also show a promising future holds for a renewable ammonia plant although this focuses on a smaller production scale. Future studies will perform MINLP optimization to minimize battery requirements and thus improve overall sustainability.

References

- Akinlabi, Charles O., Dimitrios I. Gerogiorgis, Michael C. Georgiadis, and Efstratios N. Pistikopoulos. 2007. 'Modelling, design and optimisation of a hybrid PSA-membrane gas separation process.' in Valentin Pleşu and Paul Şerban Agachi (eds.), *Computer Aided Chemical Engineering* (Elsevier, Amsterdam).
- Argyle, Morris D., and Calvin H. Bartholomew. 2015. 'Heterogeneous Catalyst Deactivation and Regeneration: A Review', *Catalysts*, 5: 145-269.
- Bard, Allen J, and Larry R Faulkner. 2002. 'Electrochemical Methods: Fundamentals and Applications'.
- Ciobanu, Madalina, Jeremy P. Wilburn, Morgan L. Krim, and David E. Cliffel. 2007. '1 - Fundamentals.' in Cynthia G. Zoski (ed.), *Handbook of Electrochemistry* (Elsevier: Amsterdam).
- Hao, Pingjiao, J. G. Wijmans, Jay Kniep, and Richard W. Baker. 2014. 'Gas/gas membrane contactors – An emerging membrane unit operation', *Journal of Membrane Science*, 462: 131-38.
- Himma, Nurul F., Anita K. Wardani, Nicholas Prasetya, Putu T.P. Aryanti, and I Gede Wenten. 2019. 'Recent progress and challenges in membrane-based O₂/N₂ separation', *Reviews in Chemical Engineering*, 35: 591-625.
- Martín, Mariano Martín. 2016. *Industrial chemical process analysis and design* Elsevier, Amsterdam.
- Murali, R. Surya, T. Sankarshana, and S. Sridhar. 2013. 'Air Separation by Polymer-based Membrane Technology', *Separation & Purification Reviews*, 42: 130-86.
- M. Guezgouz, J. Jurasz, B. Bekkouche, T. Ma, M.S. Javed, A. Kies Optimal hybrid pumped hydro-battery storage scheme for off-grid renewable energy systems112046 *Energy Conservation Management*, 199 (2019)
- Palys, Matthew J., Alon McCormick, E. L. Cussler, and Prodromos Daoutidis. 2018. 'Modeling and Optimal Design of Absorbent Enhanced Ammonia Synthesis', *Processes*, 6: 91.
- Wu, Yiqian, Yanlei Xiang, Lei Cai, Haitian Liu, and Ying Liang. 2020. 'Optimization of a novel cryogenic air separation process based on cold energy recovery of LNG with exergoeconomic analysis', *Journal of Cleaner Production*, 275: 123027.

Optimization for sustainable hydrogen production path

Patience B. Shamaki,^a G.A.C Le Roux^a

^a*Department of Chemical Engineering, Politechnic School, University of São Paulo, Av. Prof. Luciano Gualberto, trav.3,n380, São Paulo,Brazil*
galaroux@usp.br

Abstract

Hydrogen, a clean multipurpose energy carrier with application in various industrial and chemical processes has been identified as a key player in meeting the global clean energy sources target. There are several ways to produce hydrogen, however, 95% commercial hydrogen is produced via steam methane reforming (SMR) which emits CO₂ as by-product making it environmentally unfriendly. Although carbon capture sequestration (CCS) can be applied, only about 90% CO₂ capture has been reported in the literature, and the burgeoning water electrolysis technology which enables zero GHG emission depending on the source of energy applied is quite expensive and limited. Hence the increasing research attention on Methane pyrolysis (MP). In this study, the different TDM technologies are explored to develop an optimization model that considers energy efficiency, CO₂ emission and water consumption in production of hydrogen at a minimum cost, selecting a sustainable MP path that is comparable to prevalent technologies– SMR with and without CCS as well as water electrolysis- for further research. An MINLP optimization model for selecting a sustainable pathway to produce hydrogen in commercial quantity is formulated using associated data from the literature with respect to the different technologies considered. Sensitivity analysis is also performed to analyze the impact of the by-products produced from each technology on the cost of production. The model is implemented in GAMS. The findings from this work selected TDM with Plasma and TDM with metal catalyst as the optimal hydrogen production pathways that can compete with prevailing technology for sustainable hydrogen production based on the process performance index considered. It also shows that the quality of carbon could greatly influence the commercialization of MP. Moreover, MP is flexible and can attain carbon-neutrality depending on the feedstock and energy source. However, the results show that this transition depends on the cost and availability of renewable energy sources. With the supposition that MP Low to zero CO₂ emission and valuable solid carbon product features makes it sustainable, TDM technology should be explored to compete with the current technology to meet hydrogen demand, and this study provides a pointer to the optimal TDM technology as well as precedes a process simulation that explores the kinetics and thermodynamics for optimization.

Keywords: Methane pyrolysis, Hydrogen production, Optimization, Carbon.

1. Introduction

There is a global burden to find sustainable ways to reduce greenhouse gases (GHG) emissions while meeting increasing energy demand. Hydrogen, a multipurpose energy carrier with application in various industrial and chemical processes has been identified as one solution to de-carbonization, given that it does not directly emit CO₂. However,

the current techniques for commercial hydrogen production to meet global demand is heavily dependent on fossil fuels (mainly Natural Gas and Coal) either directly as feedstock and fuel for the process or indirectly for electricity generation (Muradov 2017). Utilizing natural gas which is composed mainly of methane will considerably reduce CO₂ emission in comparison to its fossil fuel counterparts because it is more environmentally friendly.

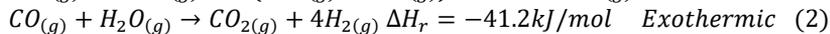
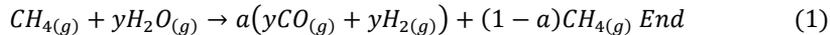
Methane pyrolysis (MP) is the thermal decomposition of methane (TDM) to produce hydrogen and solid carbon. It is an endothermic reaction like the dominant SMR but it has a smaller heat of reaction as well as only one reaction step, making it much simpler than the SMR that requires the water gas shift reaction. There are several techno-economic analysis and reports on the relevance of MP towards the hydrogen economy, as well as process simulation and experiments on the different MP technologies available to push its competitive potentials. However, to the best of the authors' knowledge as of the time of this work there are quite few research works, if any, that applied modelling and optimization to decisively select an optimal pathway.

In this work, we adopt and formulate an optimization model that selects the sustainable MP technology that can be comparable to prevalent technologies— SMR with and without CCS as well as water electrolysis for commercial hydrogen production considering energy efficiency, CO₂ emission and water consumption in production of hydrogen at a minimum cost. We analyze different scenarios of energy source combination as well as the influence of the price of valuable carbon and CO₂ emission tax/credit. The work is structured as follows: section 2 describes the different hydrogen production technologies currently used or researched. Section 3 describes the process modelling. Section 4 describes the model and optimization implementation as well as discussion of results obtained. Section 5 concludes with the finding and further works considered.

2. Hydrogen Production Methodologies

There are several technologies for producing hydrogen, the hydrogen produced from these technologies are mostly identified in color codes - based on dependence of the production process on fossil fuel or clean energy sources. These mainstream technologies include:

- Grey Hydrogen (SMR, Partial Oxidation, Auto-thermal reforming, Steam-oxygen gasification of coal, flue gases) (Muradov 2017; Keipi et al. 2016; Nikolaidis and Poullikkas 2017)



- Blue hydrogen (SMR with CCS)
- Green hydrogen (Electrolysis) (Muradov 2017).



- Turquoise Hydrogen (Methane Pyrolysis, TDM) (Keipi et al. 2016).



3. Process Modeling

To select the optimal Hydrogen production route, we employ synthesis-design processing route modelling approach proposed by Bertran et al. (2017). A superstructure based optimization problem is formulated to choose among processing alternatives. The Methodology is a three-step approach design for achieving an optimal production pathway. The three stages involve synthesis (superstructure optimization process synthesis), Design (detailed design and simulation), and Innovation (optimization,

modification or improved and sustainable process integration design) with output of each stage as the input of the next (Quaglia et al. 2012). The complicated and data intensive process is decomposed into these stages to simplify the problems and finally integrate the solution obtained from each stage to form a sustainable solution. For this work, we implement the first step to choose the sustainable hydrogen processing pathway, this is a precursor to the subsequent 2 steps which would explore in detail in future works.

In modelling, a CO_2 emission target is included to assess the total CO_2 emitted from the process and energy sources of each technology, this constraint is essential for the technology selection. Figure 1 shows a representation of individual process and its components to be modelled which is repeated for the different technologies considered. Where $f_{i,k}^{IN}$, $f_{p,k}^{out}$, $f_{p,k}^{out,1}$, $f_{p,k}^{out,1}$ represents the input and output (product and by-product) variables respectively, $E_{j,k}$ are the energy consumed with respect to the source for each technology, $\text{CO}_2^{emit}_{j,k}$ accounts for the CO_2 emitted from each energy source for each technology, given that there is no direct CO_2 emission from MP and electrolysis, and $SW_{i,k}$ accounts for process waste or unconverted feedstock.

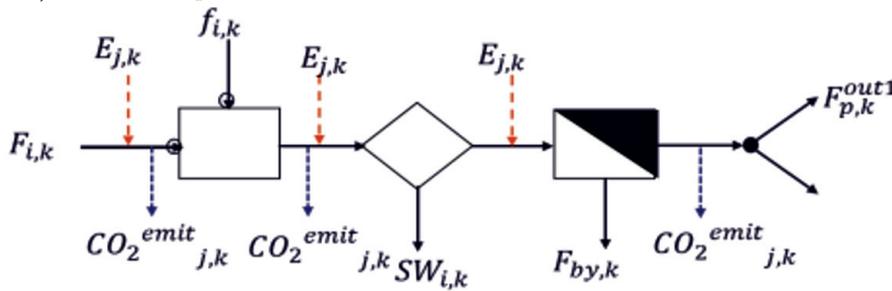


Figure 1: stages of hydrogen production process

4. Simulation, Results and Discussion

The synthesis-design processing route model as adopted from by Bertran et al. (2017) was implemented in GAMS using data from the literature on the current and competing hydrogen technologies. The objective is to select hydrogen production technology with the minimum cost given the constraints and considering the scenarios provided. The decision variables \mathbf{y}_k for the choice of technology, \mathbf{y}_a and \mathbf{y}_b for the emissions constraint. We used disjunctive programming for the modelling in GAMS and the optimization problem was solved using the BARON solver. For this work five (5) MP technologies were considered and compared to the matured technologies – SMR, SMR_CCS and Alkaline Electrolysis. The technologies considered are represented as: T1-conventional gas reactor, T2- plasma, T3-Heat exchanger reactor, T4-catalytic decomposition with metal based catalyst, T5- catalytic decomposition with carbon catalyst, T6-SMR, T7-SMR_CCS and T8-Alkaline Electrolysis. The data used for the simulation was obtained from Agency (2019); Keipi et al. (2016); Keitz (2021); Timmerberg, Kaltschmitt, and Finkbeiner (2020); Muradov 2017; Keipi, Tolvanen, and Kontinen (2018)).

4.1. Simulation

The simulation analysis was based on three (3) scenarios:

- Impact of energy sources applied: this scenario considers three sources of energy required for heating namely: Combusted natural gas (NG), electricity from the grid (EG) and renewable energy (RN) sources. A simulation for the application of

different combinations of these sources for heat as well as a single source for all technologies was carried out and the results analyzed discussed

- Impact of process by-products: considers the effect of the price of carbon on the choice MP technology to compete with prevalent technologies. Likewise, the analysis considers the influence of total CO₂ emission of each technology in the choice of technology; the carbon credit, carbon tax and carbon target were changed respectively and the effects analyzed;
- Impact of feed to product conversion ratio: this scenario considers the effect of the conversion rate of the process on the choice of technology.

4.2. Results

Table 1 shows the key variables obtained from the simulation as well as significant results obtained, the third scenario was not presented as it did not show a significant difference. The CO₂ emission constraint included in the process model is essential in the choice of technology, as it intends to meet the global emissions objective. It provides a solution space within which a decision among multiple options is made, as well as revealing parameters. It also provides a desired range with a condition for alternative decision when there is a violation, which in the case of this work influences the path selection.

The result in table 1 shows that the constraints were considered in the choice of technology for each of the scenarios assumed. It can be seen that the decision for the feed flowrate is such that it can obtain a product quantity that satisfies the CO₂ emission target, however in the case where the target is exceeded and yet a technology is selected, it points to the influence of the parameters, noting that the optimizer only tries to find the minimum. A global optimum was obtained with the given scenarios. For instance in the cases where T7- SMR with CCS is selected, the CO₂ target could be exceeded but the emission tax could easily be compensated with the cost of technology or energy cost compared to choosing electrolysis with almost zero emission but with high technology cost.

Table 1: Simulation analysis result

Variables	$f_{i,k}^{IN}$ (kg)	$f_{p,k}^{out}$ (kg)		y_k	Z	$CO_{2,k}^{em}$
	CH ₄ /H ₂ O	H	C/CO ₂			
Base_1	36.422	3.788	10.781	T4	783.9	100
Scenario 1	21.381	1.871	5.613	T2	750.672	100
	24.510	1.838	5.515	T3	736.861	100
	36.944/ 5.000	2.112	16.762	T7	706.346	100
	24.510	1.838	5.515	T3	736.861	100
Scenario 2	21.381	1.871	5.613	T2	1555.9	100
	30.941	3.094	9.282	T5	1542.7	100
	21.381	1.871	5.613	T2	1550.3	100
	242.469	11.361	100.000	T7	70945.5	100
Emission target	72.844	7.576	21.562	T4	3149.2	100
	26.489	3.788	10.781	T4	783.7	100

Scenario 1: Energy consumption

The simulation results shows, that the combination of NG and EL provided the least objective function selecting T4 i.e. the thermo-catalytic decomposition using metal catalyst using electricity from the Grid. However, the CO₂ emission was above the given emission target. This can be because of the separation requirement for the catalyst and the carbon product, given that using catalyst lowers the energy consumption thereby reducing possible emission associated with this technology. Furthermore, in the case where only electricity is applied across all the MP technologies, it is still able to select one of the MP technologies over the SMR with CCS technology. On the other hand, using the other sources of energy independently for all the technologies leads the optimizer to select the SMR process which is cheaper and has the capacity to offset for the CO₂ emission either through enough profit to pay the given tax or with the CCS technology to avoid the tax.

Scenario 2: Effects of Technology by-products

From Table 2, we observe that the byproducts of each technology significantly influence the final cost of hydrogen production. In the MP technologies there is no CO₂ produced from the process, the only emission is from the energy source giving it an advantage over the matured SMR,- given that even with CCS only about 60% have been reported practically and 90% theoretically- hence the higher chance of obtaining carbon credits than tax. Moreover, the different MP technologies produce different grades of carbon, which have different prices based on their qualities. Therefore, the price of these different types of carbon can influence the cost of production. This can be observed from the simulation result with choice T2 and T4 which have the highest prices of carbons (graphene and carbon nanotubes CNT).

The results also show that a more strict carbon emission tax leads to the selection of MP over SMR, although the increase in carbon credit had no significant effect on the selecting SMR with CCS over the MP technologies.

By using the 3-stage synthesis-design processing route approach, it gives a more structured and informed process design and optimization. In addition, the results obtained demonstrates the advantages of MP such as:

- Low CO₂ emission,
- Valuable carbon by-product
- No water required
- Ability to apply different sources of heat at different stages of the process for optimization purposes.

These attributes significantly influence the choice of the hydrogen technology.

5. Conclusion

A synthesis-design processing route approach for superstructure optimization was implemented to select the optimal hydrogen processing technology, this technique is the first of three steps which serves as a precursor to a more detailed process design, simulation and optimizations. The result of the optimization provided a direction as to the parameters that can significantly affect the optimal hydrogen production process route as well open more questions for the author to decipher in the course of the research. The result of the simulation demonstrates the advantages of Methane pyrolysis over the prevailing technologies by showing that the use of electricity from the grid as the main source of heat can lead to a more optimal MP process that can compete with the current SMR and SMR with CCS technology. The result also show that while MP has the potential to be emission free when renewable energy sources are applied, it is currently not optimal due to the maturity stage and cost of renewable energy sources.

The results obtained also showed that the quality and price of carbon product can significantly make MP a competing hydrogen production technology. Furthermore, MP technologies can enable hydrogen to have a better chance in the world energy mix with more countries adopting more strict Carbon emission regulations as the world move toward a low-zero emission. With the result of this work a detailed process simulation will be carried out with the best technology selected which in this case is between T2 (TDM with Plasma) and T4 TCDM (with metal catalyst).

The results also open room for more question for research exploration such as: The thermodynamics of the process that can lead to optimal quality carbon production, the properties of carbon produced and their modification for different applications to enable more demand, energy mix in the different process steps that can enable optimal processing. What are the main distinctions between T2 and T4, and how can they be improved to compete with the SMR technology while more research on the possibility of making MP technologies emission free is ongoing.

Acknowledgement

The authors would like to express their acknowledgement to CAPES (001), CNPq (312049/2018-8 fellowship) and RCGI/FAPESP (2020/15230-5) for their financial support.

References

- Z. Allen, Year, Article or Chapter Title, Journal or Book Title, Volume, Issue, Pages
Agency International Energy (IEA),2019, “The Future of Hydrogen”, G20 Japan
- Bertran, Maria Ona, Rebecca Frauzem, Ana Sofia Sanchez-Arcilla, Lei Zhang, John M. Woodley and Rafiqul Gani, 2017, A Generic Methodology for Processing Route Synthesis and Design Based on Superstructure Optimization, *Computer and Chemical Engineering*, 106:892-910.
- Keipi, Tiina, Ville Hankalin, Jaakko Nummelin, and Risto Raiko. 2016. “Techno-Economic Analysis of Four Concepts for Thermal Decomposition of Methane: Reduction of CO₂ Emissions in Natural Gas Combustion.” *Energy Conversion and Management* 110: 1–12.
- Keipi, Tiina, Henrik Tolvanen, and Jukka Konttinen. 2018. “Economic Analysis of Hydrogen Production by Methane Thermal Decomposition: Comparison to Competing Technologies.” *Energy Conversion and Management* 159 (December 2017): 264–73.
- Keitz, Marc von. 2021. “Methane Pyrolysis for Hydrogen– Opportunities and Challenges.
- Muradov, Nazim. 2017. “Low to Near-Zero CO₂ Production of Hydrogen from Fossil Fuels: Status and Perspectives.” *International Journal of Hydrogen Energy* 42 (20): 14058–88.
- Parkinson, Brett, Mojgan Tabatabaei, David C. Upham, Benjamin Ballinger, Chris Greig, Simon Smart, and Eric McFarland. 2018. “Hydrogen Production Using Methane: Techno-Economics of Decarbonizing Fuels and Chemicals.” *International Journal of Hydrogen Energy* 43 (5): 2540–55. <https://doi.org/10.1016/j.ijhyd>
- Quaglia, Alberto, Bent Sarup, Gürkan Sin, and Rafiqul Gani. 2012. “Integrated Business and Engineering Framework for Synthesis and Design of Enterprise-Wide Processing Networks.” *Computers and Chemical Engineering* 38: 213–23.
- Timmerberg, Sebastian, Martin Kaltschmitt, and Matthias Finkbeiner. 2020. “Hydrogen and Hydrogen-Derived Fuels through Methane Decomposition of Natural Gas – GHG Emissions and Costs.” *Energy Conversion and Management: X* 7 (April): 100043.

Development of lactic acid evaporation process model with multi effect evaporator and mechanical vapor recompression system

Sujin Cha,^{a,b} Jonghun Lim,^{a,c} Yurim Kim,^{a,b} Hyungtae Cho,^a Il Moon^c, Junghwan Kim,^{a,*}

^a*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Jonga-ro, Ulsan 44413, Korea*

^b*Department of Chemical and Biochemical Engineering, Dongguk University, 1, Pildong-ro, Seoul 04620, Korea*

^c*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul 03722, Korea*

Abstract

Lactic acid (LA) is a raw material for poly lactic acid, which is used in food, pharmaceutical, textile, and various chemical industries. Evaporation is an essential process to produce highly pure LA after the fermentation process. Notably, the evaporation process is energy intensive, and the energy consumption corresponds to approximately 50% of the total production cost. Consequently, the energy efficiency of the LA evaporation process must be enhanced. To this end, this study was aimed at establishing a process model based on a multi-effect evaporator (MEE) and mechanical vapor recompression (MVR) system. The MEE process model involved a distillation column, virtual reboiler, and flash drum. The steam consumption was reduced by recovering the latent heat of secondary vapor. Furthermore, the steam and condensate were recovered from the final evaporator. The MVR process model involved a compressor that substituted the steam consumed in the evaporation with electricity by recompressing the vapor in the evaporator. The energy efficiency of the two processes was analyzed by comparing the steam and electricity consumption values. The steam consumption decreased from 99.381 to 37.913 ton/h with the increase in the number of effects (from one to eight), and the electricity consumption of the MVR system was 31.32 kW.

Keywords: Evaporation process, Multi-effect evaporator, Mechanical vapor recompression, Process modeling

1. Introduction

Lactic acid (LA) is an eco-friendly material produced by fermenting sugar and lignocellulosic substrates. LA is a raw material for poly LA (PLA), which is a bioplastic that can replace petrochemical polymers. Consequently, the demand for LA has increased considerably in recent years. To produce the final LA product, several purification steps must be implemented after the fermentation process, with evaporation representing the most commonly applied approach for purification. However, the evaporation process is energy intensive, with the energy consumption corresponding to approximately 50% of the total production cost. To reduce costs, the energy efficiency of the evaporation process must be enhanced. Many researchers have investigated LA purification as a potential

approach to increase the energy efficiency. Oscar et al. recommended the use of a reactive distillation column for LA purification, which could achieve 99.89% recovery of LA. Tong et al. increased the yield of LA to 82.6% through purification with paper sludge as a cellulosic feedstock by using an anion exchanger, Amberlite IRA-92. Madzingaidzo et al. demonstrated the concentration of LA through monopolar electro dialysis and obtained an average yield of more than 98% after carbon treatment.

Despite the significant efforts to increase the energy efficiency of the evaporation process, several challenges remain to be solved, specifically, the high equipment cost, reduced stability owing to the addition of chemical substances, and low economic feasibility because continuous operation cannot be realized.

To address these process, in this study, two alternative process models were designed using a multi-effect evaporator (MEE) and mechanical vapor recompression (MVR) system. The MEE can reduce the steam consumption in the evaporation process by reusing the heat of the secondary steam discharged from the evaporators arranged in series under different conditions as a heat source. In the case of MVR, the steam consumption can be reduced by replacing a part of the steam consumed during the evaporation process with electricity. The energy efficiency of these processes was analyzed by comparing the steam and electricity consumption.

2. Process description

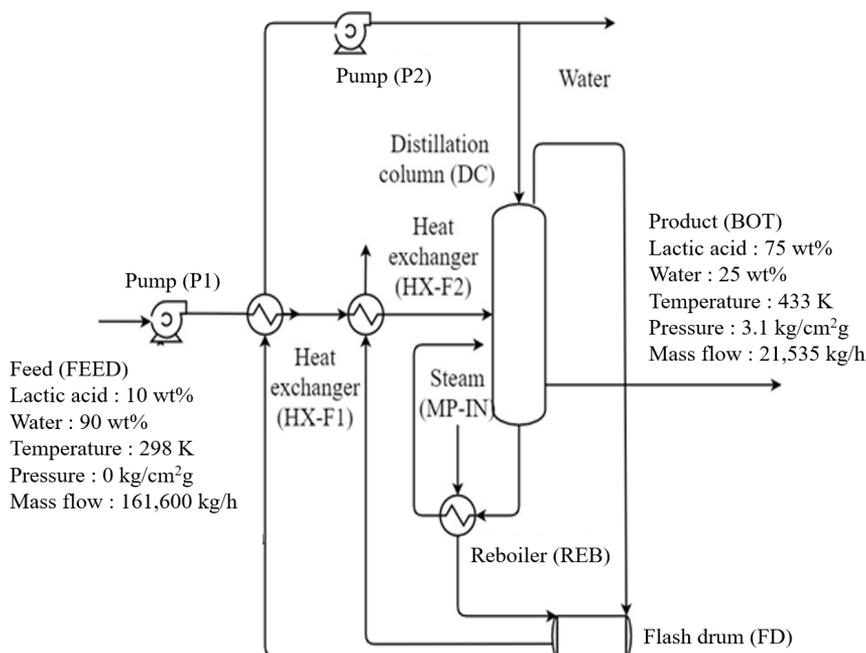


Figure 1. Process flow diagram of single-column evaporation

Among separation processes, distillation is the most effective method for evaporation in the biorefinery process. Distillation based on a single column is a conventional evaporation process for refining LA from a low concentration feed. Figure 1 shows the

process flow of single-column evaporation. After the feed is preheated through a preheater, it is supplied to the evaporator. After distillation, LA with a relatively high boiling point is discharged to the bottom of the column together with undistilled water, and water, the boiling point of which is lower than that of LA is discharged to the top of the column in the form of a gas.

3. Model development

Aspen plus V10.0 was used to simulate the process, and the considered thermodynamic equation was the UNIQUAC activity coefficient model. This equation can represent the vapor–liquid equilibrium behavior and is commonly used for simulating evaporation and condensation processes. The following assumptions were implemented.

- ✓ The feed composition is as follows: H_2O (90 wt%) and LA (10 wt%).
- ✓ The product composition is as follows: H_2O (25 wt%) and LA (75 wt%).
- ✓ Heat exchange occurs through the latent heat of steam input to the reboiler.
- ✓ The steam pressure is $9 \text{ kg/cm}^2 \cdot \text{g}$, and the temperature is $180 \text{ }^\circ\text{C}$.
- ✓ When the heat duties of the virtual reboiler and reboiler of the column are identical, the steam consumption is the optimal point.

3.1. Multiple-effect evaporation process

Multiple-effect evaporation is one of the oldest processes in desalination frameworks. In this study, the MEE process was used to increase the efficiency of the LA purification process. Figure 2 shows the process flow for a MEE.

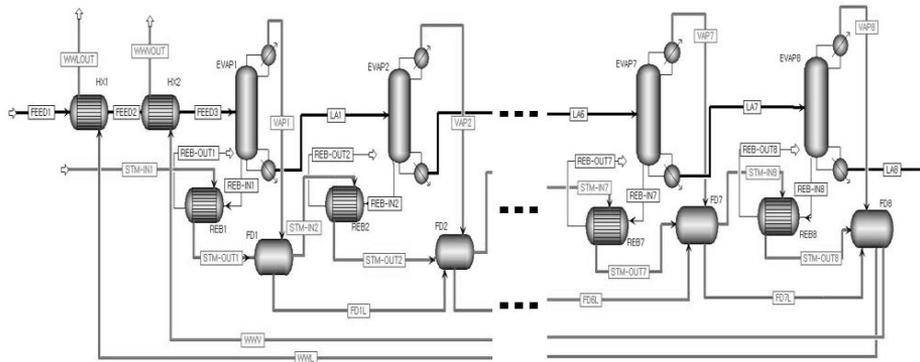


Figure 2. Schematic of multi-effect evaporator (MEE) process

The evaporator was divided into a distillation column, virtual reboiler, and flash drum in the process modeling. The virtual reboiler was used to predict the steam consumption of the distillation column. The heat calculated through the virtual reboiler and heat consumed in the distillation column were set equal to estimate the steam consumption. The impurities in the feed were ignored. The feed was preheated in the preheater and fed to the distillation column with increased pressure and temperature. The temperature and pressure of the feed and vapor continuously decreased from the first effect to the last

effect, and the least temperature and pressure corresponded to the last effect. Steam from the outside was used to evaporate the feed water in the first effect. The steam from the distillation column was used to evaporate the water/solvent in the second effect, and the steam from the second effect was used to evaporate the water/solvent in the third effect. This process was repeated until the last effect. After heat exchange at the last effect, the discharged water was separated through a flash drum and used to preheat the feed. By repeating this process until the last column, the waste heat could be effectively reused, thereby reducing the amount of steam used, which could help decrease the energy consumption. The distillation column involved 15 stages, the feed was supplied in stage 11, and the reflux stream at the top of the column was input in stage 1. The steam used as the heat source was medium-pressure steam.

3.2. MVR

MVR is a commonly used technique to recover waste heat sources as useful energy sources through mechanical compression. MVR uses only electricity as the energy source as it concentrates LA using only the secondary steam compressed through a centrifugal compressor. In general, the MVR can significantly reduce utility costs throughout the evaporation process because electricity is cheaper than steam. In the developed model, the heat of the compressed secondary vapor was set equal to the heat required by the evaporator to predict the power consumption of the centrifugal compressor.

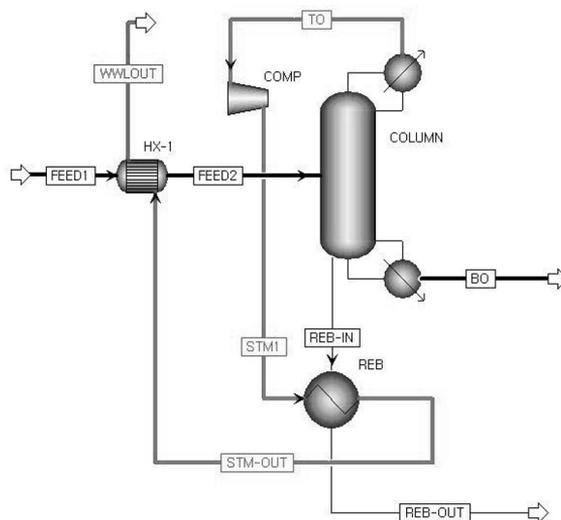


Figure 3. Schematic of mechanical vapor recompression (MVR)

Figure 3 shows the process flow for the MVR framework. In the process modeling, the evaporator was divided into a distillation column, virtual reboiler, and compressor. Furthermore, a virtual reboiler was added to predict the steam consumption of the distillation column. Feed entered the column through preheating, and the mass fraction of the product at the top and bottom of the column was the same as that for the MEE model.

4. Results

Figure 4 shows the steam consumption for each equipment in this process. The model with one effect corresponded to the conventional process. In the MEE model, the amount of steam consumption gradually decreased as the number of effects increased to eight. This phenomenon occurred because the waste heat of the steam at the top of the distillation column was reused. In particular, when eight effects were used, the steam consumption reduced from 161,656 kg/h to 37,913 kg/h. Moreover, the steam consumption for the MVR model (11,083 kg/h) was significantly smaller than that for the MEE model. This finding could be attributed to the use of electricity as the energy source in the MVR model.

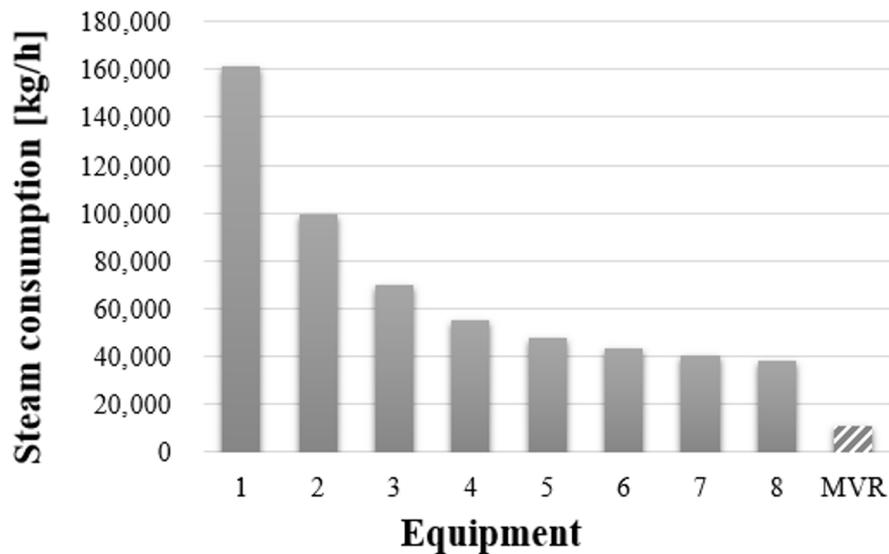


Figure 4. Steam consumption according to equipment, with the number of MEE effects ranging from 1–8

Figure 5 shows the electricity consumption for each equipment in this process. In the MEE model, the electricity consumption increased as the number of effects increased. Specifically, when eight effects were used, the electricity consumption increased from 52.88 kW to 78.899 kW. However, in the case of the MVR model, the electricity consumption reduced from 52.88 kW to 31.32 kW.

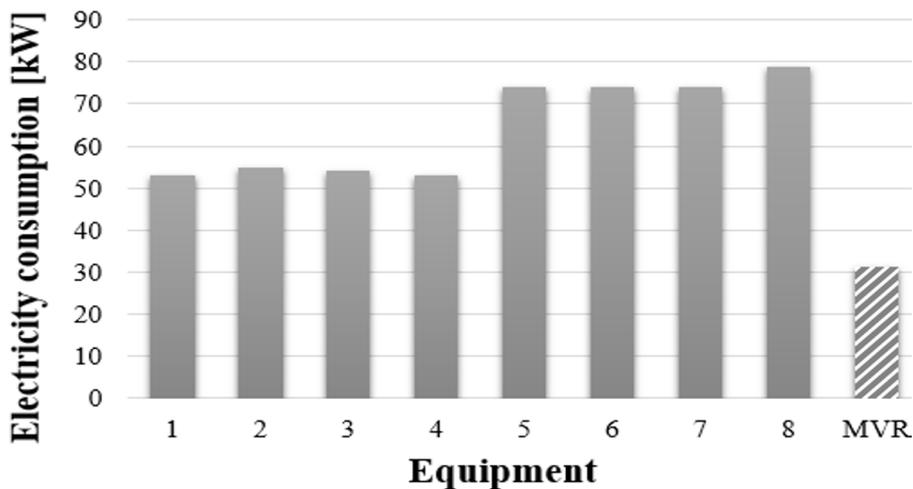


Figure 5. Electricity consumption according to equipment, with the number of MEE effects ranging from 1–8

5. Conclusions

Two process models, MEE and MVR, were established to enhance the energy efficiency of the LA evaporation process. The LA mass fractions for the feed and product were 10 wt% and 75 wt%, respectively, for both models. The energy consumption was compared by estimating the amount of steam and electricity consumed for heat exchange in each process. In the conventional process, 161,656 kg/h of steam was used. However, when eight effects were introduced, the steam consumption of the MEE model reduced by 76.55%. The MVR model, which used electricity as the main energy source, corresponded to a 93.14% smaller steam consumption. The electricity consumption of the MEE and MVR models was 49.14% higher and 40.77% lower than that of the baseline, respectively. In future work, the economic feasibility of the models can be compared by performing a techno-economic analysis considering the investment cost and operating cost.

References

- W. Y. Tong, X.Y. Fu, S.M. Lee, J. Yu, J.W. Liu, D.Z. Wei, Y.M. Koo, 2004, Purification of L(+)-lactic acid from fermentation broth with paper sludge as a cellulosic feedstock using weak anion exchanger Amberlite IRA-92, *Biochemical Engineering Journal*, 18, 89-96
- L. Madzingaidzo, H. Danner, R. Braun, 2002, Process development and optimization of lactic acid purification using electro dialysis, *Journal of Biotechnology*, 96, 223-239
- Oscar A, R. Gasca, J. Fontalvo, Fernando I, Eduardo S, R. Morales, 2020, Design and evaluation of intensified downstream technologies towards feasible lactic acid bioproduction, *Chemical Engineering and Processing-Process Intensification*, 158, 108174

Optimization-Based Framework for Robust Modeling and Design of Kinetic Systems

Eduardo Sánchez-Ramírez^a, Brenda Huerta-Rosas^a, Juan José Quiroz-Ramírez^b,
Víctor Alejandro Suárez-Toriello^b, Juan Gabriel Segovia-Hernández^a

^a*Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, Mexico,*
eduardo.sanchez@ugto.mx

^b*CIATEC, Omega No.201 Colonia Industrial Delta, Leon 37545, Mexico*

Abstract

To understand the mechanism of a catalytic system, computational analysis is essential. Once a potential reaction mechanism has been identified, it typically involves (a) evaluating the energetics for the species and transition states using a computational chemistry method such as density functional theory (DFT), (b) relating these microscopic properties to macroscopic quantities such as reaction time. Microkinetic models may not match experimental data due to (1) inaccuracies in the energetic estimation, (2) inadequate catalyst models that are not representative of the surface environment. As a result of using experimental data, a more accurate model may be developed. If the model-experiment mismatch is resolved, a microkinetic model solution compatible with DFT assumptions may be identified. This work presents a generic optimization framework for solving parameter estimation and catalyst design problems in catalysis. Using a stochastic optimization method, Differential Evolution with Tabu List in conjunction with Aspen Plus, and considering experimental data, various activation energy and kinetic constants values were predicted. A sequential approach is a traditional approach to solving parameter estimation problems. Issues with the stiffness of the microkinetic model and the optimizer's capacity to tackle such highly nonlinear systems are common challenges. This proposal has the potential to use all the reactor models present in Aspen Plus, as well as be able to use all the kinetic models and solvers, avoiding numerical difficulty in optimization solutions. This method has several advantages, including ease of implementation, which leads to physically realizable steady-state solutions, and a reduced overall optimization problem in terms of the number of variables involved. To validate this method, a dimerization of isobutane to produce isooctane is used as base line. The data considered as base case were previously worked considering experimental and simulation work. Once the parameter estimation was performed, the error produced was almost zero, and it was possible to generate the same kinetic data, concentration profile and molar flow produced.

Keywords: optimization, kinetic data prediction, oligomerization, hydrogenation, biojet fuel.

1. Introduction

In many techniques to understand a catalytic system and for ultimate design for enhanced catalysis (Nørskov et al. 2009), computational analysis is a vital tool. Incorporating experimental data, on the other hand, enables more appropriate model construction. This activity necessitates the estimate of parameters that will allow the model-experiment

connection to be reconciled and a kinetic solution model to be identified. Nonlinear optimization challenges include parameter estimates and catalyst design. In such cases, objectives must be maximized or minimized while being constrained by the catalytic system. A traditional approach to solving the parameter estimation problem is to use a sequential approach (Biegler 2010), in which the optimizer obtains necessary information about the reaction rate, surface coverage, and gradients at each step of the optimization process, looking for a simulation of the kinetic model. The stiffness of the kinetic model and the optimizer's capacity to deal with the system's nonlinearities make this technique problematic. The equations, for example, may discover many solutions that would be undetectable under normal physical conditions. As a result, a successful solution identification may be predicated on a solver-proposed starting assumption. Furthermore, when the kinetic system exhibits a temporal variation, the NLP formulation must be modified to account for the dynamics, and orthogonal collocation methods are required. Due to the simplicity of use, access, and rigorous resolution of the models connected with each piece of equipment, computer programs to duplicate the model of diverse unit activities are currently a highly important tool. The Aspen Plus simulator was created to represent chemical processes such as non-ideal phases, processes involving recycling currents, chemical reactions, adiabatic operations, and so on. It's made to handle a wide range of operating conditions in process design. Using experimental data to construct kinetic models might be a way to improve equilibrium-based models. Kinetic models vary from equilibrium-based models in that they may depict chemical reaction conversion, which is an important stage in process design, as well as the assessment and research of the findings obtained under various operating circumstances. Because of the intricacy with which kinetic models are treated, they are more precise and comprehensive than equilibrium-based models. As a result, there is a pressing need to build kinetic rate-based simulations capable of accurately forecasting the products of numerous processes that are difficult to describe, simulate, and operate under multiple conditions. In light of the foregoing, this research proposes a generic sequential optimization framework for solving kinetic parameter estimates for the reproduction of chemical processes in the Aspen Plus process simulator. The novelty of this proposed framework is that it allows for: a) a comprehensive exploration that allows for a robust explanation of experimental observations, b) comparing and constraining solutions through automated data analysis that allows for hypothesizing competing for dominant chemistries, c) proposing objectives that maximize the expected performance of kinetic systems, and d) identifying sparse solutions that avoid overfitting by selecting the fewest number of parameters that allow for a robust explanation of experimental observations. The article begins with a description of the methodology for determining the kinetic parameters, followed by a validation example.

2. Framework description: Methodology

The models solved by the framework in this study will be described in general in this part. It will begin with the kinetic model being solved in the simulator, the goal function being solved, and the parameter estimation solution approach.

2.2 Kinetic Model

Another reason to study reaction rates is to learn more about the reaction's mechanism and fundamental stages. A kinetic of any reaction represented by a rate equation should

be thoroughly understood by studying and investigating the characteristics of a rate equation, such as the order of the reaction, rate constant, and kinetic parameters such as A (pre-exponential factor) and E. (activation energy). The rate equation is used to quantitatively characterize the rate of a reaction and to explain its functional dependence on temperature and species concentrations. In symbolic form, the rate equation is as follows:

$$r_A = r_A(T, \text{all } C_i) \quad (1)$$

The term "all Ci" is used to emphasize that the concentrations of the reactant(s), product(s), and any other compounds present, even if they do not participate in the reaction, can affect the reaction rate. The partial pressure of a material pi is another sort of rate equation, especially for gases. The coefficient k in the reaction rate equation is the rate constant (as in Eq. (2)). This rate constant is changed by temperature, but not by the concentration of the species involved in the reaction. In the rate equation, the units of k are always used as a conversion factor for the product of species concentrations and units given as a change in concentration divided by time. The temperature-dependent term, or reaction rate constant, is clearly defined by Arrhenius' law:

$$k = k_0 * \exp(-E/RT) \quad (2)$$

The frequency or pre-exponential factor and activation energy, respectively, are the two parameters of Arrhenius' law. These parameters are also known as Arrhenius parameters. These figures are obtained by plotting lnk (rate constant) against 1/T (absolute temperature) and drawing a straight line. Experiments have shown that Arrhenius' law holds across a wide temperature range.

2.3 Objective function

The fundamental problem to address is to determine which kinetic parameters E and K best characterize the empirically observed behavior. In this case, the goal of the optimization issue is to decrease the observed error between experimental data and aspen plus simulator results when different values of E and K are represented. The following is a formula for the optimization/parameter estimation issue.

$$Z_{param,est} = \min \sum_{c \in C} (y_c - \bar{y}_c)^T W (y_c - \bar{y}_c) + \delta(\pi - \pi_{nom})^T (\pi - \pi_{nom}) \quad (3)$$

Subject to

$$\frac{dx_c}{dt} = f(x_c, k_c, u_c) \quad (4)$$

$$g(x_c, k_c, u_c) = 0 \quad (5)$$

$$y_c = Bx_c(t \rightarrow \infty); x_c(0) = x_c(u_c);$$

$$u_c = \{F_{in,c}, P_{tot,c}, T_c\} \quad (6)$$

$$h(y_c, u_c, k_c) \geq 0 \quad (7)$$

$$k_c = \psi(p, \pi_c, u_c) \quad (8)$$

For each experimental condition c C, the collection of all experimental circumstances has limitations that must be met. x, y, and u are referred to as x_c, y_c, and u_c, respectively since they are dependent on the experimental circumstances c. Because the objective function can be used in a variety of contexts, the subscript c and a summation have been included. All terms in equations 3–8 are depending on the experimental circumstances. The weighted least-squares error Z_{param,est} represents the model-experiment mismatch. y c is

used to represent the experimental data. The goal to be minimized is subject to the constraints relating to the kinetic model, as well as the restrictions, where p is the collection of constraints and is the set of decision variables, and h is the set of inequalities assuring thermodynamic feasibility and the choice variable bounds. The kinetic reaction of the system (function of ψ) can be described as an Arrhenius pre-exponential constant.

2.3 Optimization Strategy

This case study was solved using Differential Evolution with Tabu List, a metaheuristic optimization approach based on natural selection theory. This evolutionary approach employs the usual Differential Evolution stages. On the other hand, the Tabu List (TL) concept improves its actions. This hybrid optimization technique was created to improve DE's computing efficiency by trying tough phase equilibrium computations followed by parameter estimation challenges in dynamic systems with numerous minimums. A modified and expanded DETL approach was then applied to handle constrained multi-objective optimization problems (Sharma and Rangaiah 2013), and the resulting algorithm was named an integrated multi-objective differential evolution algorithm. Visual Basic is used to implement the hybrid algorithm. The numerical approach generates input vectors that are evaluated using dynamic data exchange (DDE) in the process model (Aspen Plus). After that, the model evaluates the input vector before generating output vectors. After that, the hybrid approach assesses the input and output data before iterating the process depending on the evolutionary nature of the algorithm. The stochastic method's operational settings were 120 individuals, 200 generations, 50% of the number of persons as tabu list, 0.6 for the crossover, and a taboo radius of 1 1006. These values were derived from a prior calculation's tweaking procedure. The optimization approach works in such a manner that it proposes E and K values for the kinetic model in each iteration. The reactor represented in aspen will generate an output as a result of this idea.

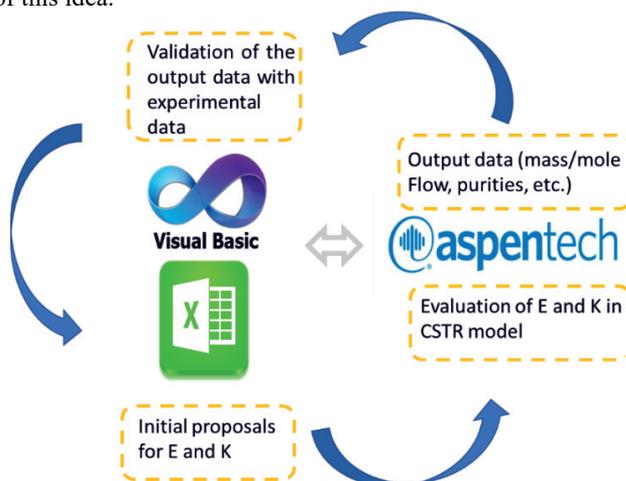


Figure 1. Determination of parameters E and K with DETL and Aspen Plus

3. Results: Isobutene dimerization.

In the 1930s, researchers looked at the dimerization of isobutylene caused by the dehydration of 2-methylpropan-2-ol. This reaction is frequently presented as an example

of alkene dimerization, and it is the subject of a laboratory experiment that allows students to check the predicted outcomes. The fact that two isomeric trimethylpentenes are created and evaluated as a mixture is a limitation of this experiment as a preparative experiment. Because of its high octane number, low vapor pressure, and lack of aromatics and sulfur, isooctane has become one of the most popular gasoline additions. Because of its high compression ratio and knock resistance, isooctane is also the principal component of aviation gasoline (Avgas). Due to the MTBE phase-out, isooctane synthesis technology allows refineries to use surplus isobutene and boost refinery earnings by converting light olefins into gasoline blends. In the not-too-distant future, isooctane use and production are predicted to skyrocket. To avoid catalyst sintering due to the highly exothermic reaction and the formation of higher oligomers that could result from the increased concentration of dimer in the reactor, most conventional isooctane production processes operate at low conversions (20-60 percent, with conversions varying depending on the catalyst and reactor design). Because of these limits, conversions of more than 60% each pass are frequently impossible and uncommon in the industry. Goortani et al. published a prior study in 2015. They propose utilizing Aspen Plus to mimic a typical commercial flowsheet that includes a dimerization reactor, distillation column, and hydrogenation reactor. They presented a concentration profile as well as kinetic data in their paper. The dimerization of isobutene to isooctane is the reaction system chosen for isooctane synthesis, as shown in the equation.



Previous experimental experiments in our laboratory (Goortani et al. 2015) employing a nickel sulfate and a Pd catalyst independently supported on Al₂O₃ were used to combine reaction kinetic data for the dimerization and hydrogenation processes in the isooctane synthesis process. These kinetic data were acquired in a 300 mL Parr autoclave reactor and pertained to liquid-phase oligomerization and hydrogenation of isobutene. The Arrhenius equation, $k = Ae^{(E/RT)}$, was used to determine the relationship between the reaction rate constant and the reaction temperature. The reaction constants k_1 for dimerization are listed below. The reaction of Dimerization:

$$r_{C_4H_8} = \frac{d[C_4H_8]}{dt} = -2k_1C_{C_4H_8} \quad (10)$$

Table 1 shows the k and E optimization limits considered for this work.

Table 1. Optimization limits

Type of Variable	Search Range	
k (mol/s/kg-cat)	Continuous	100-200
E ((kJ/kmol))	Continuous	20000-25000

Figure 2 shows the results obtained. It shows the concentration profile obtained at the reactor outlet of the reference case, and the same graph shows the concentration profile obtained by the optimization exercise proposed in this work.

Table 2. Optimization results

k (mol/s/g-cat)	E (kJ/kmol)	ISOBUT		ISOPENTA	% error	% error	% error
		(mol/s)	ISOCT (mol/hr)	(mol/hr)	isobut	isooct	isopent

Baseline	175.1818	-22184	0	21.05588	138.2018	0	0	1.23009E-07
Calculated	175.18803	-22184.0003	0	21.05588	138.201783			

The simulation was performed using the RPlug module of aspen plus as considered for the base case. A reactor length of 0.5 meters was considered, with a constant temperature of 410°K.

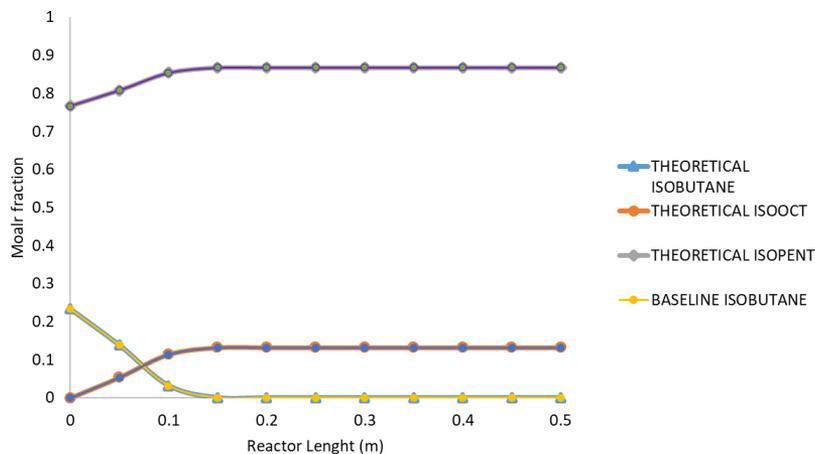


Figure 2. Concentration profiles for both the case base and this work

4. Conclusions

In the present work, a metaheuristic optimization technique was used for the prediction of kinetic data. As a case study, the dimerization of isobutane to produce isoctane was considered. Once the adjustment/optimization tool was used, it was possible to obtain k and E data, as well as concentration profiles very close to those reported, with an error percentage of less than 0.1%. The use of this tool presents a considerable numerical advantage in the prediction of kinetics for the adequate use of simulators, based on experimental data or similar situations.

References

- Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* 2009, 1 (1), 37–46
- Biegler, L. T. *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*. SIAM - Society for Industrial and Applied Mathematics 2010, 415
- Sharma, S.; Rangaiah, G. P. An Improved Multi-Objective Differential Evolution with a Termination Criterion for Optimizing Chemical Processes. *Comput. Chem. Eng.* 2013, 56, 155–173. <https://doi.org/10.1016/j.compchemeng.2013.05.004>
- Goortani, B. M., Gaurav, A., Deshpande, A., Ng, F. T., & Rempel, G. L. (2015). Production of isoctane from isobutene: Energy integration and carbon dioxide abatement via catalytic distillation. *Industrial & Engineering Chemistry Research*, 54(14), 3570-3581.

Optimal Control System for Products Quality from a Deethanizer Column

Cristian Patrascioiu, Nawwar A. Rahman, Marian Popescu

Automatic Control, Computers and Electronics Department, Petroleum-Gas University of Ploiesti, Bd. Bucuresti, 39, Ploiesti, 100680, Romania
cpatrascioiu@upg-ploiesti.ro

Abstract

Automatic control systems have a strong development and diversification. The research undertaken by the authors considered the development of an automatic system for optimal control of products quality from an industrial deethanizer column. The research aimed the following aspects: modelling the process using the resources of the Unisim Design simulator; process sensitivity analysis on disturbance/control agents - quality of separated products channels; identification of dynamic models associated with input-output channels; design of the dynamic simulator and controllers' tuning; design of the feedforward control system; design of the optimal control algorithm and the structure of the automatic control system; validation by simulation of the proposed control structure.

Keywords: deethanizer column, optimal control, product quality, simulation.

1. Introduction

Natural gas processing is a current global requirement. The structure of processing plants, installations components, used equipment, products and operating parameters depend on the natural gas composition and flowrate but also on the market requirements and price. One of the finished products of the natural gas processing plant is ethane.

The specialized literature dedicated for ethane product deals with three research directions. The first direction is associated with studies on natural gas processing techniques (Almeland et al., 2009, Shimekit and Mukhtar, 2012).

The second research direction is represented by the modelling and simulation of the distillation process. This category includes modelling and simulation of liquefied petroleum gas production (Elbadawy et al., 2017), studies on the modelling, simulation and optimization of an ethane and propane recovery plant (Mukherjee et al., 2013), simulation and optimization of deethanizer tower (Esfeh and Mohammadi, 2011).

The third research direction is dedicated to distillation control processes. Many papers in this category are educational or industrial manuals (Lipták and Venczel, 2016, Hughes, 2006). A small category of papers is dedicated to ethane distillation control system (Hori and Skogestad, 2007, Luyben, 2013).

Given the presented literature study, the authors researched and developed an optimal control system for products quality from a deethanizer column.

2. The Deethanizer Column

The studied deethanizer column is part of a natural gas processing plant (Halafawi et al., 2020). The deethanizer column is a classic column with partial condenser. The column feed comes from the bottom of the demethanizer column, this flow being previously

calculated by the authors (Patrascioiu et al., 2021). The quality conditions imposed for the two products are presented in Table 1. The control structure of the industrial column has as controls agents the reflux flowrate and the reboiler steam flowrate.

Table 1. Quality conditions for the deethanizer column

Product	Component	Restriction type	Value [mole %]
Distillate	C ₁	max	5.00
	C ₂	min	90.00
	C ₃₊	max	5.00
Bottom product	C ₂	max	0.25

The necessary steps to design the automatic system for optimal control of products quality are as follows:

- a) Design of steady-state simulator for the deethanizer column.
- b) Process sensitivity analysis.
- c) Basic control system.
- d) Dynamic simulation of the basic control structure for the deethanizer column.
- e) Design of the feedforward control system.
- f) Testing of the proposed feedforward control system.
- g) Design of the optimal control system.
- h) Testing of the proposed optimal control system.

3. Modelling and simulation of the deethanizer column

The first four steps listed above contain elements for modelling and simulation of the fractionation column. The Unisim Design simulator was used to perform the four stages, a simulator that contains mathematical models of various processes and allows the steady-state and dynamic simulation of the deethanizer column model. In the following will be presented the defining elements of the modelling, the results and the conclusions of these modelling stages.

3.1. Design of steady-state simulator for the deethanizer column

The Distillation Column model from Unisim Design simulator was used to simulate the deethanizer column. This model, configured for a column with partial condenser and reboiler, has three degrees of freedom and consequently three material or energy flows associated with the column must be set. Considering the structure of the column, the existing automatic system but also the requirements of the optimal control system, the authors opted for the following specifications: non-condensable gas flowrate, reflux flowrate and bottom product flowrate.

3.2. Process sensitivity analysis

The deethanizer column is characterized by the input-output variable shown in Figure 2. The sensitivity analysis aims to determine the influence of disturbances and control agents on products quality, to select control agents for products quality control. The process output variables are the concentration of ethane in the overhead product (subject to maximization), the concentration of propane in the overhead product (subject to minimization), and the concentration of ethane in the bottom stream. The most important disturbance is the feed flowrate.

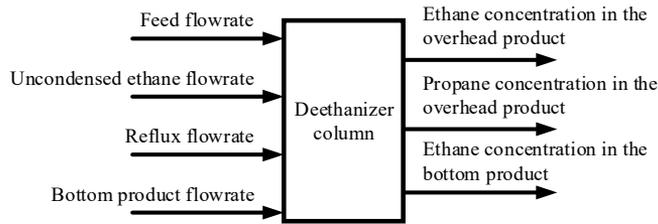


Figure 1. Block diagram of the deethanizer process

The sensitivity analysis was performed for the variation of the three control agents presented in Figure 1.

3.2.1. Process sensitivity to the variation of the non-condensable gas flowrate

From the study it was observed that a change in the non-condensable gas flowrate does not lead to significant variations in the quality of the distillate. This situation is primarily due to the high value of reflux, the usefulness of the non-condensable gas flowrate as a control agent being very low. A negative influence of the non-condensable gas flowrate is represented by the financial losses generated by the ethane lost in the non-condensable gases.

3.2.2. Process sensitivity to the variation of the reflux flowrate

The reflux flowrate is a powerful control agent for control of the separated products quality in a fractionation column. The influence of the reflux flowrate on the quality of the two separate products in the deethanizer column is as follows:

- a) The variation of the reflux flowrate changes the concentration of ethane in the distillate by 0.05% / kgmole/h, which gives importance to the reflux flowrate as control agent. At the same time, the variation of the ethane concentration in the distillate with the reflux flowrate is nonlinear, at values higher than 340 kgmole/h the variation of the ethane concentration in the distillate being small. The increase of the reflux flowrate also implies the increase of the energy consumption, both for the reboiler (steam consumption) and for the condenser (propane refrigerant consumption).
- b) For the same range of variation of the reflux flowrate, the concentration of propane in the distillate also decreases by 0.05% / kgmole/h, the nonlinear character being present in this case as well.
- c) Sensitivity analysis on the influence of reflux flowrate on the concentration of ethane in the distillate indicates the need to design an optimal system for control of the impurity concentration in the distillate (5 mole% propane) while minimizing the thermal load of the reboiler and compliance with the minimum of the quality condition for ethane (90 mole%).

3.2.3. Process sensitivity to the variation of the bottom product flowrate

The influence of the bottom product flowrate on the quality of the two separated products in the deethanizer column is as follows:

- a) The variation of the bottom product flowrate produces the change of the ethane concentration in the distillate by 0.86% / kgmole/h, 17 times higher than the gain generated by the variation of the reflux flowrate.
- b) The ethane is an impurity in the bottom product. The variation of the ethane concentration in the bottom product, depending on the bottom product flowrate, is 0.13% / kgmole/h.

The sensitivity analysis performed by the authors leads to the following conclusions:

- a) The flowrate of non-condensed ethane only influences the concentration of methane in the distillate, being a control agent only usable in the situation when the upstream column, the demethanizer column, does not work according to the specifications.
- b) To control the quality of the two separated products, two control agents are available, the reflux flowrate and the flowrate of the bottom product.
- c) The influence of the two control agents on the products quality is nonlinear.
- d) The reflux flowrate influences the energy consumption.

3.3. Dynamic simulation of the basic control structure for the deethanizer column

The development of the dynamic simulation program represented a substantial effort, being carried out according to the methodology in the literature (Patrascioiu et al., 2014). Given the lack of industrial data, the authors had to size both the flowrate measurement systems and the control valves. After the elaboration of the dynamic simulation program, the controllers were tuned, and the mathematical models associated with the input - output channels were determined. Of these models, the reflux flowrate - ethane concentration in the distillate channel and the bottom product flowrate - ethane concentration in the bottom product channel present special interest.

4. Design of the feedforward control system

The disturbances of the fractionation process are the feed flowrate and composition. Of these disturbances, the feed flowrate is the most important disturbing component, the feedforward control system being designed to reject the effect of this disturbance. The central element of the control system is the controller. Given the multivariable nature of the process (2x2), the controller will have the same characteristic, respectively 2 set-points (x_D^i and x_B^i) and two control signals ($L^i(t)$ and $B^i(t)$), Figure 2. The two controller outputs represent set-points for the reflux flowrate control system and bottom product flowrate control system respectively.

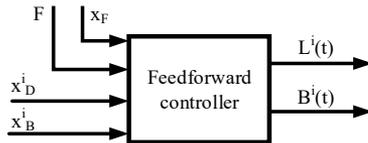


Figure 2. Block diagram of the feedforward controller

The feedforward control algorithm has two components (Marinoiu et al., 1986, Patrascioiu et al., 2015):

- a steady-state component for the calculation of the control signals values in steady-state;
- a dynamic component for the compatibility of the disturbance dynamics with that of the control signals.

4.1. The steady-state component

This component is based on the Fenske-Gilliland-Underwood model of the fractionation process and is applicable over a wide range of feed flowrate variation (the main process disturbance). The constants of the Fenske-Gilliland-Underwood model are the relative average volatility of the light component in relation to the heavy one, the number of theoretical plates. The controller has been designed with the following features:

- a) The average relative volatility α was calculated based on the equilibrium constants of ethane and propane.

- b) The feed concentration x_F was calculated considering only the ethane component.
- c) The controller set-points x_D^i and x_B^i refer only to the ethane component.

The implementation in the Unisim Design environment of the Fenske-Gilliland-Underwood relations, an extremely laborious operation, was realized using the SPRDSHT module from the environment library.

4.2. The dynamic component

The dynamic component associated with the controller outputs is materialized by 1st order transfer function for the feed flowrate - ethane concentration in the distillate channel and 2nd order transfer function for the feed flowrate - ethane concentration in the bottom product channel.

5. Optimal control system

The quality conditions of the two products (ethane and C_{3+} stream) are expressed by the restriction system

$$\begin{cases} x_{C_{2,D}} \geq 0.9 \\ x_{C_{3,D}} < 0.05 \\ x_{C_{2,B}} < 0.0025 \end{cases}$$

The performed simulations confirmed the obtaining of a $x_{C_{2,B}} < 0.0025$ ethane concentration in bottom product regardless of the set-point for ethane concentration in distillate, which validates the choice of the type and the value of the set-point x_B^i .

For the distillate quality the $x_{C_{2,D}} \geq 0.9$ condition was taken into account so the set-point for the distillate quality was considered $x_D^i = x_{C_{2,D}}$. Maximizing the ethane concentration in the distillate is achieved by increasing the set-point value.

Because the concentration of propane in the distillate must comply with the condition $x_{C_{3,D}} < 0.05$, the graph in Figure 3 shows the area where the column cannot be operated.

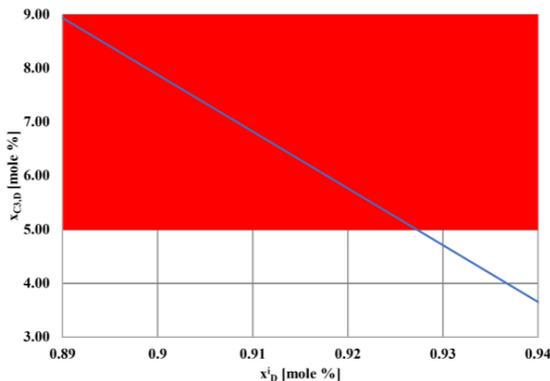


Figure 3. Variation of $x_{C_{3,D}}$ concentration depending on controller set-point x_D^i

As the energy effort of operating the column depends on the value of the ethane concentration in the distillate, it is proposed to operate the column at a minimum value of the x_D^i set-point for which the $x_{C_{2,D}} \geq 0.9$ and $x_{C_{3,D}} < 0.05$ conditions are met.

The dependency shown in Figure 3 has the equation $x_{C_{3,D}} = 103.102 - 105.8 \cdot x_D^i$.

By imposing $x_{C_{3,D}} = 0.05$, the minimum value of the controller set-point x_D^i is obtained, $x_D^i = 0.9272$. The pair of values $x_D^i = 0.9272$ and $x_B^i = 0.0025$ represent the optimal set-

point of the feedforward controller. The average deviation of the ethane concentration in the distillate in relation to the set-point is 0.85%, which validates both the calculation relations implemented in the concentration controller and the selection of the two set-points of the controller, x^i_D and x^i_B .

6. Conclusions

The research carried out by the authors led to the development of a structure for optimal control of the quality of the separated products in the deethanizer column. Steady-state and dynamic simulations validated the objective function and the proposed automation solution. The limitations of the structure developed by the authors were generated by the lack of industrial operation data and by the complexity of the industrial plant. Under these conditions, some of the elements developed in the research can only be theoretical. Future access to industrial operating data may lead to industrially applicable control structures.

References

- S.K. Almeland, K.A. Meland, D.G. Edvardsen, 2009, Process Design and Economical Assessment of a Methanol Plant, Norwegian University of Science and Technology, Trondheim, Norway.
- K.E. Elbadawy, M.A. Teamah, A.I. Shehata, A.A. Hanfy, 2017, Simulation of Liquefied Petroleum Gas (LPG) Production from Natural Gas Using Fractionation Towers, International Journal of Advanced Scientific and Technical Research, vol. 6, is. 7, 148-153.
- H.K. Esfeh, I. Mohammadi, 2011, Simulation and Optimization of DE-Ethanizer Tower, International Conference on Chemistry and Chemical Process, vol. 10, 173-177.
- M. Halafawi, A. Mhanna, M. Charif, L. Avram, 2020, Troubleshooting Identification and Upset Conditions Effect of LNG Plant, Pet Coal. 62(1), 220-23.
- E.S. Hori, S. Skogestad, 2007, Control structure selection of a deethanizer column with partial condenser, In Proceedings of European Congress of Chemical Engineering (ECCE-6), 16-20.
- T.A. Hughes, 2006, Measurement and Control Basics (4th ed.), International Society of Automation.
- B.G. Lipták, K. Venczel (eds), 2016, Instrument and Automation Engineers' Handbook (5th Ed.): Process Measurement and Analysis, CRC Press.
- W.L. Luyben, 2013, Control of a Train of Distillation Columns for the Separation of Natural Gas Liquid, Ind. Eng. Chem. Res., vol. 52, is. 31, 10741-10753.
- V. Marinoiu, N. Paraschiv, C. Patrascioiu, 1986, Conducerea cu calculatorul a procesului de separare a propenei, Revista de chimie, No. 11, 990-994.
- A. Mukherjee, V.B. Vedula, S. Medipally, U. Makeswari, 2013, Ethane Propane Recovery Process – Simulation and Optimization, 66th Annual Session of Indian Institute of Chemical Engineers, Mumbai, India, 247.
- C. Patrascioiu, M.A. Cao, M. Popescu, 2015, Control of propylene - propane distillation process using Unisim®design, Proc. of 19th International Conference on System Theory, Control and Computing, October 14 - 16, Cheile Gradistei - Fundata Resort, Romania, 747 – 752.
- C. Patrascioiu, M. Popescu, N.A. Rahman, 2021, Modelling and Simulation of Products Quality Control System for Demethanizer Column from a Natural Gas Treatment Plant, Proceedings of the 23rd International Conference on Control Systems and Computer Science, May 26-28, Bucharest, 27-32.
- C. Patrascioiu, M. Popescu, N. Paraschiv, 2014, Specific Problems of Using Unisim Design® in the Dynamic Simulation of the Propylene-Propane Distillation Column, REV. CHIM. (Bucharest), 65, No. 9.
- B. Shimekit, H. Mukhtar, 2012, Natural Gas Purification Technologies – Major Advances for CO₂ Separation and Future Directions, Advances in Natural Gas Technology, InTechOpen, 235-270.

Importance of detailed experimentation in the model based design and scale up of pharmaceutical spray dryers for heat sensitive products

P. Martin-Salvador^{a,b}, T. De Beer^b, A. Kumar^{a,*}

^a*Pharmaceutical Engineering Research Group (PharmaEng), Ghent University, Ottergemsesteenweg 460, 9000 Ghent, Belgium*

^b*Laboratory of Pharmaceutical Process Analytical Technology (LPPAT), Ghent University, Ottergemsesteenweg 460, 9000 Ghent, Belgium*

**Ashish.Kumar@UGent.be*

Abstract

Spray drying of sensitive pharmaceuticals is challenging due to thermal and other stresses which can have large effects on product quality. In addition, the design and scale up of such processes is a difficult which oftentimes requires many experiments at different scales. By using model based design, the number of experiments can be reduced. To that end, experimental data of single droplet drying experiments is used to fit a drying model and degradation kinetics of a model heat sensitive protein. The calibrated model is used to propose large scale designs. However, it is uncommon that the literature offers such data, and such detailed experimentation requires many resources. Therefore, the large scale designs from the models calibrated using the whole data set are compared to using only inlet/outlet values, to find the importance of the intermediate data points. The results show that designing with a model fully calibrated can produce designs that consider product degradation, while using models calibrated with only start/end values poses the risk of unwanted product degradation. Consequently, full range experimental data of droplet drying is necessary to use lumped parameter drying models in process design.

Keywords: Spray drying, droplet drying, pharmaceutical design

1. Introduction

Spray drying has been a robust and widely used manufacturing process in the pharmaceutical and other industries used to produce particulate products from liquid solutions or dispersions. Pharmaceutical spray dryers are available in lab, pilot, and production scales, with capabilities for processing grams to tons per day (Poozesh and Bilgili, 2019). While the design of large production-scale spray drying processes is usually the labor-intensive experimental based result of scale up from lab or pilot scale tests, current modeling trends are trying to improve this by reducing the number of experiments required and project lead times (Dobry et al., 2009).

Protein pharmaceuticals are usually formulated in aqueous forms, however, this poses limited shelf life even under refrigerated conditions and, therefore, dry formulations can be very advantageous (Pinto et al., 2021). Although several pharmaceutical commercial protein products produced via spray drying are on the market, their drying is challenging due to the potential loss of their physicochemical properties by spray drying related stresses. Thermal stresses have been pointed out

to be the major source of protein loss of activity during spray drying of a monoclonal antibody (Schüle et al., 2007), but shear and surface effects are also present (Das et al., 2021).

Mathematical modelling can be applied to understand and predict the drying kinetics of the sprayed starting solution (droplets) and resulting particles under varying operating conditions, and relate them to product quality. Several models for the drying of droplets are available in the literature. The characteristic drying curve (CDC) concept is used here. This is a lumped parameter model that offers a good balance between complexity and prediction power when combined with parameter fitting to experimental data. Even when using simplified models, mathematical modelling of the spray drying requires fitting the predictions to experimental results. Published experimentation in spray drying is focused on proof of concept studies, process optimization studies, or model validation; however, the effect that different experimental studies would have on the fitted model are most often not the focus.

2. Mathematical model

In this study the drying of droplets into particles is considered, and other steps of a typical spray drying process such as the atomization step and the separation step are ignored. The presentation of the model used is split between the drying kinetics and the rest of the spray drying model.

2.1. Drying kinetics

Drying kinetics are at the core of the drying model for spray drying. Heat and mass transfer between the droplet and surrounding air are considered, which determines the change in droplet temperature, moisture and diameter. Analytically, the change in moisture can be expressed as in equation 1 (Parti and Palancz, 1974). The rate of evaporation \dot{m}_v is typically expressed as a mass transfer coefficient times a driving force between the bulk of the drying medium and the droplet surface, for example in terms of humidity in equation 2. Substitution of the humidity at the surface Y_s by the saturation humidity Y_{sat} would overestimate the drying rate for a dissolved product compared to a pure liquid droplet, thus in the CDC approach an empirical function f is included in equation 4 (Langrish and Kockel, 2001). f varies from 0 to 1 to account for the reduction of the drying rate when dissolved or dispersed proteins enrich the droplet surface, eventually forming a crust through which water must diffuse before it can evaporate. f is a function of a non-dimensional moisture content, typically taken as a power function as shown in equation 3.

$$\frac{dW_p}{dt} = \dot{m}_v A / m_s \quad (1) \quad f = \left(\frac{W_p - W_{p,eq}}{W_{p,crit} - W_{p,eq}} \right)^{n_{CDC}} \quad (3)$$

$$\dot{m}_v = k_Y (Y_s - Y_b) \quad (2) \quad \dot{m}_v = f k_Y (Y_{sat} - Y_b) \quad (4)$$

2.2. The drying model

The model used is a 1-dimensional plug flow for both the gas and condensed phases. The space coordinate is 0 at the tip of the atomizer and is positive in the downward direction. Equation 1 translates to equation 5 by the variable change $z = v_p t$. Particle diameter is evaluated from the balloon shrinkage assumption in equation 6, for which the density is evaluated using equation 7 (Parti and Palancz, 1974). However, equations 6-7 are only used until the moment of crust formation, which happens at moisture content $W_{p,crust}$. Further drying from this point does not change the diameter, only the density decreases with equation 8 (Cotabarren et al., 2018). Mass balance with surrounding air yields the change in air humidity shown in equation 9, heat balances yield equations 10-11 (Cotabarren et al., 2018) and momentum balance yields the particle velocity equation 12 (Truong et al., 2005). Finally, the protein residual activity ra , which varies between 1 (no degradation) and 0 (complete degradation), is calculated using eq. 13, which shows the importance of water as a stabilizer for this protein as hypothesized by Lorenzen and Lee (2012).

$$\frac{dW_p}{dz} = -\frac{\pi d_p^2 \dot{m}_v}{v_p m_s} \quad (5)$$

$$\frac{dT_p}{dz} = \frac{\pi d_p^2 [h(T_a - T_p) - \dot{m}_v \Delta_v H]}{v_p m_s (c p_s + W_p c p_w)} \quad (10)$$

$$d_p = d_{p0} \sqrt[3]{\frac{\rho_{p0} - \rho_w}{\rho_p - \rho_w}} \quad (6)$$

$$\frac{dT_a}{dz} = \frac{-N_d \pi d_p^2 (\dot{m}_v \bar{c} p_{ww} + h)(T_a - T_p) - v_p U \pi D (T_a - T_{amb})}{v_p m_a (c p_a + Y_b c p_{ww})} \quad (11)$$

$$\rho_p = \rho_s \frac{1 + W_p}{1 + \frac{\rho_s}{\rho_w} W_p} \quad (7)$$

$$\frac{dv_p}{dz} = \left(1 - \frac{\rho_a}{\rho_p}\right) \frac{g}{v_p} - \frac{3\rho_a C_d (v_p - v_a)^2}{4\rho_p d_p v_p} \quad (12)$$

$$\frac{d\rho_p}{dz} = \frac{-6\dot{m}_v}{d_p v_p} \quad (8)$$

$$\frac{dra}{dz} = \begin{cases} k_0/v_p & W_p > W_{p,r} \\ k_1 \exp\left(\frac{E_a}{RT}\right)/v_p & W_p \leq W_{p,r} \end{cases} \quad (13)$$

$$\frac{dY_b}{dz} = N_d \frac{\pi d_p^2}{v_p m_a} \dot{m}_v \quad (9)$$

3. Methods

3.1. Modelled system

The spray drying of L-glutamate dehydrogenase (GDH) in triethanolamine hydrochloride buffer is modelled. Single droplet drying experiments are reported by Lorenzen and Lee (2012). GDH was chosen as model system due to its high heat sensitivity, in addition to practical requirements such as high enzymatic activity and high water solubility. Lorenzen and Lee (2012) continuously measured droplet surface temperature and droplet radius for several minutes while the droplets dried in an acoustic droplet levitator. In addition, they measured the protein's residual activity at several time points in order to follow degradation kinetics.

3.2. Model calibration

The results of Lorenzen and Lee (2012) are used to fit the model parameters in two ways.

First, the totality of the data is considered and unknown parameters are fitted: $W_{p,crit}$, $W_{p,crust}$, $W_{p,r}$, n_{CDC} , k_0 , k_1 , E_a , SEF . These parameters are as defined in section 2., in addition we define the SEF parameter empirically as follows. First the Nusselt number is evaluated from the Ranz and Marshall correlation for drying droplets (Ranz, 1952), from which the heat transfer coefficient can be obtained. Then, the SEF factor is used to account for the elevated temperatures seen in the acoustic droplet levitator in the form $Nu_{AL} = SEF \times Nu$. Note that this parameter is only required to reproduce the experimental results in the acoustic droplet levitator and that it is removed when using the model to produce actual spray drying results. Equilibrium moisture is included in the form of the GAB desorption isotherm of skim milk taken from Lin et al. (2005) for the GDH system since they have similarities, and the data captured by single droplet drying experiments is not enough to improve on this. Further parameters and boundary conditions applied to equations 5-13 are taken from the experimental conditions of Lorenzen and Lee (2012), $W_{p0} = 33\text{kg/kg}$, $V_{drop} = 2\mu\text{L}$, $T_{a0} = T_{p0} = 35\text{-}60\text{ }^\circ\text{C}$, $RH_0 = 10\text{-}50\%$, $ra_0 = 1$. U is assumed 0 in the acoustic levitator. The fitting method is based on minimization of the sum of absolute errors computed for temperature, droplet diameter and protein residual activity at 1s intervals, for the 6 experimental runs of Lorenzen and Lee (2012), implemented in python.

Second, only starting and ending points are considered. Typically, spray dryers used in laboratories or pilot facilities have temperature sensors for the air before entering the drying chamber and in the collection pipe after the drying chamber, and similarly, droplet and particle sizes can be measured at the atomizer and after the product is collected. Therefore, it is common to see reported in literature the initial and final values of these variables for a multitude of spray drying equipment and (pharmaceutical) products, e.g. Jiang et al. (2021), and this second model tries to mimic this.

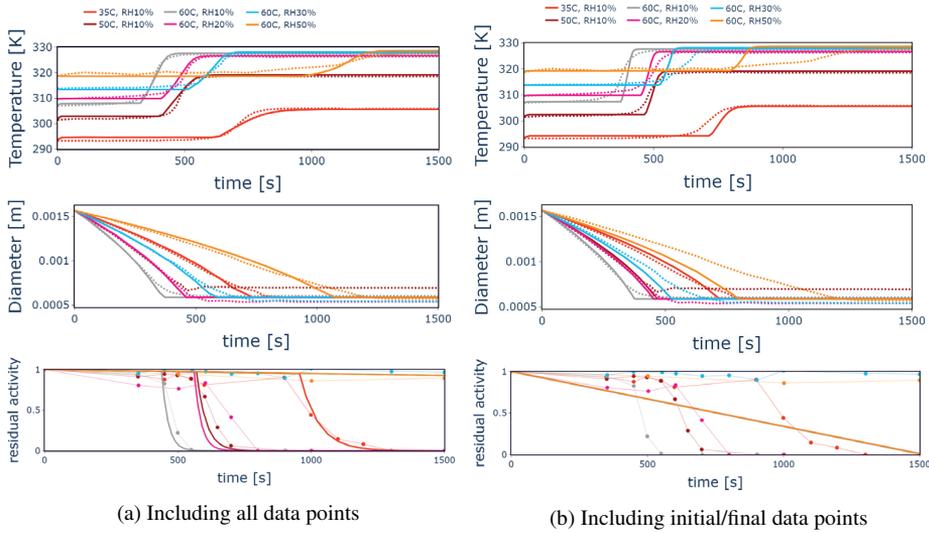


Figure 1: Parameter fit results at different temperatures ($^{\circ}\text{C}$) and relative moisture (RH%). Lines indicate model prediction, broken lines and dots indicate experimental values.

Parameters and boundary conditions used are the same as in the previous case. $W_{p,crust}$ and SEF are fitted by minimizing the error of both temperature and droplet diameter at time point 0 and end time, while n_{CDC} is assumed to be 1. The linear assumption of the characteristic drying curve is relatively common, used by Langrish and Kockel (2001) and later for example by Truong et al. (2005). In turn, $W_{p,crit}$ is taken to be equal to $W_{p,crust}$, and only the parameter k_0 is fitted in the residual activity equation 13 due to the lack of data, achieved in the model by setting $W_{p,r}$ to 0.

3.3. Methods for comparison of the fitted models in a design operation

A spray dryer design is given for a hypothetical GDH production facility. In this case the parameters and initial values are as follows. The production rate is fixed at 300L/h of liquid and the initial droplet size is fixed at 70 micron, values common in the industry. For simplicity all droplets are kept at the same size. The fitted model parameters from each of the calibrated models are used, namely $W_{p,crit}$, $W_{p,crust}$, $W_{p,r}$, n_{CDC} , k_0 , k_1 , and E_a . Inlet air humidity RH_0 is set to 0 and the assumed coefficient for heat loss to ambient $U = 5\text{W/m}^2\text{K}$. The remaining parameters are taken as the design variables, namely the chamber diameter D , inlet air temperature T_{a0} and air flow rate m_{a0} .

The design approach used here is similar to the thermodynamic design space (Dobry et al., 2009), which is based on drawing acceptable limits of variables calculated from a thermodynamic model on a flow-temperature diagram. By drawing boundaries it is easy to visualize which conditions remain inside all boundaries, which leads to an easy design process.

4. Results

4.1. Model calibration results

Figure 1 shows the predicted temperature and diameter profiles together with the experimental data. The parameters fitting with complete data are $W_{p,crit} = 2.8$, $W_{p,crust} = 1.1$, $n_{CDC} = 1.5$, $SEF = (3.2, 3.2, 3.2, 3, 2.7, 2.1)$, $W_{p,r} = 8\text{e-}2$, $k_0 = 5.0\text{e-}5$, $k_1 = 8.0\text{e}4$, $E_a = 4.0\text{e}4\text{ J/mol}$. The parameters

Table 1: Criteria used to define boundaries in figure 2

Parameter	Limit	Color (fig 2)	Reasoning
Residual moisture	5%	Red if > 5%	Product shelf life
Protein residual activity	95%	Pink if < 95%	Product quality
Outlet air humidity	0.2	Black if < 0.2	Economic optimization

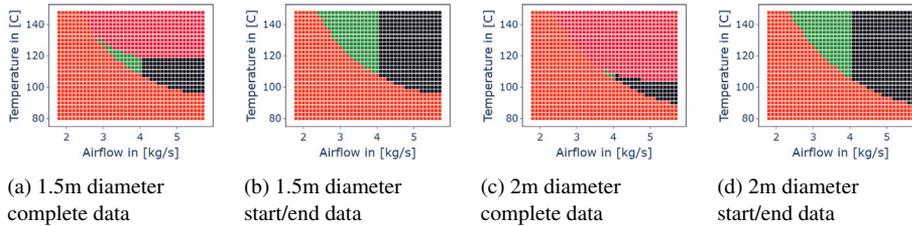


Figure 2: Design spaces predicted for different diameter dryers with both calibrated models

fitting with starting/final values are $W_{p,crust} = 1.1$, $SEF = (3, 3, 3, 3, 3, 3)$, $W_{p,r} = 0$, $k_0 = 6.6e-4$. Experimental data and predictions agree in the case when all data points are included in the calibration. However, in the case of only using starting and final values, differences are notable. This is specially true in the case of the protein residual activity, in view that the only logical assumption is to set $W_{p,r} = 0$ (equation 13).

4.2. Comparison of spray drying design space using the fitted models

Figure 2 shows the predicted design spaces for a 1.5 m diameter and 3 m length dryer, as well as for an increased diameter of 2 m. The boundaries have been defined based on the criteria of table 1. The points remaining in green after all other colors removed are all valid choices according to the defined criteria. However, it becomes apparent that the windows predicted from the different models are different: while the predicted boundary of maximum residual moisture is not largely different, the model without the intermediate data fails to predict that GDH will degrade at higher temperatures. In addition, while from an economic standpoint it would be advantageous to operate the dryer at lower air flows, the model with intermediate data points suggests using higher flows and lower temperatures. More importantly, when the diameter is increased the model without intermediate data suggests positive outcomes on a large window (figure 2d), the model calibrated with intermediate data predicts protein degradation due to particles having too long residence time (figure 2c). This comparison showcases the importance of having good lab scale experimentation before attempting to scale up the process. Protein degradation is one of the major sources of restrictions when scaling up, but data is not easily obtained, and this comparison shows that measurement of only spray dryer outputs might not be enough to predict it.

5. Conclusions

Experimental data for drying of acoustically levitated single droplets has been used to calibrate the drying model in two ways; using the full data set to try to provide the best reproduction of reality, and using only starting and ending values, to try to reproduce the steps that might be used when trying to fit the model to lab or pilot scale process data. The calibration of both models reproduces reasonably well the complete experimental data. However, both models perform very differently in terms of protein degradation.

The models have then been used to calculate design spaces for a commercial scale spray dryer. The acceptable ranges are defined by 3 criteria: product sufficiently dry, product not degraded, and air moisture in the outlet above a minimum value to justify operating costs. As expected both give similar results in terms of residual moisture and dryer performance. However, they differ when predicting protein degradation. Therefore single droplet drying experimental data is crucial when using lumped parameter models for the design of spray drying of heat sensitive products.

List of Symbols

C_d Drag coefficient [-]	k_Y Mass transfer coefficient [kg/m ² s]	T_a Temperature of air [K]
cp_a Heat capacity air [J/kg K]	m_a Mass flow of gas [kg/s]	T_{amb} Temperature ambient [K]
cp_s Heat capacity solid [J/kg K]	m_s Mass of solids in a droplet [kg]	T_p Temperature of particles [K]
cp_w Heat cap. water liq. [J/kg K]	\dot{m}_v Evaporation rate [kg/m ² s]	U Heat trans. coef. ambient [W/m ²]
cp_{wv} Heat cap. water vap. [J/kg K]	n_{CDC} Character. drying curve exp. [-]	v_a Velocity air [m/s]
\overline{cp}_{wv} Heat cap. water vapor averaged between T_p and T_a [J/kg K]	N_d Number of droplets [drops/s]	V_{drop} Droplet volume [m ³]
D Diameter drying chamber [m]	Nu Nusselt number [-]	v_p Velocity particles [m/s]
$\Delta_v H$ Heat of vaporiz. of water [J/kg]	Nu_{AL} Nusselt in acoustic levitator	W_p Moisture cont. [kg-wat/kg-solid]
d_p Diameter droplet [m]	ra Residual activity protein [-]	$W_{p,crit}$ Moisture at start hindering
E_a Reaction rate parameter [J/mol]	RH Relative moisture air [%]	$W_{p,crust}$ Moisture at crust formation
f Characteristic drying curve [-]	ρ_a Density air [kg/m ³]	$W_{p,eq}$ Moisture at equilibrium
h Heat trans. coef air-droplet [W/m ²]	ρ_p Density droplet [kg/m ³]	$W_{p,r}$ Moisture at start degradation
k_0 Reaction rate parameter [s ⁻¹]	ρ_{p0} Density droplet initial [kg/m ³]	Y_b Bulk humidity [kg-wat/kg-air]
k_1 Reaction rate parameter [s ⁻¹]	ρ_s Density dry solid [kg/m ³]	Y_s Surface humidity [kg-wat/kg-air]
	ρ_w Density liquid water [kg/m ³]	Y_{sat} Saturation hum. [kg-wat/kg-air]
	SEF Sonic energy factor [-]	z Position vertical coordinate [m]

References

- I. M. Cotabarren, D. Bertín, M. Razuc, M. V. Ramírez-Rigo, J. Piña, 2018. Modelling of the spray drying process for particle design. *Chemical Engineering Research and Design* 132, 1091–1104.
- T. K. Das, A. Sreedhara, J. Colandene, D. K. Chou, V. Filipe, C. Grapentin, J. Searles, T. Christian, 2021. Stress Factors in Protein Drug Product Manufacturing and Their Impact on Product Quality. *Journal of Pharmaceutical Sciences*.
- D. E. Dobry, D. M. Settell, J. M. Baumann, R. J. Ray, L. J. Graham, R. A. Beyerinck, 2009. A model-based methodology for spray-drying process development. *Journal of Pharmaceutical Innovation* 4 (3), 133–142.
- B. Jiang, D. Yu, Y. Zhang, H. Yu, H. Feng, S. W. Hoag, 2021. Spray Layering of Human Immunoglobulin G: Optimization of Formulation and Process Parameters. *International Journal of Pharmaceutics*, 121238.
- T. A. Langrish, T. K. Kockel, 2001. The assessment of a characteristic drying curve for milk powder for use in computational fluid dynamics modelling. *Chemical Engineering Journal* 84 (1), 69–74.
- S. X. Lin, X. D. Chen, D. L. Pearce, 2005. Desorption isotherm of milk powders at elevated temperatures and over a wide range of relative humidity. *Journal of Food Engineering* 68 (2), 257–264.
- E. Lorenzen, G. Lee, jun 2012. Slow motion picture of protein inactivation during single-droplet drying: A study of inactivation kinetics of l-glutamate dehydrogenase. *Journal of Pharmaceutical Sciences* 101 (6), 2239–2249.
- M. Parti, B. Palancz, 1974. Mathematical model for spray drying. *Chemical Engineering Science* 29 (7), 355–362.
- J. T. Pinto, E. Faulhammer, J. Dieplinger, M. Dekner, C. Makert, M. Nieder, A. Paudel, 2021. Progress in spray-drying of protein pharmaceuticals: Literature analysis of trends in formulation. *Drying Technology* 39 (11), 1415–1446.
- S. Poozesh, E. Bilgili, may 2019. Scale-up of pharmaceutical spray drying using scale-up rules: A review. *International Journal of Pharmaceutics* 562 (March), 271–292.
- W. E. Ranz, 1952. Evaporation from drops, parts i & ii. *Chem Eng Prog.* 48, 141–146.
- S. Schüle, W. Frieb, K. Bechtold-Peters, P. Garidel, jan 2007. Conformational analysis of protein secondary structure during spray-drying of antibody/mannitol. *European Journal of Pharmaceutics and Biopharmaceutics* 65 (1), 1–9.
- V. Truong, B. R. Bhandari, T. Howes, nov 2005. Optimization of co-current spray drying process of sugar-rich foods. Part I—Moisture and glass transition temperature profile during drying. *Journal of Food Engineering* 71 (1), 55–65.

Process modelling of Direct Air Capture (DAC) of CO₂ using solid amine sorbents

So-mang Kim^a, Grégoire Léonard^a

^a*Chemical Engineering, University of Liège, Belgium*

Abstract

This paper presents a fixed bed reactor system to capture CO₂ from ambient air via Temperature Vacuum Swing Adsorption (TVSA) cycles employing Lewatit® VP OC 1065 amine-functionalized adsorbents. Firstly, a model describing CO₂ mixture flow, mass, and heat transfer with adsorption isotherm model is developed in Aspen Adsorption environment to study the performance of DAC processes with TVSA. Then, a single adsorption column is simulated to study both adsorption and breakthrough curves, and the model is validated with literature results. Secondly, a full DAC model with TVSA cycle is developed and the results are compared to TVSA cycle with a steam purge in literature to validate the reliability and simulation capabilities of the software tool. Lastly, a model of a process capturing 1.1 kilograms of CO₂ per day is developed as a base case, and studies on the energy requirements, capital, and operating costs of the DAC system are conducted. In addition, sensitivity studies on the model parameters are carried out. This work stands as one of the first identified modelling studies of TVSA-based Direct Air Capture and serves as a basis to analyze the feasibility of DAC system deployment and its relevance in comparison to CO₂ capture from point sources.

Keywords: Direct air capture, Carbon capture, TVSA, Adsorption, Process modelling

1. Introduction

The idea of CO₂ capture from ambient air, also referred to as Direct Air Capture (DAC), was first suggested by Lackner and coauthors in 1999 to mitigate climate change. CO₂ adsorption processes are often considered as the most suitable option for DAC and many laboratory and pilot scaled units have been developed over the last decade. In particular, there are 15 operational DAC plants globally capturing more than 9000 tCO₂/year (IEA, 2020). Some of the industrial pioneers are Climeworks (Switzerland) and Global Thermostat (USA) where solid sorbents are used to capture CO₂ while Carbon Engineering (Canada) has developed a liquid solvent-based DAC system. Also, some examples of active academia are Twente University and ETH Zurich where DAC systems with amine solid sorbents are studied while Arizona State University is developing Moisture Swing Adsorption (MSA). However, there is currently a scarcity of literature on simulation studies of adsorption processes in the DAC context. To address this gap, the present work studied a TVSA DAC process in an attempt to bring a new light on process system design and DAC performances under cyclic operation.

2. Modelling and validations

In this paper, a one-dimensional model for gas adsorption in a fixed bed is developed in Aspen Adsorption. The adsorption isotherm and model assumptions are described below.

2.1. 2 Adsorption Isotherms

Toth Isotherm parameters described in Bos et al. (2019a) are used to describe an amine-based solid sorbent, Lewatit® VP OC 1065 that numerous authors including Veneman et al. (2012), Bos et al. (2019a&b), Young et al. (2021) have studied extensively. The Toth isotherm is presented in Equation 1 and the fitting parameters can be found in the work of Bos et al. (2019a).

$$q_e = \frac{q_s b P_{CO_2}}{(1 + (b P_{CO_2})^t)^{\frac{1}{t}}} \quad (1)$$

The parameters q_s , b and t in the Toth isotherm described in Equation 1 are the saturation loading, the affinity of the sorbent to CO_2 and the heterogeneity respectively and each parameter is a function of temperature. These parameters can be used to calculate CO_2 equilibrium loading (q_e) at various CO_2 partial pressures (P_{CO_2}) and operating temperatures. More details of the parameters and related equations can be found in the work of Bos et al. (2019a).

2.2. Fixed bed model validation

2.2.1. Model Assumptions

The model consists of mass, momentum, and energy balances of the gas and the adsorbent phases. The balance equations form a set of partial differential equations (PDEs) that are solved using a software package, achieving a cyclic steady state in Aspen Adsorption. The process model is based on the following assumptions:

- The gaseous mixture obeys the ideal gas law.
- Only carbon dioxide is adsorbed.
- The resistance of mass transfer in the gas phase is negligible; the kinetics of mass transfer within a particle is approximated by the linear driving force (LDF) model.
- A constant heat of adsorption is used.
- The adsorbent is considered as a homogenous phase and its physical properties are assumed to be constant.

The first-order upwind differencing scheme (UDS1) spatial discretization method was selected for its comparatively higher stability. A default number of nodes (20) was used and the system is assumed to be isothermal for the model validation steps in section 2.2.2.

2.2.2. Breakthrough curve

The experimental breakthrough curves of Yu et al. (2017) at air inlet conditions of 25 °C, at atmospheric pressure, and 400 ppm on a clean adsorption bed are compared with the results from the Aspen Adsorption fixed bed model in Figure 1. Table 2 presents modelling parameters and operating conditions used in the fixed column model.

Table 1. Parameters and operating conditions for breakthrough modelling

Parameter	Description	Value	Unit	Source
V_g	Superficial velocity	0.14 and 0.27	m s^{-1}	Yu et al. (2017)
$C_{p,s}$	Sorbent heat capacity	1580	$\text{J. mol}^{-1}\text{K}^{-1}$	Veneman et al., (2012)
R_p	Particle diameter	0.26	mm	Bos et al., (2019b)
ρ_p	Particle density	1580	kg m^{-3}	Bos et al., (2019b)
ϵ	Particle voidage	0.23	$\text{m}^3 \text{m}^{-3}$	Lanxess
MTC	Mass transfer coefficient	3.26×10^{-4}	s^{-1}	Yu et al. (2017)
L	Column height	0.5	m	Yu et al. (2017)
D	Column diameter	0.016	m	Yu et al. (2017)

Figure 1 shows the model results (solid lines) obtained at the operation conditions presented in Table 2 and experimental breakthrough curves (dotted lines) from Yu et al. (2017) at 0.14 m/s and 0.27 m/s superficial velocities of inlet feed. The model lines show a slightly steeper increase at the beginning of the breakthrough curve which may be due to the existence of dead spaces in the actual experimental column.

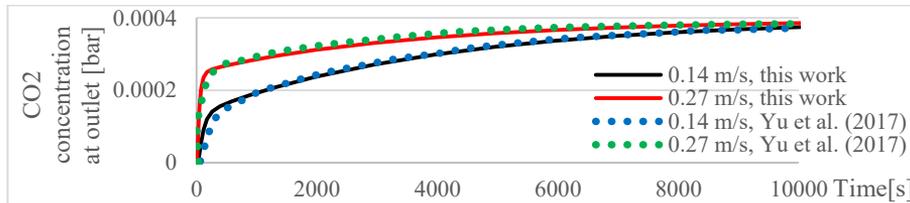


Figure 1. Fixed bed operations breakthrough curves at the superficial velocity of 0.14 m/s and 0.27 m/s with 1 g of Lewatit® VP OC 106 solid sorbent at 25 °C.

Yu et al. (2017) also computed adsorption half-time ($t_{1/2}$), which is the time required to adsorb 50% of the maximum adsorption capacity, at different V_g . In this study, $t_{1/2}$ were simulated and the model and literature values are compared in Table 3.

Table 2. Adsorption half-time comparison between this work and the results of Yu et al. (2017).

V_g [m/s]	$t_{1/2,model}$ [min]	$t_{1/2,Yu\ et\ al.,(2017)}$ [min]	Difference [%]
0.05	54.5	55	0.92
0.1	42.83	43	0.39
0.2	38.6	38	1.55
0.3	37.5	37	1.33

The results were found to be consistent with the experimental works of Yu. Also, an equilibrium loading of 1.06 mol/kg of CO₂ was obtained from the fully saturated model at the feed CO₂ concentration of 400 ppm as calculated using the Toth isotherm presented in Equation 1, and this simulation value agrees with Bos et al. (2019a) and Yu et al. (2017).

2.2.3. Temperature Vacuum Swing Adsorption (TVSA) cycle

In order to further validate the reliability and simulation capabilities of the software tool, a full TVSA cycle using a different adsorbent, an amine nanofiber, was simulated using Aspen Adsorption based on the work of Stampi-Bombelli et al. (2020). The change of materials is due to insufficient cycling results identified in the literature for Lewatit. The Toth isotherm equations, parameters, and operating conditions can be found in the aforementioned literature. Note that the value of one parameter (\mathbf{b}_0) needed to be retrieved from Gebald et al. (2014) which is one of the preceding studies on the nanofiber adsorbent, as the value reported by Stampi-Bombelli et al. (2020) for this parameter led to aberrant results. The cycle validation is presented in Figure 2.

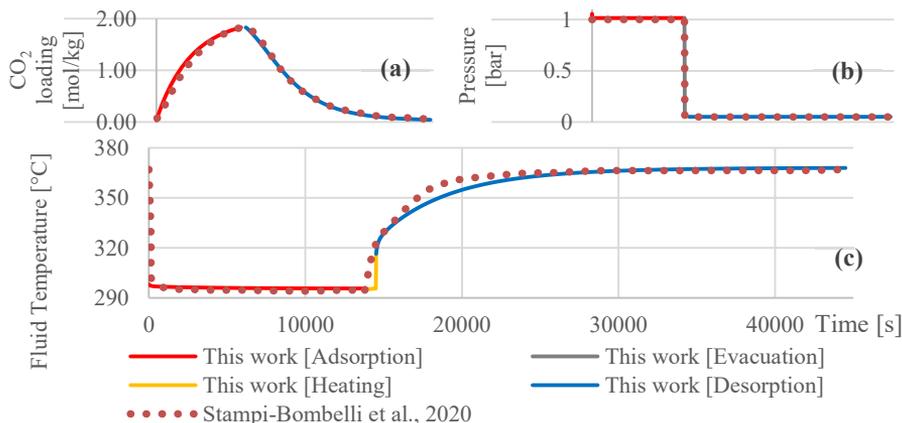


Figure 2. Cyclic profiles (solid lines) from Aspen Adsorption during the adsorption (red), evacuation (grey), heating (yellow), and desorption (blue) steps. The dotted lines show the cyclic profiles reported in Stampi-Bombelli et al. (2020). (a) CO₂ loading profile, (b) pressure profile and (c) temperature profile.

The TVSA cycle (Stampi-Bombelli et al., 2020) consists of four steps which are adsorption (20 °C, 1bar, 13772 s), evacuation (0.05 bar, 30 s), heating (via heating jacket, 704 s), and desorption (95 °C) with purge steam. The difference observed in the temperature profile could be due to the absence of a wall energy balance and the usage of the overall heat transfer coefficient in Aspen Adsorption. In general, the results from the simulation showed the same trends as reported in the literature. The next section will present a kilogram scaled DAC model as a basis to study the operating and economic performances.

3. DAC model of 1.1 kg/day results and discussion

3.1. DAC model CO₂ capture rate, purity and recovery

In the recent paper of Schellevis et al. (2021), a complete kilogram scale TVSA fixed bed model utilizing Lewatit® VP OC 1065 was presented based on experimental results collected on a small-scale DAC pilot. In the present paper, an Aspen Adsorption model is developed to study a similar process with a steam purge. Details of the bed geometry and operating conditions were retrieved from Schellevis et al. (2021). The cycle involves four steps which are the adsorption, evacuation, desorption, and cooling, and in this study, the step time of 3200 s, 30 s, 4700 s, and 2300 s were assumed respectively since these parameters were not reported in the literature. Figure 3 presents a cyclic CO₂ loading at the end of the adsorption bed over a period of 24 hours.

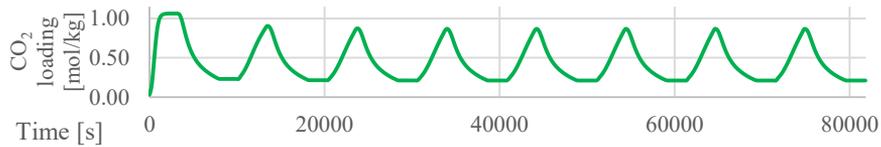


Figure 3. Cyclic CO₂ loading results at the end of column obtained from Aspen Adsorption

Based on our simulation results, 0.0339 kg_{CO₂}/cycle could be obtained, leading to 0.2712 kg_{CO₂}/day. Indeed, as can be seen in Figure 3, a total of 8 cycles could be achieved within a day. To achieve 1.1 kg CO₂ capture per day, Schellevis et al. (2021) proposed a DAC process with several parallel fixed bed reactors. In our case, four parallel reactors are needed to capture 1.1 kg of CO₂ per day. Finally, assuming all the purge steam will be condensed out of the CO₂ product stream, purity of 99.98% was achieved while a recovery of 71.70% was obtained.

3.2. Economic performances of the 1.1 kg/day DAC process

The feasibility of the DAC system was studied where such system consists of thermal and electrical energy requirements. Main thermal energy uses arise from the reaction and sensible heats of CO₂, H₂O, and steam while pumps and compressors contribute to the electrical energies. The distribution of exergy assuming an ideal heat transfer is presented in Figure 4 and the magnitude of distributions calculated in the current study are in the same magnitude as reported in Schellevis et al., (2021).

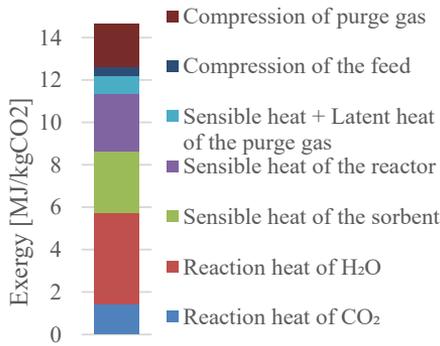


Figure 4. Distribution of exergy use

Table 3 Capital and operating expenditures

CAPEX	Value [€/yr]
Contactors	21
Fan	26
Vacuum pumps	72
OPEX	Value [€/yr]
Sorbents	0.77
Electricity	114.7
Steam	11.0
Maintenance	7.9
Labor	2.3

Annualized capital and operating expenditures are summarized in Table 4 where contactors, fans, and vacuum pumps are considered for CAPEX in this study. The costs of equipment are obtained from the per-unit costs from the manufacturers and annualized assuming a lifetime of 10 years. Since the unit size is smaller ($D = 0.4$ m, $L = 0.024$ m) than the usual industrial size, the price of each unit could be more expensive to manufacture. In the case of operating costs, sorbents are expected to have a lifespan of 2.6 years (NASEM, 2019), and electricity and steam costs are included while maintenance and labor are assumed to be 3% of total capital requirement and 30% of maintenance cost respectively. Based on the aforementioned information, a preliminary capture cost of 942 €/tonne of CO₂ was obtained for this non-optimized process, which is slightly higher than the figure reported by IEA (2020).

3.3. A preliminary sensitivity analysis

The effects of adsorption time on recovery and CO₂ capture rate per day are examined under the same operating conditions. It was observed that as the adsorption time increases, recovery decreases. This is due to the adsorption bed being fully saturated and CO₂ is being lost in the outlet stream. In the case of capture rate, by varying the adsorption time, the total number of cycles per day changes. For example, when the adsorption time is 1200 s instead of 3200 s, it is possible to achieve 10 cycles per day. However, the CO₂ capture rate per cycle (0.0163 kg_{CO₂}/cycle) is very low and requires more DAC units to achieve the same output per day when compared to the adsorption time of 3200 s which involves 8 cycles with 0.0339 kg_{CO₂}/cycle output. Therefore, there is a tradeoff between the adsorption time, recovery, and the capture rate. From the sensitivity analysis, it was found that the optimal capture rate can be achieved at the adsorption time of 3200 s as presented in Figure 5

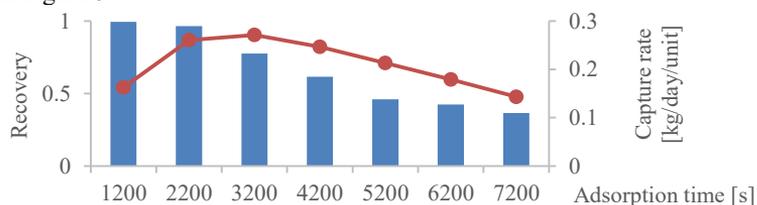


Figure 5. CO₂ recovery (bars) and CO₂ capture rate per day (line) for different adsorption times.

4. Conclusion and future work

This study has developed a fixed-bed DAC process model and presented validation for adsorption and breakthrough curves of Lewatit® VP OC 1065 amine-functionalized solid sorbents under DAC conditions with the literature. Also, a full TVSA model for DAC is developed to analyze cyclic operations as well as feasibility and economic performances. Further work needs to be done to include co-adsorption of H₂O and CO₂ on solid sorbents to better understand the influence of humidity on the adsorption stability which requires more work for accurately assessing the impact of operating conditions on sorbent performances and lifetime. Finally, these elements will lead to a more detailed techno-economic analysis and provide more insights into the relevance and possible deployment of DAC in comparison to CO₂ capture from point sources.

References

- C. Gebald, J.A. Wurzbacher, A. Borgschulte, T. Zimmermann and A. Steinfeld, 2014. Environ. Sci. Technol., 48(4), pp.2497-2504.
- IEA, 2020, Direct Air Capture, IEA, Paris <https://www.iea.org/reports/direct-air-capture>
- J. Young et al., 2021. SSRN 3814942
- K. Lackner, H. J. Ziock, and P. Grimes, 1999, (No. LA-UR-99-583). LANL.
- Lanxess. Lewatit VP OC 1065 Technical Data Sheet. Tech.
- M. J. Bos et al., 2019a, Chem. Eng. J. 377, 120374
- M. J. Bos, S. Pietersen, and D. W. F. Brilman, 2019b, Chem. Eng. Sci.: X, 2, p.100020.
- M. Schellevis, T. van Schagen, W. Brilman. 2021, SSRN 3821508.
- Q. Yu and D.W. F. Brilman, 2017. Energy Procedia, 114, 6102-6114.
- R. Veneman et al., 2012. Chem. Eng. J. 207-208, 18–26.
- V. Stampi-Bombelli, M. van der Spek, M. Mazzotti, 2020. Adsorption 26: 1183-1197
- National Academies of Sciences, Engineering, and Medicine (NASEM), 2019. Negative emissions technologies and reliable sequestration: A research agenda

Mechanistic modelling for thrips incidence in organic banana

Jean C. Campos^a, José Manrique-Silupú^{a,*}, William Ipanaqué^a, Bogdan Dorneanu^{b,c}
and Harvey Arellano-García^{b,c}

^aUniversidad de Piura, Av. Ramón Mujica 131, Piura 20009, Perú

^bUniversity of Surrey, Stag Hill University Campus, GU2 7HX Guildford, UK

^cLS- Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus D-03044, Germany

jose.manrique@udep.edu.pe

Abstract

This contribution introduces a data acquisition and modelling framework for the prediction of banana pests' incidence. An IoT sensors-based system collects weather and micro-climate variables, such as temperature, relative humidity, and wind speed, which are uploaded in real time to a cloud storage space. The incidence of the red rust thrips (*Chaetanaphothrips signipennis*) is collected "manually" by periodic inspection. The mathematical model is adapted from population growth functions and a model of insect species development and allows predictions to be made at various time intervals with an accuracy greater than 80%, improving decision-making capacity for agro-producers and enabling the improvement of pest management actions.

Keywords: Mechanistic modelling; IoT sensors; Precision agriculture; Organic banana; Red rust thrips.

1. Introduction

Peru, with an extension of organic agriculture of 197,837 ha (Sánchez Castañeda, 2017), is one of the main suppliers of organic banana to the American market (Machovina and Feeley, 2013), representing approximately 3% of the world production. However, like other crops produced in this country and worldwide, the presence of pests and diseases, influenced by climate change, is one of the main problems that afflict farmers (Dadrasnia et al., 2020; Gaitán, 2020), directly affecting the productivity and quality of the product.

Mathematical representation of the phenomena taking place on the farm (*e.g.*, plant growth, pest incidence and its impact on the plant development, etc.) are of paramount importance for improving the knowledge and decision-making for banana production. The data available shows that, in the case of the banana pests incidence, an exponential growth can be observed. Moreover, pest control measures have an effect of reducing the insect number in a similar way that antibiotics reduce bacterial populations. Various mathematical representations have been developed to describe the behaviour of bacterial population growth, which consider the effect of antibiotics on the development (Romero Leiton et al., 2011).

The problem of inefficient pest management affects small and medium agro-producers in the Piura region, impacting on the quantity and quality of their produce, and preventing them from competing on the international market. Thus, the development of a tool that improves the pest management will bring along benefits. However, to achieve this, improved representation of the pest

behaviour is required. Therefore, in the following sections, a mathematical model for predicting the level of pest incidence in banana crops is proposed, based on first principles of population growth in insect species, with roots in bacterial population growth models. The objective is to develop a tool to facilitate decision-making for organic banana agro-producers worldwide, enabling the adjustment of the model parameters to account for environments where the atmospheric and microclimate conditions differ from those studied in the current work.

2. Methodology

2.1. Data acquisition

A network of IoT sensors, Figure 1, is used to measure the weather, micro-climate and soil variables listed in Table 1.

Table 1: Features of the IoT sensors

	Level	Sensor	Variables	Manufacturer
Weather station	Climate	Vantage Pro 2	- Temperature - Relative humidity - Atmospheric pressure - Wind speed - Rainfall rate	Davis Instruments
Node 1 and Node 2	Micro-climate	Davis Instruments 6830	- Temperature - Relative humidity	Davis Instruments
	Soil	GS3 decagon	- Water content - Bulk electrical conductivity - Temperature	Decagon Devices

The data acquisition system measures these variables every minutes, while its gateway device stores it at 15 minutes intervals and uploads it to a web server. For this duration, the average, the maximum and minimum value, and variables that are not directly measured (*e.g.*, dew point, evapotranspiration, heat index, etc.) are calculated automatically. Thus, a total of 96 items are stored in the database every day. Finally, the data acquisition interface takes these values and calculates a daily average for each of the variables stored.

Based on previous studies (Elbehri et al., 2015), (Zhang et al., 2019), the variables with the highest impact on the efficiency of the process are the temperature, the relative humidity, the rainfall rate, and the wind speed. In the following, the average daily atmospheric temperature will be used, since it was found that a correlation exists between this and the growth rate of the pests.

2.2. Pest incidence

To assess the level of infection in the banana crop, the pest incidence will be used. The pest incidence is a statistical metric resulting from the inspection of the crop and is calculated by randomly choosing a defined number of plants per hectare and examining each of them for the presence of



Figure 1: Nodes and weather station distribution, dimensions in meters.

the respective pest in the stem and leaves. The following equation is used for determining the pest incidence (α) (Pasapera Cordova, 2013):

$$\alpha = \frac{\text{Total insects}}{\text{Total of plants inspected}} \quad (1)$$

Table 2: Thrips incidence level classification

Criterion	Incidence level	Control action	Cost
$0 < \alpha \leq 0.5$	Low	Leaf fumigation	18.64 USD/ha
$0.5 < \alpha \leq 1$	Medium	Stem fumigation	35.13 USD/ha
$1 < \alpha$	High	Stem fumigation and cleaning	54.04 USD/ha

Based on this measure, the corresponding control action is decided according to the criteria illustrated in Table 2.

2.3. Mathematical Modelling

Figure 2 shows the data of the incidence of red rust thrips (*Chaetanaphothrips signipennis*) from December 2019 to April 2021. During the summer (January, February and March), there is an accelerated growth of the thrips incidence, which confirms the information collected from the farmers on the field on the seasonal behaviour of the pests.

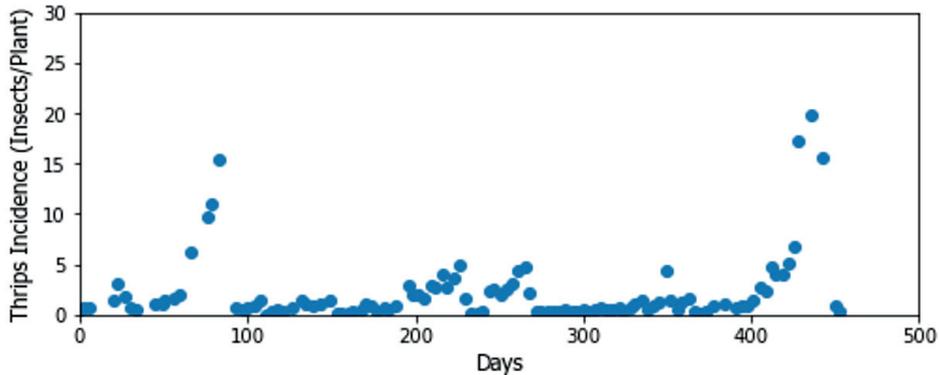


Figure 2: Thrips incidence over time.

As the behaviour in Figure 2 resembles an exponential function, a relationship can be written for the incidence of the thrips in differential form as:

$$\frac{dI_T(t)}{dt} = r \cdot I_T(t) \quad (2)$$

$$r = \beta - \mu - \alpha_1 F_1 - \alpha_2 F_2 - \gamma C \quad (3)$$

Where t is the time (day), r is the effective growth rate (1/day), and $I_T(t)$ is the incidence of Thrips as a function of time (insects/plant). The growth rate is a function of: the natural population growth rate, β (1/day), the mortality rate by natural causes, μ (1/day), the mortality rate by fumigation of thrips, α_1 (1/day), and by fumigation of other pests, α_2 (1/day), the mortality rate by cleaning γ (1/day), the concentration of the fumigation chemical for thrips, F_1 , and for other pests, F_2 , and the theoretical concentration of the plot cleaning, C .

For the modelling of the thrips development, a relationship based on temperature, following the Sharpe Schoolfield Ikemoto (SSI) function is used (Prasad et al., 2021):

$$\beta = \frac{T - T_{min}}{k} \quad (4)$$

Where T is the average daily temperature ($^{\circ}\text{C}$), T_{min} is the minimum temperature at which the species develops ($^{\circ}\text{C}$), and k is the thermal constant in degree days ($^{\circ}\text{C day}$).

To calculate the concentration of the fumigation chemicals F_i and the theoretical cleanup C a modification of the equation used to calculate the antibiotic concentration in (Esteva et al., 2011) is used, which after integration results in:

$$F_i = \int_{t_1}^{t_2} -\Phi_{F_i} F_i dt = e^{-\Phi_{F_i}(t_2-t_1)} \quad (5)$$

$$C = \int_{t_{1c}}^{t_{2c}} -\Phi_C C dt = e^{-\Phi_C(t_{2c}-t_{1c})} \quad (6)$$

Where Φ_{F_i} is the degradation rate corresponding to the fumigation chemical i , Φ_C is the degradation rate of the cleaning, t_1 is the date of the last fumigation, t_2 is the date pest incidence is evaluated after the fumigation, and $(t_2 - t_1)$ is represented in days.

2.4. Training and testing algorithm

The model is implemented as a multi-period integration, so that the next value of the thrips incidence can be predicted from the previous one. The experimental data is split into training and test sets, and the Group Shuffled Split (GSS) method of the Sci-kit learn Python library (Ojala and Garriga, 2010) is used to perform cross validation. The metrics used for the analysis of the model performance are the coefficient of determination (R^2) and the root mean square error (RMSE), calculated for each set generated with the GSS algorithm. The process is detailed in Figure 3.

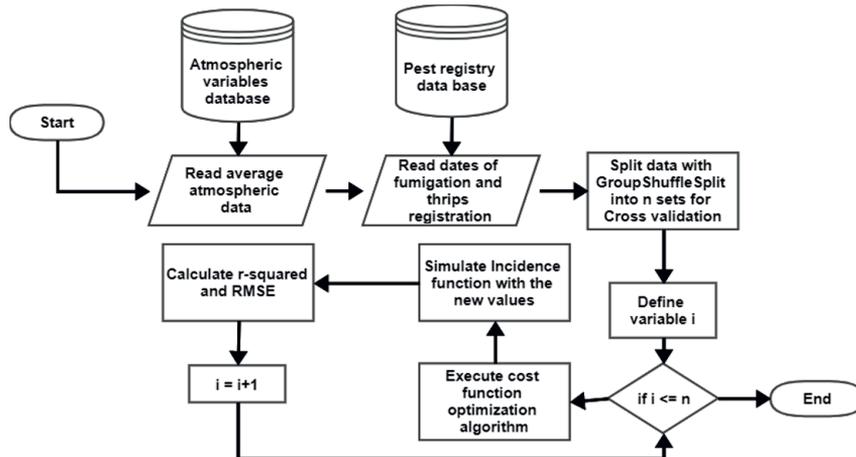


Figure 3: Flow diagram for training and testing algorithm.

3. Results

A total of 500 sets are used in the cross validation to obtain the histograms shown in Figure 4. Analysis shows that using a higher number of sets does not have a significant impact on the

results. In the training data sets, 1.56 and 0.77 are obtained as average values of RMSE and R^2 , respectively, while for the test sets the values of 1.60 and 0.68 are obtained.

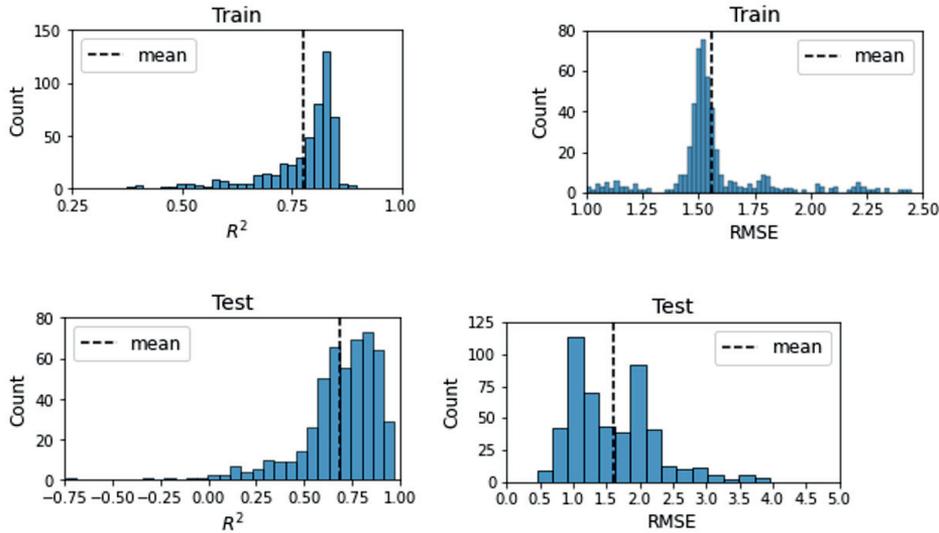


Figure 4: R^2 and RMSE histogram for training and testing sets

Table 3: Estimated parameters

Variable	Name	Value
μ	Mortality rate by natural causes	0.150195
α_1	Mortality rate by fumigation of thrips	0.529771
α_2	Mortality rate by fumigation of other pests	0.05
γ	Mortality rate by cleaning	0.0
Φ_{F_1}	Degradation rate corresponding to the thrips fumigation chemical	0.045171
Φ_{F_2}	Degradation rate corresponding to the other pests fumigation chemical	0.598177
Φ_C	Degradation rate corresponding to the cleaning	0.0496006
T_{min}	Minimum temperature at which the species develops	10.98
k	Thermal constant in degree days	0.012956

Furthermore, looking into more detail at the data set for which the best results are obtained during training and testing, the prediction of the total number of points collected is performed. The Figure 5 shows a good accuracy of the model for the prediction of the low and medium values of thrips incidence, while for large values there is a greater margin of error. This may be caused by the fact that the SSI model is linear and fails to mimic the real behaviour of population growth. However, the selected model achieves a coefficient of determination of 0.83 and an RMSE of 1.46.

4. Conclusions

This work is a starting point for the development of new techniques and models that promote artisanal agriculture, and can compete with the industrialized agricultural sector. Based on growth models with roots in bacterial population growth, a mathematical model for predicting the level

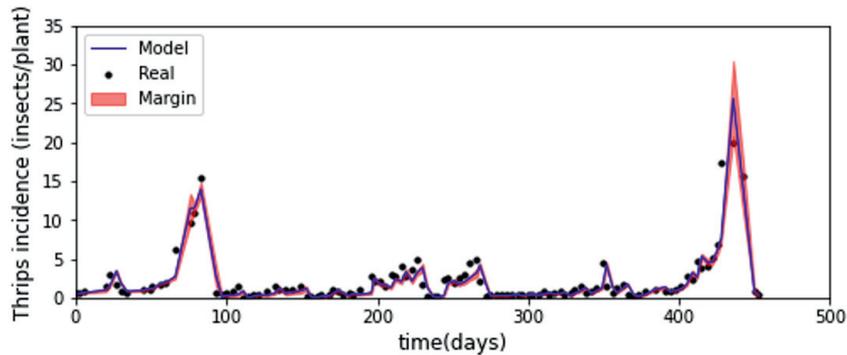


Figure 5: Comparison of model prediction and experimental data

of pest incidence in an organic banana crop is obtained. The results show that the implementation of a linear relationship based on temperature does not correctly describe the pest behaviour across the whole range of temperatures.

Future work will focus on considering more realistic, nonlinear representations of the growth rate, in order to improve the accuracy of the results. However, it should be emphasised that the current linear model gives a real physical meaning to the considered variables and offers sufficient understanding of the abiotic components and spread of the pests. Thus, it can prove a sufficiently accurate solution for implementation of fumigation or cleaning schedules, that may require iterative computations of the pest incidence model.

This will enable the farmers to improve their performance in pest control, increasing the quality of the organic product and reducing costs of spraying and maintenance of the plot.

References

- A. Dadrasnia, M. M. Usman, R. Omar, S. Ismail, R. Abdullah, 2020. Potential use of bacillus genus to control of bananas diseases: Approaches toward high yield production and sustainable management. *Journal of King Saud University - Science* 32 (4), 2336–2342.
- A. Elbehri, G. Calberto, C. Staver, A. Hospido, D. Skully, L. Ignacio Sotomayor, A. Bustamante, 2015. Cambio climático y sostenibilidad del banano en el Ecuador: Evaluación de impacto y directrices de política. Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO).
- L. Esteva, E. Ibargüen Mondragon, J. Romero Leyton, 2011. Un modelo matemático sobre bacterias sensibles y resistentes a antibióticos. *Matemáticas: Enseñanza Universitaria* 19 (2), 55–73.
- C. F. Gaitán, 2020. Chapter 7 - machine learning applications for agricultural impacts under extreme events. In: J. Sillmann, S. Sippel, S. Russo (Eds.), *Climate Extremes and Their Implications for Impact and Risk Assessment*. Elsevier, pp. 119–138.
- B. Machovina, K. J. Feeley, 2013. Climate change driven shifts in the extent and location of areas suitable for export banana production. *Ecological Economics* 95, 83–95.
- M. Ojala, G. C. Garriga, 2010. Permutation tests for studying classifier performance. *Journal of Machine Learning Research* 11 (6).
- E. Pasapera Cordova, 2013. Manejo Integrado de plagas en manejo orgánico. Ministerio de Agricultura y Mina, Lima.
- Y. G. Prasad, M. Gayathri, V. Sailaja, M. Prabhakar, G. Ramachandra Rao, G. Rekha, S. Vennila, Jun. 2021. Evaluation of linear and nonlinear models for temperature driven development of *spodoptera litura* (fabricius) on soybean crop. *Journal of Agrometeorology* 23 (2), 169–175.
- J. Romero Leiton, E. Ibargüen-Mondragon, L. Esteva, 01 2011. Un modelo matemático sobre bacterias sensibles y resistentes a antibióticos. *Matemáticas: Enseñanza Universitaria* 20, 55–73.
- J. Sánchez Castañeda, 2017. Mercado de productos agrícolas ecológicos en colombia. *Suma de Negocios* 8 (18), 156–163.
- J. Zhang, Y. Huang, R. Pu, P. Gonzalez-Moreno, L. Yuan, K. Wu, W. Huang, 2019. Monitoring plant diseases and pests through remote sensing technology: A review. *Computers and Electronics in Agriculture* 165, 104943.

A novel approach to modelling trickle bed reactors

Bogdan Dorneanu, Norbert Heinzelmann, Klaus Schnitzlein, Harvey Arellano-Garcia

LS- Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus D-03044, Germany

Abstract

In this contribution, the development of a toolbox for the simulation of trickle bed reactors based on a model able to account for the local properties of the liquid and gas flow in a packed bed at particle scale is introduced. The implementation uses a modular and flexible setup, with local liquid distribution considered as a function of the operating conditions and the physical properties of the three phases. Moreover, the impact of the local incomplete wetting on the conversion, as well as the mass transport and kinetics at both particle and reactor scale are accounted for. Furthermore, different particle geometries are considered, and the model is able to reliably predict the performance of the catalytic trickle bed reactors.

Keywords: Trickle bed reactors; Multiphase reactors; Modelling; Multiscale analysis.

1. Introduction

The trickle bed reactors (TBRs) are a frequently used solution for industrial multiphase exothermic catalytic reactions between gaseous and liquid components (e.g., hydrogenation, oxidation), in which the gas and the liquid flow downward through packed beds and undergo chemical reactions (Guo et al., 2008). The feedstock, in liquid phase forms a film around the solid catalyst, while the reactant, in a continuous gaseous phase, fills the remaining space of the catalyst bed, flowing separately (Fan et al., 2020). TBRs provide simple and safer operation at high temperature and pressure, higher conversion and selectivity due to the low axial dispersion, high interfacial area, less catalyst cost per operation run, and often lower energy consumption compared to other reactor types (Degirmenci & Rebrov, 2016; Markthaler et al., 2020; Zhao et al., 2020). On the other hand, the TBR's behavior is very complex and depends on mass and heat transfer, as well as on hydrodynamics (Qi et al., 2020). Its performance largely depends on the complex wetting morphology, which consists of film flow, rivulet flow, liquid pendulum and liquid pocket (Tang et al., 2022), with the hydrodynamic parameters (e.g., liquid holdup, gas holdup, liquid distribution, catalyst particle size, particle porosity, bed porosity) having a significant impact (Azarpour et al., 2021). However, TBRs show unacceptable levels of performance when moving from laboratory to pilot or commercial scale due to the maldistribution of the gas/liquid flow and the difficult control of the temperature (Muharam et al., 2020). Thus, the proper understanding of the TBRs remains a challenge, since numerous transport and reaction phenomena occur simultaneously inside the reactor, with uncertainties in the catalyst heterogeneity, packing, fluid flow, and transport parameters elevating the complexity (Azarpour et al. 2021). In this sense, different approaches have been developed over time to deal with their highly coupled multiphysics and multiscale nature, ranging from experimental techniques that enable global characterization of hydrodynamic behavior and reaction characteristics, to modelling and

simulation studies, with different levels of description, developed depending on the objective to satisfy, and the phenomena of interest (Qi et al., 2020). As a result, varied assumptions accompany these models, such as considering constant velocity profiles and no pressure drop for multiphase plug flow reactor (PFR) models, neglecting mass and energy radial gradients, which often result in reduced accuracy of the predictions (Uribe et al., 2019). More rigorous, computational fluid dynamics (CFD) models to simulate the multiphase flow inside the TBRs are often computationally expensive and do not always capture the wetting characteristics of the gas-liquid interface (Deng et al., 2020). Furthermore, they are not always capable of handling the multiscale coupling and the establishment of two-way communication schemes between the different scales (Uribe et al., 2019).

The focus of the paper is the improvement of a novel and intuitive tool for the design and analysis of trickle bed reactors operating in the low-interaction-regime, based on the work in Schwidder & Schnitzlein (2012). In that case, the modelling of the liquid distribution, the two-phase flow, and the reaction are modelled using an Euler-Lagrange approach. Furthermore, due to the complexity of the phenomena taking place inside the TBR, only isothermal processes are considered. To enable the extension of the tool to any type of particle geometry (*e.g.*, cylinders, Raschig rings, Pall rings, etc.), not possible using the previous model version, a new description of the surface is proposed.

2. BasMo – A simulation toolbox for trickle bed reactors

2.1. An interactive toolbox for the simulation of trickle bed reactors

The toolbox, *BasMo* (Basis Model), an interactive solution for the simulation of TBR, is implemented in C++ using a modular setup, mirroring the multiscale structure of the TBR, from reactor level to particle and fluid film levels, including reaction kinetics information, as illustrated in Figure 1.

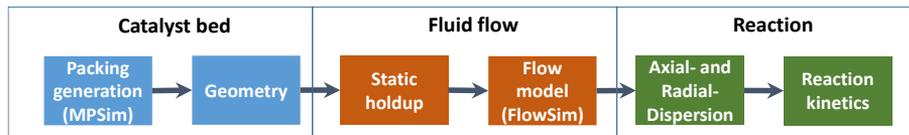


Figure 1: *BasMo* modular structure (adapted from Schwidder, 2012)

The starting point for the model development is the generation of random-loose catalyst packing bed. A layer of particles is placed at the bottom of the reactor and every particle is assumed to move with small increments in every direction. In the resulting Monte Carlos simulation, should the movement of a particle come in conflict with another, its movement will be skipped. To achieve a close packing, a “raining” mechanism is added, which assumes that a particle drops inside the vessel until it is fixed, together with a compressing algorithm based on simulated annealing, and a shacking procedure for the packing, to ensure a particle is fixed inside the bed (Schnitzlein, 2001). The resulting packing can be used directly, or saved for later simulations. During the following stage, the packing model is used to extract all the necessary global and local geometric information, to be used for the further steps, *e.g.*, the determination of system-relevant contact points, and the volume of every single bond existing at this points of contact (Schwidder & Schnitzlein, 2010). A distinction is made for the cases in which the contact point is between two packing particles or between the particle and the wall, due to the difference in wettability and curvature of the bulk or wall materials.

In the second step, using this geometric information (particle distribution, number of active and inactive particles, bed diameter and height, position of a single particle, etc.), the physical parameters and the static holdup can be predicted, as described in Schwidder & Schnitzlein (2012).

Subsequently, during the third step, based on the specified operating parameters, the liquid flow is determined. This can be used to predict the axial and radial dispersion. Furthermore, with the addition of the kinetic model for the specific heterogeneous catalytic system under consideration, other measures such as the conversion inside the reactor, or the residence time distribution can be predicted.

The modular structure of the simulation toolbox offers flexibility in adding new modules for improved representation of the internal geometries as well as of the different phenomena (e.g., holdup, fluid flow, kinetics, etc.) taking place inside the TBR, as well as expansion to energy and momentum balance.

2.2. Graph network model

The toolbox utilizes the approach introduced by Schwidder & Schnitzlein (2012) for the

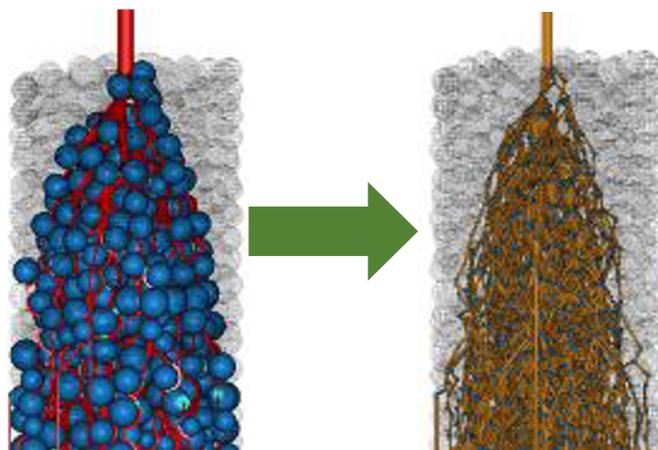


Figure 2: Three-dimensional graph network representation (right) based on the fluid distribution through the packing (left)

representation of the movement of the fluid through the bed, which is considered as a combination of two effects: the complete mixing in the cavities between the packing elements, and the displacement in the channels that connect these cavities.

A three-dimensional network (Figure 2) of edges and nodes (vertices) is created based on the knowledge on the local distribution of

the static and dynamic holdups. Furthermore, a steady-state approach is used to model the liquid distribution, while the dispersion and the reaction are modelled dynamically.

It is considered that the liquid flows in the form of rivulets on the particle surface. These rivulets are mapped by *Edges*, while the liquid within the inter-particle spaces is mapped by *Vertices*. Furthermore, an ideal PFR is used to model the rivulets, while the inter-particle liquid is modelled as a continuous-stirred tank reactor (CSTR), as

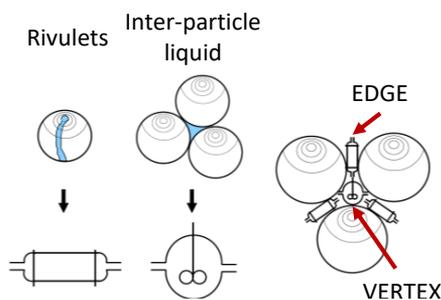


Figure 3: Representation of the liquid flow through the packing (adapted from Schwidder, 2012)

illustrated in Figure 3. Dynamic mass balances are used to describe the concentration of the two phases, accounting for both gas-solid and liquid-solid mass transfer. By connecting the two reactor types, the behavior of the liquid inside the reactor can be modelled, with experimental measurements used to account for the characteristic parameters of the reactors, mass transport, etc. As such, by using this simplified representation of the liquid as a set of PFRs and CSTRs, the computational time is significantly reduced compared to a full CFD simulation.

2.3. Extension of the particle representation inside the packed bed

An important disadvantage of the current implementation of the particle representation inside the packed bed is the fact that only simple spherical shapes can be used, which limits the applicability of the toolbox to more complex particle geometries. The geometry is accounted based on the symmetry properties of the spherical shape. To enable the use in the investigation of more realistic scenarios, methods such as the Lattice

Boltzmann (LB) or the volume of fluid (VOF) methods can be applied (Rong et al. 2020; Tang et al. 2022). However, the main drawback would be the fact that the accurate modelling of a realistic TBR implies simulation of several thousands of particles, which leads to high computational times. Furthermore, the discrete element method (DEM) has been investigated, and deemed inappropriate for the calculation of the contact points. For the improvement of the toolbox towards applicability to any arbitrary particle geometries, a new approach is considered for the representation of the surface, which will be called

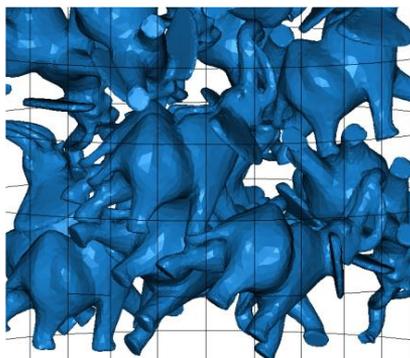


Figure 5: Representation of complex geometries

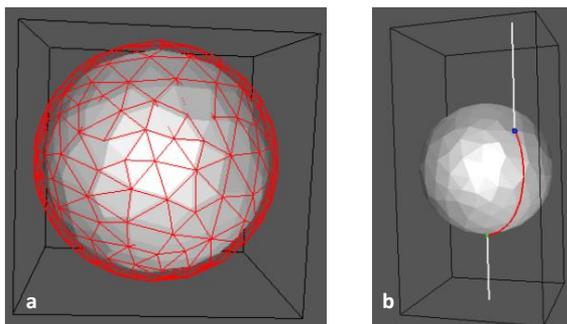


Figure 4: a) Particle surface mesh and b) liquid path flow representation using the LET method

further the *liquid element tracking* (LET) method. In the LET method, the particle surface is discretized over a finite number of triangles (Figure 4a). The path of the liquid flow (for example a rivulet) on the particle surface is then described as a pointwise sequence of motion of the fluid over individual partial surfaces, based on the applied forces. Thus, the position of impact on the particle surface can be determined (Figure 4b). The resolution of the discretization can be changed depending on the desired accuracy, although fine meshing will increase the model complexity, and lead to higher CPU times. Thus, complex particles can be represented, for example (although not a realistic catalyst shape) elephants (Figure 5).

2.4. Validation of the LET surface mesh representation

Since the spherical particle representation existing in the previous version of BasMo has been validated experimentally in the precedent work (Schwidder & Schnitzlein, 2010; 2012), in the following, the results based on the LET-approach will be validated against its predictions for spherical particle shapes.

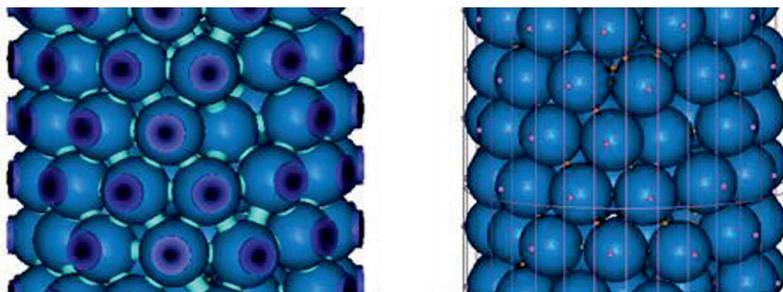


Figure 6: Prediction of static holdup based on BasMo v1 (left) and v2 (right) particle representation

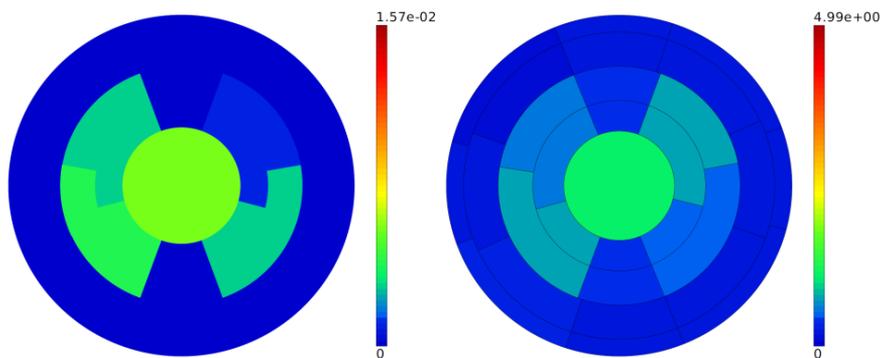


Figure 7: Prediction of the radial flow distribution

Once the information on the geometry is available, the static holdup can be calculated, with acceptable accuracy (Figure 6). The static holdup is determined from the volume of every single bond existing at a contact point, as described in Schwidder & Schnitzlein (2010). Finally, the radial flow distribution can be predicted, as shown in Figure 7, using the models developed in Schwidder & Schnitzlein (2012).

3. Conclusions

This contribution introduces improvements done in the implementation of a modular toolbox for the simulation of trickle bed reactors. The implementation of BasisModel (BasMo) in C++ mirrors the multiscale nature of the phenomena taking place in the reactor, moving from the large scale of the reactor to the medium and low scale of the particle bed, fluid flow, as well as the fluid-solid and fluid-fluid interactions, including the chemical reactions. The toolbox allows for the implementation of complex geometries for the catalyst particles, enabled by the new representation of the surface mesh. A graph network model representation is used for the liquid flow, which enables the calculation of the properties based only on the local geometry, reducing significantly the

computational effort, without hindering the accuracy. Although the new implementation of the surface mesh shows in general good agreement with the previous version of the toolbox, validated experimentally in the previous work, significant experimental work is still ongoing to validate the models for complex, non-spherical geometries, for which new models for the static holdup or the radial dispersion may need to be derived.

The knowledge gathered from the modelling of the liquid phase, and the liquid-solid interactions, as well as improvements on the particle surface representation enable new features to be displayed on the simulation toolbox in order to represent accurately the TBR under real operating conditions. Further improvements are envisaged which include the addition of enhanced modules for the continuous gas flow, liquid-gas interactions, as well as momentum balances, to improve the representation of the fluid pathways inside the bed.

References

- A. Azarpour et al., 2021, Performance analysis and modelling of catalytic trickle-bed reactors: A comprehensive review, *Journal of Industrial and Engineering Chemistry* 103, pp. 1-41
- V. Degirmenci & E.V. Rebrov, 2016, Design of catalytic micro trickle bed reactors, *Physical Sciences Reviews* 1 (4), pp. 174-219
- H. Deng et al., 2020, Computational investigation of liquid holdup and wetting efficiency inside trickle bed reactors with different catalyst particle shapes, *Applied Sciences* 10, 1436
- X. Fan et al., 2020, Modelling and simulation of industrial trickle bed reactor hydrotreating for whole fraction low-temperature coal tar simultaneous hydrodesulfurisation and hydrodenitrification, *Fuel* 269, 117362
- J. Guo et al., 2008, Modelling of trickle-bed reactors with exothermic reactions using cell network approach, *Chemical Engineering Science* 63, pp. 751-764
- S. Markthaler et al., 2020, Numerical simulation of trickle bed reactors for biological methanation, *Chemical Engineering Science* 226, 115847
- Y. Muharam et al., 2020, Modelling of partial hydrogenation of polyunsaturated fatty acid methyl esters in a trickle bed reactor, *Engineering Journal* 24 (4), pp. 1-10
- B. Qi et al., 2020, Development of a hybrid pressure drop and liquid holdup phenomenological model for trickle bed reactors based on two-phase volume averaged equations, *The Canadian Journal of Chemical Engineering* 99 (8), pp. 1811-1823
- L.W. Rong et al., 2020, Lattice-Boltzmann computation of hydraulic pore-to-pore conductance in packed beds of uniform spheres, *Chemical Engineering Science* 224, 115798
- S. Schwidder & K. Schnitzlein, 2010, Predicting the static liquid holdup for cylindrical packings of spheres in terms of the local structure of the packed bed, *Chemical Engineering Science* 65 (23), pp. 6181-6189
- S. Schwidder & K. Schnitzlein, 2012, A new model for the design and analysis of trickle bed reactors, *Chemical Engineering Journal* 207-208, pp. 758-765
- S. Schwidder, 2012, Reaktionstechnische Modellierung von Rieselbettreaktoren auf der Basis der lokalen Packungsstruktur, PhD Thesis
- K. Schnitzlein, 2001, Modelling of radial dispersion in terms of the local structure of packed beds, *Chemical Engineering Science* 56, pp. 579-585
- P.L. Spedding & R.M. Spencer, 1995, Simulation of packing density and liquid flow in fixed beds, *Computers and Chemical Engineering* 19 (1), pp. 43-73
- Y. Tang et al., 2022, Effects of stagnant zone on the effectiveness factor in a trickle bed, *Chemical Engineering Science* 248, 117211
- S. Uribe et al., 2019, Multiscale CFD modelling and analysis of TBR behaviour for an HDS process: Deviations from ideal behaviours, *Fuel* 239, pp. 1162-1172
- X. Zhao et al., 2020, Coupling non-isothermal trickle-bed reactor with catalyst pellet models to understand the reaction and diffusion in gas oil hydrodesulfurization, *Chinese Journal of Chemical Engineering* 28, pp. 1095-1106

Alkaline Water Electrolysis Model to Purify GMP grade NaOH Solutions for Biopharmaceutical Manufacturing Processes

^aYeonghyun Kim, ^aYoungjin Kim, ^aJae Hyun Cho, ^aIl Moon

^aYonsei University 50 Yonsei-ro Seodaemun-gu, Seoul, 03722, Republic of Korea

Abstract

It is essential that the Cl⁻ ions purity contained in NaOH solutions for the use of the biopharmaceutical production processes would be below 30 ppm to meet Good Manufacturing Practice (GMP) guidelines (Gilleskie, G. et al.). The development of a commercially efficient process for purifying NaOH 50% aqueous solution of such purity is of interest in the related industry. This study describes a lab-scale electrochemical mathematical model for producing GMP grade NaOH aqueous solutions and a design of electrochemically efficient lab-scale production configuration. The model validation is carried out with experimental data for a different set of membrane configurations (Marangio, F. et al.). The model calculates the theoretical open-circuit voltage via a thermodynamic analysis of the electrochemical purifying process and then outputs the expected voltage during operation by applying the Butler-Volmer equation (Jang, D et al.). A set of physical properties of the solution are obtained from well-known theoretical equations: concentrations from Henry's law and membrane diffusivity, permeability, and ionic conductivity from the Arrhenius equation, respectively. An experimental data fitting makes it possible to obtain estimated values of critical process parameters and their data tendency at different temperatures. The proposed model shows about 90% accuracy at an electrochemical lab-scale, to produce GMP grade (Cl⁻ ions below 30 ppm) 50% NaOH aqueous solution, and an appropriate cation exchange membrane-based experimental configuration turns out to be of electrochemically high efficiency. Future work is to enhance the experimental scale model to scale up to a commercial scale, with an emphasis on the optimization of electrochemical reaction time, power and voltages.

Keywords: Alkaline Water Electrolysis (AWE), prediction model, NaOH purification.

1. Introduction

With the ongoing globally pandemic COVID-19, the production of biopharmaceuticals is getting more attention as vaccines are highly demanded worldwide. In particular, efficient purification of NaOH up to GMP grade NaOH solutions becomes vital in biopharmaceutical manufacturing processes as it effectively removes proteins and nucleic acids and inactivates most viruses, bacteria, yeasts, fungi, and endotoxins (Giro, P et al.). As such, it is essential to produce GMP-grade NaOH of high purity. In this study, Alkaline Water Electrolysis (AWE) is proposed as an experimental configuration for purifying a 50% aqueous NaOH solution at GMP grade, specifically targeting below 30 ppm of Cl⁻ ion in 50% NaOH solution, and its electrochemical model is developed and verified with experimental data.

2. Experiment

The proposed experimental framework comprises two electrodes immersed in a liquid alkaline electrolyte (See Fig. 1), and its specifications are given in Table 1 (Zeng, K et al.). As the NaOH solution in an ionic state undergoes electrolysis, the electrochemical reactions described in Eqs. 1 and 2 take place where hydroxide ions react to produce oxygen and water at the anode, and hydroxide and hydrogen are generated at the cathode. The characteristic of the membrane, which selectively allows Na^+ ions passing through from the anolyte to the catholyte, leads Cl^- ions to get diluted in the catholyte. As the reaction time goes by running the experiment, the concentration of the Cl^- ion-free catholyte increases, and that of the anolyte decreases. It takes around 200 hours to reach 19M of catholyte.

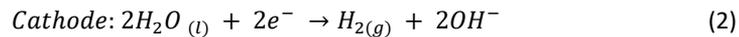
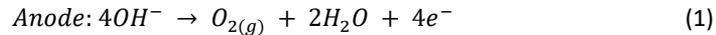


Table 1. Operating conditions of AWE

Parameters	Unit	Value
Operating temperature	K	313.15
Operating pressure	atm	1
Anode electrode size/shape	cm^2	4/Rectangular
Anode electrode material	-	Ni foam
Anode electrode thickness	mm	1
Anolyte concentration	wt% NaOH	50 (=19M)
Anolyte volume	ml	200
Cathode electrode size/shape	cm^2	4/Rectangular
Cathode electrode material	-	Ni foam
Cathode electrode thickness	mm	1
Catholyte concentration	wt% NaOH	5.3 (=2M)
Catholyte volume	ml	150
Membrane(Nafion117)	μm	183

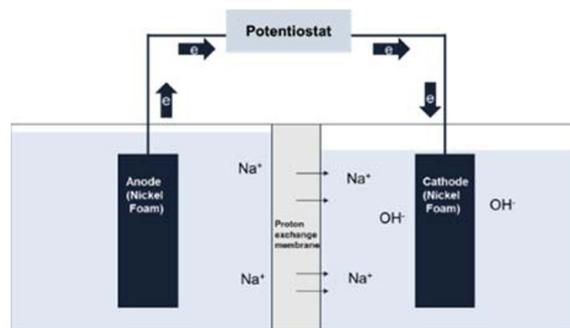


Fig. 1. Reaction mechanism of Alkaline Water Electrolysis (AWE)

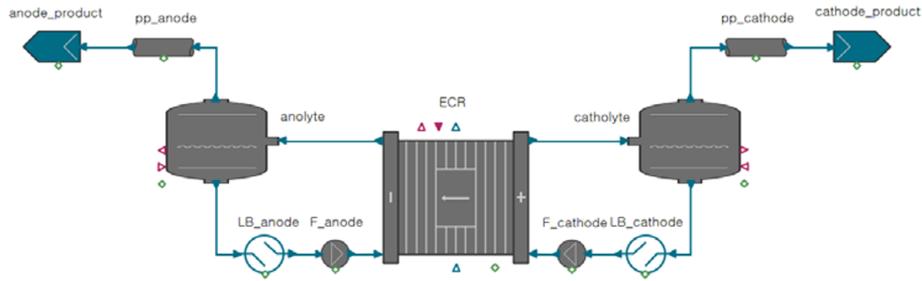


Fig. 2. Conceptual framework of the AWE model

3. Model and its Verification

3.1. Alkaline Water Electrolysis (AWE) Model

An AWE model is developed to derive the dynamics of the AWE based on gPROMS PROCESS™, which is operated in batch mode (See Fig. 2.) The initial amounts of anolyte and catholyte in the tanks are 200 mL of anolyte, 150 mL of catholyte, respectively. It is assumed that the fluids are circulated between the tanks (anolyte and catholyte) and the Electrochemical Cell Reactor (ECR) in Fig. 2 simply as a way to make sure the fluid in the tanks interacts with the membrane and electrodes in order to represent batch operation in reality. 0.5A of current is kept constant in the entire experiment time. g-SAFT (gPROMS physical property package) is selected as its physical property model for materials involved in the experiment. The operating temperature of the current reactor is 40°C, and the temperature change was found to be less than 1°C. Therefore, the reaction is assumed to be isothermal. Since there is no change in pressure as operated in the atmosphere, the reaction is isobaric. The AWE model shows the trends of electrolytic cell voltage as well as the catholyte concentration. The total voltage of the electrolysis is determined by summing up the open-circuit voltage, the activation and ohmic over-voltages as Eq. 3 below (Han, B et al.):

$$V = E_{oc} + \eta_{act} + \eta_{ohm} \quad (3)$$

The open-circuit voltage is calculated based on the Nernst equation.

$$E_{oc} = E_o + \frac{RT}{zF} \ln\left(\frac{p_{H_2} p_{O_2}^{0.5}}{p_{H_2O}}\right) \quad (4)$$

where R, T, p, and F are the universal gas constant, operating temperature, partial pressure, and Faraday constant, respectively; p_{H_2} & p_{O_2} i the partial pressure of hydrogen and oxygen gases near the electrode under Henry's law. E_o , the standard electrode potential is calculated by the Eq. 5 (Han, B et al. & Prentice, G.):

$$E_o = 1.229 + 0.9 \times 10^{-3}(T - 298.15) \quad (5)$$

The activation over-voltage is the consumed voltage caused by the electrochemical reaction (Jang, D et al.). It occurs at both electrodes, and the over-voltage on the anode appears larger than the cathode in most cases (Hammoudi, M et al.). The activation over-voltages can be determined by the Butler-Volmer equation (Prentice, G):

$$\eta_{act} = \frac{RT}{\alpha_{an}F} \sinh\left(\frac{i}{2i_{o,an}}\right) + \frac{RT}{\alpha_{ca}F} \sinh\left(\frac{i}{2i_{o,ca}}\right) \quad (6)$$

where α_{an} and α_{ca} represent the charge transfer coefficients of anode and cathode, respectively. The exchange current densities of anode and cathode are expressed as $i_{o,an}$ and $i_{o,ca}$, respectively. The exchange current density has a significant influence on the activation over-voltage (Zhang, Z. et al.). The exchange current density coefficients of the Ni electrodes used for the anode and cathode are calculated, using Eqs. 7 and 8, with data provided in (Henao Diaz, C).

$$i_{o,ca} = 1.5 \times 10^{-4} \left(\frac{P}{P_{ref}} \right)^{0.1} \exp \left[-\frac{23000}{RT} \right] \left(1 - \frac{T}{T_{ref}} \right) \quad (7)$$

$$i_{o,an} = 0.9 \times 10^{-4} \left(\frac{P}{P_{ref}} \right)^{0.1} \exp \left[-\frac{42000}{RT} \right] \left(1 - \frac{T}{T_{ref}} \right) \quad (8)$$

The charge transfer coefficients are calculated using Eqs. 9 and 10 with data provided in Ref. [9].

$$\alpha_{an} = 0.07835 + 0.0017T \quad (9)$$

$$\alpha_{ca} = 0.1175 + 0.00095T \quad (10)$$

In general, the transfer of ions or electrons causes an ohmic over-voltage which is present at electrodes, electrolytes, and membranes, comprised of the electrolysis cell. The ohmic over-voltage is generated by the resistance encountered by the ions to the flow through the membrane (Marangio, F et al.). The obstruction to the flow of ions could be due to the ionic conductivity of the membrane and gas bubble coverage of the active area (Jang, D et al.). Choi et al. express this mechanism as below:

$$\eta_{ohm} = \delta_m \frac{i}{\sigma_m} \quad (11)$$

The conductivity, σ_m in Eq. 11, is calculated by Springer et al., which also considers the membrane hydration.

$$\sigma_m = (0.005139\lambda - 0.00326) \exp \left(1268 \left(\frac{1}{303} - \frac{1}{T} \right) \right) \quad (12)$$

Membrane diffusivity, permeability and ionic conductivity are calculated based on the Arrhenius equations, respectively.

$$k_i^{diff} = k_{i,ref}^{diff} \exp \left(-E_i^{diff} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right) \quad (13)$$

$$k_i^p = k_{i,ref}^p \exp \left(-E_a^p \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right) \quad (14)$$

$$\sigma_i = \sigma_{i,ref} \exp \left(-k_{\sigma_i} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right) \quad (15)$$

where *diff*, *p*, *ref*, *E*, σ_i and *T* indicate the diffusivity coefficient, permeability coefficient, reference, activation energy, ionic conductivity and temperature.

3.2. Na⁺ ions mass balance

The mass balance of Na⁺ represents the dynamic of Na⁺ at the anode, which includes its migration to the cathode through the membrane. Eq. 16 illustrates the molar flow rate of Na⁺ corresponds to that of OH⁻ at the anode and cathode. The molar flow rate of Na⁺ by diffusion ($\dot{N}_{Na^+,d}$) and electroosmotic drag ($\dot{N}_{Na^+,eo}$) account for that of Na⁺ via the membrane in Eq. 17. The flux of water accompanied by the flow of Na⁺ is characterized by the electroosmotic drag coefficient (n_d), which is an inherent property of the membrane. Eq. 19 describes the Na⁺ concentration gradient present in the NaOH solution across the membrane according to Fick's law of diffusion, where A, D_w and C are membrane area, diffusion coefficient and concentration. The following equations apply:

$$\dot{N}_{Na^+,m} = \dot{N}_{OH^-,an} = -\dot{N}_{OH^-,cat} \quad (16)$$

$$\dot{N}_{Na^+,m} = \dot{N}_{Na^+,d} + \dot{N}_{Na^+,eo} \quad (17)$$

$$\dot{N}_{Na^+,eo} = \frac{n_d i}{F} \quad (18)$$

$$\dot{N}_{Na^+,d} = \frac{AD_w}{\partial m} (C_{Na^+,cat} - C_{Na^+,an}) \quad (19)$$

3.3. Model Verification

Fig. 3 shows the transient variation of NaOH concentration in the catholyte to verify the proposed model as compared to the experimental data. In the case of the Nafion 117, the experimental data show 240 hours of the reaction time at the cathode to reach 19M, and the reaction time calculated by the model is approximately 235 hours. This yields 5.69 of a root mean square deviation. In the case of the thinner Nafion 212, the reaction times of the experiment and models are 171 and 165 hours, respectively, showing 6.15 of the root mean square deviation. The proposed model shows good agreement with the experimental data at our laboratory scale.

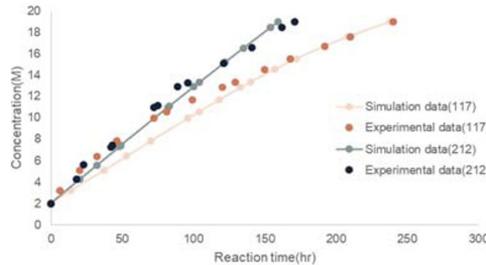


Fig. 3. Transient variation of NaOH concentration in the catholyte

4. Conclusion and Future Work

In this paper, the cation membraned-based Alkaline Water Electrolysis (AWE) model to purify GMP grade NaOH solutions (below 30 ppm Cl⁻ contained in 50% NaOH solutions)

is proposed, using gPROMS PROCESSTM and verified with experimental data, showing 90% accuracy. As future work, the proposed model at a laboratory scale will be developed further for commercial biopharmaceutical manufacturing processes.

Acknowledgement

This work was supported by Bio Industrial Technology Development Program (20009668, Development of Sodium hydroxide (NaOH) purification technology for biomedicine production) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea)

References

- Marangio, F., Santarelli, M., & Cali, M. (2009). Theoretical model and experimental analysis of a high pressure PEM water electrolyser for hydrogen production. *International journal of hydrogen energy*, 34(3), 1143-1158.
- Zhang, Z., & Xing, X. (2020). Simulation and experiment of heat and mass transfer in a proton exchange membrane electrolysis cell. *International Journal of Hydrogen Energy*, 45(39), 20184-20193.
- Jang, D., Cho, H. S., & Kang, S. (2021). Numerical modeling and analysis of the effect of pressure on the performance of an alkaline water electrolysis system. *Applied Energy*, 287, 116554.
- Zeng, K., & Zhang, D. (2010). Recent progress in alkaline water electrolysis for hydrogen production and applications. *Progress in energy and combustion science*, 36(3), 307-326.
- Ogumerem, G. S., & Pistikopoulos, E. N. (2020). Parametric optimization and control for a smart Proton Exchange Membrane Water Electrolysis (PEMWE) system. *Journal of Process Control*, 91, 37-49.
- Giroto, P., Moroux, Y., Duteil, X. P., Nguyen, C., & Boschetti, E. (1990). Composite affinity sorbents and their cleaning in place. *Journal of Chromatography A*, 510, 213-223.
- Hammoudi, M., Henao, C., Agbossou, K., Dubé, Y., & Doumbia, M. L. (2012). New multi-physics approach for modelling and design of alkaline electrolyzers. *International Journal of Hydrogen Energy*, 37(19), 13895-13913.
- Han, B., Mo, J., Kang, Z., Yang, G., Barnhill, W., & Zhang, F. Y. (2017). Modeling of two-phase transport in proton exchange membrane electrolyzer cells for hydrogen energy. *International Journal of Hydrogen Energy*, 42(7), 4478-4489.
- Henao Diaz, C. C. (2011). *Modélisation multi-physique et électrique d'un électrolyseur alcalin* (Doctoral dissertation, Université du Québec à Trois-Rivières).
- Linstrom, P. (1998). Nist chemistry webbook, nist standard reference database number 69. J. Phys. Chem. Ref. Data, Monograph, 9, 1-1951.
- Prentice, G. (1991). *Electrochemical engineering principles* (Vol. 1). Englewood Cliffs, NJ: Prentice Hall.
- Gilleskie, G., Rutter, C., & McCuen, B. (2021). *Biopharmaceutical Manufacturing: Principles, Processes, and Practices*. Walter de Gruyter GmbH & Co KG.
- Choi, P., Bessarabov, D. G., & Datta, R. (2004). A simple model for solid polymer electrolyte (SPE) water electrolysis. *Solid State Ionics*, 175(1-4), 535-539.
- Springer, T. E., Zawodzinski, T. A., & Gottesfeld, S. (1991). Polymer electrolyte fuel cell model. *Journal of the electrochemical society*, 138(8), 2334.

Effect of air dynamics on the discharge of a pharmaceutical powder using the discrete element method

L. Naranjo^{a,c}, I. Nopens^b, T. De Beer^c, A. Kumar^{a,*}

^a*Pharmaceutical Engineering Research Group (PharmaEng), Department Pharmaceutical Analysis, Ghent University, Ottergemsesteenweg 460, Ghent, 9000, Belgium*

^b*BIOMATH, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure Links 653, Ghent, 9000, Belgium*

^c*Laboratory of Pharmaceutical Process Analytical Technology (LPPAT), Department Pharmaceutical Analysis, Ghent University, Ottergemsesteenweg 460, Ghent, 9000, Belgium*

**ashish.kumar@ugent.be*

Abstract

In the pharmaceutical industry, the flow of fine powders between different manufacturing steps can be challenging due to both the solid and liquid-like behavior of bulk solids. To better understand flow behavior, the use of predictive models, such as the discrete element method (DEM) has increased in recent years. However, a general model for fine bulk solids is not yet available. This is mainly due to the long computational time requirements and the complexity of the flow, as the discharge of fine particles is influenced both by cohesive forces and interaction with the surrounding air.

In this study, a DEM model was used to predict the flow dynamics of a free-flowing powder SuperTab[®] 11SD and to determine the extent to which the model can predict the discharge behavior of fine particles from a silo. Flow behavior was comprehensively characterized in terms of flow profile development over time, mass flow index, and residence time distribution. Furthermore, mass flow rate predictions were verified considering experimental data of a laboratory-scale cylindrical silo for different aperture sizes. Lastly, in order to obtain an accurate DEM prediction, a leaner approach was implemented by introducing a relation that takes into account the aerodynamic drag effect on the system.

Keywords: Fine particles, Discrete Element Model, Computational Fluid Dynamics

1. Introduction

In continuous manufacturing of solid dosage pharmaceuticals, powder flow in silos or bins is critical to the process performance and end-product quality. The flow properties of powders vary with each material's chemical and physical characteristics, which include particle size and shape, bulk density, compressibility, cohesive strength, moisture, and the material's behaviour under vacuum, atmospheric, and loading conditions. Often, due to cohesive and poorly flowing powders, it can be challenging to accurately feed, blend, and transfer between different manufacturing steps, despite using tooling such as agitators and rotating bridge breakers to obtain uniform flow. Moreover, in pharmaceutical applications where fine particles are commonly handled, in addition to the relevant effect of the cohesive forces that affect flow behavior, the interaction between interstitial air and particles also plays an important role, since the backflow of air can cause inconsistent discharge rate.

Owing to the degree of challenge and despite the increasing research in particle technologies for pharmaceutical applications, a general model to predict such powder flow behavior is still missing and the development of first-principles models to obtain mechanistic insights is limited. This is often attributed to the exhibition of both solid and liquid-like behavior of the bulk solids. Commonly used models are based on empirical approximations for cohesionless coarse particles that are not influenced by aerodynamic drag (e.g. Beverloo, Rose and Tanaka) Gu (1991), multivariate predictions that require collection of a large amount of experimental data to generate a model, and more recently, computational models that use fundamental first principles to capture the underlying physical phenomena and predict the overall bulk behaviour, such as finite element method (FEM), computational fluid dynamics (CFD), and Discrete Element Method (DEM). There is a growing interest within the pharmaceutical industry to develop predictive models for fine powder discharge to avoid non-uniform flows and discharge rate variations, which can further affect downstream operations. While DEM has been extensively used to predict powder discharge, it remains underutilized in the pharmaceutical industry due to the required computational time to simulate particle dynamics in the micron size range, stringent process constraints, the wide variability of powder and bulk properties, and the failure to capture the interstitial air effect on fine particles. This last aspect is mostly resolved by the computationally intensive DEM-CFD coupling approach (Hesse et al., 2020). In this study, a leaner approach is proposed by using scaled-up particles and introducing a relation that can be directly applied to the DEM results to obtain an accurate prediction of discharge rates in silos when handling fine particles. Therefore, a preliminary basis for the application of DEM models to study discharge dynamics in a silo is provided for a free-flowing powder (SuperTab[®] 11SD spray-dried lactose monohydrate) and determine to what extent the model can predict the discharge behavior of fine particles.

2. Methodology and implementation

The DEM methodology is exemplified by an application case of discharge dynamics in a laboratory silo for a free-flowing powder (SuperTab 11SD[®]). The implementation stages included bulk calibration of DEM input parameters for the scaled particles. Subsequently, a qualitative comparison of powder flow patterns in two different silos (flat bottom and converging wall), determination of flow uniformity, and quantitative comparison between the residence time of the particles located in the upper layer of the powder bed. Then, a verification of the model results was performed by comparing the experimental and numerical discharge rate values of a cylindrical laboratory-scale silo for a set of aperture sizes was performed. Lastly, a relation was implemented to include the effect of interstitial air and rectify the mismatch between the discharge rate prediction by DEM and experimental values.

2.1. Numerical method

This contribution focuses on one of the most commonly used mechanistic models for granular flow, DEM. In this numerical method, the macroscopic material behavior is predicted based on the mechanical dynamics of discrete particles. Each particle is treated as a discrete element with a specific position and velocity which is estimated for each time step based on the tangential and normal contact forces exerted on the particles and Newton's second law. This process is repeated iteratively to track the dynamic motion of particles at any time. The rotational and translational motion are calculated based on the following Eq. 1 and Eq. 2 respectively. Where I , ω , M_p , t , v , m , F_g , F_c , F_{nc} , and F_f respectively correspond to moment of inertia, angular velocity, contact torque, time, translational velocity, mass of the particle, gravitational force, contact forces (e.g. elastic, plastic), non contact forces (e.g., Van der Waals, electrostatic) and particle–fluid interaction forces (e.g. pressure gradient, drag).

$$I \frac{d\omega}{dt} = M_p \quad (1) \quad m \frac{dv}{dt} = F_g + F_c + F_{nc} + F_f \quad (2)$$

This approach has been implemented in various pharmaceutical manufacturing processes such as blending, coating, and tableting, and it has proven to be an effective tool to improve the understanding of the underlying interaction of material properties, process settings, and equipment design Yeom et al. (2019). Despite its benefits, DEM application in the pharmaceutical industry is still limited as it can be significantly computationally intensive to simulate particle dynamics in the micron size range.

2.1.1. Contact model

The commonly used soft-sphere discrete simulation approach is considered in this study. In this method, discrete particles are allowed to overlap to represent the deformation that occurs during collisions. The contact models use the amount of overlap between particles to calculate the magnitude of forces acting on normal and tangential directions. In pharmaceutical application, the selection of a contact model that considers the elastic and plastic properties for non-cohesive and cohesive materials is highly relevant to adequately represent the interaction phenomena and overall bulk behavior. In this study, the commonly used Hertz-Mindlin contact model with JKR (HM + JKR) is implemented. This is a non-linear elastic model that takes into account cohesion, allowing the representation of materials as dry powders. This model has been applied in several pharmaceutical manufacturing applications such as blending and granulation Yeom et al. (2019). A detailed description of the model and governing equations can be found in Johnson et al. (1971).

2.2. Model parameter determination

The model parameters used in this study are subdivided into two types, intrinsic parameters and contact parameters. The intrinsic parameters that depend on the material properties such as density, shear modulus and Poisson's ratio are fixed from literature values. The contact parameters, such as surface energy and the coefficients of static and rolling friction, are not established directly from experimental or literature values but are calibrated with virtual experiments.

Therefore, the calibration procedure is thus performed to gain confidence in the modeling results, by linking the material properties to the actual bulk behavior. The calibration methodology used in this study consists of the comparison between experimental tests carried out with the real material and the virtually replicated test. The DEM input parameters are adjusted to match the experimental bulk response with the prediction from the experiments.

The calibration test is selected to replicate the stress state and flow regimes expected of the real application case. Therefore, as low stresses and quasi-static regime flow are expected in the selected silo, the most suitable calibration test is considered to be the ring shear cell. The calibrated factors include rolling friction (particle-particle, particle boundary), static friction (particle-particle, particle boundary), and JKR surface energy, and results between the experimental bulk response and the virtual experiment prediction are compared in terms of apparent cohesion, yield locus slope, and pre-shear point. Detailed information regarding the set of experiments and its results is considered out of the scope for this study.

The DEM input parameters, including the calibrated values for SuperTab[®] 11SD, are shown in Table. 1. Since the selected set of parameters achieves a satisfactory representation of the behavior of the powder in the ring shear cell, these are later used in the silo discharge model.

2.3. Silo model setup

The discrete element method (DEM) is applied to model silo discharge flow of a free-flowing powder (SuperTab[®] 11SD spray-dried lactose monohydrate), using the software package EDEM[®],

Table 1: DEM input parameters

Parameter	Value
Particle density [kg/m^3]	1530
Particle coefficient of restitution	0.5
Shear modulus [Pa]	5×10^6
Poisson's ratio	0.25
Static friction particle-particle	0.6
Rolling friction particle-particle	0.1

considering gravity as an external force and Hertz-Mindlin with JKR as contact model. The computational domain consists of a quasi-three dimensional silo with parallel periodic boundary conditions. Particles are represented by scaled monodisperse bi-spheres. The general computational procedure for each run consist of three steps. First, based on the set of calibrated parameters particles are created, inserted into the closed silo, and allowed to settle under gravity into a static state. Second, the silo outlet is open and particles are allowed to discharge by gravity. Lastly, the simulation continued until all the particles exit the silo. The developed DEM silo model is used to qualitatively compare the changes in flow profile development (e.g., mass flow, funnel flow) depending on the silo geometry (e.g., flat bottom, converging walls) for a specific material.

In addition to the visualization of the flow patterns, the mass flow index (MFI) and the residence time distribution (RTD) are estimated using Eq. 3 and Eq. 4. The MFI is applied to more accurately classify the developed flow profile between mass flow and funnel flow, where MFI values lower than 0.3 correspond to funnel flow. The MFI estimation relates the particle average velocities near the wall (v_{wall}) and at the center of the silo ($v_{centerline}$), providing an estimation of flow uniformity within the hopper Ketterhagen et al. (2009). The RTD was used to analyze the flow behavior in an arbitrary spatial region in the silo, where $F(t)$ represents the fraction of particles that leave the silo that have spent a time t or less. In this study, the particles located in the top layer of the powder bed are tracked to verify the flow sequence.

$$\frac{v_{wall}}{v_{centerline}} = MFI \quad (3) \quad \int_0^t E(t)dt = F(t) \quad (4)$$

2.4. Model verification

As previously mentioned, during silo discharge of fine powder the air effect is highly relevant as it can be an impediment to the gravitational powder flow and consequently reduce the discharge rates. Since in the proposed DEM model no CFD coupling was involved, an overprediction of the discharge rates could be expected. Therefore, to verify the suitability of the DEM model for the prediction of discharge rates, an instrument to measure the flow rate of powders passing through apertures of various diameters (24mm, 28 mm, and 32mm) is used and the discharge rate values from the DEM model and the experiments for SuperTab[®] 11SD are compared.

The computational domain consists of a laboratory-scale cylindrical silo where the calibrated particles are allowed to be discharged by gravity. Throughout the discharge process, the mass flow rate is monitored to verify the changes over time and the average value is estimated. The verification model allows a quantitative comparison between the discharge rates obtained with the DEM model and the experimental values for different aperture sizes. Based on this comparison, the need to include a relation that account for air impediments in the system is highlighted. The relation proposed by Wikström et al. (2021) is implemented in Eq. 5, where g , ε , D_p , D , μ_A , and M respectively correspond to gravity constant, voidage, particle diameter, outlet diameter, air viscosity, and mass flow rate. The equation considers a balance between the gravitational effect and the aerodynamic resistance, neglecting the inertial effects of fine particles.

$$\frac{\pi g \varepsilon^2 \rho^2 D_p D^2}{720 \mu_A} = M \quad (5)$$

3. Results and discussion

Based on the selected set of calibrated parameters, shown in Table. 1 and the simplified silo geometry the silo application case is implemented. The results in terms of developed flow profiles in time, flow uniformity (MFI), and residence time distribution for particles located in the top layer of the powder bed are shown in Fig. 1.

The results shown in Fig. 1 indicate that the flow profile developed depends on the silo geometry. This finding is in agreement with the observations of Ketterhagen et al. (2009) for granular flow and Jenike's design methodology studies (Mehos, 2018). Jenike's required silo angle to prevent funnel flow is 30° , which matches the predicted flow pattern for the converging wall silo (angle 35°). In general, as the flow progresses in time the flat bottom silo develops a central channel with the formation of stagnant areas, while in the case of the silo with converging walls a uniform mass flow develops. These results are corroborated by the values of mass flow index, which indicate a uniform flow for the silo with converging walls, with a MFI higher than 0.3. The results regarding RTD also corroborate our earlier observations regarding the flow profile development in the silos. The RTD values showed that particles located in the top layer of the powder bed have smaller residence time values for the flat bottom silo compared to the silo with converging walls, with mean residence time values of 0.08 s and 0.17 s respectively. These findings further support the idea that funnel flow results in a first-in-last-out flow sequence, often leading to particle segregation. Moreover, this indicates potential challenge in setting-up material traceability in silos with flat bottom and their application in continuous pharmaceutical manufacturing where material traceability is a regulatory requirement (ICH Expert Working Group, 2018).

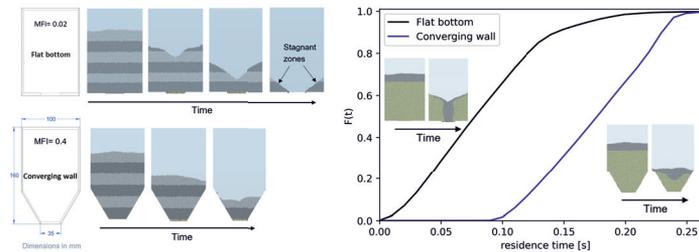


Figure 1: (A) Flow profile evolution snapshots and (B) residence time distribution of gray-colored particles, for two silo geometries

Next, experimental and numerical mass flow rates were compared for validation model as shown in Fig. 2. It is observed that the DEM model gives an overestimation of the discharge rate values for all aperture sizes considered. This result can be explained by the fact that the DEM model developed does not consider the critical air impediment effect during the fine particles discharge process. This finding, while preliminary, suggests that a relation needs to be considered to account and correct for air interaction in DEM discharge rate predictions for fine particles. The study by Wikström et al. (2021) showed that for fine particles, a model that considers the effect of the aerodynamic resistance and gravitational forces, neglecting the inertial effects, can adequately predict the discharge rate of commonly used excipients in intermediate size silos (20 L IBC). Since the assumptions in terms of the importance of forces such as drag in the discharge rate calculation are still considered valid for the current case study, the relationship implemented by

Wikström et al. (2021) is used to correct the predicted DEM discharge rates for a laboratory-scale silo. The relation proved to adequately correct the DEM results for different aperture sizes, with an average relative error of 10% with respect to the experimental values. However, the general applicability of this model for different powders and silo shapes remains to be verified. Furthermore, the assumptions in terms of the selected particle representation (size, shape) need to be further reviewed.

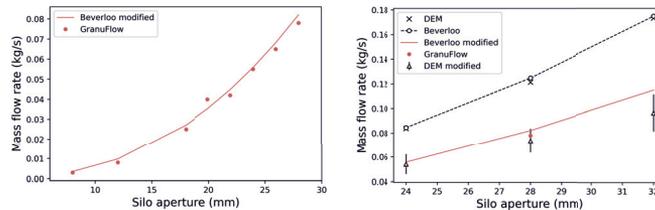


Figure 2: Comparison of numerical and experimental mass flow rate values depending on orifice size

4. Conclusions and outlook

A predictive model for powder discharge in a silo has been developed. The implemented model served as a preliminary basis in identifying the DEM's capabilities to predict flow profiles and discharge rate of fine powders. In addition, it highlighted the influence of the interaction between air and particles when discharging fine particles.

The first part of this study allowed a qualitative comparison of the flow profiles developed in two silo geometries. The results in terms of flow patterns, flow uniformity, and residence times indicated an association between silo geometry and flow profile, matching the findings of earlier studies Mehos (2018). The second part of the study showed a comparison of discharge rates for different silo apertures. The results indicated an overprediction of the DEM model in all cases, these findings may be explained by the effect of air impediment. Additionally, a relation was applied to the DEM results to accurately predict discharge rates.

The DEM model developed included assumptions in terms of the selected particle representation (size, shape, size distribution) that need to be further revised for the selected application. Moreover, additional verification is ongoing to include discharge rate comparisons for smaller silo openings.

References

- Z. H. Gu, 1991. Gravity flowrate of bulk solids from mass flow bins.
- R. Hesse, F. Krull, S. Antonyuk, 2020. Experimentally calibrated cfd-dem study of air impairment during powder discharge for varying hopper configurations. *Powder Technology* 372, 404–419.
- ICH Expert Working Group, 2018. Q13 continuous manufacturing of drug substances and drug products. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.
- K. L. Johnson, K. Kendall, a. Roberts, 1971. Surface energy and the contact of elastic solids. *Proceedings of the royal society of London. A. mathematical and physical sciences* 324 (1558), 301–313.
- W. R. Ketterhagen, J. S. Curtis, C. R. Wassgren, B. C. Hancock, 2009. Predicting the flow mode from hoppers using the discrete element method. *Powder technology* 195 (1), 1–10.
- G. Mehos, 2018. *Hopper design principles for chemical engineers*. Kearney, NE: Morris Publishing.
- H. Wikström, J. Rimmelgas, S. Solin, M. Marucci, N. Sandler, C. Boissier, P. Tajarobi, 2021. Powder flow from an intermediate bulk container—discharge predictions and experimental evaluation. *International Journal of Pharmaceutics* 597, 120309.
- S. B. Yeom, E.-S. Ha, M.-S. Kim, S. H. Jeong, S.-J. Hwang, D. H. Choi, 2019. Application of the discrete element method for manufacturing process simulation in the pharmaceutical industry. *Pharmaceutics* 11 (8), 414.

Biorefinery modelling is in tatters, and here is why

Robert Pujan^{a,b} and Heinz A. Preisig^b

^a*DBFZ Deutsches Biomasseforschungszentrum gemeinnützige GmbH,
Torgauer Straße 116, 04347 Leipzig, Germany*

^b*NTNU Norwegian University of Science and Technology,
Høgskoleringen 5, 7491 Trondheim, Norway
robert.pujan@dbfz.de*

Abstract

Today, the design of biorefineries is following the traditional chemical process design, thereby reflecting that biorefining-specific models are not established. This issue is one of the predominant reasons why most biorefinery concepts still remain limited in their commercial as well as sustainable capabilities. This situation most likely originated from the very heterogeneous, multi-scale as well as multi-disciplinary nature of biorefineries - a circumstance that often turns model design into an expensive and elaborate task. However, models are essential for understanding integrated biorefining and how feedstock and operating conditions affect performance. Correspondingly, the absence of reliable predictive models causes uncertainties, which put investors off and generate significant development bottlenecks in design, optimisation and control.

The aim of this contribution is to outline the current state of biorefinery modelling, derive the presumed root causes, and expose the preferred characteristics of the models required for condign biorefinery research.

Keywords: modelling, simulation, process control, optimisation, process design.

1. Motivation

The social support for a more sustainable process industry has never been this high yet, although the considered processes are leaping behind considerably. Most biorefinery concepts are still in the experimental design phase, thus remaining limited and far off the expected industrial standard in their commercial and sustainable capabilities. This is partly governed by the lack of proper design methods for plants and supply networks, as reported by Bauer et al. (2017), Dragone et al. (2020), Temmes & Peck (2020), and Ubando et al. (2020). Their further development and commercial implementation are hampered by high investment costs and risks, and little trust in its novel technologies, expected yields, and operating reliability.

Although many biorefinery concepts are technically feasible and can have positive environmental impacts, the costs and sustainability are determined by the effectiveness of each process step and the utilisation of the different platform products (Corona et al., 2018). Therefore, a successful design requires intensive research and development activities, and one requires rigorous models to achieve a high-quality, reliable, and optimal design (Tey et al., 2021). Experimental studies may be the conventional alternative but they are expensive, time-consuming and often not viable in all operational areas or even reliable for the exploration and testing of differing process scales. Numerical simulations are done in computers and as such not limited by physical constraints. They thus allow for the exploration of physical regions that are commonly

not accessible by experimental work. The fast development of computing in recent decades rapidly increased the capabilities of process simulation, thereby enabling the simulation of more and more complex models. However, as we will discuss in this study, large parts of biorefinery research and development do not seem to follow this trend, thereby failing to provide robust models.

Without rigorous holistic models, parameter estimates and process designs are speculative, and reliable identification of the dynamics is improbable. Correspondingly, Alzate et al. (2018) and Solarte-Toro et al. (2021) conclude that the absence of reliable predictive models causes uncertainties, which put investors off and generate significant development bottlenecks in design, optimisation, control, and implementation.

Therefore, we reason that the hampered implementation of biorefinery processes is the result of a lacklustre state of modelling in this domain. To discuss this hypothesis within the scope of this short contribution, we shall first outline the present state of biorefinery modelling. Next, we uncover the primary challenges in biorefinery design, which will indicate the presumed root causes for the lack of modelling. To conclude, we expose the preferred characteristics of mechanistic models required for biorefinery concepts and their potential benefits for research, development and process integration.

2. Outlining the current state of biorefinery modelling

Scopus[®] (Elsevier B.V., 2021) shows that less than 17 % of biorefinery-related publications in the last 20 years dealt with topics of modelling, while the overall share in chemical engineering publications is almost 24 %. This lack of modelling activities in the biorefinery community is significant and raises concerns. It implies that today biorefineries are designed on the backbone of industrial chemistry and not based on process systems engineering concepts.

Although several biorefinery-specific conversion processes have been studied, the understanding of the dynamic nature of most of these reactions remains incomplete and limited. The lack of comprehensive mechanistic knowledge combines with a lack of experimental data: Since bioconversion platform processes are in most cases still confined to pilot-scale set-ups, De Buck et al. (2020) found that many related models are highly assumption-based, and often rely on ideal kinetics. These models are only valid for the narrow processing conditions and feedstocks for which they were created and extrapolation does not produce reliable results.

Most of the published biorefinery designs follow a top-down approach, only demonstrating kinetic process models for one particular process part at a time (De Buck et al., 2020). Detailed and replicable process designs, especially ones designed for flexible operational conditions, are few, as observed by Chaturvedi et al. (2020).

Without reliable modelling, assessment studies remain speculative. Although various decision support systems have been developed, strategic, tactical and operational decision tasks are not considered together in most of these studies, even though studies like Geraili et al. (2014) attest significant interdependence between them. Aristizábal-Marulanda et al. (2020) conclude that, until now, knowledge on the best configurations remains limited, and only few studies have looked at the environmental sustainability and socio-economic impacts of specific biorefinery routes. To the best of our knowledge, none has focused on finding the most-sustainable biorefinery cascade. Most life-cycle assessment (LCA) studies on biorefineries compare feedstocks between each other and to fossil references. Some compare the impacts of different processing routes, however, Julio et al. (2017) found these evaluations to concern already designed processes, thereby considering only unique operations in the assessments. Since even

small changes in operating parameters can notably modify environmental impacts, the reliability of those studies is questionable.

Problems are best solved by attempting to correct or eliminate the root causes instead of merely addressing the immediately obvious symptoms of the current state. Therefore, we will discuss in the following section the presumed root causes for the lacklustre development of biorefinery modelling before we will list the characteristics of the models required for overcoming the present period of stagnation in the field.

3. The challenges facing biorefinery modelling

Modelling chemical and biological processes is intrinsically a multi-scale and multi-disciplinary problem that requires expert knowledge from different scientific domains. This diversity is even more obvious in biorefinery concepts, which predominantly process aqueous solutions of manifold substances. Since many of the processed chemical compounds have not notably been encountered in traditional chemistry, their thermodynamic data and reaction kinetics are widely missing today (Corazza & Trancoso, 2021). Furthermore, the widespread utilisation of solids in biorefining increases the need for robust particle and surface reaction simulations, which has not been accomplished satisfactorily so far due to the computational complexity inherent. These circumstances often turn model design into a time-consuming and elaborate task.

With models being at the top end of the overall assessment process, the correctness and completeness of the model equations are of the highest priority in modelling. Since the processes we observe are intrinsically an ensemble of various sub-processes on different time and spatial scales, the model design has to incorporate this behaviour consistently, including linking multiple scales for accuracy and integrity. Multi-scale models pose notoriously difficult theoretical and computational problems, but the knowledge and abilities in mathematics, physics, biology, chemistry, process control, and engineering are proliferating. In biorefinery processes, the multi-scale issue is amplified by these systems being critically dependent on interactions often coupled through multiple spatial and temporal domains without clear, discrete differences between scales. Consistent and reliable mathematical relations of the behaviour and interactions of the various scales are required to obtain robust and predictive models. Modern biorefinery concepts have to utilise a vast amount of scientific disciplines for reliable design and optimisation (Punnathanam & Shastri, 2020), effective process control implementations (Petre et al., 2020), and accurate assessment (Julio et al., 2017) to get the most out of any feedstock in competitive pricing ranges. One modeller cannot be an expert in all of the topics encompassed, and, more often than not, communication between experts of different scientific fields appears to be cumbersome.

Biorefinery processes' modelling, assessment, and control is challenging due to the complexity and non-linearity, uncertain kinetics, different scales of process and logistics, the heterogeneous and seasonally fluctuant solid biomass feedstocks, and high uncertainty in bioprocess kinetics (Ifrim et al. (2015); Nguyen et al. (2015)). Latter implies that understanding and modelling those processes requires taking several dynamic factors into account. However, dynamic models are often restricted to one dimension as otherwise, the computational cost would be tremendous. Nevertheless, they are critical for optimising start-up and shut-down procedures, HAZOP studies, and investigating the plant transient operation under high load change rates. Additionally, plant control systems, have to be trained on dynamic models. The cost or even lack of adequate sensors further hampers the design of reliable control structures (Petre et al., 2020).

Conversion models require fairly detailed feedstock composition, while supply network models are often limited to only predicting the expected amounts of the available feedstock during a specific time period. Additionally, conversion and logistics models are usually built on significantly different time scales (Baldea & Harjunkoski, 2014). Since many conversion processes in biorefinery concepts, like fermentation, are batch processes, it is very important to define the optimum conditions for both, logistics and operation, to achieve feasibility. Due to this and the small profit margins to be expected, the assessment of biorefineries must be accurate, utilise all biomass fractions, and encompass economic, environmental and social aspects as well as the erratic supply and pricing of feedstocks. The sustainability and economic feasibility of the required feedstock's supply to the plant is crucial when designing and operating a biorefinery (De Buck et al., 2020). The time perspective in these assessments is particularly relevant for biorefineries due to the feedstock's highly seasonal cultivation characteristics.

4. A wish list for biorefinery modelling

To address the myriad of challenges to biorefinery modelling, adequate efforts have to encompass a multitude of characteristics that can probably only be achieved by extensive interdisciplinary liaisons. In the following, some of these characteristics will be discussed. Most of the flowsheeting packages commercially available are probably not suitable for the complex needs in biorefinery modelling and not flexible enough for custom-modelling efforts. Since these packages have been developed for and traditionally been used by the petrochemical industries, their libraries lack unit operations as well as many chemical properties specific for biomass processing. Furthermore, most of these packages are designed to solve steady-state mass and energy balances on a plant-wide scale only. Quite obviously, this is rather not sufficient for the design of the dynamic, multi-scale models required for biorefineries. Fortunately, many of those issues have been addressed in non-commercial modelling suites like, for example, ModKit (Bogusch et al., 2001) and ProMo (Preisig, 2021).

Since strategic, tactical and operational decision tasks have to be considered together in the design and operation of biorefineries, jack-of-all-trades modelling efforts are a necessity, targeting not only commercial feasibility and operational safety but also social and environmental sustainability (Sikdar, 2003). This high degree of interconnectivity is further reflected by the study of Corona et al. (2018), in which maximising product yields proved to be the most important environmental optimisation parameter for green biorefineries, even more important than reducing energy consumption. The secret to sustainability hides, after all, in the optimisation of processes. Hence, LCA studies cannot rely on black-box models but must be coupled to detailed process simulation. Suggested by Julio et al. (2017), those simulations should be built on multi-objective optimisation algorithms that establish iterative corrections between operational condition changes and the effects on life-cycle aspects.

Although empirical models can offer some understanding of the processes, in-depth process understanding and predictive modelling are the only reliable prerequisites for decision-making tools and to achieve a cost-effective design of biorefineries. This emphasizes the importance of multi-scale modelling approaches from the plant scale down to the molecular level. Holistic models require a more in-depth understanding of the processes, however, once developed, these models can be used for a wide range of processing conditions. Consequently, they enable efficient process optimisation and reliable up-scaling studies.

True process parameter estimates are crucial for the development of robust models. Therefore, non-linear process dynamics that are inherent in complex bioprocesses cannot be circumvented when modelling the plant. While steady-state models can provide accurate results regarding the plant operation on specific loads and are very useful in investigating the sustainable minimum load limit, only dynamic models are reliable sources to evaluate the plant performance, optimizing its operation, and improve its efficiency under the flexible operations inherent to most biorefinery concepts.

A time scale is particularly relevant for biorefineries due to erratic feedstocks and the interdependence of processing and logistics. The feedstock fluctuations must be taken into account in the models, whilst also considering the multitude of possible feedstocks. Early considerations of fluctuant as well as substitutable feedstocks can eventually speed up the design and implementation of economically viable biorefineries.

In conversion processes, kinetic modelling constitutes a critical step. Like any chemical plant, a biorefinery consists of multiple, interconnected equipment units. Thus, any unit's performance is strongly affected by the operation of the others. This underlines the need for the implementation of proper plant-wide control concepts in such complex plants as biorefineries to ensure optimal productivity with the desired specifications. Furthermore, operational issues caused by a lack in adequate instrumentation can often be circumvented by utilising robust predicting models that estimate unavailable process variables.

The complex exigencies on biorefinery models place a high demand on the modellers. This induces an increasing interest in more effective model generation processes that incorporate the various expert's knowledge, which can additionally open up new prospects for many projects and may have the effect of building bridges between different disciplines and people (Preisig et al., 2019).

5. Conclusion

When aiming for condign biorefinery research and development efforts, modelling is not only an option but a necessity. However, the abundance of obstacles to biorefinery modelling led to a research community struggling to implement robust, holistic approaches but relying on empirical models if at all. The pulp and paper domain demonstrated the issue before: Although the significance of models in this field were identified early (Pilati & Rosen, 1978), a lack of model utilisation and a thereby caused stagnation have been reported even decades later by studies like Blanco et al. (2009). Similarly, if the biorefinery community's current state is not abandoned but prolonged into the next scientific generations, biorefinery research will be hampered for decades to come, thus failing to comply with the urgency of attaining sustainable industries.

Acknowledgements

This paper originated from ongoing work in the research projects Bio4Fuels (Norwegian Centre for Environment-friendly Energy Research (FME), project 257622), MarketPlace (Horizon 2020, project 760173), and VIPCOAT (Horizon 2020, project 952903).

References

- C.A.C. Alzate, J.C.S. Toro & Á.G. Peña, 2018, Fermentation, thermochemical and catalytic processes in the transformation of biomass through efficient biorefineries, *Catal. Today*, no. 302, 61-72
- V. Aristizábal-Marulanda, J.C. Solarte-Toro & C.A.C. Alzate, 2020, Economic and social assessment of biorefineries: The case of Coffee Cut-Stems (CCS) in Colombia, *Bioresour. Technol. Rep.*, no. 100397

- M. Baldea & I. Harjunkoski, 2014, Integrated production scheduling and process control: A systematic review, *Comput. Chem. Eng.*, no. 71, 377-390
- F. Bauer, L. Coenen, T. Hansen, K. McCormick & Y. V. Palgan, 2017, Technological innovation systems for biorefineries: a review of the literature, *Biofpr.*, 11 (3), 534-548
- A. Blanco, E. Dahlquist, J. Kappen, J. Manninen, C. Negro & R. Ritala, 2009, Use of modelling and simulation in the pulp and paper industry, *Math. Comp. Model. Dyn. Syst.*, 15:5, 409-423
- R. Bogusch, B. Lohmann & W. Marquardt, 2001, Computer-aided process modeling with modkit, *Comp. Chem. Eng.*, no. 25, 963-995
- T. Chaturvedi, A.I. Torres, G. Stephanopoulos, M.H. Thomsen & J.E. Schmidt, 2020, Developing process designs for biorefineries - definitions, categories, and unit operations, *Energies*, vol. 13 (6), no. 1493
- M.L. Corazza & J. Trancoso, 2021, Phase equilibria modeling of biorefinery-related systems: a systematic review, *Chem. Prod. Proc. Model.*, 2020-0119
- A. Corona, M. Ambye-Jensen, G.C. Vega, M.Z. Hauschild & M. Birkved, 2018, Techno-environmental assessment of the green biorefinery concept: Combining process simulation and life cycle assessment at an early design stage, *Sci. Total Environ.*, no. 635, 100-111
- V. De Buck, M. Polanska & J. Van Impe, 2020, Modeling Biowaste Biorefineries: A Review, *Front. Sustain. Food Syst.* 4:11
- G. Dragone, A.A. Kerssemakers, J.L. Driessen, C.K. Yamakawa, L.P. Brumano & S.I. Mussatto, 2020, Innovation and strategic orientations for the development of advanced biorefineries, *Bioresour. Technol.*, vol. 302, no. 122847
- Elsevier B.V., Scopus®, [Online]. Available: <https://www.scopus.com/>. [Accessed 01 09 2021].
- A. Geraili, P. Sharma & J.A. Romagnoli, 2014, A modeling framework for design of nonlinear renewable energy systems through integrated simulation modeling and metaheuristic optimization: Applications to biorefineries, *Comput. Chem. Eng.*, no. 61, 102-117
- G.A. Ifrim, M. Barbu, E. Ceanga & S. Caraman, 2015, Modeling and control of a multipurpose biotechnological plant. Photobioreactor modeling, *Proceedings of the 19th ICSTCC*
- R. Julio, J. Albet, C. Vialle, C. Vaca-Garcia & C. Sablayrolles, 2017, Sustainable design of biorefinery processes: existing practices and new methodology, *Biofpr.*, no. 11 (2), 373-395
- D. Nguyen, V. Gadhamshetty, S. Nitayavardhana & S.K. Khanal, 2015, Automatic process control in anaerobic digestion technology: A critical review, *Bioresour. Technol.*, no. 193, 513-522
- E. Petre, D. Selișteanu & M. Roman, 2020, Control schemes for a complex biorefinery plant for bioenergy and biobased products, *Bioresour. Technol.*, vol. 295, no. 122245
- D.A. Pilati & R. Rosen, 1978, The use of the pulp and paper industry process model for R&D decision making (No. BNL-50839), Brookhaven National Lab., Upton, NY (USA)
- H.A. Preisig, 2020, ProMo -a multi-disciplinary process modelling suite, *Comp. Aid. Chem. Eng.*, no. 48, 571-576
- H.A. Preisig, A. Elve & S. Karoliuss, 2019, More power to the individual modeller using ontologies, *Comp. Aid. Chem. Eng.*, no. 46, pp. 1099-1104
- V. Punnathanam & Y. Shastri, 2020, Efficient optimization of a large-scale biorefinery system using a novel decomposition based approach, *Chem. Eng. Res. Des.*, no. 160, 175-189
- S.K. Sikdar, 2003, Sustainable development and sustainability metrics, *AIChE journal*, no. 49(8), 1928-1932.
- J.C. Solarte-Toro, C.A.R. Duran, M. Ortiz-Sanchez & C.A. Cardona-Alzate, 2021, A comprehensive review on the economic assessment of biorefineries: The first step towards sustainable biomass conversion," *Bioresour. Technol. Rep.*, no. 100776
- A. Temmes & P. Peck, 2020, Do forest biorefineries fit with working principles of a circular bioeconomy? A case of Finnish and Swedish initiatives, *For. Policy Econ.*, vol. 110, no. 101896
- S.Y. Tey, S.S. Wong, J.A. Lam, N.Q. Ong, D.C. Foo & D.K. Ng, 2021, Extended hierarchical decomposition approach for the synthesis of biorefinery processes, *Chem. Eng. Res. Des.*, no. 166, 40-54
- A.T. Ubando, C.B. Felix & W.H. Chen, 2020, Biorefineries in circular bioeconomy: A comprehensive review, *Bioresour. Technol.*, vol. 299, no. 122585

Simulation of a Fischer-Tropsch reactor for jet fuel production using Aspen Custom Modeler

Alejandro Morales^a, Gregoire Leonard^a

^a*Chemical Engineering, University of Liège, Belgium.*

Abstract

The reduction of CO₂ emissions is a key aspect to the world energy transition. Fischer-Tropsch (FT) technologies can contribute to lower our CO₂ emissions to the environment by transforming captured CO₂ into a wide array of hydrocarbon chains. Under optimized conditions, the reaction yields a high concentration of long carbon chains; which in turn, can be further upgraded into jet fuel. Thus, difficult to electrify sectors of the economy such as the aviation sector could have a defossilized supply of fuel; provided that the feedstock required for the reaction originates from low-carbon energy. In this context, the objective of the present research is to design, install, operate and optimize a FT reactor to serve as core of a future Power-to-Jet Fuel pilot-scale implementation at the University of Liège (ULiège). The first step was to select from the literature the stoichiometry and the kinetics that accurately depict the FT reaction. In this study, the kinetics reported by Iglesia et al. (1993) and the stoichiometry proposed by Hillestad et al. (2014) were selected and implemented in a simulation model developed using Aspen Custom Modeler (ACM). This model was then validated by simulating a computational FT implementation reported in the literature by Tomte (2013). Once validated, the parameters are used to simulate a FT pilot reactor having an inlet of 62 mol H₂/h and 29.5 mol CO/h, in compliance to the electrolysis capacity available at the ULiège. The results portray a production of 0.3 mol/h C₁₂₋₂₀ chains that could be further upgraded to jet fuel and a conversion of almost 50% of CO towards the FT reaction. In future work, the ACM model of the FT reactor will serve as input for a more complex process model of the Power-to-Jet Fuel production chain, further improving conversion efficiency.

Keywords: Fischer-Tropsch, Jet Fuel, Aspen Custom Modeler, Simulation.

1. Introduction

The European Commission aims to lower CO₂ emissions by 55 % compared to 1990 by 2030. In 2016, the aviation sector accounted for 3.6% of overall emissions and 13.4% of transport emissions in the EU (EASA et al., 2019). The sector also has the lowest share of renewables, with only 6% (European Commission, 2020). Thus, innovative alternatives such as using green hydrogen as fuel or producing jet fuel in sustainable ways are considered to defossilize this means of transportation. Sustainable jet fuel could become an adequate substitute of its traditional counterpart by providing at least 70% life-cycle carbon reductions (Sustainable Aviation, 2019).

The present study is carried out in the framework of further improving the Power-to-Jet Fuel value chain. By utilising hydrogen and carbon monoxide as feedstock, a FT reactor synthesizes jet fuel as liquid energy vector. This chemical pathway is rendered more sustainable when the hydrogen is produced from renewable energy sources via water

electrolysis; and the carbon monoxide is obtained from the Reverse Water Gas Shift (RWGS) of said hydrogen and captured carbon dioxide.

Although the production of sustainable FT fuels has already been studied (e.g. Loewert et al., 2019; U. Pandey et al., 2021), most of these studies do not focus on the maximization of jet fuel production. Moreover, previous studies mostly focus on reaction modelling, and less work has been carried out regarding process study. To do so, the present work uses ACM. First, it allows to make use of the vast properties database that AspenTech provides and second, it allows a detailed modelling of the reactor with complex kinetics associated to different catalysts and reactor configurations. In further work, the reactor model will then be included into a complete process model developed in Aspen Plus. Thus, it will be easier to understand the challenging interactions between the synthesis section and other process steps (RWGS, product upgrade, recycling loop...).

In this paper, the main properties of the reaction and modelling assumptions are first portrayed. Then, the simulation is validated by comparison with a model developed by Tomte at the Norwegian University of Science and Technology (NTNU). Finally, the FT reactor model is scaled down in compliance to the electrolysis capacity available at the ULiège, corresponding to a pilot-scale reactor.

2. Model description

The FT reaction results in a wide array of hydrocarbon products. Jet fuel, the main focus product of this present study, has a desirable composition of 75-85 vol% consisting of paraffins, iso-paraffins and cycloparaffins and the remaining 15-25% of olefins and aromatics. Jet fuel also encompasses hydrocarbon chains of 12 to 20 carbon atoms (Doliente et al., 2020). There are many technological configurations that enable the FT reaction but to maximize the production of C₁₂₋₂₀ chains the use of low temperatures (200-230°C) as well as cobalt as catalyst is preferred (Dieterich et al., 2020).

The developed reactor model uses kinetics from Iglesia et al. (1993) further adapted by Panahi et al. (2012); and the stoichiometry proposed by Hillestad et al. (2014) based on the Anderson-Schulz-Flory (ASF) distribution. The ASF model proposes a statistical distribution of FT products based on the chain growth probability (α). The Hillestad stoichiometry makes some simplifications. It only considers the methanation reaction in parallel to the FT reaction and the production of linear paraffins by the FT reaction (Table 1). Likewise, it models all products with carbon atom number below 21 as individual units. On the other hand, components with carbon atom number from 21-30 were put together in a component designated C₂₁₊. The lumped component C₂₁₊ is modelled as C₃₀H₆₂ due to similar properties.

For low-temperature cobalt FT, a constant α value of 0.9 is assumed, as done by Hillestad. Based on Eq. 1, this leads to a H₂/CO ratio (Z) of 2.1. By using Eq. 2 and Eq. 3, the ASF coefficients for every single potential hydrocarbon that can be produced via FT are estimated. These coefficients are then used in Eq. 4 to provide the full stoichiometry of the FT reaction. Additionally, Eq. 5 provides the stoichiometry for the methanation reaction.

$$Z = \frac{H_2}{CO} = 3 - \alpha \quad (1)$$

$$ASF \text{ coefficient}_i = (1 - \alpha)^2 * \alpha^{i-1} \quad \text{for } C_i, i = 1, 2, 3 \dots 20 \quad (2)$$

$$ASF \text{ coefficient}_{21+} = (1 - \alpha) * \alpha^{20} \quad \text{for } C_{21 \rightarrow 30} \quad (3)$$

Table 1. Reactions in parallel in the FT reactor.

FT reaction:	$CO + Z H_2 \rightarrow \sum_{i=1}^{20} (ASF \text{ coefficient})_i C_i H_{2i+2} + (ASF \text{ coefficient})_{21+} C_{30} H_{62} + H_2 O \quad (4)$
Methanation reaction:	$CO + 3 H_2 \rightarrow CH_4 + H_2 O \quad (5)$

The Iglesia kinetics modified by Panahi, portrayed by Eq. 6 and Eq. 7, provide reaction rates in $\text{kmol/m}^3 \cdot \text{s}$ for the methanation (r_{CH_4}) and FT (r_{FT}) reactions respectively. With k_i being kinetic constants reported by Panahi and P_{H_2} & P_{CO} being the partial pressure for hydrogen and carbon monoxide. The kinetics are valid for a range of temperature between 200 and 210 °C, 100 to 3000 kPa and a $Z = H_2/CO = 1$ to 10.

$$r_{CH_4} = \frac{k_1 P_{H_2} P_{CO}^{0.05}}{1 + k_3 P_{CO}} \quad (6)$$

$$r_{FT} = \frac{k_2 P_{H_2}^{0.6} P_{CO}^{0.05}}{1 + k_3 P_{CO}} \quad (7)$$

For all models the reactor length is discretized in 5 sections and mass and heat balances are evaluated for each one of them.

3. Validation model

The Tomte study performed at NTNU (Tomte, 2013) is taken as main reference for the validation of the simulation. It simulated a Gas-to-Liquid (GTL) plant capable of achieving a similar production capacity to that of the Bintulu GTL plant, 14700 bbl/day. In the present study, only the FT reactor of the plant will be analyzed and simulated. Similar inlet conditions as in the reference model were chosen (Table 2). The Tomte study does not report the molar fractions of each of the components at the inlet flow of the reactor but the molar flow of groups of hydrocarbon chains (C_{1-2} , C_{3-4} , C_{5-11} , C_{12-20} , C_{30}); thus, an estimate was calculated for each one of them. It is important to highlight that there is presence of products in the inlet flow of the reactor as a recycle is used to further improve the conversion of the reactor.

Due to the highly exothermic nature of the FT reaction and to keep the reaction as close to isothermal as possible, the Tomte study used pressurized liquid water at 220°C to exchange heat with the reactor. A very large but not specified molar flow was used to keep the cooling water at a constant temperature. As no additional information is

provided, for the present study an overall heat value coefficient of $400 \text{ W/m}^2\cdot\text{K}$ was chosen for the heat balance of the system (M. Boda et al., 2017).

The Tomte study implemented a 1000 m^3 multitubular fixed bed reactor but does not provide additional information; thus, the dimensions of the reactor had to be estimated as well. A patent for a gas phase reaction multitubular fixed bed reactor (G. Olbert et al., 2007) states that these reactors have in between 10000 to 30000 catalytic tubes. These tubes have an internal diameter between 20 mm and 35 mm and a length of 1.5 to 7 m. By following these design parameters and keeping the same reaction volume as the Tomte study, a train of 16 reactors in parallel with 28500 catalytic tubes each is chosen. The catalytic tubes have an internal diameter of 20 mm and a length of 7 m each. Due to the low pressure drop reported by Tomte, these reactors are assumed isobaric. The chosen configuration achieves a similar outlet temperature and molar product flows from the reactor as the Tomte study (Table 2). It is worth highlighting that a similar configuration is used in the Pearl GTL plant in Qatar. This plant has 24 reactors, each containing 29000 catalytic tubes, and achieves a production of 140000 bbl/day (Shell, n.d).

Table 2. Inlet and outlet conditions of the validation model and comparison with the Tomte study.

Variable	Units	Inlet stream	Outlet, this work	Outlet, Tomte (2013)	Relative difference (%)
Total molar flow	kmol/h	46563.2	34895.6	34228.03	1.95
Molar flow CO		12183.5	6349.68	6015.95	5.55
Molar flow H ₂		25595.8	13073.9	12354.56	5.82
Molar flow C ₁₋₂		3472.67	3882.25	3908.84	0.68
Molar flow C ₃₋₄		132.19	215.18	223.78	3.84
Molar flow C ₅₋₁₁		195.56	383.68	385.27	0.41
Molar flow C ₁₂₋₂₀		0	110.65	112.55	1.69
Molar flow C ₃₀		0	66.39	71.07	6.59
Molar flow H ₂ O		4988.41	10822	11156.01	2.99
Temperature		°C	210	221.4	221.8
Pressure	Bar	20	20	19.4	3.09

Due to the close proximity of the results (less than 7% deviation for partial flowrates), it was possible to confirm that the simulation was validated. The slight differences between the results could be attributed to the fact that the Tomte study does not report the exact compositions of each of the components at the inlet flow of the reactor, as previously

stated. It can also be attributed to neglecting pressure drops in the reactors. Likewise, the temperature profile of the Tomte reactor is not presented in the study.

4. Case study - ULiège model

Once the FT kinetics and stoichiometry have been validated in the previous section, the model is scaled down to integrate the operating conditions of a pilot-scale project at ULiège. The feed of the ULiège model is composed only of hydrogen and carbon monoxide; although, it is planned that future work will explore the influence of a recycle into the reactor. The hydrogen feed is based on the total hydrogen production capacity of the 3 electrolyzers at ULiège, 62 mol H₂/h. Based on the Z ratio used so far, a molar flow of 29.5 mol CO/h is required. As previously mentioned, the Iglesia kinetics work better in a range of 200 - 210 °C; therefore, a fixed bed reactor with a cooling jacket having a high inlet flow of water at 210°C is chosen (instead of 220°C as implemented in the validation model). The same overall heat coefficient, inlet temperature and inlet pressure are used as in the validation model. No pressure drops are considered in this model either.

A reactor volume of 2.5 L (having an internal diameter of 0.05 m and a length of 1.275 m) is chosen. This volume yields a 49.01% conversion of the inlet CO through the FT reaction and 2.57% through methanation. The rest remains unreacted, 13.95 mol/h. These conversion values are similar to the ones reported by Tomte and the validation model. Figure 1 displays the production of FT products inside the reactor. A molar flow of 0.3 mol/h of C₁₂₋₂₀ chains that could be further upgraded to jet fuel is achieved. Figure 2 displays the temperature profile, the temperature is kept close to 210°C as intended.

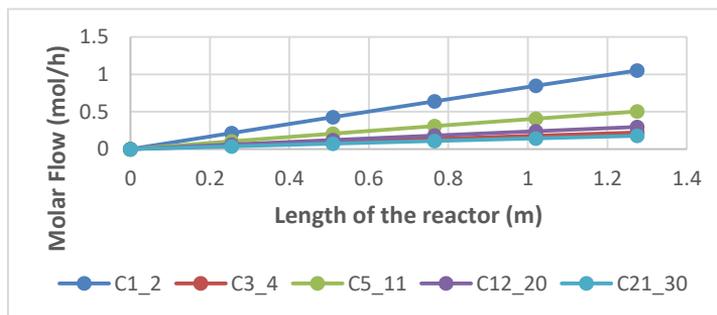


Figure 1. Molar flow of FT products of the ULiège model.

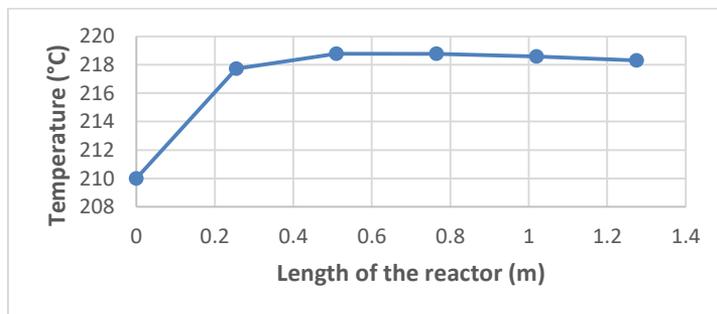


Figure 2. Temperature profile of the ULiège model.

5. Conclusions

The kinetic model of Iglesia (Iglesia et al., 1993) and the stoichiometry proposed by Hillestad (Hillestad, 2014) for the FT reaction were validated using ACM by obtaining the same results as the Tomte study (Tomte, 2013), which in turn modelled a GTL plant with a capacity of 14700 bbl/day. The hydrogen production capacity available at ULiège was then used as guideline for a scaled-down design of a FT pilot reactor.

Future work will focus on further improving the accuracy and the performance of the simulation for jet fuel production maximization. A sensibility analysis will be performed on multiple process variables to determine their influence on jet fuel production. Some of the variables to consider are: the temperature, volume and pressure of the reactor; the cooling medium and its temperature; the catalyst and its configuration; and the recycle ratio of the outlet flow of the reactor. In parallel, new models will be developed using more detailed kinetics. For example, they will comprise other products other than paraffins (e.g. olefins) and will evaluate α as a function of the temperature of the reactor and other operation conditions. Later phases will focus on improving the integration of the carbon capture, electrolysis, RWGS, recycle and post-treatment sub-processes to the reactor.

References

- E. Iglesia et al., 1993, Reaction-Transportation selectivity models and the design of Fischer-Tropsch catalysts, *Computer-Aided Design of Catalysts*, 199-257.
- M. Hillestad, 2014, Modelling the Fischer-Tropsch distribution and model implementation, *Chem. Prod. Process Model*, Vol 10, 147-159.
- K. Tomte, 2013, Modelling and optimization of a Gas-to-Liquid plant.
- EASA et al., 2019, European Aviation Environmental Report 2019.
- European Commission, 2020, Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions. Sustainable Aviation, 2019, Sustainable aviation fuels road-map.
- M. Loewert et al., 2019, Microstructured Fischer-Tropsch reactor scale-up and opportunities for decentralized application, *Chemical Engineering & Technology*, Vol 42, 2202-2214.
- U. Pandey et al., 2021, Modeling Fischer-Tropsch kinetics and product distribution over a cobalt catalyst, *AIChE Journal*, 7, Vol 67, e17234.
- S. Doliente et al., 2021, Bio-aviation fuel: A comprehensive review and analysis of the supply chain components.
- V. Dieterich et al., 2020, Power-to-Liquid via synthesis of methanol, DME or Fischer-Tropsch-fuels: a review, *Energy & Environmental Science*, 10.
- M. Panahi et al., 2012, A natural gas to liquids process model for optimal operation, *Industrial & Engineering Chemistry*, 51, 425-433.
- A. Rafice, 2012, Optimal design issues of a gas-to-liquid process.
- H. Becker et al., 2016, Enhancing internal mass transport in Fischer-Tropsch catalyst layers utilizing transport pores, *Catal. Sci. Technol.*, Vol 6, 275-287.
- M. Boda et al., 2017, Design and development of parallel-counter flow heat exchanger, *International Journal of Innovative Research in Advanced Engineering*, 2, Vol 4.
- G. Olbert et al., 2007, Multi-tube fixed-bed reactor, especially for catalytic gas phase reactions, US 7226567 B1.
- Shell, n.d, Pearl GTL, Accessed date: November 28, 2021, from https://www.shell.com.qa/en_qa/about-us/projects-and-sites/pearl-gtl.html.

Techno-economic-environmental analysis of a microbial oil production integrated into a bioethanol sugarcane biorefinery

Andreza Aparecida Longati,^a Felipe Fernando Furlan,^b Roberto de Campos Giordano,^b Everson Alves Miranda^a

^aUniversity of Campinas, School of Chemical Engineering, Department of Materials and Bioprocess Engineering, Campinas 13083-852, Brazil

^bFederal University of São Carlos, Chemical Engineering Graduate Program, São Carlos 13565-905, São Carlos, Brazil
andrezaalongati@gmail.com

Abstract

Biodiesel and bioethanol play an important role as renewable liquid fuels. The production of these biofuels generates low-value by-products, such as sugarcane bagasse (usually used for heat generation). This bagasse can be processed to produce a sugarcane bagasse hemicellulose hydrolysate (SCBH) that contains fermentable sugars, mainly xylose. Oleaginous yeasts (eg., *Rhodotorula toruloides*) can grow in this SCBH producing microbial oil (MO) which can be used as raw material for biodiesel production. This strategy arises as a promising approach to exploit a synergy between bioethanol and biodiesel production processes within a biorefinery. Since techno-economic-environmental analysis (TEEA) can identify bottlenecks, providing insights concerning feasibility and sustainability of processes, in this work TEEA was applied to a study case consisting of a MO production from SCBH integrated into a first-generation (1G) bioethanol plant. The 1G section of the plant – processing 4×10^6 t of sugarcane per season (240 days) - produces bioethanol and bioelectricity generating a surplus bagasse, which is sent to the MO production. Net present value and internal rate of return were chosen as the economic metrics and global warming potential as environmental one for evaluating the process options simulations. The main result of this study had the plant producing 74.8 m³/h of bioethanol, 2.8 t/h of MO, and 89.3 MW of bioelectricity. The integrated process exhibited a positive economic performance (net present value equals to US\$ 34.5×10^6 and internal rate of return of 12.4%) indicating that the microbial oil production from sugarcane is feasible. From the environmental perspective, a decrease of 13% in GHG emissions was obtained.

Keywords: microbial oil, biorefinery, techno-economic analysis, modelling and simulation.

1. Introduction

The diversification of energy matrices is of vital importance in slowing down climate change effects caused by use of fossil fuels. In order to simultaneously mitigate climate change in transport and chemical sectors and reduce the dependence on oil, alternative

production chains are necessary. Biomass is considered as the primary emerging option to replace the fossil fuel resources, since it can deliver energy and multiple products through of the biorefineries. These facilities are integral to supporting a knowledge-driven and environmentally robust circular bio-economy (Awasthi et al. 2020).

The established sugarcane mills in Brazil are biorefineries, since they produce bioethanol, sugar, power, fusel oil, yeast biomass, and other products. However, there still a potential for improvement of this process. One example is the integrated production of bioethanol and biodiesel, since these biofuels play an important role as renewable sources of energy for transport and they are an attractive solution for reducing global warming (GWP), energy crisis, and fossil fuel depletion (Souza and Seabra 2014). Sugarcane bagasse, a by-product of the sucrose and ethanol production, can be processed to generate a sugarcane bagasse hemicellulose hydrolysate (SCBH) that contains fermentable sugars, mainly xylose. Oleaginous yeasts (eg., *Rhodotorula toruloides*) can grow in this SCBH producing microbial oil (MO) which can be used as raw material for biodiesel production. Therefore, such MO production does not compete with food supply and arable land is not required (Bonturi et al. 2017).

The integration of MO production into a first-generation (1G) bioethanol production plant could decrease this lipid cost due to the easier availability of raw material, utilities (steam and electricity), effluent treatment, and disposal and logistic distribution. Therefore, the objective of this study was to applied the techno-economic-environmental analysis (TEEA) to a study case consisting of a MO production from SCBH integrated into a first-generation (1G) bioethanol plant in order to evaluate the feasibility of MO production from byproducts generated in the bioethanol process.

2. Methods

2.1. Process modelling and simulation

EMSO (Environment for Modelling, Simulation and Optimization, Soares and Secchi (2003)), an equation-oriented simulator, was used to carry out the simulations. The TEEA was performed considering the implementation of the MO production process integrated into a 1G bioethanol plant. The modelled biorefinery (Figure 1) can produce 1G bioethanol from sugarcane juice, MO from bagasse, and electricity from bagasse and residues in an integrated process.

The process data used in the 1G simulation were based on Elias et al. (2021). The bagasse obtained after milling is diverted into two fractions. The first one is fed in the boiler (Elias et al. 2021) and the second one is used for MO production. The bagasse diverted to MO production undergoes a diluted acid pretreatment to solubilize the hemicellulose fraction. This is carried out at solid to liquid ratio 1:10 (10% w/v of bagasse in 1.5% (w/v) H_2SO_4) for 60 min at 121 °C. The pretreated mixture cooled to 80 °C is filtered. The solid fraction is sent to the combined heat and power sector and the liquid fraction is concentrated to 55 g/L of xylose in an evaporator at 80 °C. The obtained SCBH is directed to the MO production by the oleaginous yeast *R. toruloides* in a bioreactor for 120 h at 30 °C with the operating conditions described in Lopes et al. (2021).

Yeast suspension from the bioreactor is concentrated through centrifuge operations. The moisture content of the microbial cells is reduced to 1% using a spray dryer. The dried cells are mixed in a tank with hexane (25% w/w) and directed to the high-pressure homogenizer in order to disrupt cells and release the intracellular lipids (Koutinas et al. 2014). The homogenate is then suspended in hexane and sent to a decanter, where the

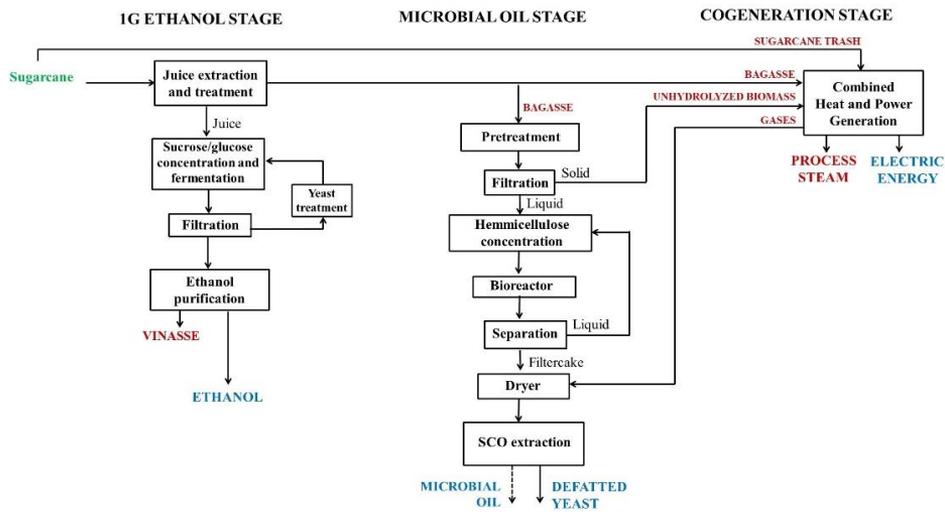


Figure 1. Simplified process diagram of the integrated bioethanol-bioelectricity-MO sugarcane biorefinery

defatted biomass is separated from the MO containing hexane. The MO is recovered from the organic solvent in an evaporator (Vieira et al. 2016) and the hexane is recycled.

2.2. Environmental analysis

The life cycle assessment (LCA) implemented in EMSO considered the climate change impact category using the GWP under a horizon of 100 years – GWP100. This metric is indicated in Renovabio, a Brazilian federal biofuels policy (Grassi and Pereira 2019). It was performed an inventory of inputs, outputs and emissions based on process simulation (item 2.1), agricultural data (Potrich et al. 2020), transporting and burning steps under a cradle to grave scope. The functional unit was defined as 1 MJ of bioethanol and impacts are allocated based on energy. SimaPro 8.4 software and Ecoinvent database 3.0 (Wernet et al. 2016) were used to for raw materials inputs. The average logistical distance of transport in the road modal was considered to be 700 km for bioethanol resulting in transport emissions of 0.80 CO_{2eq}/MJ. The emissions of 0.44 CO_{2eq}/MJ were considered for the biofuel burning step (Matsuura et al., 2018).

2.3. Economic analysis

The economic analysis was also implemented in EMSO coupled to the process model. The integrated plant processes 4 million tons of sugarcane per year operating 5040 h per year (Elias et al. 2021). Capital Expenditure (capex) were estimated from industrial information (personal communication) and data reported in the literature (Koutinas et al. 2014; Vieira et al. 2016) adjusted for inflation (Peters, Timmerhaus, and West 2002). The final investment included the costs of equipment and their installation, civil construction, electricity installation, instrumentation, land, and engineering. The price of the MO is US\$ 1.077/kg (Vieira et al. 2016). Other relevant costs were US\$ 0.535/L of ethanol, US\$ 68.758/MWh of electricity, US\$ 0.352/kg of defatted biomass, US\$ 17.853 US\$/t of sugarcane, US\$ 7.712/t of sugarcane straw. The CO_{2eq} saved, the carbon credit, was estimated. The price of US\$ 5.203 US\$ per t of CO_{2eq} saved was assumed (B3, 2021). Brazilian foreign trade statistics platform (ComexStat 2020) was used for others inputs

and outputs costs and Oliveira et al. (2018) was used for utilities price. Results of the economic analysis were reported in terms of net present value (NPV) and internal rate of return (IRR) (Peters, Timmerhaus, and West 2002). The economic evaluation is based on the following hypothesis: construction and startup period of 24 months; project lifetime of 25 years; discount rate of 11%; 100% of outlay of investment in the first year; linear depreciation over 10 years; no salvage value of direct fixed capital; maintenance of 3.0% of capex per year; tax rate of 34%; working capital of 10% of capex, and values at an exchange rate from Brazilian currency real to US dollar of 3.89 R\$/USD (BCB, 2020).

3. Results and discussion

The integrated plant produces 74.8 m³/h of bioethanol (333.1 million of L/year of bioethanol), 89.3 MW of electric energy and 2.8 t/h of MO (13,001 t/year of MO). Considering that about 1 t of biodiesel is produced from 1 t of lipid (Pimentel and Patzek 2008), this amount of MO has the potential to generate 13,001 t of biodiesel. This biofuel can be used to substitute the diesel in the agricultural operations without engine modification integrating and connecting the agricultural and industrial fields (Ogunkunle and Ahmed 2019). Considering an average specific consumption of 4 L of diesel per ton of sugarcane (Seabra et al. 2011), the biodiesel from MO could replace about 80% of the total diesel consumption in the sugarcane plant.

The total emissions of 23.38 g CO_{2eq}/MJ of ethanol were obtained from biorefinery assessed. This represent a decrease of 6.6% in GHG emissions when comparing with the emissions from traditional 1G bioethanol process estimated by Junqueira et al. (2017) (25.03 g CO_{2eq}/MJ of ethanol considering the cradle-to-grave analysis of the 1G-base scenario). When compared with the emission of the gasoline process (emissions of 87.4 g CO_{2eq}/MJ of gasoline) (Grassi and Pereira 2019) this value represented a reduction by over 73% , saving about 540 · 10³ t of CO_{2eq} per year. The reduction in the emissions is observed because the integrated plants present higher energy yields (higher energy production per unit of biomass). The decrease in climate change impacts when MO is produced integrated into a 1G plant indicated that energy efficiency improvements in the industrial process are imperative for the improvement in the sustainability of sugarcane biorefineries in Brazil.

The economic assessment (Table 1) was performed in order to evaluate the feasibility of the integrated bioethanol-bioelectricity-MO production from sugarcane. It can be realized that the process is economically feasible, since the NPV is positive and the IRR is higher than the discount rate applied (11%). However, the feasibility of the integrated bioethanol-bioelectricity-MO production is achieved mainly because 1G bioethanol process is a traditional, established, and highly profitable industry. It should be highlighted the importance of the carbon credits representing 1.02% of the annual sales of the plant, and indicating that the environmental performance have a direct impact on the economic profitability of the biorefinery. As MO production systems from *R. toruloides* are still at an early stage of development, the TEEA results can be further improved with the development of each step of the process. Therefore, further improvement of the MO process performance could enforce the potential for the MO use as a sustainable resource for biodiesel production.

Table 1. Economic assessment of the integrated biorefinery

Parameter	Value
Capex	
Total (US\$ · 10 ⁶)	318
Sugarcane reception, juice extraction, and evaporation (%) ^a	14.4
Combined heat and power (steam and electricity) (%) ^a	32.3
Ethanol production (fermentation, distillation/dehydration) (%) ^a	18.2
Pretreatment, hemicellulose separation and concentration (%) ^a	3.01
MO production (bioreaction and separation) (%) ^a	20.03
MO extraction and purification (%) ^a	0.52
Others costs (%) ^a	11.45
Annual operating costs (US\$ * 10 ⁶ /year)	198
Annual sales	
Total (US\$ · 10 ⁶)	274
Bioethanol (%) ^b	73.7
Bioelectricity (%) ^b	11.3
MO (%) ^b	5.57
Defatted yeast (%) ^b	8.41
Carbon credit (%) ^b	1.02
NPV (US\$ * 10 ⁶ /year)	34.5
IRR (%)	12.4

^a Percentage of the total industrial investment; ^b Percentage of the annual sales.

4. Conclusions

A techno-economic and environmental assessment was performed considering the implementation of the MO production process integrated to a 1G bioethanol plant. The integrated plant was able to produce 74.8 m³/h of bioethanol, 2.8 t/h of MO, and 89.3 MW of bioelectricity. The integrated process exhibited a positive economic performance (NPV equals to US\$ 34.5x10⁶ and IRR of 12.4%) indicating that the microbial oil production from sugarcane is feasible. From the environmental perspective, the decrease in GHG emissions was 13%.

Acknowledgments

The authors thank to FAPESP (São Paulo State Research Funding Agency, Brazil), for the financial support (grant #2016/10636-8 and #2019/15851-2) and CNPq (National Council for Scientific and Technological Development, Brazil). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

- M.K. Awasthi, S. Sarsaiya, A. Patel, A. Juneja, R.P.Singh, B. Yan, S.K.Awasthi, A. Jain, T. Liu, Y. Duan, A. Pandey, Z. Zhang, M.J. Taherzadeh, 2020, Refining biomass residues for sustainable energy and bio-products: An assessment of technology, its importance, and strategic applications in circular bio-economy, *Renewable and Sustainable Energy Reviews*, 127, 109876
- B3, 2021, Crédito de Descarbonização (CBIO). Available in: http://www.b3.com.br/pt_br/produtos-e-servicos/outros-servicos/servicos-de-natureza-informacional/credito-de-descarbonizacao-cbio/ (accessed 9.26.21)

- N. Bonturi, A. Crucello, A.J.C. Viana, E.A. Miranda, 2017, Microbial Oil Production in Sugarcane Bagasse Hemicellulosic Hydrolysate without Nutrient Supplementation by a *Rhodospiridium Toruloides* Adapted Strain, *Process Biochemistry*, 57, 16–25
- ComexStat, 2020, Comex Stat: Exportação e Importação Geral.
- A.M. Elias, A.A. Longati, R.C. Giordano, F.F. Furlan, 2021, Retro-Techno-Economic-Environmental Analysis Improves the Operation Efficiency of 1G-2G Bioethanol and Bioelectricity Facilities, *Applied Energy*, 282, 116133
- M.C.B. Grassi, G.A.G. Pereira, 2019, Energy-Cane and RenovaBio : Brazilian Vectors to Boost the Development of Biofuels, *Industrial Crops & Products*, 129, 201–205.
- T.L. Junqueira, M.F. Chagas, V.L.R. Gouveia, M.C.A.F. Rezende, M.D.B. Watanabe, C.D.F. Jesus, O. Cavalett, A.Y. Milanez, A. Bonomi, 2017, Techno-Economic Analysis and Climate Change Impacts of Sugarcane Biorefineries Considering Different Time Horizons, *Biotechnology for Biofuels*, 10, 1–12
- A.A. Koutinas, A. Chatzifragkou, N. Kopsahelis, S. Papanikolaou, I.K. Kookos, 2014, Design and Techno-Economic Evaluation of Microbial Oil Production as a Renewable Resource for Biodiesel and Oleochemical Production, *Fuel*, 116, 566–77
- H.J.S. Lopes, N. Bonturi, E.A. Miranda, 2021, Induction of Resistance Mechanisms in *Rhodotorula Toruloides* for Growth in Sugarcane Hydrolysate with High Inhibitor Content, *Applied Microbiology and Biotechnology*, <https://doi.org/10.1007/s00253-021-11687-z>
- O. Ogunkunle, N.A. Ahmed, 2019, A Review of Global Current Scenario of Biodiesel Adoption and Combustion in Vehicular Diesel Engines, *Energy Reports*, 5, 1560–1580
- C. Oliveira, L.V. Pavão, M.A.S.S. Ravagnani, A.J.G. Cruz, C.B.B. Costa, 2018, Process Integration of a Multiperiod Sugarcane Biorefinery, *Applied Energy*, 213, 520–539
- M.S. Peters, K.D. Timmerhaus, R.E. West, 2002, Plant design and economics for chemical engineers, 5th Edition, Vol. 5, McGraw-Hill
- D. Pimentel, T. Patzek, 2008, Ethanol Production Using Corn, Switchgrass and Wood; Biodiesel Production Using Soybean, In *Biofuels, Solar and Wind as Renewable Systems: Benefits and Risks*, Springer, edited by David (Cornell University) Pimentel, 373–94
- E. Potrich, S.C. Miyoshi, P.F.S. Machado, F.F. Furlan, M.P.A. Ribeiro, P.W. Tardioli, R.L.C. Giordano, A.J.G. Cruz, R.C. Giordano, 2020, Replacing Hexane by Ethanol for Soybean Oil Extraction : Modeling , Simulation , and Techno-Economic-Environmental Analysis, *Journal of Cleaner Production*, 244, 118660
- J.E.A. Seabra, I. C. Macedo, H.L. Chum, C.E. Faroni, C.A. Sarto, 2011, Life Cycle Assessment of Brazilian Sugarcane Products : GHG Emissions and Energy Use, *Biofuels, Bioproducts and Biorefining*, 5, 519–32
- R. P. Soares, A. Secchi, 2003, EMSO: A new environment for modelling, simulation and optimisation, *Computer Aided Chemical Engineering*, 14, 947-952
- S.P. Souza, J.E.A. Seabra, 2014, Integrated Production of Sugarcane Ethanol and Soybean Biodiesel : Environmental and Economic Implications of Fossil Diesel Displacement, *Energy Conversion and Management Journal*, 87, 1170–79
- J.P.F. Vieira, J.L. Ienczak, P.S. Costa, C.E.V. Rossell, T.T. Franco, J.G.C. Pradella, 2016, Single Cell Oil Production Integrated to a Sugarcane-Mill : Conceptual Design , Process Specifications and Economic Analysis Using Molasses as Raw Material, 89, 478–485
- G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, B. Weidema, 2016, The ecoinvent database version 3 (part I): overview and methodology, *The International Journal of Life Cycle Assessment*, 21, 1218-1230

Modeling the hydrodynamic sizing and rating of reactive packing in Aspen Plus

Andressa Neves Marchesan, Ingrid Lopes Motta, Rubens Maciel Filho and Maria Regina Wolf Maciel

School of Chemical Engineering, University of Campinas, Av. Albert Einstein 500, Campinas, 13083-852, Brazil
marchesan.andressa@gmail.com

Abstract

Reactive distillation (RD) is the integration of reaction and thermal separation in one piece of equipment. Structured catalytic packings have been developed for this application, but their simulation and design are often limited to manufacturer software. Studies in the literature that rely on process simulation using Aspen Plus typically assume that the RD equipment has trays partially filled with catalyst, and do not take into consideration the pressure drop across the column. This simplification, in turn, affects the temperature profile and the reaction kinetics of the column, which might lead to simulation results that would be unfeasible, due to catalyst degradation. In this work, a custom model was developed for hydrodynamic modeling of the Katapak-S packing in Aspen Plus. The model was validated using literature data on pressure drop and holdup for the water-air system and the esterification of lactic acid with ethanol over Amberlyst-15, showing a good fit between the literature data and the model results. Therefore, the model offers a more representative description of RD columns and serves as a tool for the holdup, pressure drop, and diameter sizing in Aspen Plus simulations.

Keywords: Aspen Plus, Structured packing, Katapak, Reactive distillation

1. Introduction

Reactive distillation (RD) is a successful application of process intensification concepts, in which catalyzed reactions and distillation take place in a single equipment. Such a combination may be beneficial to overcome equilibrium limitations of reactions, and separate azeotropes based on the differences between the reaction rates of the substances forming the azeotrope with another reaction (Kiss et al., 2019). RD can be held in structured catalytic packings, which are modular hybrid structures composed of separation and reaction elements assembled in an alternate sequence. The commercial Katapak-S packing, commonly used in RD, is characterized by separation elements made of corrugated sheets of distillation layers, and reaction elements consisting of catalytic baskets filled with catalyst particles (Viva et al., 2011).

The hybrid structure of the catalytic packing influences the flow development in the packed bed, as well as the reactive and separation performances of the RD column. Parameters such as the holdup inside the catalytic baskets and corrugated sheets and the pressure drop across the column (Viva et al., 2011) must be monitored and represented via process simulation, but several works in the literature adopt simplifications that affect the temperature profiles and reaction kinetics of the column, leading to unfeasible simulation results. For example, Gudena et al. (2013) used the Mellapak-250Y properties in Aspen Plus to calculate column pressure drop and holdup using the

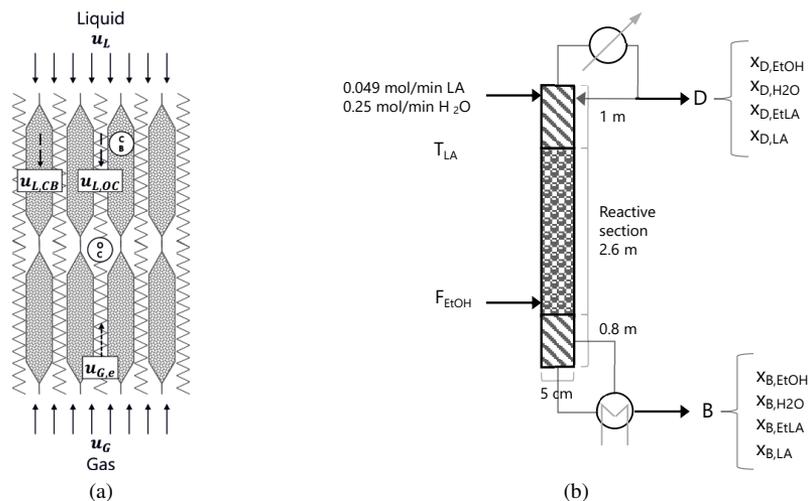


Figure 1: (a) Packing representation. Adapted from Hoffmann et al. (2004); (b) Pilot-scale RD column configuration. Experimental setup from Asthana et al. (2005)

Stichlmair method to simulate a hybrid reactive stripper-membrane process for methyl lactate hydrolysis, which has resulted in the overestimation of pressure drop (Ratheesh and Kannan, 2004). Su et al. (2013) performed simulation on lactic acid esterification in Aspen Plus assuming that the column trays are partially filled with catalyst, but they do not provide enough information regarding reaction holdup in their simulations. Therefore, the development of more rigorous models for RD simulation is required.

This work developed a custom model in Aspen Plus for the rigorous hydrodynamic simulation of the Katapak-S packing in an RD column, which was further validated against experimental data. Model validation was conducted with literature data on pressure drop and holdup for the water-air system and the esterification of lactic acid with ethanol over Amberlyst-15. The good fit between the literature data and the model results demonstrate that the custom model developed in this work offers a more representative description of RD columns and serves as a tool for RD column sizing and performance evaluation in the Aspen Plus platform.

2. Methods

To estimate the liquid holdup, pressure drop, and column diameter, the Katapak SP-12 packing characteristics were applied in a user model subroutine in the rigorous distillation model Radfrac in Aspen Plus v10. Table 1 and Table 2 show the equations and packing parameters used in the Aspen Plus packing user model, respectively. The modelling is based on the work of Ratheesh and Kannan (2004), who obtained the experimental correlations to determine the loading point and the pressure drop profile for liquid velocities in the range between 0.001 and 0.008 m/s. Figure 1a shows a representation of the catalytic packing, where CB represents the catalyst bags and, OC, the open channels.

The gas is assumed to flow only in the open channels, while the liquid flow is distributed in the open channels and the catalyst bags. An important variable in the design of packed columns is the gas loading velocity ($u_{G,LP}$, Equation 11), i.e., the gas flow rate that causes the holdup to increase from a constant value at a given liquid flow rate. The loading point distinguishes two different flow regimes, that are treated separately. Below the loading point, the liquid holdup (h_{LP} , Equation 13)

Table 1: Equations used in the Aspen Plus packing user model

Parameter Ref.	Definition	
Effective gas velocity ^a	$u_{G,e} = \frac{u_G}{\epsilon_{OC} \sin \alpha}$	(1)
Liquid load; Liquid superficial velocity ^b	$u_L = u_{L,CB} \phi + u_{L,OC} \epsilon_{OC}$	(2)
Maximum liquid velocity inside the catalyst bags ^b	$u_{L,CBmax} = \sqrt{\frac{d_p g}{\psi} \frac{\epsilon_{CB}^3}{(1 - \epsilon_{CB})}}$	(3)
Friction factor ^b	$\psi = \frac{160}{Re_{CBmax}} + \frac{3.1}{Re_{CBmax}^{0.1}}$	(4)
Effective Reynolds number inside the catalyst bags ^b	$Re_{CBmax} = \frac{u_{L,CBmax} \rho_L d_p}{(1 - \epsilon_{CB}) \mu_L}$	(5)
Flow distribution parameter ^c	$\gamma = 156.65 u_L - 0.1512$	(6)
Liquid velocity inside the catalyst bags ^c	$u_{L,CB} = \frac{u_L}{\phi(\gamma + 1)}, u_{L,CB} < u_{L,CBmax}$	(7)
	$u_{L,CB} = u_{L,CBmax}, u_{L,CB} \geq u_{L,CBmax}$	(8)
Liquid velocity in the open channels ^c	$u_{L,OC} = \frac{u_L}{\epsilon_{OC}(\gamma + 1)}, u_{L,CB} < u_{L,CBmax}$	(9)
	$u_{L,OC} = \frac{u_L - \phi u_{L,CB}}{\epsilon_{OC}}, u_{L,CB} \geq u_{L,CBmax}$	(10)
Gas loading velocity ^c	$u_{G,LP} = 0.276 u_L^{-0.285}$	(11)
F-factor	$F = u_G \sqrt{\rho_G}$	(12)
Liquid holdup below the loading point ^c	$h_{LP} = h_{L,OC} + h_{L,CB}$	(13)
Liquid holdup in the open channels ^d	$h_{L,OC} = 0.2 \left(\frac{u_{L,OC}^2 A_{SP}}{g \sin \alpha} \right)^{0.25}$	(14)
Liquid holdup inside the catalyst bags ^d	$h_{L,CB} = \phi \epsilon_{CB} \left[1 - 0.5 \left(1 - \frac{u_L}{u_{L,CBmax} (\epsilon_{OC} + \phi)} \right)^2 \right]$	(15)
Liquid holdup above the gas loading point ^c	$h_L = h_{LP} + 1.108 \left(\frac{(\Delta P / \Delta Z)_{irr}}{\rho_L g} \right)^{1.89}$	(16)
Dry gas pressure drop ^c	$\left(\frac{\Delta P}{\Delta Z} \right)_0 = f \frac{\rho_g u_{G,e}^2}{d_H}$	(17)
Friction factor ^c	$f = \frac{2.293}{Re_g^{0.308}}, \text{ for } Re_g < 1500$	(18)
	$f = \frac{0.628}{Re_g^{0.131}}, \text{ for } Re_g > 1500$	(19)
Gas phase Reynolds number ^c	$Re_g = \frac{d_H u_{G,e} \rho_g}{\mu_g}$	(20)
Irrigated pressure drop below the loading point ^c	$\left(\frac{\Delta P}{\Delta Z} \right)_{irr} = 1.941 (h_{LP})^{0.446} \left(\frac{\Delta P}{\Delta Z} \right)_0^{1.12}$	(21)
Irrigated pressure drop above the loading point ^c	$\left(\frac{\Delta P}{\Delta Z} \right)_{irr} = 0.6208 \rho_L g u_L^{0.815} (u_g \sqrt{\rho_g})^{4.183}$	(22)

^a Ellenberger and Krishna (1999) ^b Moritz and Hasse (1999) ^c Ratheesh and Kannan (2004) ^d Hoffmann et al. (2004)

is independent of the gas velocity and pressure drop, and it has two contributions: one in the open channels ($h_{L,OC}$) and another in the catalyst bags ($h_{L,CB}$). Due to difficulties in reproducing the equations reported by Ratheesh and Kannan (2004) for the liquid holdup, Equations 14 and 15 were based on the work of Hoffmann et al. (2004) for MULTIPAK packing. The pressure drop is calculated in terms of gas and liquid flow contributions, where the irrigated packing pressure drop is determined as an enhancement of the dry gas pressure drop $\left(\frac{\Delta P}{\Delta Z} \right)_0$, which is measured when only air flows upwards in the packing (Ratheesh and Kannan, 2004).

Above the loading point, the pressure drop is correlated in terms of liquid and gas superficial velocities (Equation 22) and the liquid holdup (h_L , Equation 16) is calculated as a function of the pre-loading holdup (h_{LP}) and the pressure drop. For design and rating purposes, the maximum capacity is defined as the operating point at which a pressure drop of 1200 Pa/m of packing is obtained. So, the approach to the maximum capacity is defined as the ratio between the superficial gas velocity and the velocity that would result in the 1200 Pa/m pressure drop. Also, an iterative procedure is used in the *Design Spec* tool of Aspen Plus. The reaction holdup in the reactive stages is recalculated based on the holdup in the catalyst bags determined from the hydrodynamic model ($h_{L,CB}$) and the calculated diameter to satisfy the approach to maximum capacity.

The lactic acid esterification (Equation 23) in an RD column (Figure 1b) was simulated using the pseudo-homogeneous kinetic model (Equation 24) for the liquid-phase reaction using the Amberlyst-15 ion-exchange resin (Delgado et al., 2007) for an assumed catalyst bag void fraction (ϵ_{CB}) of 0.35 and an Amberlyst-15 density of 608 kg/m³ (Yu et al., 2004). The temperature dependence of the kinetic constant k_e was determined in the range from 328 to 360 K (Equation 25), and the activity-based equilibrium constant K_{eq} estimated by Pereira et al. (2008) (Equation 26) was used. To model the vapor-liquid equilibrium, the non-random two-liquid (NRTL) activity coefficient model was used for the liquid phase and the Hayden-O'Connell equation of state was used for the vapor phase, using the parameters from Marchesan et al. (2021).



$$r = k_e \left(a_{LA} a_{EtOH} - \frac{a_{EtLA} a_{H_2O}}{K_{eq}} \right) \quad (24)$$

$$k_e = 1.88 \times 10^8 \left[\frac{\text{kmol}}{\text{s} \cdot \text{m}^3 \text{holdup}} \right] \exp \left(- \frac{58.47 \left[\frac{\text{kJ}}{\text{mol}} \right]}{RT} \right) \quad (25)$$

$$K_{eq} = \exp \left(2.9625 - \frac{515.13}{T} \right) \quad (26)$$

3. Results and Discussion

The validation of the model was performed with the comparison with experimental data for two RD systems, showing good agreement between experimental and simulated results. Figure 2 shows the comparison for air-water systems reported by Ratheesh and Kannan (2004); Götze et al. (2001) as a function of the gas load (expressed in terms of the F-factor, Equation 12). The results show that the model offers a reasonable estimate of the hydrodynamic behavior of the packing, namely the holdup and pressure drop across the column for varying liquid and gas loads. Despite its limitations, it offers a more suitable approximation of industrial applications compared to other models based on tray columns with an estimate of the volume occupied by the catalyst in each tray (Su et al., 2013; Dai et al., 2019; Mo et al., 2011).

Table 2: Packing parameters used in the hydrodynamic model

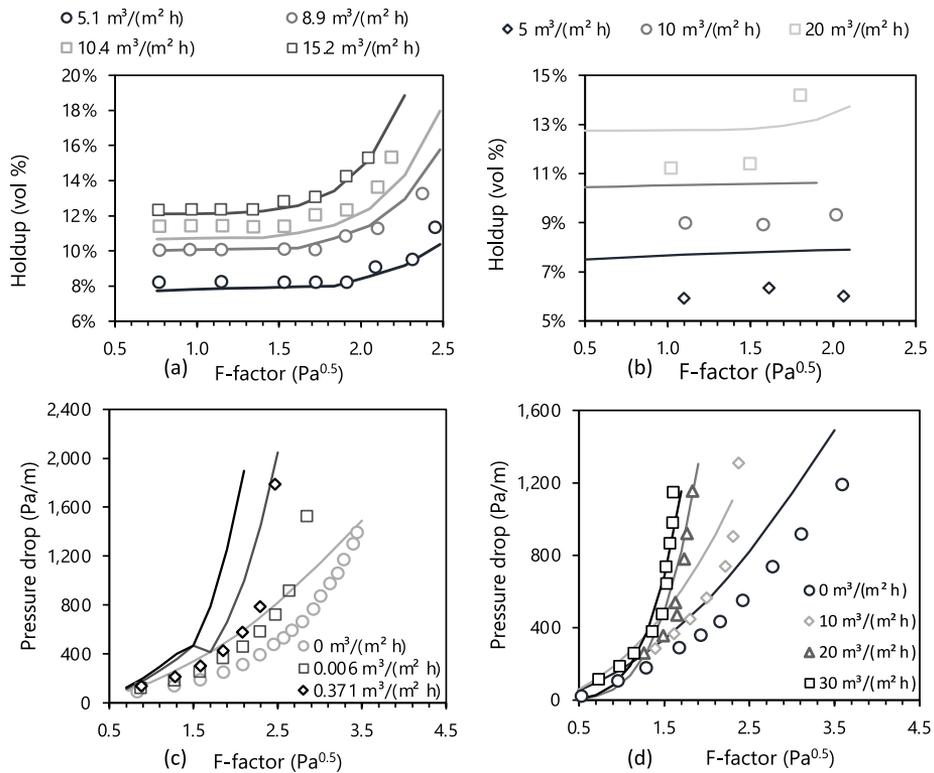
Packing properties ^{Ref.}	Value	Packing properties ^{Ref.}	Value
d_p – Catalyst particle diameter (m) ^a	0.0009	ϵ_{CB} – Catalyst bag void fraction ^a	0.35
ϕ – Catalyst volume fraction in the packing element ^a	0.242	ϵ_{OC} – Open channel void fraction ^a	0.738
A_{SP} – Specific surface area in the open channel (m ² /m ³) ^b	454	d_H – Equivalent diameter (m)	0.0065
α – Channel angle relative to the vertical axis ^a	45°		

^a Ratheesh and Kannan (2004) ^b value for MellapakPlus 752.Y, Viva et al. (2011)

Table 3: Simulation and experimental results comparison for lactic acid esterification. Experimental results from Asthana et al. (2005)

Run	F_{EiOH} (mol/min)	T_{LA} (°C)	D (mol/min)	Experimental								
				$x_{B,EiOH}$	$x_{B,H2O}$	$x_{B,EiLA}$	$x_{B,LA}$	$x_{B,EiOH}$	$x_{B,H2O}$	$x_{B,EiLA}$	$x_{B,LA}$	X_{LA}
1	0.35	25	0.36	0.730	0.095	0.140	0.030	0.320	0.680	0.003	0.000	79%
2	0.5	100	0.39	0.760	0.038	0.170	0.027	0.310	0.690	0.005	0.000	82%
3	0.5	100	0.37	0.810	0.033	0.130	0.020	0.320	0.680	0.005	0.000	83%

Run	F_{EiOH} (mol/min)	T_{LA} (°C)	D (mol/min)	Simulation								
				$x_{B,EiOH}$	$x_{B,H2O}$	$x_{B,EiLA}$	$x_{B,LA}$	$x_{B,EiOH}$	$x_{B,H2O}$	$x_{B,EiLA}$	$x_{B,LA}$	X_{LA}
1	0.35	25	0.36	0.672	0.159	0.133	0.036	0.322	0.675	0.002	0.001	78.5%
2	0.5	100	0.39	0.733	0.079	0.161	0.027	0.300	0.697	0.002	0.001	85.1%
3	0.5	100	0.37	0.817	0.070	0.097	0.016	0.287	0.710	0.002	0.001	85.3%


 Figure 2: Holdup (a-b) and pressure drop (c-d) at various liquid loads (m³/(m² h)) in comparison with experimental results from (a, c) Ratheesh and Kannan (2004) and (b, d) Götze et al. (2001).

A further validation of the custom model was performed with experimental data of the esterification of lactic acid with ethanol over Amberlyst-15 in the RD column depicted in Figure 1b for three scenarios under varying operational conditions (ethanol molar flow rate, lactic acid feed temperature, and distillate molar flow rate), based on the results reported by Asthana et al. (2005). Table 3 shows that, except for the overrepresentation of the water mole fraction in the bottom stream, the mole fraction estimations produced by the custom model are in accordance with the experimental results. Also, the temperature of the reboiler (close to 80 °C) and the calculated conversions (around 80%) in the three scenarios obtained via simulation were close to the values reported

experimentally. The differences in the composition of the bottom streams can be explained by the formation of lactic acid oligomers in small amounts for the chosen scenarios, which were not included in this study for simplification. Furthermore, Asthana et al. (2005) reported material balances closure of $\pm 7\%$.

4. Conclusion

This work developed a rigorous custom model in Aspen Plus for the hydrodynamic modeling of the Katapak-S packing in reactive distillation columns, considering the effect of performance parameters such as the holdup and the pressure drop. The model was validated with experimental datasets of two different applications (water-air system and esterification of lactic acid with ethanol). This showed that it successfully represents the performance of RD columns with Katapak-S packings for varying applications and process conditions. Therefore, the custom model serves as a reliable tool for future studies on the sizing and performance evaluation of RD columns in Aspen Plus.

5. Acknowledgements

This work was supported by grant #141584/2017-3 CNPq, National Council for Scientific and Technological Development – Brazil, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and grant #2015/20630-4 São Paulo Research Foundation (FAPESP).

References

- N. Asthana, A. Kolah, D. T. Vu, C. T. Lira, D. J. Miller, 2005. A continuous reactive separation process for ethyl lactate formation. *Organic Process Research and Development* 9, 599–607.
- S. B. Dai, H. Y. Lee, C. L. Chen, 2019. Design and economic evaluation for the production of ethyl lactate via reactive distillation combined with various separation configurations. *Industrial and Engineering Chemistry Research* 58, 6121–6132.
- P. Delgado, M. T. Sanz, S. Beltrán, 2007. Kinetic study for esterification of lactic acid with ethanol and hydrolysis of ethyl lactate using an ion-exchange resin catalyst. *Chemical Engineering Journal* 126, 111–118.
- J. Ellenberger, R. Krishna, 1999. Counter-current operation of structured catalytically packed distillation columns: Pressure drop, holdup and mixing. *Chemical Engineering Science* 54, 1339–1345.
- K. Gudena, G. P. Rangaiah, S. Lakshminarayanan, 2013. Modeling and analysis of hybrid reactive stripper-membrane process for lactic acid recovery. *Industrial & Engineering Chemistry Research* 52, 2907–2916.
- L. Götz, O. Bailer, P. Moritz, C. V. Scala, 2001. Reactive distillation with katapak®. *Catalysis Today* 69, 201–208.
- A. Hoffmann, C. Noeres, A. Górák, 2004. Scale-up of reactive distillation columns with catalytic packings. *Chemical Engineering and Processing: Process Intensification* 43, 383–395.
- A. A. Kiss, M. Jobson, X. Gao, 2019. Reactive Distillation: Stepping Up to the Next Level of Process Intensification. *Industrial and Engineering Chemistry Research* 58 (15), 5909–5918.
- A. N. Marchesan, J. F. L. Silva, R. M. Filho, M. R. W. Maciel, 2021. Techno-economic analysis of alternative designs for low-ph lactic acid production. *ACS Sustainable Chemistry & Engineering* 9, 12120–12131.
- L. Mo, J. Shao-Tong, P. Li-Jun, Z. Zhi, L. Shui-Zhong, 2011. Design and control of reactive distillation for hydrolysis of methyl lactate. *Chemical Engineering Research and Design* 89, 2199–2206.
- P. Moritz, H. Hasse, 1999. Fluid dynamics in reactive distillation packing katapak®-s. *Chemical Engineering Science* 54, 1367–1374.
- C. S. Pereira, S. P. Pinho, V. M. Silva, A. E. Rodrigues, 2008. Thermodynamic equilibrium and reaction kinetics for the esterification of lactic acid with ethanol catalyzed by acid ion-exchange resin. *Industrial and Engineering Chemistry Research* 47, 1453–1463.
- S. Ratheesh, A. Kannan, 2004. Holdup and pressure drop studies in structured packings with catalysts. *Chemical Engineering Journal* 104, 45–54.
- C. Y. Su, C. C. Yu, I. L. Chien, J. D. Ward, 2013. Plant-wide economic comparison of lactic acid recovery processes by reactive distillation with different alcohols. *Industrial & Engineering Chemistry Research* 52, 11070–11083.
- A. Viva, S. Aferka, E. Brunazzi, P. Marchot, M. Crine, D. Toye, 2011. Processing of x-ray tomographic images: A procedure adapted for the analysis of phase distribution in mellapakplus 752.y and katapak-sp packings. *Flow Measurement and Instrumentation* 22, 279–290.
- W. Yu, K. Hidajat, A. K. Ray, 2004. Determination of adsorption and kinetic parameters for methyl acetate esterification and hydrolysis reaction catalyzed by amberlyst 15. *Applied Catalysis A: General* 260, 191–205.

Simplified Model-based Design of Plate-fin Microdevices with Uniform Flow Distribution at High Flow Rates

Osamu Tonomura, Kaori Maenaka, Shinji Hasebe

*Dept. of Chem. Eng., Kyoto University, Nishikyo, Kyoto 615-8510, Japan
tonomura@cheme.kyoto-u.ac.jp (O. Tonomura)*

Abstract

Computational fluid dynamics (CFD) provides the detailed information about flow and pressure distributions in microdevices, whose importance in the chemical process industry has increased in recent years. However, since CFD requires a considerable computational time, the use of a simple model different from CFD would be desirable to efficiently investigate the influence of design parameters on the flow and pressure distributions in microdevices. In this study, the previously proposed simple model, which is based on the predictive equation for the pressure loss due to viscous forces, is extended by taking into account inertial forces. A case study on plate-fin microdevices shows that the extended simple model can predict the flow distribution among microchannels with almost the same accuracy as CFD not only under low flow rates but also under high flow rates. In addition, another case study demonstrates that the optimal shape of manifolds of stacked/unstacked plate-fin microdevices with uniform flow distribution is efficiently derived using the extended simple model.

Keywords: Shape design, Microdevices, Simple model, CFD model.

1. Introduction

Microdevices, which are composed of channels on the order of micrometers or millimeters, have large specific surface area and short diffusion distance. Therefore, microdevices are characterized by high-efficiency heat exchange and rapid mixing. There are many reports that these characteristics have improved the reaction performance of organic synthesis, polymerization, particle synthesis, and so on. While expectations are growing that microdevices are used for mass production of chemical materials, it is necessary to develop models for analyzing flow and transport phenomena, optimal design and control methodologies, and sensing and monitoring technologies. The throughput of microdevices is increased by numbering-up, that is, parallelization of channels. At this time, if the flow distribution among the channels is not uniform, the product quality may deteriorate (Delsman et al., 2005). Since the channels are small, it is not realistic to install sensors and actuators in all the channels. Therefore, it is important to appropriately design a flow distributor that bundles the parallelized channels. So far, bifurcation-type (Amador et al., 2004), manifold-type (Commenge et al., 2002), and split-and-recombine-type (Tanaka et al., 2011) flow distributors have been reported. Computational fluid dynamics (CFD) model is often used to predict the flow patterns in the flow distributors and derive optimal design conditions (Tonomura et al., 2004). By comparison with experiments, it has been reported that CFD achieves high prediction accuracy of flow and transport phenomena in microdevices. (Bothe et al., 2006; Zhendong et al., 2012).

The use of CFD model enables rigorous flow simulation, but enormous time and effort is associated with geometric model creation, mesh generation, and calculations. Therefore, an optimal design method based solely on the CFD model may not be realistic. So far, a simple model has been proposed in place of the CFD model (Commengé et al., 2002). This simple model, which is called the conventional simple model in this paper, is based on the formula for estimating the pressure loss of laminar flow in a pipe from the average flow velocity and the pipe size. The pressure and flow velocity distributions over the device are approximated by connecting the compartments to which the conventional simple model is applied. However, since the conventional simple model does not consider the influence of fluid inertia forces, there is a problem that the prediction accuracy decreases for high Re (Reynolds) flow. In this study, the conventional simple model is extended by taking into account inertial forces. The usefulness of the extended simple model is verified through case studies on plate-fin microdevices.

2. Simple Model Development

A schematic diagram of the plate-fin microdevice is shown in Fig. 1. A reactant is fed to the inlet manifold, distributed to the parallelized microchannels, and rejoined at the outlet manifold. The model described in this chapter is used to efficiently design the shape of the device so that the flow distribution among the parallelized microchannels is uniform.

As shown in Fig. 2, the plate-fin microdevice is divided by compartments. Each of microchannels is considered as one compartment, and the inlet and outlet manifolds are divided into as many compartments as the number of microchannels (N [-]). The compartments assigned to the inlet manifold, outlet manifold, and microchannels are called the distribution, junction, and channel compartments, respectively, which are numbered in order from the upstream side. In the distribution compartments, the flow velocity and pressure at the inlet of the i -th compartment are expressed as U_1^i and P_1^i ,

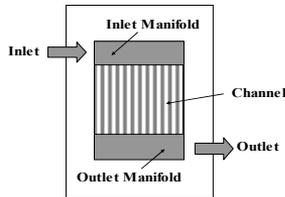


Fig. 1 Plate-fin microdevice.

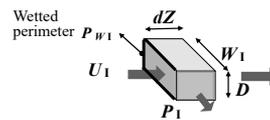


Fig.3 Small volume element for a compartment.

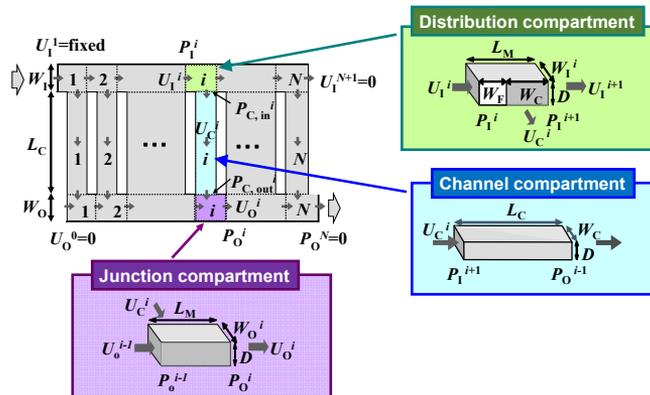


Fig. 2 Compartment-based simple model for plate-fin microdevice.

respectively. In the junction compartments, the flow velocity and pressure at the outlet (after mixing) of the i -th compartment are expressed as U_o^i and P_o^i , respectively. In the channel compartments, the flow velocity, inlet pressure, and outlet pressure of the i -th compartment are expressed as U_c^i , $P_{c, in}^i$, and $P_{c, out}^i$, respectively. Each compartment is assumed to be a rectangular channel having length L [m], width W [m], and height D [m]. The subscript C for each variable is the microchannel, I is the inlet manifold, O is the outlet manifold, and M is both manifolds. W_F [m] is the distance between parallel channels, that is, the fin width. In this study, D is fixed to be constant in all compartments and the z axis is defined in the direction of flow at the inlet. The flow velocity and pressure distributions over the device are predicted on the basis of the mass balance and pressure balance equations among the compartments. The mass balance is expressed by

$$W_1^i U_1^{i+1} - W_1^i U_1^i = -W_c U_c^i \quad 1 \leq i \leq N, \quad (1)$$

$$W_o^i U_o^i - W_o^{i-1} U_o^{i-1} = W_c U_c^i \quad 1 \leq i \leq N. \quad (2)$$

The pressure balance among adjacent compartments is expressed by

$$\Delta P_c^i + \Delta P_o^i = \Delta P_1^{i+1} + \Delta P_c^{i+1} \quad 1 \leq i \leq N - 1. \quad (3)$$

ΔP_c^i is given by the following equation that estimates the pressure drop of a fully developed laminar flow in a rectangular channel.

$$\Delta P_c^i = P_{c, in}^i - P_{c, out}^i = \frac{32\lambda_c \mu L_c}{D_{Hc}^2} U_c^i \quad 1 \leq i \leq N \quad (4)$$

Here, λ_c is the correction coefficient for the cross-sectional shape of the channel, μ is the fluid viscosity [Pa·s], and D_H is the hydraulic equivalent diameter [m]. ΔP_1^i and ΔP_o^i are given by the following equations:

$$\Delta P_1^i = P_1^i - P_1^{i+1} = 8\lambda_1 \frac{(D + 2W_1^i)W_c^i + 2(D + W_1^i)W_F}{DW_1^i D_{H1}^i} \mu U_1^i + 2\rho U_1^i (U_1^{i+1} - U_1^i) \quad (5)$$

$$\Delta P_o^i = P_o^{i-1} - P_o^i = 8\lambda_o \frac{(D + 2W_o^i)W_c^i + 2(D + W_o^i)W_F}{DW_o^i D_{Ho}^i} \mu U_o^i + 2\rho U_o^i (U_o^i - U_o^{i-1}) \quad (6)$$

These equations were derived by formulating the momentum change in the z -axis direction due to pressure loss and shear stress for a small volume element as shown in Fig. 3. On the right-hand side of each equation, the first and second terms represent the viscous and inertia terms, respectively. The feature of this study is that the pressure loss and pressure balance equations are established in consideration of both viscous and inertia terms. The conventional simple model ignores the inertia term, and its application will be limited to low flow rates. When the dimensions of all compartments and the following

Table 1 Design conditions and physical properties.

	Symbol	Value	Unit
Number of channels	N	10	-
Depth	D	500	μm
Manifold width	W_M	1000	μm
Fin width	W_F	200	μm
Fluid density	ρ	1000	kg/m^3
Fluid viscosity	μ	0.001	Pa s

Table 2. Design conditions.

	Symbol	Value	Unit
Number of channels	N	10	-
Channel depth	D	500	μm
Channel length	L_c	40	mm
Channel width	W_c	500	μm
Fin width	W_F	200	μm
Device inlet width	W_1^1	2000	μm
Device outlet width	W_o^N	2000	μm
Inlet flow velocity	U_1^1	0.05	m/s
Fluid density	ρ	1000	kg/m^3
Fluid viscosity	μ	0.001	Pa s

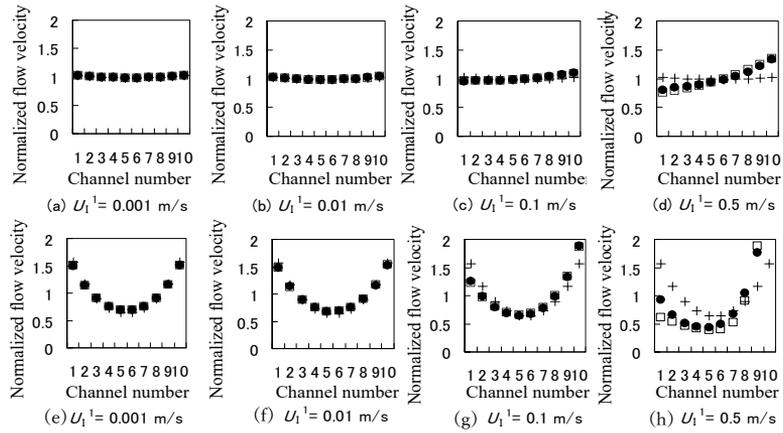


Fig. 4 Flow velocity distribution among channels. The results from (a) to (d) were obtained when $(L_c, W_c) = (20 \text{ mm}, 300 \mu\text{m})$ and the results from (e) to (h) were obtained when $(L_c, W_c) = (20 \text{ mm}, 1000 \mu\text{m})$.

boundary conditions,

$$U_1^1 = \text{fixed}, \quad U_0^0 = 0, \quad U_i^{N+1} = 0, \quad (7)$$

are given to the mass balance and pressure balance equations, the number of unknown variables and the number of independent equations are equal to $4N$, that is, the degrees of freedom are 0, and as a result, U_i^i , U_c^i , and U_0^i can be derived.

3. Evaluation of Prediction Accuracy of the Developed Model

A case study on the plate-fin microdevice is conducted to compare the prediction accuracy of the conventional simple model, the developed simple model, and the CFD model. Table 1 shows the design conditions and fluid physical properties. The width of all compartments, W_i^i and W_0^i , is set to be constant at W_M . The remaining design conditions are as follows: $(L_c, W_c) = (20 \text{ mm}, 300 \mu\text{m})$ and $(20 \text{ mm}, 1000 \mu\text{m})$. Under these conditions, flow simulation based on each model is performed for $U_1^1 = 0.001$ m/s, 0.01 m/s, 0.1 m/s, and 0.5 m/s. Ansys Fluent® and gPROMS® are used for calculations based on the CFD model and the simple models, respectively. Comparing the calculation times, the CFD model took tens of minutes to several hours, but the conventional and developed simple models took only 1 to 2 seconds. The flow velocity of each microchannel normalized by flow velocity averaged by all microchannels ($U_{c, \text{ave}}$) is plotted in Fig. 4, where +, □, and ● represents the results of conventional, developed, and CFD models, respectively. The results of CFD model show that as the inlet flow velocity increases, more fluid flows into the microchannels far from the inlet due to the influence of flow inertia. The results of the conventional simple model differ from those of the CFD model, especially when the flow velocity is high. On the other hand, the results of the developed simple model are close to those of the CFD model under any flow velocity condition.

4. Design of Plate-fin Microdevice

The shape design of manifolds of a plate-fin microdevice is carried out. The design conditions are shown in Table 2. The design objective is to minimize the manifold volumes. The manifold shape is optimized under the following constraints: 1) in each microchannel, the relative deviation of the flow velocity from the value under uniform flow distribution is 0.25% or less, 2) the upper limit of the pressure loss of the entire

Table 3 Flow velocity in each microchannel. [m/s]

Channel	1	2	3	4	5	6	7	8	9	10
Conventional model (CM)	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200
CFD for the result by CM	0.0189	0.0191	0.0193	0.0195	0.0198	0.0201	0.0204	0.0207	0.0209	0.0212
Developed model (DM)	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200	0.0200
CFD for the result by DM	0.0197	0.0198	0.0200	0.0200	0.0200	0.0201	0.0201	0.0201	0.0201	0.0201

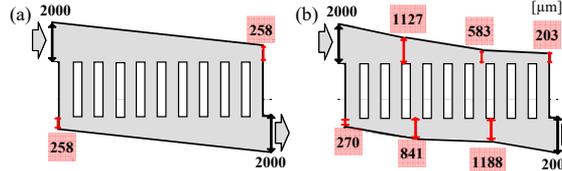


Fig. 5 Manifold design results by the conventional model (a) and the developed model (b).

device is 500 Pa, and 3) the upper and lower limits of the width of the manifolds are 2 mm and 0.2 mm, respectively. The shape of the inlet and outlet manifolds is represented by a piecewise linear function. The design procedure is as follows. First, assuming that the device has a point-symmetrical shape, optimization is performed with the number of sections of the piecewise linear function in both manifolds being 1. If no solution is obtained, the device is considered to have a non-point symmetric shape and optimization is performed. If no solution is still obtained, the design optimization is repeated while increasing the number of sections by 1. As a result of this optimization, the manifold shapes shown in Fig. 5 were derived. Table 3 shows the flow velocity distribution among the microchannels by the simple models and the CFD model. The conventional simple model predicts the achievement of the uniform flow distribution, but this prediction result is not achieved when CFD is executed for the manifold shape derived by the conventional simple model. On the other hand, it is shown that there is almost no difference between the results of the developed simple model and the CFD model.

5. Design of Stacked Plate-fin Microdevices

In addition to parallelizing the channels in one device, it is possible to increase the production amount by stacking the device. In this chapter, the microdevice designed in the previous chapter is stacked. A schematic diagram of the entire apparatus when the number of stacked devices is 10 is shown in Fig. 6. The purpose of this chapter is to design the inlet and outlet headers that realize equal fluid distribution to the stacked devices. The design conditions are as follows: 1) the shape and dimensions of the channels and manifolds of each plate-fin microdevice are equal to the design result in the previous chapter, 2) the thickness of each plate-fin microdevice is 1 mm, 3) the header is a circular tube with a constant diameter, 4) the header and each microdevice are connected by a rectangular duct with depth of 500 μm , width of 2 mm and length of 4 mm, 5) the number of stacked microdevices is 10, 6) total flow rate is 0.5 cm^3/s , 7) assuming that the pressure loss of each microdevice are proportional to flow rate, and 8) the fluid is water (298K). The design objective is to minimize the header volumes. The diameters of inlet and outlet headers are optimized under the following constraints: 1) in each microdevice, the relative deviation of the flow rate from the value under uniform flow distribution is 0.25% or less, and 2) the upper and lower limits of the header diameter are 10 mm and 2 mm, respectively. This optimum design problem is solved using the developed simple model. That is, the stacked microdevices and the headers are regarded as the parallelized microchannels and the manifolds in the developed simple model, respectively, as shown in Fig. 6. As a result of optimization, D_{IN} and D_{OUT} were determined to be 5.24 mm and 5.96 mm, respectively. To verify this design result, CFD simulation was performed. Table

Table 4 Flow rate in each microdevice.

Device number		1	2	3	4	5	6	7	8	9	10
The developed model	(after optimization)	0.0499	0.0499	0.0499	0.0500	0.0500	0.0500	0.0500	0.0501	0.0501	0.0501
	(after optimization)	0.0497	0.0500	0.0500	0.0500	0.0500	0.0500	0.0501	0.0501	0.0501	0.0500
CFD	(before optimization)	0.0435	0.0450	0.0458	0.0468	0.0481	0.0496	0.0514	0.0534	0.0559	0.0594

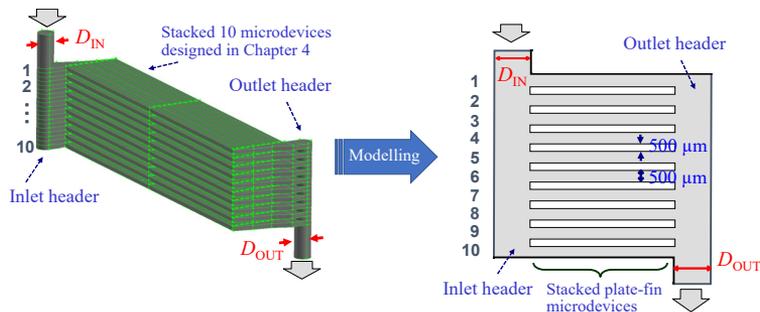


Fig. 6 Stacked plate-fin microdevice (left) and application of the developed model to it (right).

4 shows the flow rate of each microdevice. It was shown that a more uniform flow distribution was achieved after optimization, compared to before optimization ($D_{IN} = D_{OUT} = 2$ mm), and that the predicted flow rates after optimization of the developed simple model and the CFD model were almost the same. Therefore, it was confirmed that the developed simple model is useful not only for the optimal design of the plate-fin microdevice itself but also for the optimal design when it is stacked.

6. Conclusions

The application of microdevices in the chemical process industry has gained significant importance in recent years. The production capacity of microdevices can be increased by numbering-up. Such microdevices often have manifolds that help to guide the fluid into many microchannels. The flow uniformity depends on pressure and flow distributions inherent to the design of manifolds and others. In this study, the compartment-based simple model was developed to realize efficient design. The developed model is composed of mass balance and pressure balance equations among compartments. When establishing pressure balance equation, both viscous and inertia terms were considered in the developed model. The conventional model ignored the inertia term, and as a result, its application was limited to low flow rates. The usefulness of the developed simple model was verified through case studies on plate-fin microdevices at high flow rates.

References

- C. Amador, et al., 2004, Flow Distribution in Different Microreactor Scale-out Geometries and the Effect of Manufacturing Tolerances and Channel Blockage, *Chem. Eng. J.*, 101, 379-390.
- D. Bothe, et al., 2006, Fluid Mixing in a T-shaped Micro-mixer, *Chem. Eng. Sci.*, 61, 2950-2958.
- J.M. Commenge, et al., 2002, Optimal Design for Flow Uniformity in Microchannel Reactors, *AIChE J.*, 48, 345-358.
- E.R. Delsman, et al., 2005, The Influence of Differences Between Microchannels on Micro Reactor Performance, *Chem. Eng. Technol.*, 28, 367-375.
- Y. Tanaka, et al., 2011, Detection and Diagnosis of Blockage in Parallelized Microreactors, *Chem. Eng. J.*, 167, 483-489.
- O. Tonomura, et al., 2004, Detection and Diagnosis of Blockage in Parallelized Microreactors, *Chem. Eng. J.*, 101, 1-3, 397-402.
- L. Zhendong, et al., 2012, Mixing Characterization and Scaling-up Analysis of Asymmetrical T-shaped Micromixer: Experiment and CFD Simulation, *Chem. Eng. J.*, 181-182, 597-606.

CFD-based study of fluid flow and transport phenomena in fixed bed compact reactors

Osamu Tonomura, Akihiro Kitagawa, Kazuki Kato, Taisuke Maki, Ken-ichiro Sotowa

*Dept. of Chem. Eng., Kyoto University, Nishikyo, Kyoto 615-8510, Japan
tonomura@cheme.kyoto-u.ac.jp (O. Tonomura)*

Abstract

The fixed bed reactor design is traditionally based on a model that assumes a piston flow and regards the fixed bed as homogeneous porosity. This kind of model is often unsuitable for designing fixed bed compact reactors in which wall and local phenomena dominate. In this work, three-dimensional steady-state computational fluid dynamics (CFD) simulation was performed to analyze how the geometric structured packing with spherical particles, reactor diameters, and particle sizes, which are important design variables, influence the conversion and selectivity of a model reaction. The simulation results show that the generation of flow perpendicular to the axis of the reactor leads to the improvement of the reactor performance. In addition, it was suggested that the honeycomb type fixed bed compact reactor is superior to the packing type one in realizing a uniform reaction temperature field and reducing the pressure loss.

Keywords: CFD simulation, Compact reactor, Fixed bed, Fluid dynamics, Heat transfer.

1. Introduction

The market for functional chemicals, including pharmaceutical intermediates, is growing. Currently, functional chemicals are manufactured exclusively by the batch method, and a large amount of waste containing organic solvents is discharged, and energy is consumed to dispose of the waste. In recent years, it has been reported that functional chemicals are synthesized by a flow method utilizing an immobilized catalyst having high selectivity (Tsubogo et al., 2015). It is expected that the use of such a flow method will increase in the future in place of the batch method. In a reaction process with an immobilized catalyst, compact reactors such as microreactors, which are expected to enable rapid mass and heat transfer due to the large specific surface area between different phases, are often used. However, research on fixed bed compact reactors is not sufficient compared to conventional reactors, and their design methods have not been established. Although catalyst development is being actively carried out, the performance of the developed catalyst will not be fully exhibited without consideration of appropriate reactor design. It is important to rationally solve the engineering problems of the fixed bed compact reactors such as temperature control and pressure loss reduction, based on the models, which can quantitatively express mass and heat transfer as well as fluid dynamics.

When analyzing the flow and transport in a conventional reactor, a model that assumes a piston flow and regards the fixed bed as homogeneous porosity is often used. On the other hand, in the case of a fixed bed compact reactor, which is often used for reaction processes involving a large amount of heat of reaction, a model with high prediction accuracy of heat transfer is required to realize strict reaction control. Considering that heat transfer

depends on flow, a model that can analyze local flow and transport is required. Therefore, it is useful to analyze the fixed bed compact reactor based on CFD simulation, which can express the discrete porosity distribution in consideration of an individual shape of catalyst carriers. In CFD simulations of fixed bed conventional reactors, CFD has been considered to be a powerful tool for predicting the hydrodynamics and mass transfer from microscale perspectives (Haroun et al., 2012; Hamidipour et al., 2013; Bouras et al., 2021). CFD simulation will also be useful for fixed bed compact reactors. In this study, the flow and transport characteristics of a fixed bed compact reactor were analyzed by CFD simulation. How the packing methods of spherical catalyst carriers, reactor diameters, and catalyst carrier sizes, which are important design variables, affect the conversion and selectivity of a model reaction was investigated.

2. Simulation settings

This chapter describes the settings for CFD simulation of the reactor.

2.1. Reaction system

A reaction system, its rate equation, and the physical properties of fluid and catalyst were given with reference to the reaction examples dealt with in NEDO project (see Ref.). The parallel reaction was assumed to proceed on the surface of the catalyst particles. A is the raw material, B is the main product, and C is the by-product. The temperature dependence of each reaction rate follows the Arrhenius equation. The reaction parameters including standard reaction enthalpy ($\Delta_r H_i^\circ$) are shown in Table 1. The higher the reaction temperature, the more by-product is produced. The physical properties of the reaction fluid and catalyst are shown in Table 2. They were all assumed to be constant.



2.2. Design and operation conditions

Figure 1 shows an example of packing type fixed bed compact reactors used in this study. Assuming that spherical particles are regularly packed as a carrier, two typical of packing methods were examined as shown in Fig. 1. Packing method 1 (PM1) has a fundamental structure where one particle is placed between adjacent particles located below it. Packing method 2 (PM2) has a fundamental structure where one particle is placed directly above another particle located below it. Assuming that the shape of the reactor is tubular, the inner diameter of the reactor (d_r) was set to 10 mm and 40 mm, and the ratio of the inner diameter of the reactor to the carrier particle size (N) was set to 10 and 4, referring to the reactors developed in the NEDO project. The reaction fluid was supplied from the reactor inlet at a temperature of 300 K, a linear velocity of 0.08 m/s, and a raw material concentration of 1 kmol/m³. The outlet pressure of the reactor was assumed to be equal to atmospheric pressure. The reactor wall temperature was constant at 300 K.

Table 1 Reaction parameters.

k_{10} [1/s]	4.78×10 ⁶
k_{20} [1/s]	4.78×10 ⁶
E_1 [kJ/mol]	31.0
E_2 [kJ/mol]	36.5
$\Delta_r H_1^\circ$ [kJ/mol]	10
$\Delta_r H_2^\circ$ [kJ/mol]	20

Table 2 Physical properties of reaction fluid and catalyst.

Density [kg/m ³]	700 (fluid), 3940 (cat.)
Specific heat [J/kg/ K]	2000 (fluid), 780 (cat.)
Thermal conductivity [W/m/ K]	0.1 (fluid), 300 (cat.)
Viscosity [kg/m/s]	0.0003 (fluid)
Diffusivity [m ² /s]	1×10 ⁻⁵ (fluid)
Molar mass [g/mol]	50 (fluid)

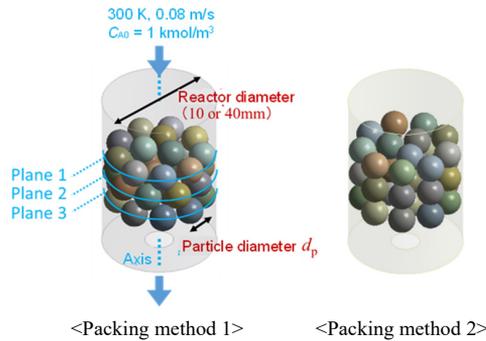


Fig. 1 A fixed bed compact reactor for $N = 4$.

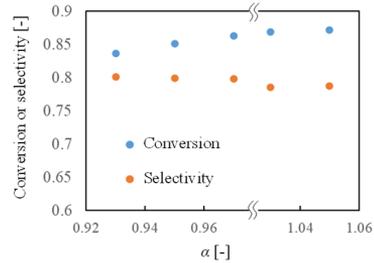


Fig. 2 Effect of α on reaction results
(Case 7 in Table 3).

Table 3 Simulation conditions and results in each case

Case	α	d_r	N	PM	X_r	X_s
1	0.95	10	10	1	1.000	0.812
2				0.997	0.810	
3			4	1	0.986	0.833
4				2	0.915	0.834
5		40	10	1	0.963	0.781
6				2	0.905	0.781
7			4	1	0.851	0.799
8				2	0.712	0.806

In a preliminary study to evaluate the flow velocity distribution over the packed bed while changing the number of layers from 3 to 10, the result was that there was almost no effect of the number of layers on the flow velocity distribution. Therefore, in consideration of the computational load, the number of layers was fixed at 3. In addition, regarding mesh generation by CFD, the mesh quality deteriorates near the contact points between particles and the contact points between particles and reactor wall, and as a result, mesh generation cannot be performed. To avoid this problem, the particle volume change coefficient α was introduced and set to 0.95 (Bai et al., 2009). This means that each particle is slightly shrunk in size but its position remains unchanged. It was examined in advance that the relationship between the reaction results and the particle shrinkage or expansion is as shown in Fig. 2. In this figure, it was judged that α does not have a large effect on the reaction results.

3. Simulation results and discussion

CFD simulation was performed using Ansys Fluent[®] software to investigate how the packing methods (PMs), reactor diameters (d_r), and catalyst carrier sizes ($d_p = d_r/N$), which are important design variables, affect the reaction conversion and selectivity, which are represented by X_r and X_s , respectively. The simulation results of eight cases, where the diffusion coefficient D was set to 1.0×10^{-5} m²/s and the catalyst area in the reactor was constant, are summarized in Table 3. In addition, the cross section of the packed bed (Planes 1-3) and the cross section along the axis were analyzed to evaluate the distributions of velocity, concentration, and temperature. The following sections show and discuss the obtained CFD simulation results in detail.

3.1. Packing methods: PM1 vs. PM2

Focusing on Cases 7 and 8, the effect of the PMs on the reaction results is discussed. The X_r of PM1 was 0.139 higher than that of PM2. The tendency of this result was also

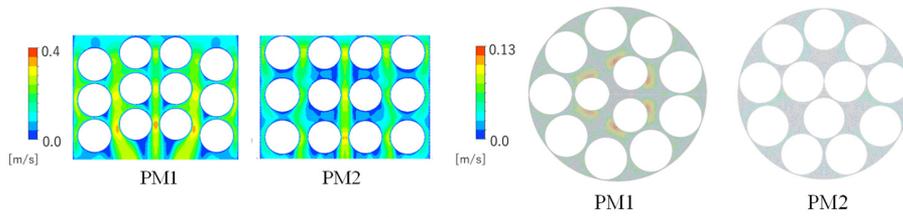


Fig. 3 Flow velocity distribution on the axis plane.

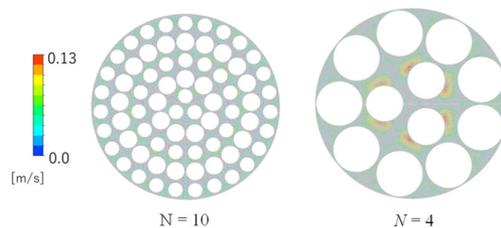
Fig. 4 Radial flow velocity distribution on plane 3.

observed in other similar cases. It was also shown that the difference in X_r was larger when N was smaller and when d_r was larger. As shown in Fig. 3, PM2 has more regions where the velocity is 0 above and below each catalyst particle, that is, stagnant portions, compared to PM1, and as a result, the raw material is not efficiently supplied to the particle surface. On the other hand, in the case of PM1, it can be seen that radial flows occur as shown in Fig. 4. The maximum radial flow velocity corresponds to 0.63 times the axial flow velocity in the bed. It can be estimated from the CFD analysis results that the effective area of the catalyst particles used for the reaction is about 1.3 times larger in Case 7 than in Case 8, and it was reflected in the reaction results.

3.2. Reactor diameter and catalyst carrier size

The effect of $N (=d_r/d_p)$ on the reaction results is discussed. Focusing on Cases 5 and 7, it was shown that X_r was 0.112 higher and X_s was 0.018 lower for $N = 10$ than for $N = 4$. X_r became higher for $N = 10$ than for $N = 4$, regardless of cases. The increase in X_r may be due to the fact that the catalyst surface area per unit volume of $N = 10$ is 2.5 times that of $N = 4$. However, if the heat of reaction is not removed efficiently, it is expected that the catalyst surface temperature will rise, resulting in a decrease in X_s . This expected result was actually seen in Cases 5 and 7. When PM1 is adopted as the catalyst packing method, the efficient heat removal by convective heat transfer of radial flow may be achieved. As shown in Fig. 5, the radial flow was observed in both Cases 5 and 7. However, it can be seen that the maximum velocity of the radial flow in the case of $N = 10$ is about half that in the case of $N = 4$. Therefore, it was difficult to expect a great heat removal effect by the radial flow in the case of $N = 10$.

The effect of d_r on the reaction results is discussed. Focusing on Cases 3 and 7, it was shown that when d_r was changed from 40 mm to 10 mm, the reaction rate increased by 0.135 and the selectivity increased by 0.034. This is because as d_r becomes smaller, mixing by diffusion is promoted, resulting in a more uniform concentration field, as shown in Fig. 6, and in addition, the efficiency in heat removal through the reactor wall is improved, resulting in a more uniform temperature field.

Fig. 5 Radial flow velocity distribution on plane 3 ($N=4, 10$).

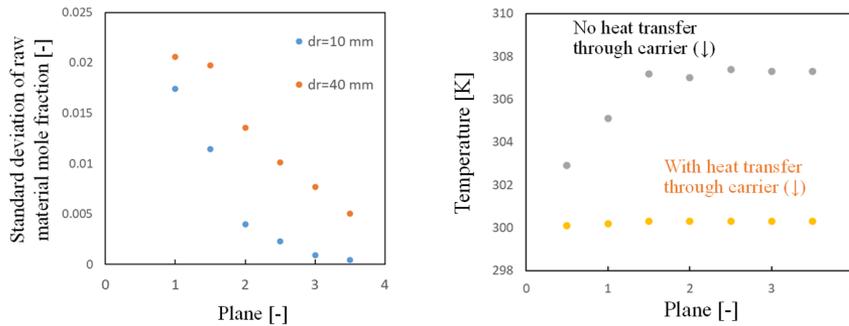


Fig. 6 Raw material dispersion on planes. Fig. 7 Average cross-sectional temperature in the reactor.

3.3. Heat transfer through catalyst carriers

The effect of heat transfer through catalyst carriers on the reaction results is discussed. CFD simulation was performed for case 7 under the condition of $\alpha = 1.05$. This value of α means that the size of each particle was increased by 5% but its position has not changed. As a result of CFD simulation, it is shown that X_r increased by 0.004 very slightly but X_s increased by 0.109. The large increase in X_s is largely due to the improvement in the ability to remove reaction heat by conduction heat transfer through the catalyst carriers, which partially overlap each other. Figure 7 shows average fluid temperatures on cross sections in the reactor with/without heat transfer through catalyst carriers. It can be seen that an isothermal reaction field was realized in the reactor with heat transfer through catalyst carriers, and as a result, X_s was improved.

3.4. Honeycomb type catalytic reactor

The result in the previous section suggests that promotion of conduction heat transfer through the catalyst carriers achieves a more uniform temperature field. However, in an actual particle-packing type fixed bed compact reactors, it may be difficult to realize the structure that connects particles by planes instead of points. Therefore, the usefulness of a honeycomb type compact reactor was examined by CFD. Figure 8 shows the honeycomb type compact reactor designed in this study. The cross section of the reactor and the cross section of each channel were assumed to be square, and the wall thickness and the channel width were set to 1.1 mm and 1.3 mm, respectively. This channel width was determined by dividing the channel volume by the catalyst surface area. Since 14 channels were arranged in both the vertical and horizontal directions, the total number of channels was 196. The CFD simulation results for packing type and honeycomb type

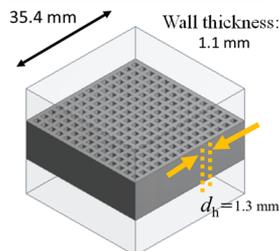


Table 4 CFD-based evaluation of honeycomb type fixed bed compact reactor.

	X_r [-]	X_s [-]	ΔP [Pa]
Packing type	0.979	0.879	101
Honeycomb type	0.989	0.901	60

Fig. 8 Honeycomb type fixed bed compact reactor.

fixed bed compact reactors are summarized in Table 4. It can be seen that the honeycomb type can achieve the same reaction result as the packing type shown in the previous section. Furthermore, it can be seen that the pressure loss of the honeycomb type was reduced by about 40% compared to the packing type.

4. Conclusions

In this study, CFD simulation was used to analyze the flow and transport phenomena of the packing type fixed bed compact reactors, which are currently used for synthesizing functional chemicals, and to evaluate the reaction results. The simulation results showed that the generation of flow perpendicular to the axis of the reactor leads to the improvement of the reaction results because the surface area of each catalyst can be effectively used. In addition, it was suggested that the honeycomb type is superior to the packing type in controlling the reaction time, realizing a uniform reaction temperature field, and reducing the pressure loss.

Acknowledgement

This work was partially supported by the Grant-in-Aid for Scientific Research (C) (No. 19K05140) and a project, Development of Continuous Production and Process Technologies of Fine Chemicals, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

- T. Tsubogo, et al., 2015, Multistep Continuous-flow Synthesis of (R)- and (S)-rolipram Using Heterogeneous Catalysts, *Nature*, 520, 7547, 329-332
- NEDO Project, https://www.nedo.go.jp/english/activities/activities_ZZJP_100152.html [accessed 1 Dec 2021]
- H. Bai, J. Theuerkauf, P.A. Gillis, P.M. Witt, 2009, A Coupled DEM and CFD Simulation of Flow Field and Pressure Drop in Fixed Bed Reactor with Randomly Packed Catalyst Particles, *Ind. Eng. Chem. Res.*, 48, 4060-4074.
- H. Bouras, Y. Haroun, R. Philippe, F. Augier, P. Fongarland, 2021, CFD Modeling of Mass Transfer in Gas-Liquid-Solid Catalytic Reactors, *Chem. Eng. Sci.*, 233, 16378.
- M. Hamidipour, J. Chen, F. Larachi, 2013, CFD Study and Experimental Validation of Trickle Bed Hydrodynamics under Gas, Liquid and Gas/Liquid Alternating Cyclic Operations, *Chem. Eng. Sci.*, 89, 158-170.
- Y. Haroun, L. Raynal, D. Legendre, 2012, Mass Transfer and Liquid Hold-up Determination in Structured Packing by CFD, *Chem. Eng. Sci.*, 75, 342-348.

Multicomponent, nonisothermal VOC adsorption modelling for pharmaceutical effluent purification: effect of operating conditions on bed performance

Vasiliki E. Tzanakopoulou,^a Alexandra Costa,^b Daniel Castro-Rodriguez,^b Dimitrios Gerogiorgis^{a,*}

^a*School of Engineering (IMP), University of Edinburgh, Edinburgh, EH9 3FB, UK*

^b*GlaxoSmithKline (GSK), Montrose, Angus, DD10 8EA, UK*

D.Gerogiorgis@ed.ac.uk

Abstract

Volatile organic compounds (VOC) use is ubiquitous in the pharma industry, posing great risks as potential organic gas releases are harmful to both environment and human health. Fixed-bed columns containing activated carbon or other adsorbents selectively remove VOCs from gas effluent streams. Nevertheless, they can be quickly and/or irregularly saturated, due to a great, simultaneous variation of flowrates and mixture compositions. This paper presents the development and implementation of a dynamic, non-isothermal adsorption model (under both adiabatic as well as non-adiabatic conditions) for the study of multicomponent alkane adsorption. A scenario-based investigation of binary mixture (heptane-decane) behaviour examines the effect of key parameter changes and reveals the preferential adsorption of decane on beaded activated carbon (BAC) under different flowrate, mixture composition and column length. Modeling of heat transfer boundary conditions (BC) combined with critical parameter variations improves our operational understanding, towards comprehensive technoeconomic VOC abatement optimisation.

Keywords: Volatile Organic Compound (VOC), dynamic simulation, adsorption.

1. Introduction

The fragile interconnected relationship between human health and physical environment is increasingly highlighted, with air pollution heavily affecting both (Dobre, 2014). Volatile organic compounds (VOC) represent a class of solvents commonly employed for industrial process use, contributing to ca. 40% of VOC emissions in the EU for 2017 (EEA, 2019). In an effort to mitigate climate impact, pharma industries turned to VOC abatement and its process optimisation. Adsorption, as less capital- and energy-intensive vs. other technological options, is established practice. Nevertheless, quick and irregular bed saturation due to feed variability impedes process efficiency and increases costs.

This work presents the development and application of a dynamic, nonisothermal, VOC adsorption model under both adiabatic and non-adiabatic conditions. Recent literature studies (Fournel et al., 2010; Tefera et al., 2013 & 2014; Knox, 2016) forego the detailed investigation of heat balance boundary conditions. Our aim here is to highlight the effect of key operating parameter changes (flowrate, composition, column length) under varying heat transfer scenarios, to establish a framework for industrial column usage optimisation.

2. Dynamic Model Development: Assumptions and Structure

The mathematical model used in the present study relies on the following assumptions:

1. The temperature difference between particles and the gas phase as well as carrier gas adsorption are considered negligible, while the ideal gas law also applies (Suzuki, 1990).
2. Equilibrium obeys the Extended Langmuir model/binary mixtures (Tefera et al., 2014).
3. Mass transfer in the solid phase is approximated by the LDF model, Eq. (Sircar, 2000).
4. Adiabatic conditions modelled by eliminating last term in Eq. (22) (Ruthven, 1984).

The model equations are derived from these sources, with Eqs. (17-21) from Knox (2016). The Bosanquet formula, Eq. (12) thus Eq. (5), is verified (Krishna & van Baten, 2012).

$$\frac{\partial C_i}{\partial t} = D_{z,i} \frac{\partial^2 C_i}{\partial z^2} - \frac{\partial(uC_i)}{\partial z} - \frac{(1-\varepsilon_b)}{\varepsilon_b} \rho_p \frac{\partial q_i}{\partial t} \quad (1) \quad k_{eff} = k_g \left(\frac{k_p}{k_g}\right)^n \quad (17)$$

$$D_{z,i} = \left(\alpha_0 + \frac{Sc_i Re_p}{2}\right) \frac{D_{AB,i}}{\varepsilon_b} \quad (2) \quad n = 0.28 - 0.757 \log_{10} \varepsilon_b - 0.057 \log_{10} \left(\frac{k_p}{k_g}\right) \quad (18)$$

$$D_{AB,i} = 10^{-3} T^{1.75} \frac{\sqrt{\left(\frac{M_A + M_B}{M_A M_B}\right)}}{P((\Sigma v)_A^{0.33} + (\Sigma v)_B^{0.33})^2} \quad (3) \quad k_{ez} = k_g \left(\frac{k_{eff}}{k_g} + 0.75 Pr Re\right) \quad (19)$$

$$\frac{\partial q_i}{\partial t} = k_{LDF} (q_{e,i} - q_i) \quad (4) \quad \frac{1}{h_{od}} = \frac{1}{dh_i} + \frac{x}{k_w d_{tm}} \quad (20)$$

$$k_{LDF} = \frac{60 \varepsilon_p C_{o,i} D_{eff,i}}{\tau_p C_{s0,i} d_p^2} \quad (5) \quad h_i = \frac{k_g}{2R} \left[2.03 Re^{0.8} \exp\left(-6 \frac{R_p}{R}\right) \right] \quad (21)$$

$$\rho_p = \frac{\rho_b}{1-\varepsilon_b} \quad (6) \quad \left(\rho_g C_{pg} + \frac{(1-\varepsilon_b)}{\varepsilon_b} \rho_p C_{pp}\right) \frac{\partial T}{\partial t} = k_{ez} \frac{\partial^2 T}{\partial z^2} - \rho_g C_{pg} \frac{\partial(uT)}{\partial z} + \frac{(1-\varepsilon_b)}{\varepsilon_b} \sum_{i=1}^n \Delta H_{ad,i} \frac{\partial q_i}{\partial t} - \frac{2h_o}{\varepsilon_b R_p} (T - T_w) \quad (22)$$

$$\varepsilon_b = 0.379 + \frac{0.078}{\left(\frac{D}{d_p}\right)^{-1.8}} \quad (7) \quad -\frac{\partial P}{\partial z} = 150 \mu u \frac{(1-\varepsilon_b)^2}{\varepsilon_b^2 d_p^2} + 1.75 \rho_g u^2 \frac{(1-\varepsilon_b)}{\varepsilon_b d_p} \quad (23)$$

$$\varepsilon_p = V_{pore} \rho_p \quad (8) \quad D_{z,i} \frac{\partial C_i(z=0,t)}{\partial z} = -u(C_{o,i} - C_i) \quad (24)$$

$$\tau_p = \frac{1}{\varepsilon_p^2} \quad (9) \quad C_i(z, t = 0) = C_{o,i} \quad (25)$$

$$C_{s0,i} = \rho_b q_{e,i} \quad (10) \quad q_i(z, t = 0) = 0 \quad (26)$$

$$D_{k,i} = 97 r_p \sqrt{\frac{T}{M_A}} \quad (11) \quad \frac{\partial C_i(z=L,t)}{\partial z} = 0 \quad (27)$$

$$\frac{1}{D_{eff,i}} = \frac{1}{D_{AB,i}} + \frac{1}{D_{k,i}} \quad (12) \quad k_{z,i} \frac{\partial T(z=0,t)}{\partial z} = -u C_{pg} \rho_g (T_g - T) \quad (28)$$

$$\frac{\partial C_t}{\partial t} = -\frac{\partial(uC_t)}{\partial z} - \frac{(1-\varepsilon_b)}{\varepsilon_b} \rho_p \sum \frac{\partial q_i}{\partial t} \quad (13) \quad \frac{\partial T(z=L,t)}{\partial z} = 0 \quad (29)$$

$$q_{e,i} = \frac{q_m b_i C_i}{1 + \sum b_i C_i} \quad (14) \quad u(0) = \frac{V_s}{\varepsilon_b} \quad (30)$$

$$b_i = b_{o,i} \exp\left(\frac{-\Delta H_{ad,i}}{RT}\right) \quad (15) \quad \frac{\partial u(L)}{\partial z} = 0 \quad (31)$$

$$-\Delta H_{ad,i} = 103.2 + 1.16 \alpha_i + 0.76 \Delta H_{vap,i} - 3.87(IP_i) - 0.7 \gamma_i - 26.19 w_{mic} \quad (16) \quad Q = V_s A \quad (32)$$

A short model description follows (i: component): in Eq. (1), C , $D_{z,i}$, u , ε_b , R and q are the gas phase VOC concentration (mol m^{-3}), the axial dispersion coefficient ($\text{m}^2 \text{s}^{-1}$), the interstitial velocity (m s^{-1}), the bulk bed porosity, the column inner radius (m) and the adsorbed phase VOC concentration (mol m^{-3}), respectively. Eq. (2) further introduces Sc_i , Re_p , $D_{AB,i}$, and α_0 as the Schmidt number of i, the Reynolds number (adsorbent particle), the molecular diffusivity ($\text{m}^2 \text{s}^{-1}$) and the empirical mass diffusion correction factor (20) respectively. In Eq. (3), Σv is the atomic diffusion volume (A: VOC, B: carrier), T is temperature (K), P is pressure (atm) and M is molecular weight (g mol^{-1}). On Eq. (4), ρ_p , $k_{LDF,i}$ are the particle density (kg m^{-3}) and LDF mass transfer coefficient (s^{-1}) respectively and $q_{e,i}$ is the inlet PT adsorbent equilibrium capacity (mol kg^{-1}). In Eq. (5), ε_p is the particle porosity, $C_{0,i}$ is the inlet concentration of i (mol m^{-3}), $D_{eff,i}$ is the effective diffusivity of i ($\text{m}^2 \text{s}^{-1}$), τ_p is particle tortuosity, $C_{s0,i}$ is the adsorbed phase concentration at equilibrium with $C_{0,i}$ (mol m^{-3}) and d_p the particle diameter (0.75 mm). Eqs. (6)-(13) present ρ_b , D , V_{pore} , $D_{k,i}$, r_p as bed density (kg m^{-3}), bed inner diameter (m), adsorbent pore volume ($5.7 \cdot 10^{-4} \text{ m}^3 \text{ kg}^{-1}$), Knudsen diffusivity ($\text{m}^2 \text{s}^{-1}$) and pore radius ($1.1 \cdot 10^{-9} \text{ m}$). Eqs. (14)-(15) introduce $q_{e,i}$, $q_{m,i}$, b_i , $b_{o,i}$ and $\Delta H_{ad,i}$ as the equilibrium and maximum adsorption capacity of i (mol kg^{-1}), the Langmuir affinity coefficient ($\text{m}^3 \text{ mol}^{-1}$), the pre-exponential constant ($\text{m}^3 \text{ mol}^{-1}$) and the heat of adsorption (J mol^{-1}), respectively. Then, Eq. (16) presents α_i , $\Delta H_{vap,i}$, IP_i , γ_i and w_{mic} as polarizability (10^{-24} cm^3), heat of vaporization (kJ mol^{-1}), ionization potential (eV), VOC surface tension (mN m^{-1}) and average micropore width (nm), respectively. Eqs. (17)-(22) introduce T_w , ρ_g , C_{pg} , C_{pp} , k_{ez} , k_{eff} , k_g , k_p , h_o , h_i , k_{ew} , k_w , x , d_{lm} as the wall temperature (K), gas density (kg m^{-3}), specific heat capacity of gas/particle ($\text{J kg}^{-1} \text{ K}^{-1}$), effective axial thermal conductivity/effective /gas/particle thermal conductivity ($\text{W m}^{-1} \text{ K}^{-1}$), overall/internal heat transfer coefficient ($\text{W m}^{-2} \text{ K}^{-1}$), effective wall/wall thermal conductivity ($\text{W m}^{-1} \text{ K}^{-1}$), column wall thickness (m), mean logarithmic column diameter, respectively. Finally, μ is gas viscosity (Pa s^{-1}), Pr is the Prandtl number, Q the volumetric flowrate ($\text{m}^3 \text{ s}^{-1}$), and A the bed area (m^2).

3. Dynamic Model Parameters for Adsorption Systems

The adsorption of a binary (heptane–decane) mixture, with air as the carrier, has been examined on BAC (Tefera et al., 2014) under both isothermal and adiabatic BC (CS0), different flowrate, composition, bed length (CS 1–3). The set of PDEs is solved using orthogonal collocation on finite elements using the gPROMS[®] Process 2.0 software suite. Viscosities are computed from Wilke’s equation and densities through pure component data via mixing rules ($T=300 \text{ K}$, $P=1 \text{ atm}$) (NIST, 2021). Air is assumed a binary mixture ($\text{N}_2:\text{O}_2=79:21 \text{ \% v/v}$). Tables 1, 2 and 3 present key parameter values for the simulations.

Table 1: Parameters for heptane-decane scenarios and axial dispersion coefficient calculation.

Case Study	ρ (kg m^{-3})	μ (Pa s^{-1})	Re	Sc	$D_{AB,i}$ ($\text{m}^2 \text{ s}^{-1}$)	R (m)
0	1.52	$2.19 \cdot 10^{-5}$	47.64	1.96	$7.35 \cdot 10^{-6}$	0.0076
				2.37	$6.08 \cdot 10^{-6}$	0.0076
1	1.52	$2.19 \cdot 10^{-5}$	23.82	1.96	$7.35 \cdot 10^{-6}$	0.0076
				2.37	$6.08 \cdot 10^{-6}$	0.0076
2	1.53	$2.20 \cdot 10^{-5}$	47.74	1.96	$7.35 \cdot 10^{-6}$	0.0076
				2.36	$6.08 \cdot 10^{-6}$	0.0076
3	1.52	$2.19 \cdot 10^{-5}$	47.64	1.96	$7.35 \cdot 10^{-6}$	0.0076
				2.37	$6.08 \cdot 10^{-6}$	0.0076

Table 2: Thermal properties parameters for heptane-decane mixture case studies.

Case Study	k_{ez} ($W m^{-2} K^{-1}$)	T_{in} (K)	h_o ($W m^{-2} K^{-1}$)	k_p ($W m^{-1} K^{-1}$)	C_{pp} ($J kg^{-1} K^{-1}$)	Pr	x (m)	ρ_b ($kg m^{-3}$)
0	0.85	300	43.57	0.17	706.7	1.1	0.001	606
1	0.46	300	25.07	0.17	706.7	1.1	0.001	606
2	0.86	300	43.65	0.17	706.7	1.1	0.001	606
3	0.85	300	43.57	0.17	706.7	1.1	0.001	606

Table 3: Main simulation parameters for heptane-decane mixture adsorption case studies.

Case Study	$D_{z,i}$ ($m^2 s^{-1}$)	T_w (K)	L (m)	x_{VOC}	V_s ($m s^{-1}$)	q_m ($mol kg^{-1}$)	ϵ	k_{LDF} (s^{-1})	b_0 ($m^3 mol^{-1}$)	Fig. 1
CS0	0.00129	295	0.0650	0.000250	0.914	3.69	0.38	$6.88 \cdot 10^{-5}$	$5.02 \cdot 10^{-13}$	(a)-(d)
	0.00122			0.000250		3.16		$2.26 \cdot 10^{-5}$	$8.65 \cdot 10^{-14}$	
CS1	0.00084	295	0.0650	0.000250	0.457	3.69	0.38	$6.88 \cdot 10^{-5}$	$5.02 \cdot 10^{-13}$	(e)-(f)
	0.00077			0.000250		3.16		$2.26 \cdot 10^{-5}$	$8.65 \cdot 10^{-14}$	
CS2	0.00129	295	0.0650	0.000125	0.914	3.69	0.38	$8.44 \cdot 10^{-5}$	$5.02 \cdot 10^{-13}$	(g)
	0.00122			0.000375		3.16		$2.75 \cdot 10^{-5}$	$8.65 \cdot 10^{-14}$	
CS3	0.00129	295	0.0975	0.000250	0.914	3.69	0.38	$6.88 \cdot 10^{-5}$	$5.02 \cdot 10^{-13}$	(h)
	0.00122			0.000250		3.16		$2.26 \cdot 10^{-5}$	$8.65 \cdot 10^{-14}$	

4. Results and Discussion

Breakthrough curves, temperature and pressure variations of the examined systems are shown in Fig. 1. Our model is successfully validated, Fig. 1 (a), vs. published experimental breakthrough curves which dictated our inputs (CS0) in Fig. 1 (a)-(d) (Tefera et al., 2014). Fig. 1 (a) shows a breakthrough duration predicted with minimal error for decane, but a mismatch for heptane, particularly during the transient high-C phase of the latter. Discrepancies could be attributed to the limitations of our 1-D nonisothermal model, but we remark this mismatch is very similar to that in the published plot (Tefera et al., 2014).

We consider three case studies (CS 1-3) of different flowrate (CS1), composition (CS2) and bed length (CS3), depicted in Fig. 1 (e)-(h). Adiabatic BCs of the base case, Fig. 1 (d), clearly induce a heat buildup within the bed, compared to the isothermal BCs of Fig. 1 (c). A linear ΔP profile is seen in Fig. 1(b), consistent with reported trends (Tefera et al., 2013).

Fig. (e) presents the heptane-decane binary mixture behaviour at a flowrate 50% smaller than the base case (CS1). The order of affinity strength remains (decane > heptane), but the onset of breakthrough occurs later and the duration is longer compared to the base case. The lower flowrate is directly associated with a 56% pressure drop reduction presented in Fig. 1 (f), i.e. roughly half the pressure drop observed in Fig. 1 (b) for CS0.

Fig. 1(g) illustrates the effect of mixture composition variation: $C_{decane} = 3 \cdot C_{heptane}$ (CS2). Though gradual displacement of heptane persists, the plot indicates that the concentration change induces a earlier breakthrough onset for decane as well as a later one for heptane, combined with a clearly shorter duration for both VOCs compared to the base case (CS0).

Finally, Fig. 1(h) presents the effect of key design variable changes, considering a design variation of 50% increased bed length (CS3), and its effect on breakthrough behaviour. Once again, decane is adsorbed much later with the same trend of heptane displacement, but the onset of breakthrough for both occurs later in time, due to the increased bed size.

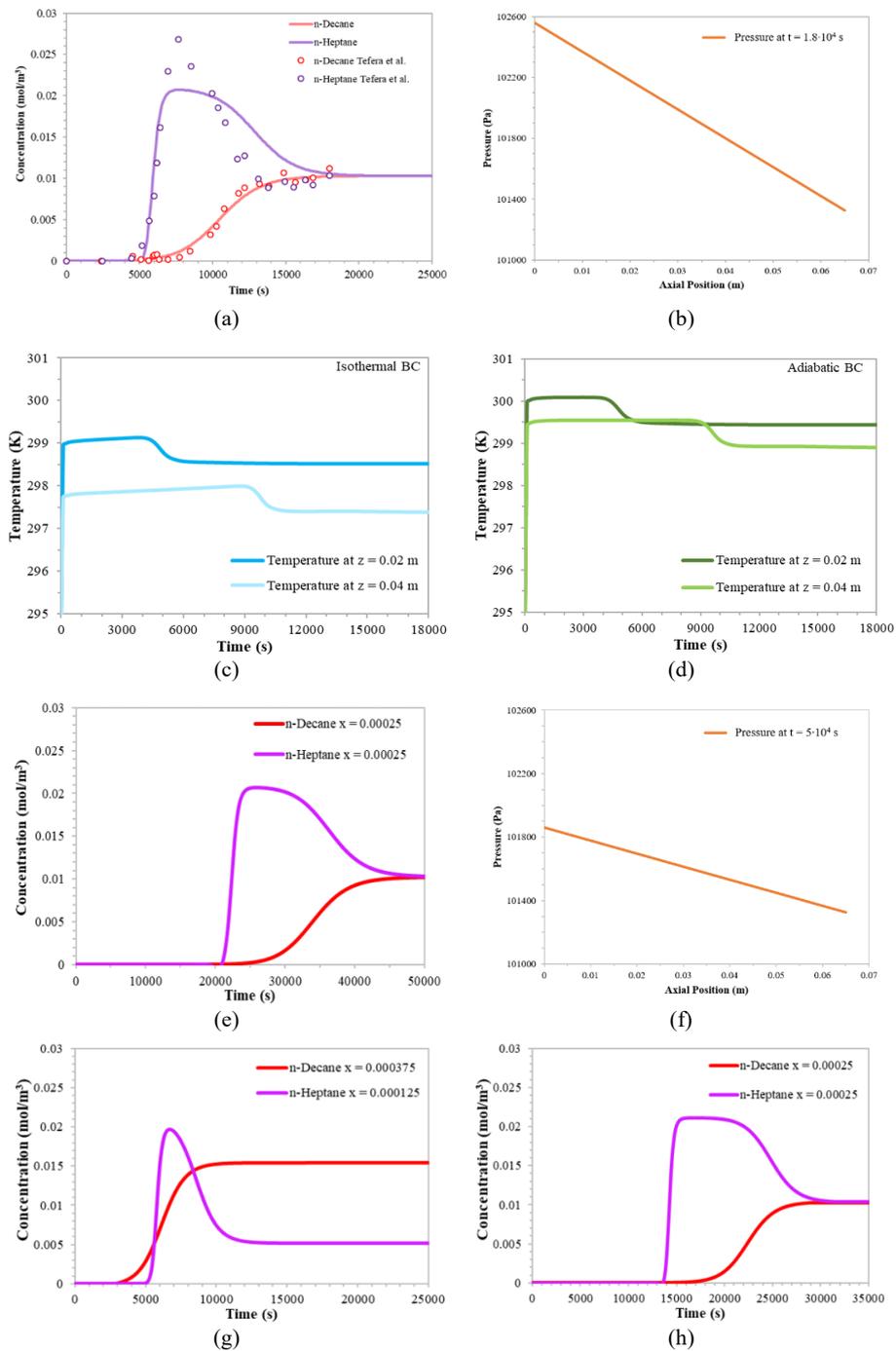


Figure 1: Model validation (a) (Tefera et al., 2014); operat. profiles (b)-(d); case studies (e)-(h).

5. Conclusions

Active Pharmaceutical Ingredient (API) manufacturing relies on the use of vast amounts of solvents, thus increasing risks on environment and public health due to VOC emissions. Cognizant of the growing environmental concern, pharma companies are shifting their focus towards drastic climate impact minimisation through process optimisation in an effort to mitigate air pollution caused by VOC emissions. Adsorption is often adopted as an industrial VOC abatement method, yet not without operational efficiency challenges. Although adsorption modelling is extensively studied in the literature, comprehensive studies examining the effect of operational decisions on VOC removal efficiency towards technoeconomic process optimisation still remain elusive (Tzanakopoulou et al., 2021).

This paper demonstrates the development and application of a one-dimensional in space, multicomponent, nonisothermal adsorption model to highlight the complex phenomena taking place inside pharma VOC adsorption columns, considering axial dispersion in the gas phase and the Linear Driving Force model for solid phase mass transport. Differences between isothermal and adiabatic BC are critical considering the exothermic nature of the process. Our simulations revealed quicker uptake of heptane under varying flowrate, composition and bed length, with decane is adsorbed later in all four cases we considered. A decrease in the flowrate resulted in a later onset, a larger duration of breakthrough and a reduction of pressure drop compared to the base case, thus implying inefficient practice. In view of industrial applications, a larger (thus costlier) bed length investigation resulted in a later onset and longer breakthrough duration. Reliable modelling of multicomponent, nonisothermal VOC adsorption paves the way for technoeconomic optimisation efforts.

Acknowledgements

The authors gratefully acknowledge the SRPe-NMIS IDP PhD Scholarship awarded to V.T., as well as a Royal Society Short Industrial Fellowship (2020-22) awarded to D.I.G.

References

- EEA, 2019, Emissions of main air pollutants by sector in EEA, www.eea.europa.eu/data-and-maps.
- T. Dobre et al., 2014, VOC removal from gas streams by adsorption onto activated carbon, *Ind. Eng. Chem. Res.*, 53, 9, 3622–3628.
- L. Fournel et al., 2010, Modeling breakthrough curves of VOC on activated carbon fibers, *Adsorption*, 16, 3, 147–153.
- J.C. Knox et al., 2016, Limitations of breakthrough curve analysis in fixed-bed adsorption, *Ind. Eng. Chem. Res.*, 55, 16, 4734–4748.
- R. Krishna and J.M. van Baten, 2012, Investigating the validity of the Bosanquet formula for estimation of diffusivities in mesopores, *Chem. Eng. Sci.*, 69, 1, 684–688.
- National Institute of Standards & Technology, *NIST Chemistry WebBook*, NIST Std. Ref. Database No 69, Gaithersburg, MD, USA.
- D.M. Ruthven, 1984, Principles of adsorption and adsorption processes, D. M. Ruthven, Wiley.
- S. Sircar, 2000, Why does the LDF model for adsorption kinetics work, *Adsorption*, 6, 2, 137–147.
- M. Suzuki, 1990, *Adsorption Engineering*, Elsevier, Amsterdam.
- D.T. Tefera et al., 2013, Two-dimensional modeling of VOC adsorption on beaded activated carbon (BAC), *Env. Sci. Tech.*, 47, 20, 11700–11710.
- D.T. Tefera et al., 2014, Modeling competitive adsorption of mixtures of VOCs in a fixed-bed of beaded activated carbon (BAC), *Env. Sci. Tech.*, 48, 9: 5108–5117.
- V.E. Tzanakopoulou et al., 2021, Process modelling and simulation of VOC recovery from pharmaceutical gas emission streams, *Comput.-Aided Chem. Eng.*, 50, 717–723.

Hydrogen Separation via Continuous Hydrate Formation

Marcelino Artur L. Fernandes^{a,b}, Mariana G. Domingos^c, Fernando G. Martins^{a,b*}, Isabel S. Fernandes^{b,d}, Marcelo F. Costa^c, Ricardo J. Santos^{b,d}, José Carlos B. Lopes^c

^a*LEPABE, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465, Porto, Portugal*

^b*ALiCE - Associate Laboratory in Chemical Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

^c*CoLAB NET4CO2, Rua Júlio de Matos, 828-882, 4200-355 Porto, Portugal*

^d*LSRE-LCM, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465, Porto, Portugal*

**Corresponding author: fgm@fe.up.pt*

Abstract

According to the International Energy Agency, hydrogen is one of the pillars to reach Net Zero CO₂ emissions by 2050. Forecasts indicate a market size of 200 Mton in 2030, of which 35% will be blue hydrogen (IEA, 2021). This work approaches the separation of hydrogen from carbon dioxide, by integrating the Python software environment with Aspen Plus software. Compression and cooling of the feed components to the hydrate formation conditions upstream the NetMIX is modelled in Aspen Plus. In the Python interface, the heat and mass balances associated with the NetMIX are performed, including the Gibbs Energy Minimization algorithm, developed by Ballard (Ballard, 2002) for the prediction of hydrate/liquid/vapor equilibrium. The model developed was applied to a case study, which considers a gas stream of 3300 m³_{PTN}·h⁻¹, containing 60 % of hydrogen and 40 % of carbon dioxide (molar percentages). The results show an improvement in the purity of H₂ in the gaseous stream of nearly 25 %, reaching a value of 84 %, while also a recovery near 62% of CO₂ via the hydrate phase is obtained, with 30 MJ_e · m³_{PTN}⁻¹ of electrical demand and 210 MJ_t · m³_{PTN}⁻¹ of cooling demand.

Key Words: **Hydrogen separation; NetMIX technology; Hydrates.**

1. Contextualization

1.1. CO₂ concerns

In 2015 in Paris, during the United Nations Framework Convention on Climate Change, the Paris Agreement was signed, where 195 nations agreed that changes were required to combat climate change and mitigate its impacts. In 2021, at the COP26 UN Climate Conference, 190 countries reinforced the targets set by the Paris Agreement to limit global warming to below 2 degrees Celsius, and as close as possible to 1.5 degrees Celsius. This objective can be achieved through the reduction of the emissions of greenhouse gases, such as carbon dioxide (CO₂) and methane (CH₄). The CO₂ concentration has increased around 40 times since preindustrial times, mainly due to its emissions by the burning process of fossil fuels. A solution to reduce these emissions urges and one possibility consists of finding fuel alternatives to fossil fuels, such as hydrogen.

1.2. H_2 as a source of energy

Hydrogen has an energy density of $120 \text{ MJ}\cdot\text{kg}^{-1}$, which is more than two times larger than methane. In the context of reaching net zero emissions by 2050, hydrogen needs to be produced from electrolysis with renewable electricity, i.e., green hydrogen, or steam methane reforming with CO_2 capture, i.e., blue hydrogen. Hydrogen can then be used as an energy source, either by producing energy in turbines or by its use in fuel cells, after going through the required purification steps.

1.3. Hydrates

Hydrates are non-stoichiometric crystalline structures, composed of hosts, typically water molecules, that are connected to each other by hydrogen bonds. These hosts create cavities, where guests, small gas molecules, are constrained. For hydrates to form, low temperatures and high pressures are required and the amount of gas in contact with the liquid must surpass the solubility limit of the gas on water. The hydrate formation conditions are specific to the gas molecule being hosted, thus providing a way for separating different gases in a gas mixture, such as flue gas (CO_2 and N_2) or mixtures for a reformer outlet (CO_2 and H_2), while allowing the storage of gases in solid form.

For hydrates to keep their stability, the necessary temperature and pressure conditions must be maintained. This is very challenging for a continuous production scenario since the hydrate formation process is highly exothermic.

A solution proposed to this challenge is the use of NetMIX technology, which consists of a network of mixing chambers, interconnected by transport channels, which enable better control of the mixing process, and result in outstanding heat transfer capabilities (Costa et. al., 2015), providing an efficient heat removal that keeps hydrates stable.

The aim of this work is to develop a framework simulation for modelling the formation of hydrates in NetMIX device (Lopes et al., 2019), as a method to separate hydrogen from a mixture with carbon dioxide, to investigate the potential of this solution.

2. Materials and Methods

2.1. Process modelling

The software used for creating the simulation model was Aspen Plus. The designed flowsheet is shown in Figure 1.

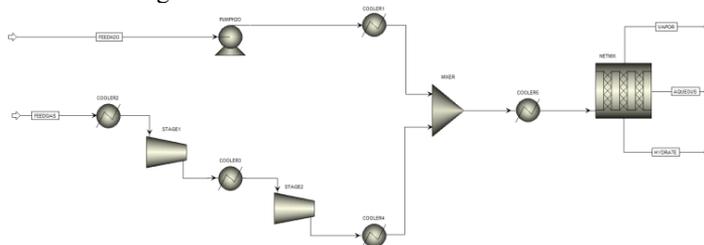


Figure 1 - Designed flowsheet for the hydrogen separation via hydrate.

The flowsheet consists of two feed streams: one for water, the other for the gas mixture. The water feed is sent to a pump, to increase the pressure to 125 bar, and through a cooler, to decrease the temperature to $0.5 \text{ }^\circ\text{C}$, the conditions for hydrate formation. The gas feed first goes through a cooler, to decrease the temperature, due to the high temperatures of the reforming stage. Then it goes through a two-stage compressor, with a cooler in between, to control the temperature increase due to the compression. The gas stream, at the operating pressure, then goes through a cooler, to decrease the temperature. The streams are then mixed, and to prevent temperature increase due to the mixing process, a

cooler is added after the mixer. This stream is then fed to the NetMIX device, where the separation occurs.

Aspen Plus does not contain a model of the NetMIX device, nor can it predict the formation of hydrates. For these purposes, a Python code was implemented. This code, named *HydrateFlash.py* (Darnell et. al., 2017) is the Python translation of another software application, named CSMGem. This software was developed by Adam Ballard (Ballard, 2002) and predicts the formation of hydrates, based on the Minimization of the Gibbs Energy. It contains a set of components in its database, but not hydrogen. Thus, the properties of hydrogen were added to the *HydrateFlash.py* code database.

2.2. The *HydrateFlash.py* simulator

The *HydrateFlash.py* simulator minimizes the Gibbs Energy, coupling with the thermodynamic equilibrium conditions, where temperature, pressure, and fugacity are equal for all phases. The objective function uses the concept of K-values. These values represent the repartition coefficient between two phases, one of those being the reference phase.

Assuming that the equilibrium of a mixture of carbon dioxide, hydrogen, and water, under certain conditions, can be restrained to a three-phase distribution, those being vapor, liquid and hydrate phases, two sets of K-values must be given. Assuming that the vapor phase is the reference one, then the simplified Raoult's law can be used to express the liquid/vapor equilibrium, while a thirteen-parameter equation, described in Ballard's work is used to express the equilibrium between the vapor and the hydrate phase.

3. Results and Discussion

3.1 Validation of *HydrateFlash.py* results

The results of *HydrateFlash.py* for predicting the thermodynamic equilibrium between the vapor-liquid-hydrate phases were validated with the results obtained with CSMGem for different scenarios for vapor-liquid-hydrate equilibrium. When possible, these results were checked with results published in the literature.

Figure 2 shows the results obtained by CSMGem and *HydrateFlash.py* for a mixture of CH₄/CO₂, since these components are present in both softwares. The objective was to compare the equilibrium line, i.e., the conditions at which hydrate is formed, for a temperature, in the range of 0.5 to 8.5 °C.

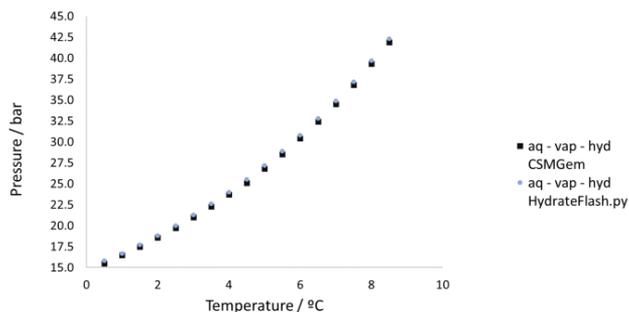


Figure 2 - Equilibrium lines for the H₂O, CH₄ and CO₂ mixture

As seen from Figure 2, the results present slight deviations of about 0.3 bar, on average. This behavior can be justified by the fact that CSMGem, at a certain range of pressures (very near the transition zone of vapor-liquid equilibrium to vapor-liquid-hydrate equilibrium), does not produce results, due to convergence problems. However,

HydrateFlash.py presented no convergence problems in these zones. Thus, the *HydrateFlash.py* code can be considered as a valid approach to simulate the equilibrium of formation of hydrates.

3.2 Vapor-Liquid equilibrium

The values obtained with *HydrateFlash.py* code for vapor-liquid equilibrium between hydrogen and carbon dioxide were validated against Aspen Plus predictions for the PSRK thermodynamic model with H₂ as a Henry component, since H₂ is at the supercritical region (Skogestad, 2009), and has a solubility in water less than 0.005 (mol:mol) at the target conditions (Rahbari et. al, 2019). The phase fraction, as well as the phase composition are shown in Figure 3 (a) and (b), respectively.

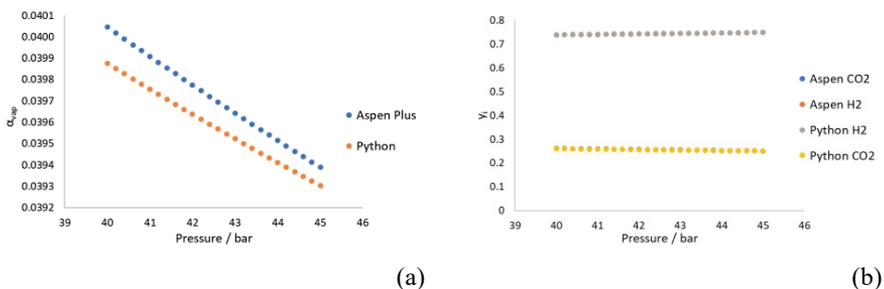


Figure 3 – (a) Vapor fraction evolution at 0.5 °C and, (b) Molar fraction on the vapor phase for CO₂ and H₂ at 0.5 °C.

The results are matching with a deviation of 0.36 % for the H₂ vapor phase composition, while the average error for the phase fraction was 0.32 %.

3.3 Validation of Vapor-Liquid-Hydrate equilibrium

The results presented in the work of Kumar et. al (2006) were compared with the equilibrium pressure of a mixture of water/hydrogen/carbon dioxide, predicted by the *HydrateFlash.py* software, for different H₂ fractions. As an example, Figure 4 shows the results concerning pressure vs. temperature for an H₂ fraction of 0.608 and 0.579.

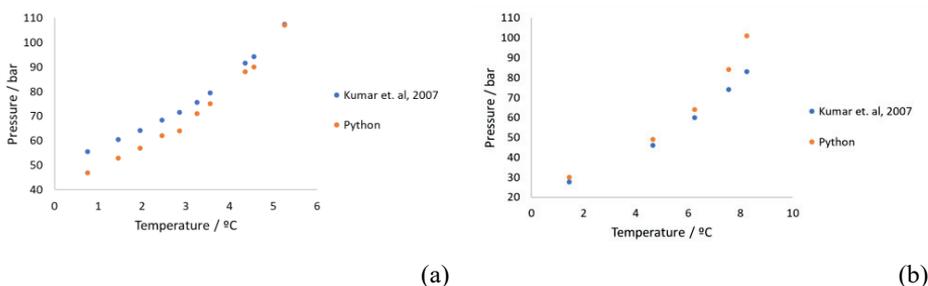


Figure 4 – Equilibrium pressure variation for $y_{H_2} = 0.608$ (a) and $y_{H_2} = 0.421$ (b).

The highest relative deviation between the published work and the model prediction was 11.3 %, obtained for the H₂ molar fraction of 0.579 (Figure 4b), being so these results were considered acceptable.

3.4 Influence of Pressure and Temperature for H₂/CO₂ hydrates formation

Figure 5 (a) shows the pressure influence on the hydrate fraction, at 0.5 °C, while Figure 5 (b) shows the temperature influence on the equilibrium pressure. For the pressure, up to the equilibrium pressure of 51.6 bar, obtained via the equilibrium model of the *HydrateFlash.py* code, no hydrates are formed. After reaching this value, an increase of pressure leads to an increase in the hydrate phase fraction. For the temperature, it was noted that the increase in temperature increases the equilibrium pressure.

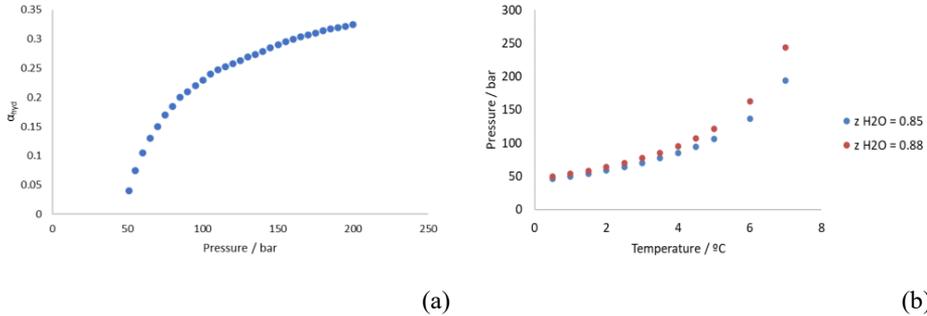


Figure 5 – (a) Pressure influence on the hydrate fraction and, (b) temperature influence on the equilibrium pressure.

3.5 Case study

The flowrate assumed for this simulation was 3300 m³_{PTN}·h⁻¹, which was composed of 60 % carbon dioxide, and 40 % hydrogen (molar percentages). For this flowrate, it would be required to feed 1000 ton·h⁻¹ of water to the reactor.

The conditions of the gas stream were a temperature of 200 °C, and a pressure of 30 bar (Iulianelli et al., 2015). For the water stream, it was considered to be at ambient conditions, those being 20 °C, and at 1 bar.

The process was then set so that the operating conditions are 0.5 °C and 125 bar, being these conditions defined to obtain a hydrate phase fraction larger than 30%. For the gas side, it was required to specify the conditions of the multi-stage compressor. For that, a compression ratio of 2.04 was used, and the outlet cooler temperature was 35 °C.

The results of the simulation indicate a gaseous stream with a molar H₂ fraction of 84 %, while the remaining 16 % is CO₂. The hydrate phase is mainly composed of water, representing 87 %, while 12.5 % is CO₂. The phase distribution obtained was 50 % aqueous phase, 37 % hydrate phase and 13 % vapor phase.

The results show an increase of nearly 25 % in the purity of the gas stream, regarding hydrogen. Additionally, 62 % of the carbon dioxide at the NetMIX inlet is recovered from the hydrate outlet stream. A separation factor, defined in equation (1), of 90 was obtained. This means that CO₂ tends to be kept in the hydrate phase more than H₂, while this last tends to remain in the vapor phase.

$$SF = \frac{n_{CO_2,hyd}/n_{CO_2,vap}}{n_{H_2,hyd}/n_{H_2,vap}} \quad (1)$$

Additionally, the energy consumption was calculated. A value of 30 MJ_e · m³_{PTN}⁻¹ was obtained, considering the pump and the compression system, while it would be required 210 MJ_t · m³_{PTN}⁻¹ for cooling, considering heat removal on the heat exchangers, the

intercooler, and the NetMIX device. Reducing the energy requirements of the process may be achieved through heat integration and the use of promoters which will be covered in a future work.

Conclusions

This work approached the production of blue hydrogen, via continuous hydrate production. The results obtained show that the gas stream enriches in H₂, reaching a purity of 84 % (mol:mol), while the hydrate produced is mainly composed of CO₂. A separation factor of 90 was achieved between the vapor and hydrate phases, translating to a preference of H₂ to be kept in the vapor phase, while CO₂ tends to go to the hydrate phase. Due to the early stage in knowledge regarding this mixture behavior under hydrate formation conditions, these results can be further improved under different circumstances, such as different operating conditions, which will be approached in a future work.

Acknowledgments

This work was financially supported by: Base Funding - LA/P/0045/2020 of the Associate Laboratory in Chemical Engineering – AliCE; Base Funding - UIDB/00511/2020 of the Laboratory for Process Engineering, Environment, Biotechnology and Energy – LEPABE, and Base-UIDB/50020/2020 and Programmatic-UIDP/50020/2020 Funding of LSRE-LCM, Project POCI-01-0145-FEDER-030445 – funded by FEDER funds through COMPETE2020 – Programa Operacional Competitividade e Internacionalização (POCI), and funded by national funds through the FCT/MCTES (PIDDAC). Also, funding from project POCI-01-0247-FEDER-069953, funded by the European Regional Development Fund - FEDER, through the Competitiveness and Internationalization Operational Program (POCI). Isabel Fernandes acknowledges the FCT scholarship SFRH/BD/143958/ 2019. José Carlos B. Lopes, Mariana G. Domingos, Marcelo F. Costa acknowledge funding from NORTE-06-3559-FSE-000069 – Highly Qualified Human Resources for CoLAB NET4CO₂ – Network for a Sustainable CO₂ Economy – funded by the European Social Fund (FSE)

References

- Ballard, A.L. (2002) “A non-ideal hydrate solid solution model for a multi-phase equilibria program.” Ph.D. Thesis, Colorado School of Mines
- Costa, M. F., Fonte, C. M., Dias, M. M., Lopes, J. C. B. (2017). Heat transfer performance of NETmix—A novel micro-meso structured mixer and reactor
- Darnell, K. N., Flemings, P. B., and DiCarlo, D. (2017), Subsurface injection of combustion power plant effluent as a solid-phase carbon dioxide storage strategy, *Geophys. Res. Lett.*, 44, 5521–5530. Github. Hydrateflash. <https://github.com/kdarnell/hydrateflash>.
- IEA (2021), Net Zero by 2050, IEA, Paris <https://www.iea.org/reports/net-zero-by-2050>
- Iulianelli, A., Pirola, C., Comazzi, A., Galli, F., Manenti, F., & Basile, A. (2015). Water gas shift membrane reactors. *Membrane Reactors for Energy Applications and Basic Chemical Production*, 3–29.
- Kumar, R., Wu, H. jie, & Englezos, P. (2006). Incipient hydrate phase equilibrium for gas mixtures containing hydrogen, carbon dioxide and propane. *Fluid Phase Equilibria*, 244(2), 167–171.
- Lopes, J.C.B., Dias, M.M.G.Q., Costa, M.F.S., Teixeira, C.A.M. & Santos, R.J.N. (2019) “Continuous production of clathrate hydrates from aqueous and hydrate-forming streams, methods and uses thereof.” EU Patent 19220274.5. Portuguese Patent 116038.
- Skogestad, S. (2009). *Chemical and energy process engineering*. Boca Raton, FL: CRC Press
- Rahbari, A., Brenkman, J., Hens, R., Ramdin, M., van den Broeke, L. J. P., Schoon, R., Henkes, R., Moutos, O. A., & Vlugt, T. J. H. (2019). Solubility of water in hydrogen at high pressures: A molecular simulation study. *Journal of Chemical and Engineering Data*, 64(9), 4103–4115. <https://doi.org/10.1021/acs.jced.9b00513>

Combined particle model and experimental approach for predicting pyrolysis with palm kernel shells

Andres Chico-Proano^{a,b}, Michelle Romero^c, Ricardo A. Narváez C.^{c,d}, Boris G. German^c, Daniel Rivadeneira^c, George Manos^e, Lazaros G. Papageorgiou^a and Eric S. Fraga^{a,*}

^a*Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London (UCL), Torrington Place, London WC1E 7JE, UK*

^b*Departamento de Ingeniería Química, Escuela Politécnica Nacional, Quito EC170525, Ecuador*

^c*Instituto de Investigación Geológico y Energético (IIGE), Quito, Ecuador*

^d*Universidad Central del Ecuador, UCE-GIIP, EC170521, Quito, Ecuador*

^e*Department of Chemical Engineering, University College London (UCL), Torrington Place, London WC1E 7JE, UK*

e.fraga@ucl.ac.uk

Abstract

This work combines a particle model and experimental data as an alternative to improve the agreement between predicted and experimental results in pyrolysis. A mechanistic model, implemented and solved in gPROMS, is used to describe biomass pyrolysis in a shrinking particle. The content of cellulose, hemicellulose and lignin in palm kernel shells (PKS) is determined with high performance liquid chromatography (HPLC) and UV-Visible spectrophotometry. A synthetic sample of cellulose, hemicellulose and lignin is prepared to resemble PKS composition. Pyrolysis of both PKS and the synthetic sample is evaluated using thermogravimetric analysis (TGA). The results highlight the complexity of pyrolysis and the importance of supporting modelling with experimental data. TGA results show the effect of biomass components interactions in pyrolysis. A 20% difference in pyrolysis mass conversion was found between PKS and the mixture. The agreement between predicted and experimental results improves when the model is modified considering PKS pyrolysis information. This work represents a step forward for understanding the role of combined modelling and experimental strategies in describing complex processes like pyrolysis.

Keywords: biomass pyrolysis, mechanistic model, TGA, palm kernel shells

1. Introduction

The production of fuels and platform chemicals from renewable sources can contribute to generate a shift from our current oil-driven economy towards more environmental-friendly alternatives. Biomass is an abundant resource that could facilitate such a shift. Pyrolysis is a key technology for transforming biomass into both fuels and platform chemicals, which could support the transformation of agro-industries into biorefineries. However, the integration of pyrolysis into biorefineries requires a better understanding of the complex interactions between biomass components. This is particularly critical for the development of reliable models for pyrolysis process design (Bridgwater, 2018). Previous pyrolysis models consider biomass as a cellulose, hemicellulose and lignin mixture (Ranzi et al., 2017; Di Blasi, 1994). However, biomass components interactions are not fully understood and they are difficult to quantify (Wu et al., 2016); thus leading to poor agreement between predicted and experimental results. There are previous experimental works that

have identified interactions between biomass components (Hosoya et al., 2007; Wu et al., 2016). However, this is not the case for modelling approaches, where there is little information regarding the effect of such interactions over pyrolysis modelling outcomes. Moreover, information is still scarce regarding the effect of such interactions in residues with a complex morphology like PKS.

2. Biomass pyrolysis and modelling strategies

Lignocellulosic biomass is an abundant resource that contains cellulose, hemicellulose, lignin, water, inorganic components and extractives. Cellulose, hemicellulose and lignin represent more than 90 wt.% of biomass constituents. The thermochemical decomposition of biomass in pyrolysis involves a complex system of primary and secondary chemical reactions. Hence, the determination of biomass pyrolysis reaction mechanisms and their associated kinetic parameters is a difficult task. This is particularly challenging for lignin because of the different composition and structures that lignin could present (Debiagi et al., 2015). Previous models have addressed pyrolysis complexity by considering biomass as a cellulose, hemicellulose and lignin only mixture (Anca-Couce and Zobel, 2012; Di Blasi, 1994). Other modelling approaches have included extractives by using additional lumped pyrolysis reactions (Debiagi et al., 2015; Ranzi et al., 2017). Inorganic components are not usually included in pyrolysis models; however, they could catalyze secondary reactions (Ranzi et al., 2017). Existing interactions between cellulose, hemicellulose and lignin have been identified in pyrolysis (Wu et al., 2016; Hosoya et al., 2007); however the effect of such interactions over pyrolysis kinetics is not fully understood. Further work is needed to incorporate such interactions into biomass pyrolysis models, for improving the agreement between predicted and experimental results (Wu et al., 2016). The combination of modelling and experimental approaches in complex thermochemical processes like pyrolysis, stands out as an opportunity to develop reliable models for future process design and optimization. This work aims to combine modelling and experimental strategies as an alternative for improving predictions in pyrolysis.

3. Experimental approach and particle model for biomass pyrolysis

3.1. Materials and experimental methods

This work considers PKS and a synthetic mixture of cellulose, hemicellulose and lignin, as materials. PKS from the *Elaeis guineensis* Jacq. variety were collected from a palm oil extracting company in Ecuador. The synthetic mixture combined cellulose (Alpha Chemika), low sulfonate lignin (Sigma-Aldrich) and D-xylose (LobaChemie) to resemble PKS composition. The characterization of the PKS includes moisture and ash content analysis, and cellulose, hemicellulose, lignin and extractives determination. Moisture and ash contents are determined by gravimetry in accordance with BS EN 18134-1:2015 and BS EN 18122:2015 standards, correspondingly. Extractives were separated from PKS with solid-liquid extraction using ethanol USP grade and water. The content of cellulose, hemicellulose and lignin is determined using high performance liquid chromatography (HPLC) and UV-Visible spectrophotometry. For this purpose, samples are prepared and hydrolyzed in accordance with the NREL/TP-510-42618 laboratory analytical procedure (Sluiter et al., 2004). Cellulose and hemicellulose contents are determined using an Agilent 1260 Infinity HPLC equipment with a refraction index detector. The acid soluble lignin is determined with a Hach DR6000 UV-Visible spectrophotometer and the acid insoluble lignin is determined by gravimetry. PKS and the synthetic mixture's pyrolysis is studied with a Shimadzu TGA-5 thermogravimetric analyzer. TGA experiments are performed by duplicate using a 50 mL min⁻¹ of nitrogen, a maximum temperature of 800 °C and a 50° C min⁻¹ heating rate.

3.2. Particle model for biomass pyrolysis

A particle model for describing PKS pyrolysis is developed using mass, energy and momentum conservation. Similar approaches exist to describe cellulose pyrolysis (Di Blasi, 1994; Chico-

Proano et al., 2021); however, they do not analyze differences between pyrolysis of ideal and real samples. For addressing this, a shrinking particle is modelled considering isometric behaviour and including only radial variation of properties. Only gases move across the particle and they follow an ideal behaviour. Moreover, gas/solid local thermal equilibrium is considered (Ranzi et al., 2017). The model includes a combination of lumped kinetic models for cellulose, hemicellulose and lignin pyrolysis from literature (Di Blasi, 1994; Miller and Bellan, 1997). Products are grouped under three categories, solids (char), condensables (tar) and gases as shown in Figure 1.

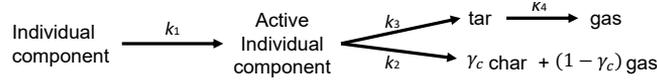


Figure 1: Scheme of the kinetic model considered for pyrolysis.

The changes in the particle's porosity, ϕ , are calculated using bulk, $\bar{\rho}$, and initial, ρ , densities. The mass balance for the solid phase, S , and for each individual solid component, S_i , includes net formation rates for the solid phase Γ_S and for individual species $\Gamma_{S,i}$ as shown in Eqs. (1) to (6).

$$\partial_t [\rho_S (1 - \phi)] = \Gamma_S \quad (1)$$

$$\partial_t [\rho_{S,i} (1 - \phi)] = \Gamma_{S,i} \quad (2)$$

$$\phi = \frac{(\bar{\rho}_{\text{char}} + \bar{\rho}_{\text{biomass}})^2}{(\rho_{\text{char}} \bar{\rho}_{\text{char}} + \rho_{\text{biomass}} \bar{\rho}_{\text{biomass}})} \quad (3)$$

$$\Gamma_S = \sum_{i=1}^{n_c} \Gamma_{S,i} \quad (4)$$

$$\Gamma_{S,i} = \sum_{j=1}^{n_r} \gamma_{i,j} r_j \quad (5)$$

$$r_j = A_j e^{-\left(\frac{E_{a,j}}{R_{\text{ideal}} T}\right)} C_{i,j} \quad (6)$$

In the previous equations, reaction rates, r_j , are determined using the stoichiometric coefficients, γ_i , of each i component in a j reaction. The Arrhenius equation is applied, using the concentration of a specie, $C_{i,j}$, the pre-exponential factors A_j and the activation energy $E_{a,j}$ from literature (Miller and Bellan, 1997). For the gas phase, G , and its individual components, G_i , the mass balance is described in terms of the velocity of the gaseous phase, \mathbf{u} , and its density, ρ_G , the net formation of the gaseous phase, Γ_G , and of individual components, $\Gamma_{G,i}$, and mass fractions $w_{G,i}$ as follows:

$$\partial_t (\rho_G \phi) = -\nabla \cdot (\rho_G \phi \mathbf{u}) + \Gamma_G \quad (7)$$

$$\partial_t (\rho_G \phi w_{G,i}) = -\nabla \cdot (\rho_{G,i} \phi w_{G,i} \mathbf{u}) - \nabla \cdot \mathbf{j}_{G,i} + \Gamma_{G,i} \quad (8)$$

The Darcy equation, Eq. (10), describes the movement of the gaseous phase, and the individual gas flux through the porous particle is calculated with the diffusive term $\nabla \cdot \mathbf{j}_{G,i}$, the effective diffusivity, $D_{\text{effective}}$, the gas phase viscosity, μ_G , the permeability of the solid phase, K_S and the pressure P .

$$\mathbf{j}_{G,i} = -D_{\text{effective}} \nabla \rho_G \phi w_{G,i} \quad (9)$$

$$\mathbf{u} = -\frac{K_S}{\phi \mu_G} \nabla P \quad (10)$$

The energy balance in Eq. (11) includes averaged specific heats for the solid, \hat{C}_{pS} , and gaseous phases, \hat{C}_{pG} . Similarly, the \mathbf{q} term includes the heat of reaction and the thermal conductivities.

$$(\rho_S \phi \hat{C}_{pS} + \rho_G (1 - \phi) \hat{C}_{pG}) \frac{DT}{Dt} = -\nabla \cdot \mathbf{q} - \left(\frac{\partial \ln \rho_G}{\ln T} \right)_P \frac{DP}{Dt} \quad (11)$$

Table 1: Properties of substances and initial conditions used in the model.

Property	Value	Units	Reference
r_0	$138 \cdot 10^{-6}$	m	Experiments
$\rho_{cellulose}$	1656.00	kg m^{-3}	(Gorensek et al., 2019)
$\rho_{hemicellulose}$	1865.42	kg m^{-3}	(Gorensek et al., 2019)
ρ_{lignin}	1600.00	kg m^{-3}	(Gorensek et al., 2019)
ρ_{char}	1540.00	kg m^{-3}	(Anca-Couce and Zobel, 2012)
μ_G	$3.0 \cdot 10^{-5}$	$\text{kg m}^{-1} \text{s}^{-1}$	(Di Blasi, 1994)
P_0	101325	Pa	-
T_0	373	K	-

Table 2: Palm kernel shell characterization.

Property	Value
Moisture content (wt.%)	10.62 ± 0.07
Extractives (wt.%)	5.78 ± 0.18
Cellulose (wt.%)	14.25 ± 1.89
Hemicellulose (wt.%)	21.44 ± 0.23
Lignin (wt.%)	45.65 ± 0.97
Ashes (wt. %)	1.93 ± 0.14

The particle's initial temperature, T_0 , radius, r_0 and pressure, P_0 , and the used properties are defined in Table 1. The particle's boundary is at temperature T_R following a $50^\circ\text{C min}^{-1}$ heating rate, like the experiments. The model is implemented and solved in gPROMS ModelBuilder® 7.0.

4. Results and Discussion

Table 2 shows that lignin is the main component of PKS, followed by hemicellulose and cellulose. Consequently, cellulose/lignin and hemicellulose/lignin interactions would be expected in pyrolysis (Hosoya et al., 2007). PKS also show a relatively low content of extractives and ashes. Regarding the pyrolysis experiments, Figure 2 a) shows that the synthetic mixture decomposes differently than the PKS during pyrolysis. Whereas PKS mainly decompose between $250 - 400^\circ\text{C}$, the synthetic mixture decomposes between $200 - 800^\circ\text{C}$. Thus, the remaining mass after pyrolysis reaches 18% for PKS and 34% for the mixture. This represents a biomass-to-gas mass conversion of 82% for the PKS and of 66% for the mixture. Such behaviour could be due to the different morphology that PKS and the mixture present. Moreover, lignin/cellulose and lignin/hemicellulose interactions would be expected for PKS, which could affect pyrolysis conversion (Hosoya et al., 2007). Figure 2 b) shows that the synthetic mixture's pyrolysis resembles that of pure lignin. On the contrary, in PKS, lignin seem to form a stable structure, which is more difficult to be pyrolyzed.

Figure 3 a) shows that the developed model reproduces cellulose pyrolysis from experiments. For both the model and the experiments, cellulose pyrolysis occurs between 300 and 400°C , and the remaining mass at 800°C is 3.3% for the experiment and 0.43% for the model. Discrepancies are found between modelled and experimental results for hemicellulose and lignin, which is evidenced in Figure 3 b). Indeed, for hemicellulose, experiments show pyrolysis occurring between $200 - 500^\circ\text{C}$, whereas the model predicts pyrolysis between $300 - 400^\circ\text{C}$. A similar behaviour was observed for lignin. Although for hemicellulose, there is a good agreement between modelled (3.1%) and experimental (5.9%) remaining mass at 800°C ; a poor agreement was found for lignin between predicted (3.0%) and experimental (43.1%) pyrolysis remaining mass. This behaviour could be explained when considering that unlike cellulose, hemicellulose and lignin are not formed by homogeneous building blocks. Whereas hemicellulose is a mixture

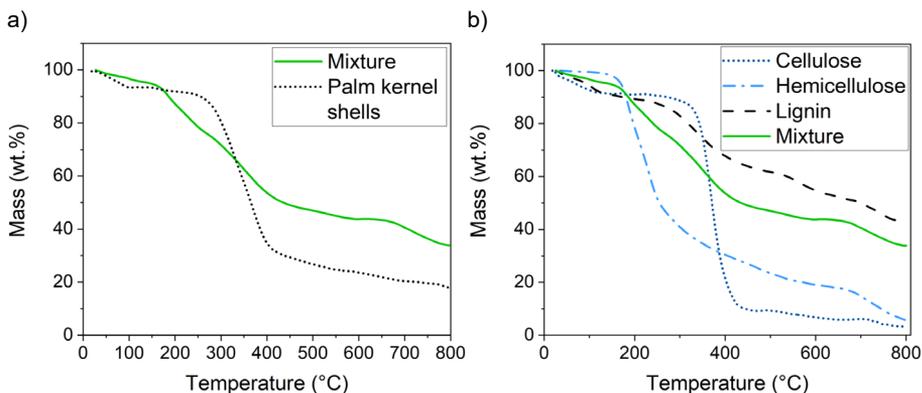


Figure 2: Comparison of the remaining mass in TGA with $50\text{ }^{\circ}\text{Cmin}^{-1}$ for pyrolysis of the synthetic mixture and PKS a), and for the synthetic mixture and its individual components b).

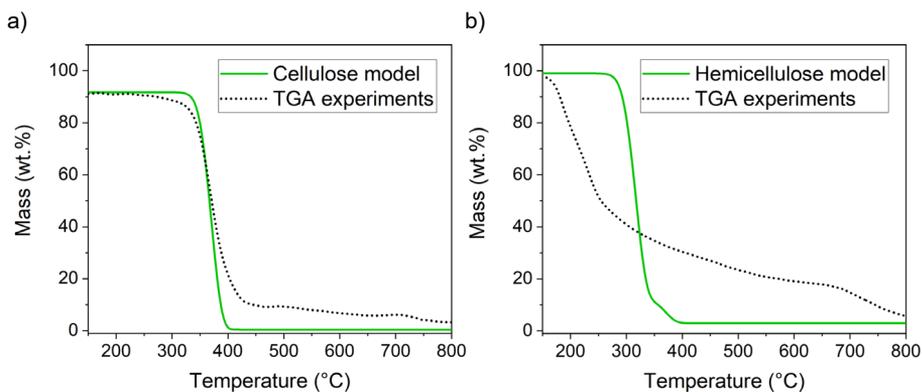


Figure 3: Comparison of experimental and modelled remaining mass for cellulose a) and hemicellulose b) pyrolysis, with a heating rate of $50\text{ }^{\circ}\text{Cmin}^{-1}$.

of pentoses, hexoses and hexuronic acids; lignin contains a variety of aromatic monomers (Debiagi et al., 2015), which are not included in the model. Consequently, modelling PKS pyrolysis as a mixture of individual components, overestimates PKS pyrolysis conversion. This is shown in Figure 4 as a lower remaining mass in the model (2.3%), compared to the experimental results (17.9%). Modified models that included additional forms of hemicellulose, and a 2-stage lignin pyrolysis model (Wang et al., 2016) were also included in the analysis. Although such modifications improve the model's prediction, there is still a discrepancy in the lignin pyrolysis region ($> 400\text{ }^{\circ}\text{C}$). Considering the temperature ranges for cellulose, hemicellulose and lignin pyrolysis, and the poor agreement between modelled and experimental results after $350\text{ }^{\circ}\text{C}$; discrepancies could be attributed to lignin/hemicellulose and lignin/cellulose interactions. Moreover, the used kinetic models consider types of hemicellulose and lignin that might not represent the PKS. Using a detailed kinetic model for PKS, considering the effect of components interactions to adjust the model, and including experimental kinetic parameters for PKS, could improve the model's predictions. The internal structure of the biomass could be also playing a role in PKS pyrolysis, and further work should be performed to understand the way morphology affects pyrolysis kinetics.

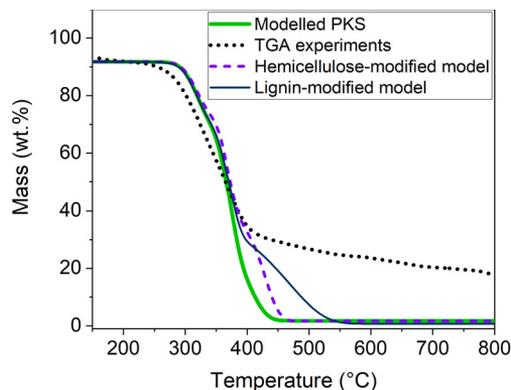


Figure 4: Comparison between experimental and modelled PKS pyrolysis at 50 C min^{-1} . Modelled PKS consider a mixture of cellulose, hemicellulose and lignin. The modified PKS models consider different kinetics for hemicellulose and lignin from Wang et al. (2016).

5. Conclusions

The studied PKS decompose in the $250 - 400 \text{ }^\circ\text{C}$ range throughout pyrolysis. The developed mechanistic model describes cellulose pyrolysis better than hemicellulose and lignin pyrolysis. The agreement between experimental and modelled results for PKS is improved by modifying the model in accordance with the experimental TGA results.

6. Acknowledgements

Secretaría de Educación Superior, Ciencia, Tecnología e Innovación-SENESCYT Ecuador.

References

- A. Anca-Couce, N. Zobel, 2012. Numerical analysis of a biomass pyrolysis particle model: Solution method optimized for the coupling to reactor models. *Fuel* 97, 80–88.
- T. Bridgwater, 2018. Challenges and Opportunities in Fast Pyrolysis of Biomass: Part I. *Johnson Matthey Technology Review* 62 (1), 118–130.
- A. Chico-Proano, G. Manos, L. G. Papageorgiou, E. S. Fraga, 2021. Heating Strategies in Cellulose Pyrolysis as an Alternative For Targeting Energy Efficient Product Distribution. *Chemical Engineering Transactions* 86, 61–66.
- P. E. A. Debiagi, C. Pecchi, G. Gentile, A. Frassoldati, A. Cuoci, T. Faravelli, E. Ranzi, 2015. Extractives Extend the Applicability of Multistep Kinetic Scheme of Biomass Pyrolysis. *Energy and Fuels* 29 (10), 6544–6555.
- C. Di Blasi, 1994. Numerical simulation of cellulose pyrolysis. *Biomass and Bioenergy* 7 (1-6), 87–98.
- M. B. Gorenssek, R. Shukre, C. C. Chen, 2019. Development of a Thermophysical Properties Model for Flowsheet Simulation of Biomass Pyrolysis Processes. *ACS Sustainable Chemistry and Engineering* 7 (9), 9017–9027.
- T. Hosoya, H. Kawamoto, S. Saka, 2007. Cellulose-hemicellulose and cellulose-lignin interactions in wood pyrolysis at gasification temperature. *Journal of Analytical and Applied Pyrolysis* 80 (1), 118–125.
- R. S. Miller, J. Bellan, 1997. A generalized biomass pyrolysis model based on superimposed cellulose, hemicellulose and lignin kinetics. *Combustion Science and Technology* 126 (1-6), 97–137.
- E. Ranzi, P. E. A. Debiagi, A. Frassoldati, 2017. Mathematical Modeling of Fast Biomass Pyrolysis and Bio-Oil Formation. Note II: Secondary Gas-Phase Reactions and Bio-Oil Formation. *ACS Sustainable Chemistry and Engineering* 5 (4), 2882–2896.
- A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker, 2004. Determination of Structural Carbohydrates and Lignin in Biomass. Tech. Rep. April 2008, National Renewable Energy Laboratory, Colorado. URL <https://www.nrel.gov/biomass/analyticalprocedures.html>
- S. Wang, H. Lin, B. Ru, G. Dai, X. Wang, G. Xiao, Z. Luo, 2016. Kinetic modeling of biomass components pyrolysis using a sequential and coupling method. *Fuel* 185, 763–771.
- S. Wu, D. Shen, J. Hu, H. Zhang, R. Xiao, 2016. Cellulose-hemicellulose interactions during fast pyrolysis with different temperatures and mixing methods. *Biomass and Bioenergy* 95, 55–63.

Methanation of CO₂ byproduct from an ammonia plant with green hydrogen

Samuel Asante^{a,b}, Mark W. Hlawitschka^a, Robert Schlesinger^b

^a*Institut für Verfahrenstechnik, Johannes Kepler Universität, Altenberger Straße 69, Linz 4040, Austria*

^b*Borealis Polyolefine GMBH, St.-Peter-Strasse 25, Linz 4021, Austria*
samuel.asante@gmx.at

Abstract

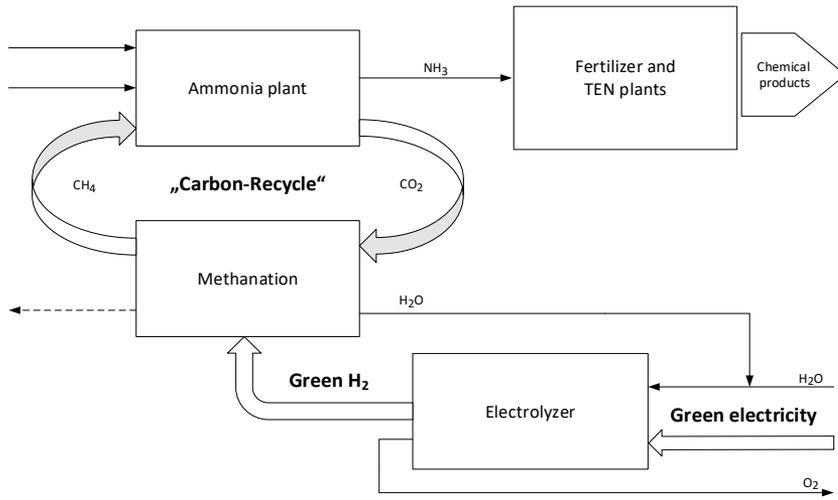
Methanation with regard to the symbiosis of the Borealis and voestalpine industries operating in the chemical park in Linz (Austria) is studied using the commercial software Aspen Plus. The methanation of carbon dioxide with hydrogen allows a reduction of carbon dioxide emissions, especially by considering green energy for electrolysis. The reaction dynamics of carbon dioxide methanation is studied to obtain a suitable range of operating parameters, and a general flow diagram is proposed. Flow schemes and process parameters are proposed based on thermodynamic equilibrium and kinetic reaction models. The effects of using coke oven gas in an ammonia synthesis-methanation complex are presented.

Keywords: CO₂ methanation, Haber-Bosch-process methanation coupling, process model design, kinetic model process simulation

1. Introduction

Borealis operates two ammonia plants at the chemical park in Linz, Austria. An integrated Haber-Bosch process is operated, which uses natural gas as feedstock for the ammonia synthesis. Ammonia is then used as feedstock for downstream fertilizer and melamine plants. Located adjacent to the chemical park are the voestalpine steelworks, a fully integrated steel mill. Coal is processed in coke ovens to produce coke as a carbon source for the blast furnace. This process emits off gas rich in hydrogen, also containing nitrogen, called coke oven gas. Coke oven gas is currently used thermally. A direct chemical utilization as hydrogen source in the methanation reaction would reduce emissions and increase the value chain. These local practicalities enable a methanation unit to run highly efficient. Abundant CO₂ sources are available as concentrated streams along with off gases, and green hydrogen can be synthesized by electrolysis with green energy sourced locally. Carbon emissions could be reduced substantially with the additional benefit of producing a valuable feedstock extensively used by local companies. The implementation enables the formation of a continuous “carbon recycle” across multiple facilities, contributing economic and environmental benefits.

Due to this favorable industrial site set up, the operation of a methanation unit coupled with the Haber-Bosch process and supported by coke oven gas input is simulated to estimate the benefits of its implementation.



Graphical abstract: Ammonia plant methanation coupling

2. Methods

In this work process simulations were performed with Aspen Plus Version 10 (36.0.1.249). Physical properties of the components involved were provided directly by the software from following databanks: APV, APESV (Aspen Plus exclusive) and NISTV (Linstrom and Mallard). Thermodynamic models used are Redlich-Kwong-Soave and Peng-Robinson equations of state. A second order rate expression kinetic model by Lunde (1973), modified and reported by Falbo et al. (2018) was used. The kinetic is:

$$r_{CO_2} = \frac{k}{1+\alpha p_{H_2O}} \left\{ [p_{CO_2}]^n [p_{H_2}]^{4n} - \frac{[p_{CH_4}]^n [p_{H_2O}]^{2n}}{(K_{eq}(T))^n} \right\} \quad (1)$$

$$K_{eq}(T) = \text{EXP} \left[\left(\frac{1}{1.987} \right) * \left(\frac{56000}{T^2} + \frac{56000}{T} - 16.4 \ln T + 0.00557T \right) + 33.165 \right] \quad (2)$$

$$k = k_0 \exp \left(-\frac{E_A}{RT} \right) \quad (3)$$

The following kinetic parameters, provided by Falbo et al. (2018), are used:

$$k_0 = 95.43 \frac{\text{mol}}{\text{s g}_{\text{cat}} \text{ata}^{5n}}$$

$$E_A = 75.3 \frac{\text{kJ}}{\text{mol}}$$

$$n = 0.152 \frac{\text{kJ}}{\text{mol}}$$

$$\alpha = 0.91 \frac{1}{\text{ata}}$$

3. Thermodynamic and kinetic baseline simulations

Several reaction simulations were performed to establish a scientific basis, concerning temperature and pressure dependence of the methanation reaction. The influence of reaction temperature at 1 bar is shown in Figure 1.

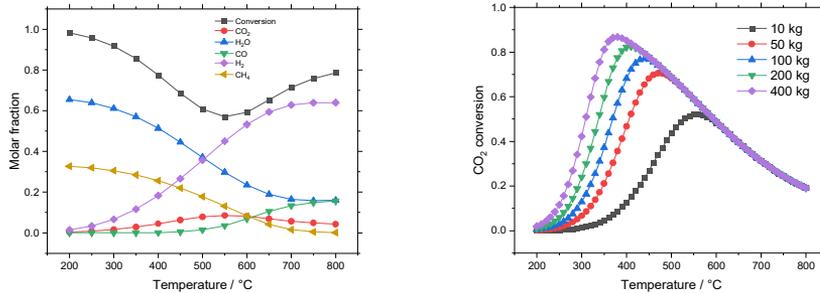


Figure 1: Influence of temperature on product composition and CO₂ conversion in methanation at 1 bar and thermodynamic equilibrium (left). Influence of temperature on CO₂ conversion in methanation at 1 bar and different catalyst loads (kinetic model) (right).

Product gas composition is highly dependent on reaction temperature and in accordance with the observed changes in the conversion of CO₂. Increasing temperature shifts the thermodynamic equilibrium of the methanation reaction to the educt side. An increase in reaction temperature dramatically boosts catalyst activity up to a maximum between 340 and 500 °C after which conversion is reduced again due to the thermodynamic limits on reaction temperature. The methanation reaction responds strongly to an increase in pressure (figures not included). Under increasing pressure higher conversions can be achieved, and higher temperatures can be employed.

4. Process model

The full plant model used is shown in Figure 2. A two-reactor setup was chosen to compensate for thermodynamic and kinetic differences concerning the influence of the reaction temperature.

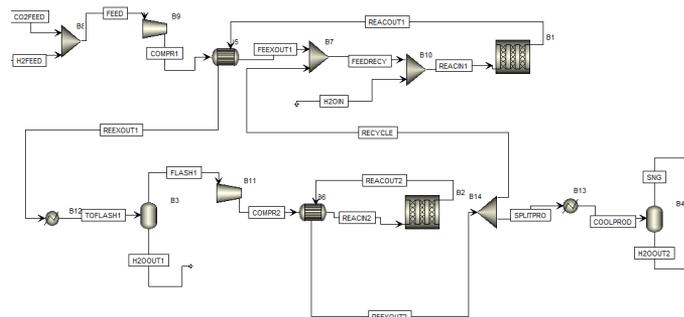


Figure 2: Full kinetic plant model as Aspen Plus flowsheet.

Reactor cooling was simulated with cooling fluid at constant temperature. The volume of reactor one and two are 2 m³ and 3 m³, respectively. For reactor 1 cooling fluid at 500 °C was chosen to simulate molten salt. The second reactor was used to perform with cooling medium at 350 °C resembling steam. The compressors are set to 10 bar for reactor 1 and 30 bar for reactor 2. The split fraction is set to 0.5. All simulations are performed with 1000 kg/h CO₂ and 181 kg/h H₂ input. Steam addition before reactor 1 is set to achieve 10 mol % of water at the reactor inlet by implementation of a design specification. Performance curves are shown in Figures 3 and 4.

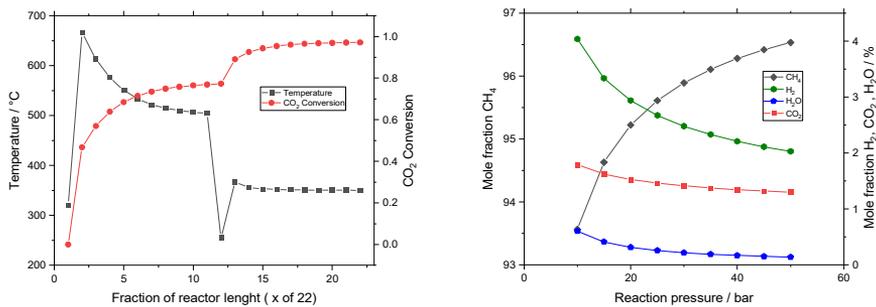


Figure 3: Reactor temperature and CO₂ conversion of both reactor B1 (1-11) and reactor B2 (12-22) (left). Product composition at changing reaction pressure in reactor B2 at 0.5 product recycle fraction (right).

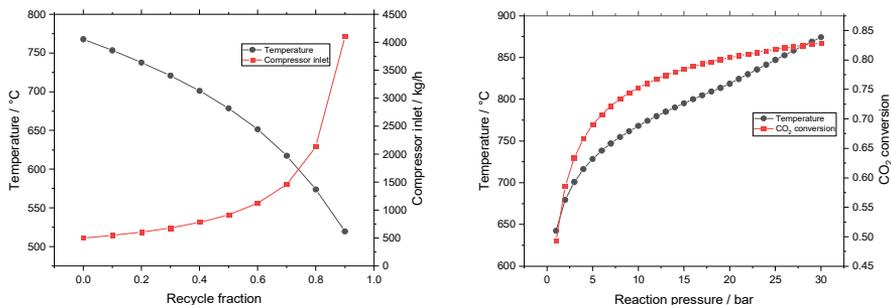


Figure 4: Change in maximum reactor temperature of B1 and compressor B1 inlet at increasing recycle fraction (left). Change in maximum reactor temperature and CO₂ conversion in reactor B1 at increasing reaction pressure and 0.5 product recycle fraction (right).

Characteristic for the methanation reaction is the severe temperature spike at the reactor inlet. This temperature spike could be successfully reduced by 150 to 200 °C if a product recycle and limitation of the reaction pressure in B1 is employed. These methods show disadvantages of higher compression work and lower conversion, respectively. It can be seen in Figure 3 that a product composition in agreement with Austrian natural gas grid quality guidelines according to ÖVGW G. 31 (2001) can be achieved if reaction pressure in B2 is set accordingly.

5. Coke oven gas utilization

Coke oven gas generally contains sizeable amounts of nitrogen. Due to the large amounts of water formed and separated in the methanation reaction the nitrogen content is further increased in the finished product. Natural gas containing nitrogen is problematic for grid injection and may also lead to difficulties in further industrial use. It would be beneficial if the nitrogen content of the synthetic natural gas does not need to be removed before injection into the steam reformer in the Haber-Bosch process. The influence of increasing nitrogen load in the reformer natural gas feed, as commercial natural gas is exchanged, is therefore studied. This was done by manual configuration of the internal Aspen Plus simulation model for the ammonia 2 plant operated by Borealis in Linz. Four Simulations were performed. The observed parameters are reaction pressure (1), feed flow (2), process air flow (3), purge gas flow (4), product flow (5) and product quality (6). Table 1 and 2 show parameter allocation and overall results and Figure 5 examples of recorded performance indicators.

Table 1: Fixed and variable parameter allocation for simulations performed.

Simulation	Fixed parameters	Variable parameters
1	(1), (2), (3), (4)	(5), (6)
2	(3), (5)	(1),(2),(4),(6)
2.5	(3), (6)	(1),(2),(4),(5)
3	(4), (6)	(1),(2),(3),(5)
4	(5), (6)	(1),(2),(3),(4)

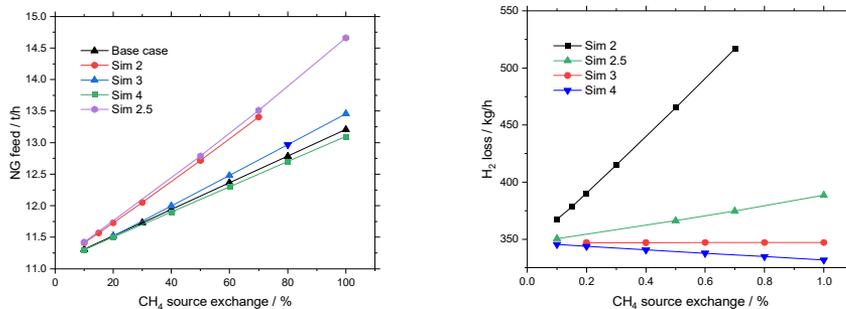


Figure 5: Comparison of NG feed flows and H₂ losses according to CH₄ source exchange at static CH₄ amount

Table 2: Comparison of synthesis gas, reaction pressure, H₂ loss, CH₄ input and NH₃ output changes of different simulations with the base case

Simulation	Change in % vs base case at maximum exchange				
	Synthesis gas	Reaction pressure	H ₂ Loss	CH ₄ Input	NH ₃ Output
1	+101.34	0.00	+0.04	-5.39	+1.27
2	+20.45	-41.67	+48.98	+6.61	+0.00
2,5	+14.25	-1.09	+12.03	+11.00	+10.43
3	+3.44	-3.70	+0.06	+1.90	+2.32
4	-0.46	-0.60	-4.33	-0.83	0.00

Simulations 1 to 3 show large changes in the observed process parameters and therefore bring large overall economic advantages and disadvantages. The setup in simulation 4 shows only direct advantages, but the magnitudes are generally smaller.

6. Conclusion

Insight into CO₂ methanation reaction characteristics is gained by detailed temperature and conversion curves at 1 bar. At thermodynamic equilibrium methanation is most effective at temperatures below 500 °C and high pressure. The kinetic model shows increasing conversion with temperature (maxima between 340 and 500 °C), due to rising catalyst activity. At temperatures exceeding 500 °C the conversion then starts to decline again as a result of the thermodynamic limitations.

A full flow sheet simulation model able to convert 1 t/h of CO₂ to SNG at high conversion and selectivity is designed and optimization paths based on temperature control and product quality are outlined. The model contains 2 cooled methanation reactors and a product recycle. A performance curve is recorded, showing 97 % CO₂ conversion under temperature requirements imposed by catalyst stability.

It is concluded that a complete exchange of external natural gas with natural gas from methanation with coke oven gas as hydrogen source is not only tolerable, but has considerable advantages concerning feed requirements, as H₂ loss, recycle gas volume and reaction pressure. Five different plant operation setups are simulated and able to reduce synthesis loop volume by 0.5 %, reaction pressure by 42 %, H₂ loss by 49 %, CH₄ input by 11 % and increase product output by 10 %. One setup showed purely positive effects of N₂ input concerning the observed parameters. Recycle volume is reduced by 0.46 %, reaction pressure by 0.6 % H₂ loss by 4 % and CH₄ input by 0.83 % at the same product output.

References

- L. Falbo, M. Martinelli, C. G. Visconti, L. Lietti, L., C. Bassano, P. Deiana, 2018, Kinetics of CO₂ methanation on a Ru-based catalyst at process conditions relevant for Power-to-Gas applications, *Applied Catalysis B: Environmental* 225, 354–363.
- P.J. Linstrom, W.G. Mallard, 2020, Eds., NIST Chemistry WebBook, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899.
- P. Lunde, 1973, Rates of methane formation from carbon dioxide and hydrogen over a ruthenium catalyst, *Journal of Catalysis* 30, 3, 423–429.
- ÖVGW G. 31 (2001). Regeln der ÖVGW; Richtlinie G 31-Erdgas in Österreich-Gasbeschaffenheit. Österreichische Vereinigung für das Gas- und Wasserfach, Vienna.

Uncertainty analysis applied to distillation columns calculations

Leonardo O. S. Santana^a, Ewerton E. S. Calixto^a, Fernando L. P. Pessoa^a

^a*Centro Universitário SENAI-CIMATEC, Avenida Orlando Gomes, 1845 Piatã, Salvador 41613-036, Brazil.*

Email: leosantana049@gmail.com

Abstract

Process simulations must involve evaluating output variables by specifying input variables and process parameters. However, in a real process, input data and parameters contain uncertainties, which requires a more realistic analysis, using for example, a stochastic simulation procedure. This technique is based on a large set of simulations using random data according to a suitable probability density function. The aim of this work is to analyze the propagation of uncertainties on the calculation of an equilibrium quantity in a distillation column through a stochastic simulation approach, a very useful tool in process system engineering. In this way, the result is not one exact value for the parameters, but a range of probable values. This method proved to be promising and means more accurate process calculations, generating energy savings in industrial units. A debutanizer column was specified based on FUG algorithm, and Monte Carlo (MC) approach has been used to quantify and calculate the uncertainty propagation on parameters. The feed vapor fraction, distillate (D), and bottom (B) temperatures and compositions were evaluated. The results demonstrate a smooth propagation and a large range in D and B temperatures.

Keywords: Uncertainty analysis, Monte Carlo method, Distillation Column calculation.

1. Introduction

Process simulations must involve evaluating output variables by specifying input variables and process parameters. However, in a real process, input data and parameters contain uncertainties, which can limit the use of simulation results to predict plant behavior. Engineers must therefore know the importance of reliability in the property data of pure substances and mixtures, and understand how small uncertainties in this information can significantly affect the technical and economic performance of an industrial plant. (Santana, 2021^a)

A common mistake made in process calculations is to treat problems using a set of parameters and input data and not as a range. literature presents works using stochastic methods to quantify and measure the propagation of uncertainties along real chemical processes. (Santana, 2021^b)

The quantification of uncertainties, in general, consists basically of three steps: (1) representing the uncertainties in the input parameters of the system; (2) propagation of uncertainty throughout the process, and (3) estimating the stochastic effect on the calculations. (Ramos, 2014) These uncertainties can be introduced mainly in two ways: the first is associated with the method used and depends, for example, on the thermodynamic model, or on the numerical method, and the second way, the error presented in the input data propagates in the result. thus, the better the quality of the data

provided and the accuracy of the prediction methods, the better the estimate of the desired property will be. (Vasquez & Whiting, 2004)

To obtain a more realistic analysis, a stochastic simulation procedure can be performed using the Monte Carlo method, the most widely used stochastic method for calculating uncertainties. This technique is based on a large set of simulations using random data according to a suitable probability density function to estimate the chance of a future event happening. Thus, several simulations are performed to calculate the probabilities of hit or miss the generation of random numbers must follow a probability density function and respect a certain range. (Ramos, 2014.)

This work presents an application of stochastic simulation procedures in the chemical process industry. The method was applied to the calculation of a distillation column (a debutanizer), under uncertainty in the Antoine parameters. The objective of this paper is to analyze the propagation of uncertainties on the calculation of an equilibrium quantity through a distillation column (i.e., debutanizer).

2. Methodology

The specifications for the debutanizer such as column operating pressure, reflux ratio, number of stages, and feed location were obtained and then the distillation column were obtained in the literature (Seader, et al., 2011.) through the FUG algorithm for multicomponent distillation on a debutanizer distillation column. And the component properties were obtained in the DIPPR data bank. (DIPPR801, 2021) and all these data were used to simulate the distillation column on Aspen Plus v12.

Microsoft excel was used to compare pseudo-experimental data for vapor pressure from the DIPPR® data bank with the calculated vapor pressure from Antoine's equation. and the least-squares method was employed using MS excel solver to minimize the uncertainty between the calculated values and the experimental ones. Thus, new parameters were defined (if necessary) and optimized for the estimated temperature range (simulated temperature $\pm 10\%$).

Using the uncertainty from Vapor pressure comparison, the MC approach was applied to quantify and determinate how these uncertainties impact the operation. The column feed, which was flashed at column pressure using the adiabatic flash procedure (Eq. 1). was simulated under uncertainty in the vapor pressure using the generation of 10000 random numbers for P^{sat} defined by Eq. 4 and Eq. 5, of each component. The K-values for vapor-liquid equilibrium were given by Raoult law (Eq. 2) once the pressure system is near to ambient pressure. The stochastic approach generates a range of values for the vapor fraction in the feed stage. Using the boxplot chart, the outliers were removed and a vapor fraction range was found.

Furthermore, simulations were performed on Aspen Plus using different vapor fractions in the feed stage. Different results were obtained on the top and bottom of the column, and MC procedure was used to estimate the probable range of temperature in the bottom and top based on the results of the under-uncertainty simulations.

These different feed conditions generate a range of different product compositions at the bottom and in the top of the column. As the outputs are in a liquid state it must be calculated the bubble point temperature (Eq. 3), using the MC approach, with uncertainty in the P^{sat} , like the feed calculations, and in the component's composition (from feed uncertainty).

The bubble temperature was calculated using 1000 MC simulations. Through the boxplot plot, the outliers were removed and a bubble temperature range was found.

$$\sum_i \frac{z_i K_i}{1 + v(K_i - 1)} = 1 \tag{Eq. 1}$$

Where z_i is the molar fraction composition of the feed, K_i is the partition coefficient defined by the K-values equations, in this case the equation chosen was Raoult Law (Eq. 2.) and v is the vapor fraction of the feed

$$K_i = \frac{P^{sat}}{P} \tag{Eq. 2}$$

For bubble point calculations it is defined that the vapor fraction of the mixture is equal to zero ($v=0$), and the composition is defined as $z_i=x_i$, so that Equation 1 is simplified in the form of Equation 3, below.

$$\sum_i K_i x_i = 1 \tag{Eq.3}$$

The K_i values estimated from Raoult's Law can be defined by $K_i = P^{sat} / P$, where P^{sat} can be calculated using the Antoine equation defined in the Equations 4 and 5 (with 5 and 3 parameters respectively). The Antoine parameters are found in Table 1.

$$P^{sat}(Pa) = e^{A+B/T(K)+C \ln T(K)+DT(K)^E} \tag{Eq. 4}$$

$$P^{sat}(Pa) = 10^{\left(A - \frac{B}{T(K)+C}\right)} * 1000 \tag{Eq. 5}$$

As can be seen, the variables temperature and vapor fraction can be easily found using numerical calculations. Microsoft Excel's VBA programming language was used to solve the equations using interactive methods through Solver and to generate the random numbers from Monte Carlo simulations.

Table 1. Antoine equation parameters (DIPPR801, 2021; L. Santana, 2021^b)

Comp	A	B	C	D	E	Eq
iC4	108.4300	-5039.9000	-15.0120	0.0227	1	4
nC4	66	-4363.2	-7	9.45E-06	2	4
iC5	3.89559	999.99096	-44,1236	N/A	N/A	5
nC5	79	-5420.3	-9	9.62E-06	2	4
C6	104.65	-6995.5	-13	1.24E-05	2	4
C7	88	-6996.4	-10	7.21E-06	2	4
C8	96	-7900.2	-11	7.18E-06	2	4
C9	109.35	-9030.4	-13	7.85E-06	2	4

3. Results

A preview of the debutanizer can be seen on the Figure 1 and the column specifications in the table 2. These results are based on FUG calculation procedure from Seader, et al., 2011. The information from figure 2 and table 2 was used in the Aspen Plus v.12 where temperature, molar fraction, and other results have been obtained.

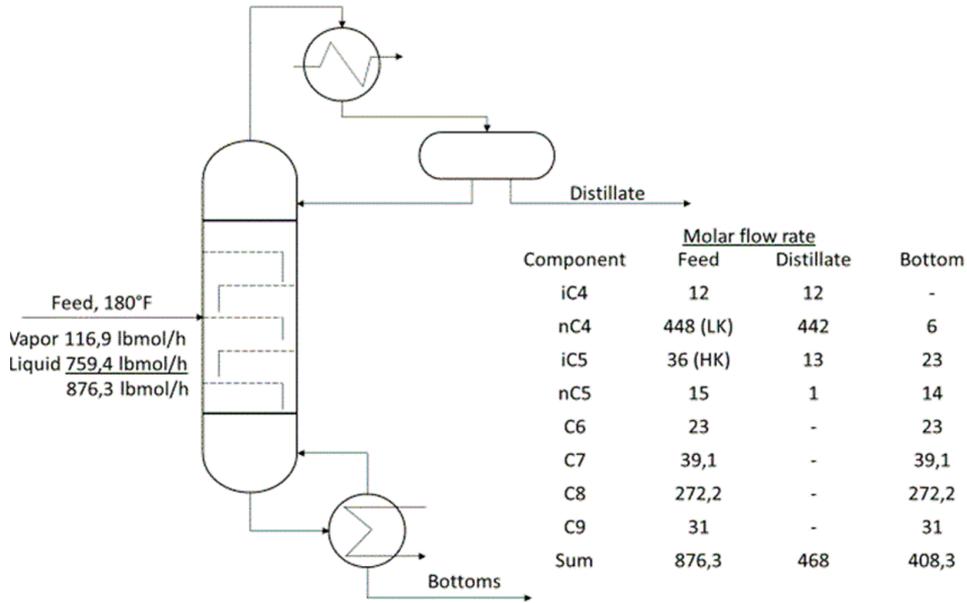


Figure 1: Debutanizer specifications

Table 2: Column specifications (J. Seader, et al., 2011.)

Number of stages	18
Reflux ratio	0.8
Drop pressure	5 Psia
Feed location	Above stage 6

Table 3 shows the first simulating results resume. Used as base case for the MC procedures.

Table 3: initial simulating results.

	Feed	Distillate	Bottom
Molar vapor fraction	0.133	0	0
Temperature	355.372	328.231	447.582
Pressure (Pa)	497565.306	558475.340	592949.127

The Antoine's parameters shown to be very accurate for the range of initial feed temperature $\pm 10\%$, then wasn't necessary to optimize the parameters values. The parameters uncertainties after this analysis can be found in Table 4 and the confident interval is based on this data. All uncertainties used in the calculations are presented in Table 4, and in Table 5 it can be seen the results of uncertainty analysis.

Table 4: Uncertainty (%) based

Component	In P^{sat} (from Literature data) (DIPPR801, 2021)	In P^{sat} (after analysis)	In distillate composition (after simulations)	In bottom composition (after simulations)

Ic4	3.00000	0.09736	0.048625	44.74108
Nc4	3.00000	0.12408	0.745863	42.80125
Ic5	3.00000	0.12408	19.81697	10.06561
Nc5	3.00000	0.12408	24.47607	7.398475
C6	3.00000	0.36381	44.33357	0.146274
C7	3.00000	0.12517	47.55897	3.09486E-03
C8	3.00000	0.19169	51.43366	3.52628E-05
C9	3.00000	0.20361	53.90483	5.45788E-07

Based on the results from Figure 3 we can conclude that the uncertainty remains smooths in the process after the outliers have been excluded, the data from table 5 shows that the results respect an almost ideal normal distribution with a very low asymmetry coefficient. With a smooth platykurtic format curve, that is a distribution with a larger standard deviation than the mesokurtic curve, and so the curve is less peaked.

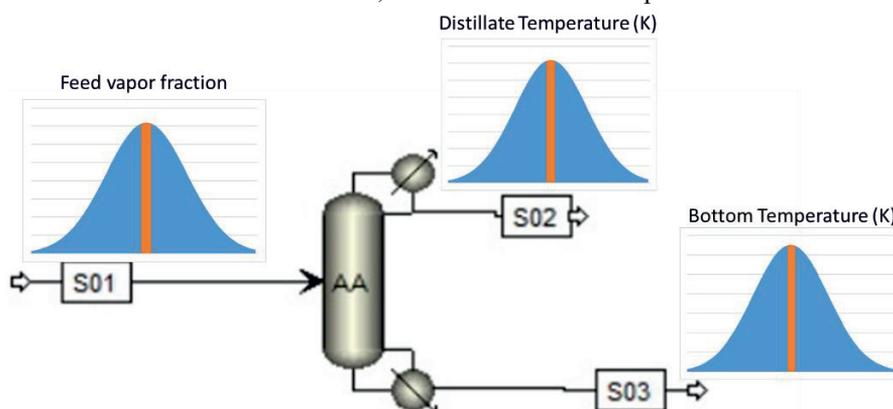


Figure 3: Monte Carlo simulation result

Table 5: Evaluated uncertainties results

	Feed (v)	Distillate (T [K])	Bottom (T [K])
Mean \pm Std. dev.	0.13256 \pm 0.02664	326.6694 \pm 1.1097	435.9215 \pm 3.0647
Uncertainty (In vapor pressure)	3%	3%	3%
Confidence interval	12%	14%	13%
Kurtosis coefficient	0.26687 (platykurtic)	0.29101 (platykurtic)	0.28090 (Platykurtic)
Asymmetry coefficient	-0.05281	-0.01199	0.03687

4. Conclusion

The stochastic approach is a good way to simulate and predict most of the feasible scenarios in an industrial plant, being an alternative to overcome the problems caused by the lack of reliability in the data for process calculations. Uncertainties are intrinsic to any process, so the approach to real physical problems must be seen as a range of probable values and not as a set of result. Therefore, it is essential that chemical engineers know the characteristics of this category of approach and know the fundamental steps to quantify process uncertainties. In this work, the method was applied to the calculation of chemical processes, the calculation of a debutanizer, the results show that a minimal

uncertainty in only one temperature-dependent property can produce larger output uncertainty.

Acknowledgements

To Yan Valdez who helped me coding in VBA and Artur Bispo who helped simulate the process. And to SENAI-CIMATEC, CNPq e FAPERJ for the financial support.

References

- DIPPR801 data bank, available in: <https://dippr.aiche.org/SampleDb>, access in: 20/09/2021
- Seader, J. et al., 2011. Separation process principles: chemical and biochemical operations, Hoboken: John Wiley & Sons, Inc., 3rd ed.
- Ramos, K. 2014. 5 Propagação de incertezas. In: Propagação de incertezas via expansão por caos polinomial em simulação de reservatórios de petróleo. Dissertação de mestrado. Departamento de engenharia elétrica. Pontifícia universidade católica do Rio de Janeiro, Rio de Janeiro.
- Santana, L. et al., 2021.a introduction of stochastic simulation of chemical processes: Properties calculations. 2021 Brazil student regional conference AIChE.
- Santana, L. et al., 2021.b Uncertainty analysis applied to the calculation of the bubble point temperature of mixtures. VII international symposium on innovation and technology SIINTEC.
- Vasquez, V & Whiting, W. 2004. Analysis of random and systematic error effects on uncertainty propagation in process design and simulation using distribution tail. Chemical Engineering Communications, vol. 191, pp. 278-301.

Gas diffusion channels in Li-O₂ batteries: a comparison of wet and flooded electrodes

Jean Felipe Leal Silva, Gustavo Doubek, Rubens Maciel Filho

*Advanced Energy Storage Division – Center for Innovation on New Energies (AES/CINE), School of Chemical Engineering, University of Campinas (UNICAMP).
Rua Michel Debrun, s/n, Campinas, SP, Brazil, ZIP code 13083-841.
jefelipe@outlook.com*

Abstract

The projected increase in the market share of renewable energy sources such as wind and solar power will demand a similar increase in the global energy storage capacity. Batteries represent one of the most flexible solutions to cope with this demand because of their easy installation and high energy density when compared to other large-scale energy storage systems. In this context, one of the most promising technologies is Li-O₂, whose energy density can exceed 1000 Wh kg⁻¹. However, many challenges need to be addressed for this battery technology before it can become available to the market. For instance, the O₂ transport inside the battery limits the power density. One of the strategies to improve the supply of O₂ to the battery is the so-called wet electrode approach, in which parts of the O₂ electrode are filled with gas instead of electrolyte, thus serving as a means to deliver O₂ to the interior of the electrode. Bearing all this in mind, this study presents the simulation of two battery models working under different current densities to demonstrate the impact of electrode geometry on the performance of the battery. The models, developed on COMSOL Multiphysics 5.6, consist of two Li-O₂ batteries with either a simple flooded O₂ electrode or a wet O₂ electrode with gas diffusion channels, with both electrodes based on properties of carbon nanotubes. According to the results, although the gas diffusion channels take space inside the battery and decrease the surface area for deposition of discharge product, this electrode design strategy increased the accessibility of O₂ throughout the length of the electrode. The increased transport capacity was fundamental to increasing the energy density of the battery. The results also indicate that this design approach was fundamental to enable increased power demand for Li-O₂ batteries without compromising their energy density because of the possibility to deliver O₂ more easily to the whole electrode. Based on this design, it is important to optimize the density of these gas diffusion channels to increase the available surface area of the O₂ electrode without compromising the transport of O₂ to the interior of the electrode. Moreover, the optimization of this design should also consider limitations related to surface properties of electrode materials such as hydrophobicity and hydrophilicity to ensure that the gas diffusion channels stay available for O₂ transport.

Keywords: energy storage, battery, lithium-air, simulation

1. Introduction

Solar and wind power have become the source of electricity with the lowest cost on a global average (Lazard, 2020). This achievement will enable the wide adoption of renewable electricity sources in the near future. However, these renewable electricity sources are intermittent, and their wide adoption will represent a challenge for grid

operators unless energy storage becomes widely available. To accomplish this, energy storage costs must decrease, and new capacity needs to be installed across the whole grid (Jülch, 2016). One technology option to meet these requirements is the Li-O₂ battery. Batteries can be installed everywhere, independent of suitable geography. Moreover, this battery has a specific energy potential above 1000 Wh kg⁻¹, higher than that of lithium-ion, and this factor has an important contribution to device cost (Tan et al., 2017).

However, the full potential of Li-O₂ batteries is still hindered by many factors, such as instability of battery components, high resistivity of the discharge product, irreversibility of discharge product, mass transport of O₂, among others (Wang et al., 2020). Sluggish O₂ transport represents a large burden because the battery reaches the cut-off potential much sooner when more O₂ transport is demanded at an increased discharge rate.

As the diffusion of O₂ is faster in gas than in liquid, one of the solutions to improve the transport of O₂ is to only partially fill the electrode with electrolyte — the wet electrode approach (Gwak & Ju, 2016). Bearing all this in mind, this work shows the simulation of Li-O₂ batteries operating with a flooded or a wet electrode. The performance of the two cases was compared in different current densities. Results demonstrate the potential of gas diffusion channels in improving the performance of Li-O₂ batteries for increased power output without compromising the energy density of the device.

2. Method

Figure 1 illustrates a Li-O₂ battery and the two approaches considered in this model: flooded electrode and wet electrode with gas diffusion channels. Both models were built on COMSOL Multiphysics® 5.6 (COMSOL AB, 2020). In both electrodes, the regions with carbon nanotubes were modeled using the Newman approach. In the wet model, gas diffusion channels with diameters of 3–9 μm were built to simulate the presence of macropores free of electrolyte inside the O₂ electrode. The models represent a slice of the O₂ electrode to account for the variability of pore shape, whose geometry was based on images of pores from other works (Su et al., 2020). The battery operates at 25 °C.

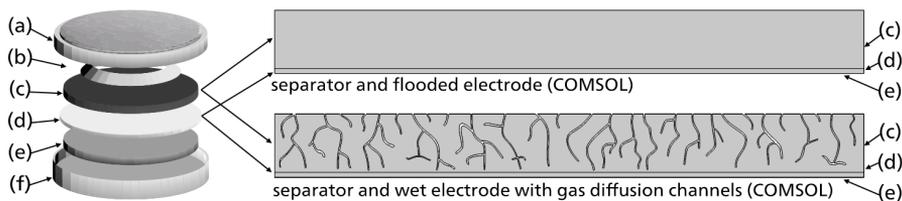


Figure 1. Exploded-view drawing of a Li-O₂ coin cell (not to scale) showing: a) perforated top cover, b) spring, c) O₂ electrode, d) separator membrane, e) Li electrode, f) bottom casing. At the right side, the equivalent 2D model on COMSOL Multiphysics® for the flooded and wet electrodes, showing the c) O₂ electrode (top domain), d) separator membrane (bottom domain), and the Li electrode (bottom boundary). O₂ is fed at the boundary at the top of the O₂ electrode.

In this battery, the reactions described in Eq. (1) and (2) take place during discharge. Li⁺ is dissolved in the electrolyte (LiClO₄ in dimethyl sulfoxide) at a concentration of 1 mol L⁻¹, which is a range suitable for the application of the concentrated solution theory (Tan et al., 2017). The concentration of a species *i* is expressed based on the Nernst-Planck equation as a function of the volume fraction of electrolyte (ε_l), concentration (C_i), molar flux (N_i), and a source or sink term (r_i), as shown in Eq. (3). The volume fraction of electrolyte is related to the electrode volume fraction (ε_s) and the volume fraction of Li₂O₂ (ε_p , the discharge product), as shown in Eq. (4). The molar flux of Li⁺ and O₂, considering

that no convection is present, is a function of the effective diffusion ($D_{Li^+}^{eff}$ or $D_{O_2}^{eff}$), concentration (C_{Li^+} or C_{O_2}), current density in the electrolyte (i_l), transference number (t_+), and the Faraday constant (F), as shown in Eq. (5) and (6).

$$Li \rightarrow Li^+ + e^- \quad (E^0 = 0.00 \text{ V}) \quad (1)$$

$$2 Li^+ + O_2 + 2e^- \rightarrow Li_2O_2 \quad (E^0 = 2.96 \text{ V}) \quad (2)$$

$$d(\varepsilon_i C_i)/dt + \nabla \cdot \mathbf{N}_i = r_i \quad (3)$$

$$1 = \varepsilon_s + \varepsilon_l(t) + \varepsilon_p(t) \quad (4)$$

$$\mathbf{N}_{Li^+} = -D_{Li^+}^{eff} \nabla C_{Li^+} + \mathbf{i}_l t_+ / F \quad (5)$$

$$\mathbf{N}_{O_2} = -D_{O_2}^{eff} \nabla C_{O_2} \quad (6)$$

Considering a binary Li⁺ electrolyte, the current density in the liquid solution is a function of the effective ionic conductivity of the electrolyte (κ_{eff}), electrolyte potential (φ_l), universal gas constant (R), temperature (T), and activity coefficient of Li⁺ (f_{\pm}), as shown in Eq. (7). The derivative of f_{\pm} is given by Eq. (8) in the range of $0.2 < C_{Li^+} < 2 \text{ mol L}^{-1}$ (Nyman et al., 2008). Transport of charges in the O₂ electrode follows Ohm's law, and it is a function of the effective electronic conductivity of the O₂ electrode (σ_{eff}) and the potential of the electrode (φ_s), as shown in Eq. (9). Because of the porous structure of the O₂ electrode, effective parameters are estimated based on the volume fraction and the tortuosity of the porous medium. According to the Bruggeman model (Tjaden et al., 2016), in the case of carbon nanotubes, the tortuosity can be estimated as the reciprocal of the porosity, thus leading to Eq. (10)–(13) to estimate the effective parameters. In this model, O₂ is fed to the battery by means of a boundary condition, as a function of the partial pressure of O₂ (p_{O_2}) and the O₂ solubility in the electrolyte (H), as seen in Eq. (14).

$$\mathbf{i}_l = -\kappa_{eff} \nabla \varphi_l - 2RT \kappa_{eff} / F (-1 + t_+) (1 + \partial \ln(f_{\pm}) / \partial \ln C_{Li^+}) \nabla \ln(C_{Li^+}) \quad (7)$$

$$\frac{\partial \ln(f_{\pm})}{\partial \ln C_{Li^+}} = \frac{(0.28687C_{Li^+}^2 - 0.74678C_{Li^+} + 0.44103)}{0.1287C_{Li^+}^3 - 0.4106C_{Li^+}^2 + 0.4717C_{Li^+} + 0.5508} - 1 \quad (8)$$

$$\mathbf{i}_s = -\sigma_{eff} \nabla \varphi_s \quad (9)$$

$$D_{Li^+,eff} = \varepsilon_l^2 D_{Li^+} \quad (10)$$

$$D_{O_2,eff} = \varepsilon_l^2 D_{O_2} \quad (11)$$

$$\kappa_{eff} = \varepsilon_l^2 \kappa \quad (12)$$

$$\sigma_{eff} = \varepsilon_s^2 \sigma \quad (13)$$

$$C_{O_2,e}^{sat} = H p_{O_2} / RT \quad (14)$$

Conservation of charges is described by Eq. (15). Transference of charge between phases is a function of available surface area (a) and local transfer current density (j), as shown in Eq. (16). Based on these parameters, the source term is calculated via Eq. (17).

$$\nabla \cdot \mathbf{i}_l + \nabla \cdot \mathbf{i}_s = 0 \quad (15)$$

$$\nabla \cdot \mathbf{i}_l = aj \quad (16)$$

$$r_i = -aj / F \quad (17)$$

The rate expression for the electrochemical reaction at the O₂ electrode can be described by the Butler-Volmer kinetics, shown in Eq. (18), which is a function of the anodic and cathodic rate constants (k_a and k_c) and the activated overpotential (η_c), described in Eq. (19). The activated overpotential is a function of the potential drop over the electrolyte, the electrode, and the film of Li₂O₂ ($\Delta\varphi_{film}$), and the equilibrium potential of the reaction (E_c°). The potential drop in the film of discharge product is a function of its thickness (Δs)

and conductivity (σ_p), as shown in Eq. (20). The volume fraction of discharge product deposited onto the O₂ electrode ($\varepsilon_{Li_2O_2}$), given by Eq. (21), is a function of its concentration ($C_{Li_2O_2}$), molar mass ($M_{Li_2O_2}$), and density ($\rho_{Li_2O_2}$). This volume fraction is used to estimate the surface coverage (θ) with Eq. (22), the available surface area with Eq. (23), and the thickness of the film with Eq. (24).

$$j/nF = k_a C_{LiO_2} \exp(F\eta_c/RT) - k_c C_{Li^+}^2 C_{O_2} \exp(-F\eta_c/RT) \quad (18)$$

$$\eta_c = \varphi_s - \varphi_l - \Delta\varphi_{film} - E_c^\circ \quad (19)$$

$$\Delta\varphi_{film} = j\Delta s/\sigma_p \quad (20)$$

$$\varepsilon_{Li_2O_2} = C_{Li_2O_2} M_{Li_2O_2}/\rho_{Li_2O_2} \quad (21)$$

$$\theta = (\varepsilon_{Li_2O_2}/\varepsilon_{l,0})^Y \quad (22)$$

$$a = a_0(1 - \theta), \quad a_0 = 2\varepsilon_s/r_0 \quad (23)$$

$$\Delta s = (\sqrt[3]{(\varepsilon_{Li_2O_2}\theta + \varepsilon_s)/\varepsilon_s} - 1)r_0 \quad (24)$$

The membrane and electrode thicknesses were set to 25 μm and 500 μm , and the slice of O₂ electrode had 3 mm. The conductivity of the O₂ electrode was 300 S m⁻¹ between planes and 5000 S m⁻¹ in-plane (Matsumoto et al., 2021). Other parameters were based on previous literature (Sahapatsombut et al., 2013). A mesh refinement study was performed to decide upon the number of mesh elements and solver tolerance. The simulations were run on an Intel Core i5 9600KF (overclocked to 4.7 GHz and with 32 GB RAM). The batteries were discharged at current densities of 0.5, 2.0, 3.5, and 5.0 A m⁻² till the stop condition of deep discharge (cut-off potential of 2.5 V) was reached.

3. Results and discussion

Figure 2 presents the results of the mesh refinement study. In the case of the flooded electrode, the mesh with 336 elements was selected because increasing the number of elements to 760 leads to a variation of 0.1% in the specific capacity with an increase of 80% in solver time. For the wet electrode, the mesh with 5135 elements was selected because increasing the number of mesh elements to 8389 leads to a change of 0.02% in the specific capacity with an increase of 82% in solver time. Relative tolerance of 10⁻³ was selected because tightening the relative tolerance from 10⁻³ to 10⁻⁴ leads to a change in the specific capacity of 0.006% in the flooded electrode and 0.004% in the wet electrode, with an increase in the solver time of 60% and 130%, respectively.

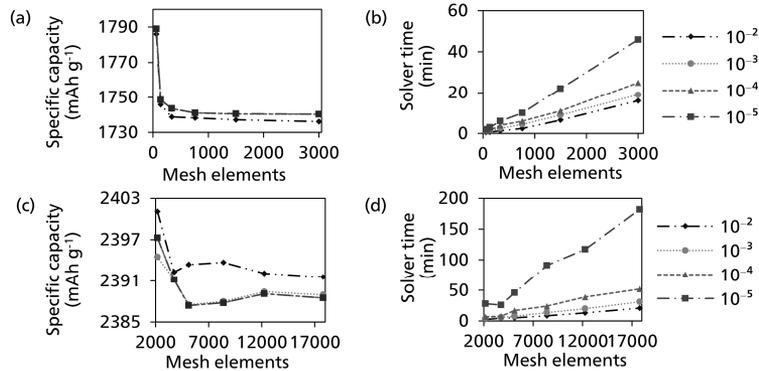


Figure 2. Mesh refinement study for the flooded (a, b) and wet (c, d) electrodes showing the specific capacity (a, c) and solver time (b, d) for different relative tolerances (10⁻²–10⁻⁵).

Figure 3 shows the specific capacity of the Li-O₂ battery for different current densities (0.5–5.0 A m⁻²) in both flooded and wet electrodes, and Figure 4 shows the distribution of Li₂O₂ after discharge at 0.5 and 5.0 A m⁻² and the concentration of O₂ after discharge at 5.0 A m⁻². Comparing graphs (a) and (b) of Figure 3, it is possible to see that, despite presenting 6.85% less area for discharge than the flooded electrode, the wet electrode reaches a specific capacity 15% larger at 0.5 A m⁻² and almost three times larger at a current density of 5 A m⁻². Also, a tenfold increase in the current density leads to a loss of capacity 21% lower in the wet electrode when compared to the flooded electrode.

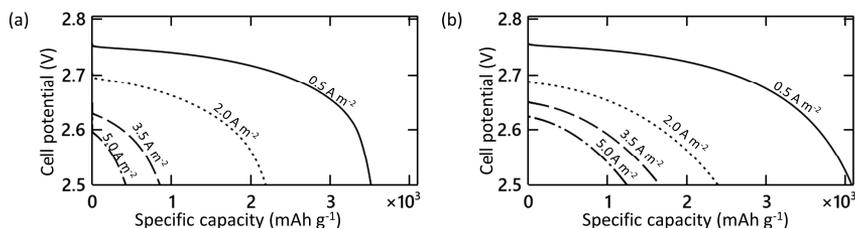


Figure 3. Specific capacity of the Li-O₂ battery for different discharge current densities (0.5–5.0 A m⁻²) for the a) flooded and b) wet electrodes.

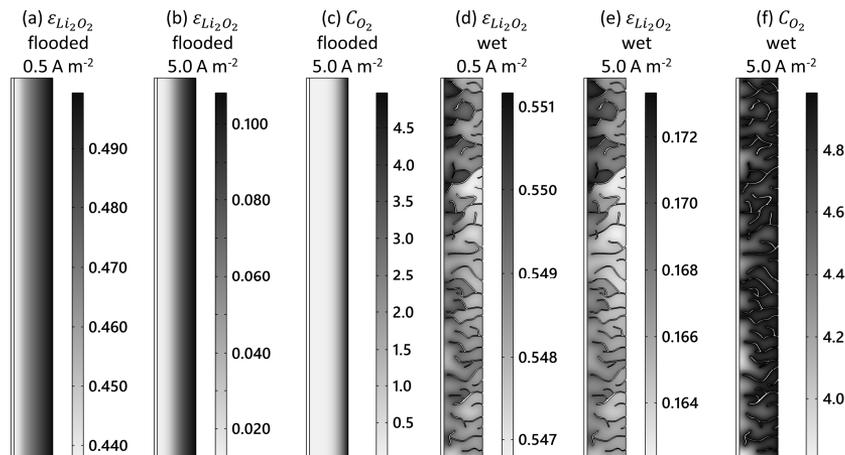


Figure 4. Distribution of discharge product represented as the volume fraction of Li₂O₂ ($\epsilon_{Li_2O_2}$ in a, b, d, e) for current densities of 0.5 (a, d) and 5.0 (b, e) A m⁻² and concentration of O₂ (C_{O_2} in c, f) at a current density of 5.0 A m⁻² for flooded (a–c) and wet (d–f) electrodes.

The decreased performance of the battery in the case of the flooded electrode is explained by the poor distribution of discharge product in the O₂ electrode, as seen in Figure 4. Observing graph (a) of Figure 4, it is possible to see a large gradient of discharge product distribution, with a significant product buildup near the O₂ feed side of the O₂ electrode. In this case, the volume fraction of discharge product in the O₂ electrode varies from 0.438 to 0.499, a range of 0.061. On the other hand, in the wet electrode, it varies from 0.547 to 0.551, a range of only 0.004. Because of the large range in the case of the flooded electrode, the effective diffusion of O₂ near the O₂ feed side decreases faster as discharge takes place because of the increased volume of solids in this region, according to Eq. (11). This problem is circumvented in the case of the wet electrode.

The challenge of delivering O₂ to the flooded electrode is better demonstrated at a current density of 5.0 A m⁻². In this case, the final volume fraction of Li₂O₂ in the flooded electrode varies from 0.012 to 0.108, a range of 0.096. In the wet electrode, the final volume fraction of Li₂O₂ varies from 0.162 to 0.173, a range of only 0.011. The O₂ diffusion bottleneck is clear in the comparison of graphs (c) and (f) of Figure 4. At the end of the discharge, O₂ concentration is below 10% of the saturated concentration in 59% of the flooded electrode, and below 1% of the saturated concentration in 41% of the flooded electrode, making it underused. However, the lowest O₂ concentration in the wet electrode is 76% of the saturated concentration, showing that the presence of gas diffusion channels greatly increases the capacity of supplying O₂ to the cell during discharge.

4. Conclusions

Past works have demonstrated that one of the main problems of Li-O₂ batteries is the transport of O₂ inside the battery. This work compared two electrodes: the flooded electrode and the wet electrode with gas diffusion channels. The results show that mass transport of O₂ is greatly assisted by the presence of gas diffusion channels despite the small space occupied by them inside the electrode (6.85%), and they allow a loss of capacity 21% lower for a tenfold increase in discharge current. The great improvement also indicates that the geometry of these gas diffusion channels can be further analyzed to find a balance between their size and performance improvement taking into consideration the limitations of materials for the manufacture of O₂ electrodes.

Acknowledgments

The authors acknowledge the support from FAPESP (the Sao Paulo Research Foundation, #2018/16663-2 and #2017/11958-1), Shell, ANP (Brazil's National Oil, Natural Gas, and Biofuels Agency) through the R&D levy regulation, and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

- G. Gwak, H. Ju, 2016, Three-dimensional transient modeling of a non-aqueous electrolyte lithium-air battery, *Electrochim Acta*, 201, 395–409.
- V. Jülch, 2016, Comparison of electricity storage options using levelized cost of storage (LCOS) method, *Appl Energy*, 183, 1594–1606.
- Lazard, 2020, Lazard's Levelized Cost of Energy Analysis – version 14.0.
- M. Matsumoto, R. Yamaguchi, K. Shima, M. Mukaida, M. Tomita, T. Watanabe, T. Ishida, T. Fujigaya, 2021, Control of anisotropic conduction of carbon nanotube sheets and their use as planar-type thermoelectric conversion materials, *Sci Technol Adv Mater*, 22(1), 272–279.
- A. Nyman, M. Behm, G. Lindbergh, 2008, Electrochemical characterisation and modelling of the mass transport phenomena in LiPF₆-EC-EMC electrolyte, *Electrochim Acta*, 53, 6356–6365.
- U. Sahapatombut, H. Cheng, K. Scott, 2013, Modelling the micro-macro homogeneous cycling behaviour of a lithium-air battery, *J Power Sources*, 227, 243–253.
- Z. Su, V. De Andrade, S. Cretu, Y. Yin, M. J. Wojcik, A. A. Franco, A. Demortière, 2020, X-ray Nanocomputed Tomography in Zernike Phase Contrast for Studying 3D Morphology of Li-O₂ Battery Electrode. *ACS Appl Energy Mater*, 3(5), 4093–4102.
- P. Tan, W. Kong, Z. Shao, M. Liu, M. Ni, 2017, Advances in modeling and simulation of Li-air batteries, *Prog Energy Combust Sci*, 62, 155–189.
- B. Tjaden, S.J. Cooper, D.J. Brett, D. Kramer, P.R. Shearing, 2016, On the origin and application of the Bruggeman correlation for analysing transport phenomena in electrochemical systems, *Curr Opin Chem Eng*, 12, 44–51.
- F. Wang, X. Li, X. Hao, J. Tan, 2020, Review and Recent Advances in Mass Transfer in Positive Electrodes of Aprotic Li-O₂ Batteries, *ACS Appl Energy Mater*, 3(3), 2258–2270.

Techno-economic analysis of flexible AP-X LNG production process under risks and uncertainties

Noor Yusuf, Rajesh Govindan, Tareq Al-Ansari*

College of Science and Engineering, Hamad Bin Khalifa University, Doha, P.O. Box: 34110, Qatar.

**talansari@hbku.edu.qa*

Abstract

Natural gas liquefaction has been deployed in different countries to produce liquefied natural gas (LNG) for the economic transportation to distant markets. However, when assessing a natural gas monetisation technology, the profitability of the project is evaluated using the economic performance indicator, net present value (NPV), under deterministic conditions. This does not reflect the real-life scenario since the risks associated with exogenous uncertainties arising from emergence of new suppliers, and changes in contractual structures are not captured under deterministic analysis. Consequently, risks associated with investing in a mega LNG plant shall be considered in the design stages to decide on an optimal strategy for dealing with possible uncertainties throughout the lifetime of the project. In this study, a techno-economic analysis is conducted for the flexible AP-X LNG production technology licensed by Air Products. A three-steps methodology is considered to assess the AP-X technology: (1) technical assessment; (2) economic evaluation; and (3) techno-economic risk assessment and management. In the technical assessment, an LNG production plant model is developed and simulated using the Aspen HYSYS software, and evaluated under deterministic conditions using the NPV. The process is then evaluated under stochastic conditions using Monte Carlo Simulation to understand the impact of different uncertainties on the profitability of the plant. A proactive mode of response strategy to deal with risks is tested to increase the responsiveness of the project to exogenous changes through embedding flexibility in production in the early design stages of the project. The results indicate that a flexible system demonstrates a better economic performance than a rigid system.

Keywords: liquefied natural gas, flexibility, uncertainty, stochastic analysis

1. Introduction

In the past few years, natural gas has been the fastest growing fossil-based fuel due to its environmental and economic characteristics. This is especially true after the Paris Agreement in 2015, where different countries shifted to cleaner energy resources such as natural gas and renewables. Natural gas can be physically or chemically monetized to increase its economic attractiveness. Amongst the different monetization options, liquefied natural gas (LNG) has been adopted by different projects worldwide for economic transportation of natural gas to distant markets. However, the deployment of LNG infrastructures is capital intensive and require signing long-term purchase agreements with consumers in the pre-final investment decision phase of the project (Pre-FID) to guarantee sales, and minimize risks associated with investment both suppliers and

consumers have been tied together with long-term contracts. However, different factors have influenced the LNG business in the last few years, such as the emergence of new suppliers, competitiveness with renewable resources and natural gas supply disruptions due to unexpected shocks, such as the latest unprecedented COVID-19 pandemic that disrupted global energy demand and halted the financing of several LNG projects worldwide due to associated investment risks.

In addition to the latest changes in the LNG business, the liquidity of the US and European natural gas markets also contributed to reshaping the contractual structures, where spot selling and short/mid-term contracts have been favored by consumers to avoid long-term sales and purchase agreements under demand and prices uncertainties. This resulted in an increase of un-contracted LNG volumes, oversupply, and prices fluctuations. Consequently, investigating uncertainties prior to starting new LNG projects is crucial to understanding future possible market trends and behavior in terms of demand and prices. This investigation is essential for investors for financing such capital-intensive projects, for decision-makers to decide on production and selling strategies, and for policymakers for policies developments.

After studying uncertainties involved in final markets, different passive and active mode of responses can be considered to deal with uncertainties. Amongst the different tools, the concept of flexibility arising from financial engineering has been trending for designing flexible engineering systems with certain characteristics and design to allow for changing the system when needed (Cardin et al., 2015; Saleh et al., 2009). Designing a flexible LNG production system provides the producer the right, but not the obligation, to adjust the production capacity based on market demand and prices to capture high demand opportunities or avoid possible losses from selling LNG at lower prices. This study evaluates the flexibility of an AP-X LNG production system under final markets prices uncertainties.

2. Methodology and Data

Natural gas liquefaction is the core of an LNG project and the most capital-intensive part of the project. After treating natural gas, the stream undergoes fractionation for removing heavy hydrocarbons (C3+) wherein by-products such as liquefied petroleum gases (LPG) and natural gas liquids (NGL) are produced. The natural gas stream then undergoes for liquefaction by cooling it down to -164 °C using the AP-X technology licensed by air products. The AP-X technology consists of three main cooling sections: pre-cooling in propane multistage coolers, main cooling in the main cryogenic heat exchanger using mixed refrigerants, and final sub-cooling in the nitrogen expander.

2.1. Process Simulation

Initially, the AP-X LNG production system is simulated using the commercial software Aspen HYSYS V.11 for producing 7.5 MTPA of LNG with methane purity of 90mol%. The economic profitability of the AP-X technology is then evaluated using the economic performance indicator, net present value (NPV), at deterministic demand and forecasted Henry Hub natural gas price in international markets. A flexible design is then proposed to allow for changing the production capacity based on the market needs. Finally, the flexible AP-X system is evaluated under stochastic natural gas prices to evaluate the economic performance of the proposed design.

2.2. Deterministic and Stochastic Economic Analysis

The economic performance indicator, net present value (NPV), is used to evaluate the profitability of the project under deterministic LNG sales price. Subsequently, the project value is evaluated using stochastic modelling under fixed capital and operational costs. Mean reverting jump diffusion Geometric Brownian Motion (GBM) (Poisson distribution) model is considered to describe the stochastic behavior of Henry Hub natural gas prices. Python programming language in Visual Studio Code enabled by Anaconda software is used to create a large sample size of 10,000 using Monte Carlo Simulation (MCS). The stochastic analysis results are essential to understand risks arising from uncertainties. This is also known as Jensen’s Law:

$$f(E[x]) \neq E[f(x)] \tag{1}$$

Where the NPV under average prices and/or demand $f(E[x])$ does not equal the expected NPV under a distribution of prices and/or demand $E[f(x)]$ (Cardin et al., 2015). Hence, time series modelling for natural gas prices is crucial to calculate the expected NPV under different price and/or demand scenarios. The mean-reverting jump diffusion GBM model was given by (Lucheroni & Mari, 2018):

$$d\log(P_t) = (\theta^{gas} - \alpha^{gas} \log(P_t))dt + \sigma^{gas} dW + JdN \tag{2}$$

Where P_t is the natural gas Henry Hub spot price at time t ; θ^{gas} and α^{gas} are the mean-reversion parameters; σ^{gas} is the volatility of gas prices; W is the standard Brownian motion Wiener Process, with zero mean and a standard deviation of 1; and N is the jump-diffusion process, Poisson process, with a jump amplitude of J and intensity λ^{Jump} . The jump-diffusion process is normally distributed with a mean of zero a standard deviation σ^{Jump} . Moreover, the jump diffusion process (N) and the standard Brownian motion Wiener Process are independent process.

Table 1: Geometric Brownian motion and jump diffusion process model parameters (Lucheroni & Mari, 2015).

Parameter	Notation	Value
Mean-Reversion parameters	θ^{gas}	0.0432
	α^{gas}	0.0292
Jumps intensity	λ^{Jump}	0.2542
Standard deviation for Poisson process	σ^{jump}	0.1258
Volatility of gas prices	σ^{gas}	0.0737
Mean for Poisson process	μ^{Jump}	0.01

The estimated natural gas Henry Hub spot prices (P_t) are then used to calculate 10,000 NPV to analyze the NPV distribution, possible risks and opportunities associated with investing under market uncertainties. The capital and operating costs for a brownfield AP-X technology with 7.5 MTPA production capacity are \$5.53 billion (year 2021) and \$1.39 billion (year 2026), respectively (Steuer, 2019). A project’s lifetime of 25 years, with 5 years of construction and 20 years of operation is considered. Revenues from NGL and LPG sales are assumed for evaluating the profitability of the fractionation section of the LNG plant. Hence, only revenues from LNG sales to Asia-Pacific markets under

different pricing/contractual systems were assumed for estimating the NPV and ENPV of the AP-X liquefaction process. LNG can be sold under long-term contracts (LTCs) at a fixed price of \$8.76/MMBTU (year 2025), or under mid-term contracts (MTCs) and/or on spot based on forecasted Henry Hub natural gas prices. Different combinations of selling strategies influence the NPV and ENPV of the project. Consequently, scenarios based on different selling strategies are investigated based. Henry Hub natural gas prices are forecasted using mean-reversion GBM jump diffusion model utilizing parameters obtained from literature based on historical data from 1990 to 2013 (Lucheroni & Mari, 2015):

2.3. Flexibility Analysis

Understanding the potential of embedding flexibility with an industrial system is essential to be evaluated in the early design stages of the project. After evaluating the baseline production system technically and economically under both deterministic and stochastic conditions. The flexibility of the AP-X technology is investigated technically and economically. The stochastic ENPV is used to evaluate the economic performance of the flexible production system under different selling strategies.

3. Results and Discussion

Investing in mega-projects is capital intensive where different entities are involved in project financing. Moreover, with the changes in the LNG business due to emergence of new suppliers, changes in contractual structures and demand disruptions due to COVID-19 pandemic, proactive risk management is required prior to starting new LNG projects. Flexibility implementation is one of the promising proactive modes of responses to be implemented in the AP-X technology. The AP-X natural gas liquefaction technology is simulated using Aspen HYSYS software as illustrated in Figure 1. After liquefying natural gas, nitrogen is removed in a nitrogen distillation column (T-300) and compressed in a series of compressors, whilst the bottom stream (S-301) is LNG produced at $-164\text{ }^{\circ}\text{C}$ and pumped to storage before prior to shipping to final markets.

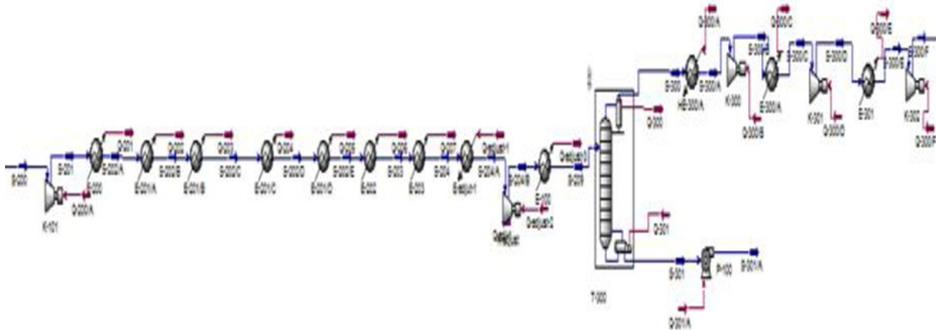


Figure 1: AP-X natural gas liquefaction process simulated using Aspen HYSYS.

The technology is further assessed under deterministic fixed capital and operating costs, and fixed selling price to Asia Pacific markets under LTCs, which results in an NPV of \$5.86. billion. However, in real-life scenarios, it is challenging to secure 100% of production capacity for a single market under LTCs. This is due to the flexibility needed by consumers in diversifying their import portfolios to minimize risks. Consequently, the AP-X technology is further assessed through dedicating the full capacity to spot selling.

The stochastic simulation of the mean-reversion jump-diffusion GBM with an initial Henry Hub natural gas price of 3.06 \$/MMBTU resulted in an ENPV of \$ -3.57 billion and an NPV distribution illustrated in Figure 2. This proves that dedicating the full capacity of a mega LNG production process to spot markets is risky and would potentially involve economic losses. Hence, a diversified selling strategy is needed to hedge against market risks. Moreover, flexibility in production can be implemented though shutting down the third cooling unit (the nitrogen expander) in the AP-X technology. This would result in reducing the production capacity from 7.8 MTPA to 5 MTPA without technical limitations (Roberts et al., 2002).

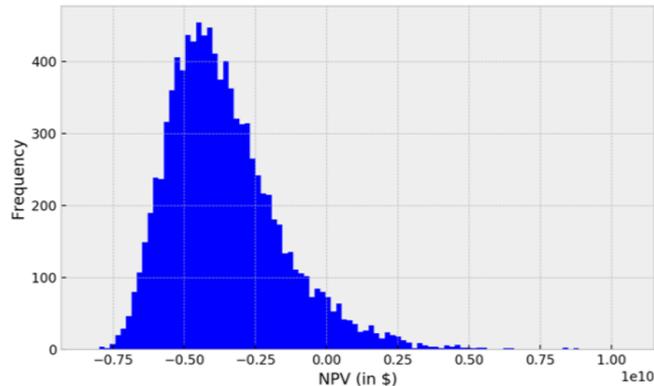


Figure 2: NPV distribution of AP-X's technology with LNG sold based on forecasted stochastic Henry Hub natural gas prices.

An investor might decide on investing on a full capacity (7.8 MTPA) in the early years of the project or start with a production capacity of 5 MTPA and introduce the nitrogen expander during the project's lifetime to increase the production capacity to 7.8 MTPA. Consequently, time-series market analysis is crucial in evaluating the optimal time to expand the production capacity. In this work, it is assumed that a full capacity AP-X technology is deployed in the early years of the project to capture any market opportunities. This is especially important due to the fact that LNG is a transitional fuel, where the market demand for fossil-based fuels is anticipated to gradually decrease with shifts to renewable resources.

When considering a flexible production, this allows the producer to increase the capacity up to 7.8 MTPA when the market conditions are favorable, or to operate at 5 MTPA when the market conditions are low. A producer might consider dedicating 5 MTPA to LTCs and selling the remaining capacity on spot based on market conditions. A case-study of multi-demand level was analyzed to evaluate the influence on considering production flexibility and selling strategies diversification on the ENPV of the project. The multi-demand scenario assumes that the process will only operate at 5 MTPA in the first 5 years of production, then the demand increases where the producers can expand the production capacity up to 7.8 MTPA and sell the remaining in the spot market. In this scenario, the Henry Hub natural gas spot prices mainly influence the returns from LNG sales in spot market and result in NPV distribution illustrated in Figure 3. When following this selling and operating strategy, an ENPV of \$2.72 billion is resulted, which indicates the profitability of considering the presented strategy over dedicating the full capacity to spot selling. A flexible LNG production system will give the producer the right to increase

the production capacity when the market is favorable, or to reduce the capacity subject to unfavorable market conditions.

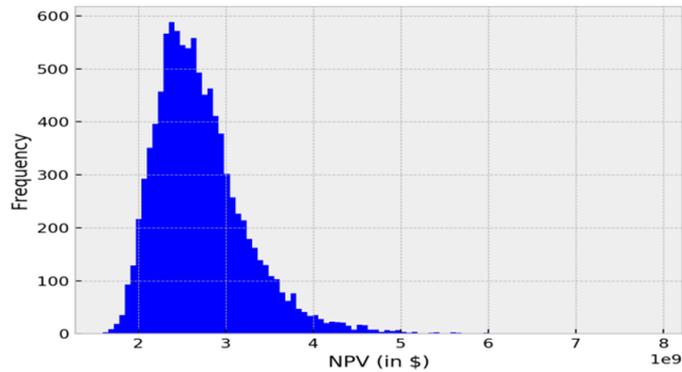


Figure 3: NPV distribution of flexible AP-X's with diversified selling strategy.

4. Conclusion

The last changes in the energy markets due to unexpected shocks, emergence of new suppliers and competitiveness with other renewable resources have influenced the natural gas and LNG markets. Moving from the pre-FID to the FID investment decision change has been challenging in new LNG projects. Consequently, this study proposes embedding flexibility in production with selling strategies diversification as a proactive mode of response to uncertainties. In fact, different combinations of selling strategies influence the ENPV of the project. Multiple demand and selling strategies scenarios can be studied in the early design stages of a mega project to understand possible future market behavior. Moreover, the concept of flexibility can be extended to multi-product value chains that consists of different products. For example, in a natural gas to multi-product value chain, the flexibility in producing different products, i.e., CNG, LNG, methanol, and hydrogen, could contribute to capturing market opportunities and/or avoiding risks based on the market performance.

References

- M. Cardin, M. Ranjbar-Bourani, & R. De Neufville, 2015, Improving the lifecycle performance of engineering projects with flexible strategies: example of on-shore LNG production design. *Systems Engineering*, 18(3), 253–268.
- C. Lucheroni, & C. Mari, 2015, *Risk Shaping of Optimal Electricity Portfolios in the Stochastic LCOE Theory*. October. <https://doi.org/10.13140/RG.2.1.3404.8729>
- C. Lucheroni, & C. Mari, 2018, Risk shaping of optimal electricity portfolios in the stochastic LCOE theory. *Computers and Operations Research*, 96, 374–385. <https://doi.org/10.1016/j.cor.2018.02.011>
- M. Roberts, J. Petrowski, Y. Liu, & J. Bronfenbrenner, 2002, Large Capacity Single Train AP-X™ Hybrid LNG Process. *Gastech Conference Proceedings*.
- J. Saleh, G. Mark, & N. Jordan, 2009, Flexibility: A multi-disciplinary literature review and a research agenda for designing flexible engineering systems. *Journal of Engineering Design*, 20(3), 307–323. <https://doi.org/10.1080/09544820701870813>
- C. Steuer, 2019, Outlook for Competitive LNG Supply. In *Oxford Institute for Energy Studies* (Issue March). <https://doi.org/10.26889/9781784671310>

Modelling and Parameter Fitting of the Dosage of Hydrogen Peroxide in a Photo-Fenton Process

Kourosh Nasr Esfahani, Montserrat Pérez-Moya, Moisès Graells*

*Chemical Engineering Department, Universitat Politècnica de Catalunya, Av. Eduard Maristany, 16, 08019 Barcelona, Spain.
moises.graells@upc.edu*

Abstract

This study addresses the modelling of the photo-Fenton processes, sensitivity analysis, and parameter fitting, focusing on the supply of hydrogen peroxide (H_2O_2) and aiming at the future exploitation in model-based optimization. A flexible H_2O_2 inlet flow (fed-batch) reported in the literature (Audino et al. 2019) was adopted and computational experiments were performed to examine the behaviour of the model and to investigate the opportunities for reducing its complexity. The fitting of the kinetic and stoichiometric parameters was first undertaken using simulated data and focusing on the experimentally available information: H_2O_2 , Total Organic Carbon (TOC), and dissolved oxygen (O_2). Hence, sensitivity analysis was conducted to assess the influence of the parameters of the model and the exclusion of those that do not contribute to explaining the process. Finally, the modified model was shown to fit to experimental data. These results provide valuable insight into the planning of the subsequent experimental phase of this research aimed at providing a reliable model of the flexible dosage of H_2O_2 in photo-Fenton processes.

Keywords: Photo-Fenton, H_2O_2 Dosage, Parameter Estimation, Sensitivity Analysis.

1. Introduction

Toxic and low biodegradable organic pollutants in wastewaters can be hardly removed through conventional biotreatment methods. An alternative to cope with such persistent contaminants is advanced oxidation processes (AOPs), based on the generation of highly oxidant hydroxyl radicals ($\cdot OH$) which could easily degrade recalcitrant compounds. In particular, the photo-Fenton process is a photocatalytic process in which UV irradiation, Fe(II) salt, and H_2O_2 generate such hydroxyl radicals. However, it is well known that an excess of H_2O_2 in the photo Fenton process can lead to unproductive reactions that reduce process performance. Although the use of H_2O_2 is essential, oversupply can be counterproductive, which causes the need to model the flexible dosage of H_2O_2 and the challenge of optimizing its supply (Ortega-Gómez et al., 2012). Several studies have experimentally addressed the improvement of the photo-Fenton process performance by proposing a reasonable hydrogen peroxide supply (Yu et al., 2020, Zazo et al., 2009), but solutions are still incomplete and far from model-based approaches. Hence, this study is limited to the development of a flexible dosage model capable to be fit to experimental data that could be later used for model-based optimization purposes.

2. Methodology

The work addresses the simulation of the operation of the photo-Fenton reactor, its parameter estimation and sensitivity analysis, and the fitting to experimental data.

2.1. Mathematical Modelling

The study adopts the kinetic model (Fig.1) first proposed by Cabrera Reina et al. (2012).

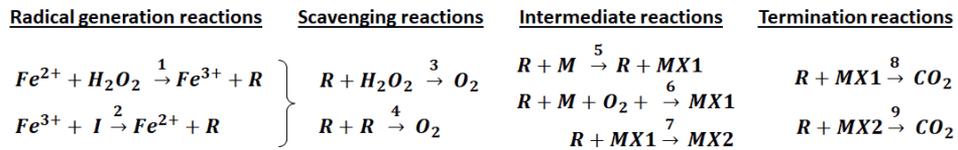


Figure 1. Kinetic model of the photo-Fenton process (Cabrera Reina et al., 2012).

The model takes into account the continuous production of hydroxyl radicals (R) from H_2O_2 , iron and irradiation (I); the consumption of radicals in scavenging reactions; and the reduction of organic matter (M) to CO_2 through some partially oxidized intermediates (MX_1 and MX_2), all of them leading to the lumped parameter measured ($TOC = M + MX_1 + MX_2$). This model gives the set of ODEs listed in Table 1.

Table 1. Equations of the mathematical model of the dosage of the photo-Fenton process.

Reaction type	Reaction rates	Kinetic constants	Eq.
Efficient reactions (Production of radicals)	$r_1 = k_1[Fe^{2+}][H_2O_2]$ $r_2 = k_2[Fe^{3+}][I]$	k_1 k_2	(1) (2)
Inefficient reactions (Decomposition of radicals)	$r_3 = k_3[R][H_2O_2]$ $r_4 = k_4[R][R]$	k_3 k_4	(3) (4)
Degradation of parent compound and intermediates	$r_5 = k_5[M][R][O_2]$ $r_6 = k_6[M][R]$ $r_7 = k_7[MX_1][R]$ $r_8 = k_8[MX_1][R]$ $r_9 = k_9[MX_2][R]$	k_5 k_6 k_7 k_8 k_9	(5) (6) (7) (8) (9)
Material balances:			
	$d[H_2O_2]/dt = (F/V) \cdot ([H_2O_2]_{in} - [H_2O_2]) - r_1 - r_3$		(10)
	$d[O_2]/dt = (F/V)([O_2]_{in} - [O_2]) + g_1r_3 + g_2r_4 - c_1r_5 + K_L a([O_2]^* - [O_2])$		(11)
	$d[TOC]/dt = d[M]/dt + d[MX_1]/dt + d[MX_2]/dt$		(12)
being:			
	$d[M]/dt = (F/V) \cdot ([M]_{in} - [M]) - r_5 - r_6$		(13)
	$d[MX_1]/dt = (F/V) \cdot ([MX_1]_{in} - [MX_1]) + r_5 + r_6 - r_7 - r_8$		(14)
	$d[MX_2]/dt = (F/V) \cdot ([MX_2]_{in} - [MX_2]) + r_7 - r_9$		(15)

The model includes the overall gas-liquid mass transfer coefficient for O_2 ($K_L a$) and three stoichiometric coefficients describing the oxygen balance (c_1 , g_1 , and g_2). Finally, the fed-batch model by Audino et al. (2019) extends the kinetic model by including an inlet flow (F) allowing the flexible supply of hydrogen peroxide ($[H_2O_2]_{in}$).

2.2. Model fitting and parameter estimation

The model has been implemented in Simulink[®]. The photo-Fenton process has been simulated and the estimation of the set of kinetic and stoichiometric parameters has been first studied through the fitting of the model to perfect data obtained by simulation for H_2O_2 , TOC, and O_2 profiles (the common experimentally measured information) using the parameter values reported by Cabrera Reina et al. (2012). This allows testing the

potential of the fitting procedure and provides a quantitative reference, given by the sum of the squared errors to be minimized (eq.1), and the root mean square error (RMSE), the normalized root-mean-square deviation (NRMSE), and the coefficient of variation of the RMSE, CV(RMSE) to quantify the goodness of fit.

$$\min Z = \sum_i([\widehat{TOC}]_i - [TOC]_i)^2 + \sum_j([\widehat{H_2O_2}]_j - [H_2O_2]_j)^2 + \sum_k([\widehat{O_2}]_k - [O_2]_k)^2 \quad (16)$$

$$RMSE = \sqrt{(\sum_{n=1}^N (\hat{y}_n - y_i)^2 / N)} \quad (17)$$

$$NRMSE = RMSE / (\hat{y}_{max} - \hat{y}_{min}) \quad (18)$$

$$CV(RMSE) = RMSE / \hat{y}_{mean} \quad (19)$$

2.3. Sensitivity analysis

Sensitivity analysis has been performed based on variances-based sensitivity analysis of model output using Simulink Design Optimization software. Sensitivity analysis was conducted to assess to which extent variations of the model parameters can affect each measured response (TOC, H₂O₂, and O₂). Hence the model is discussed regarding its hypothesis and chances for reducing complexity. The Rank Partial Correlation method was used to analyze to which extent each model parameter affects the function output. The correlation coefficients vary from -1 to 1, and a zero value means no correlation at all. The correlation coefficient between two sets of variables X and Y, R_{XY}, is given by:

$$R_{XY} = C_{XY} / \sqrt{C_{XX}C_{YY}} \quad \text{being} \quad C_{XY} = cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (20)$$

One approach to sensitivity analysis is local sensitivity analysis, which is derivative-based (one-at-a-time technique). The term local refers to the fact that all derivatives are taken at a single point. However, this approach can be infeasible for complex models, where formulating the cost function (or the partial derivatives) is nontrivial. Also, they do not provide insight into how the interactions between parameters influence the cost function. Another approach to sensitivity analysis is global sensitivity analysis, often implemented using Monte Carlo techniques. This approach uses a representative (global) set of samples to explore the design space.

Mathematically, the sensitivity of the cost function with respect to certain parameters is equal to the partial derivative of the cost function concerning those parameters. Accordingly, A sensitivity function was defined as the partial derivatives of Eq.16 with respect to each parameter. This function implies the matching of output signals (TOC, H₂O₂, and O₂) to measured (perfect) data as the requirement.

2.4. Fitting to experimental data

The experimental data is from the work by Yu et al. (2020) that measured different hydrogen peroxide dosage profiles for Fenton and photo-Fenton processes. The parameter estimation procedure will be tested by fitting the model to some of the assays reported.

3. Result and discussion

3.1. Fitting of the photo Fenton model to No dosage & Dosage cases

First, perfect data was obtained through the simulation of the model using the parameter values reported by Cabrera Reina et al. (2012). Figure 2 shows the profiles obtained for TOC, H₂O₂, and O₂ during a 4h reaction time. Simulations with (b) and without (a) dosage are also shown. The H₂O₂ dosage profile (Fig. 2b) is given by an on-off profile set to be on from 0 to 15 min and from 45 to 60 min.

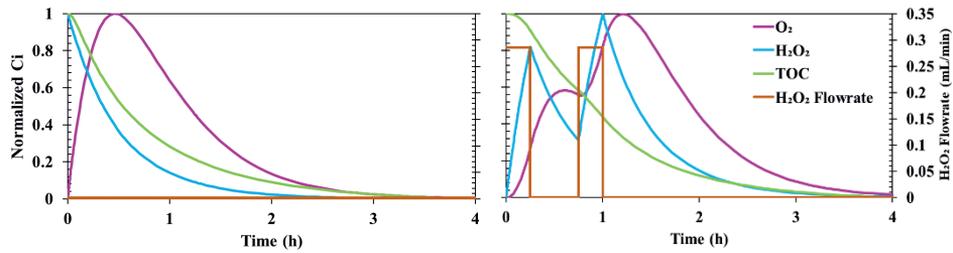


Figure 2. Profile of TOC, H_2O_2 , and O_2 : (left) No dosage, (right) Flexible dosage profile

The fitting of these simulated data using the parameter estimator tool in Simulink[®] was checked setting different sets of initial kinetic parameter values obtained randomly in the range of $\pm 50\%$ of the true values. Results (Fig.3) showed a good agreement between perfect data and predicted data in no dosage mode with the lowest values of C.V.(RMSE) resulted in 0.52%, 0.61%, and 1.04% for TOC, H_2O_2 , and O_2 , respectively. Similarly, the parameter search resulted in good agreement of the fitted data after parameter estimation with the lowest values of C.V.(RMSE) of 0.52%, 1.83%, and 1.01% for TOC, H_2O_2 , and O_2 , respectively. The C.V.(RMSE) results, as represent the simulation errors, below than the detection limits of the measurements corresponding to the capability of the given model to “fit” the datasets. This provides a reliable model of the flexible dosage of H_2O_2 in photo-Fenton processes for further experimental studies.

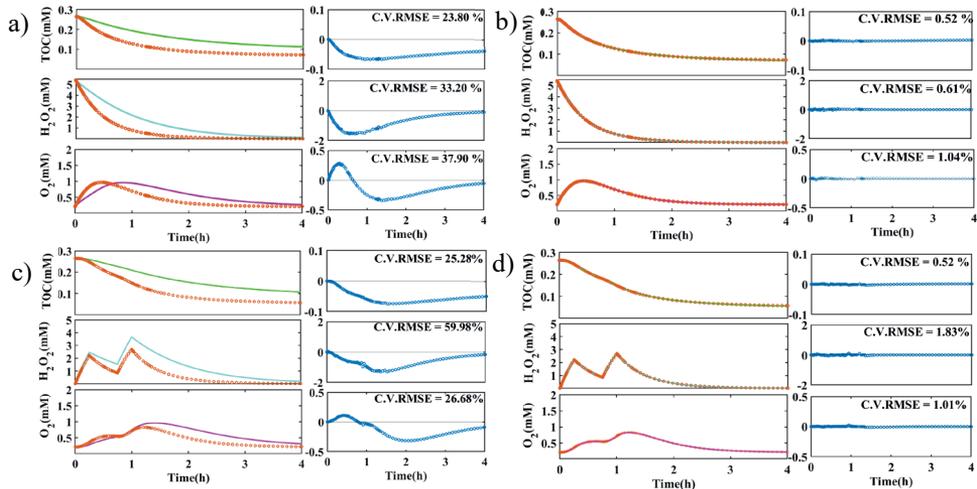


Figure 3. Fitting to simulated data and residuals: No dosage (a, b); Dosage (c, d)

3.2. Sensitivity analysis on the photo Fenton model

Global sensitivity analysis was performed to explore which model parameters (assumed independent) most influence the goodness of fit. A set of 100 parameter samples was generated randomly by varying the value of each parameter for both no dosage and dosage cases. The corresponding correlation coefficients R_{XY} are plotted in Figure 4 in order of influence on the signal matching of the simulated data to the perfect data for TOC, H_2O_2 , and O_2 in the role of the sensitivity function.

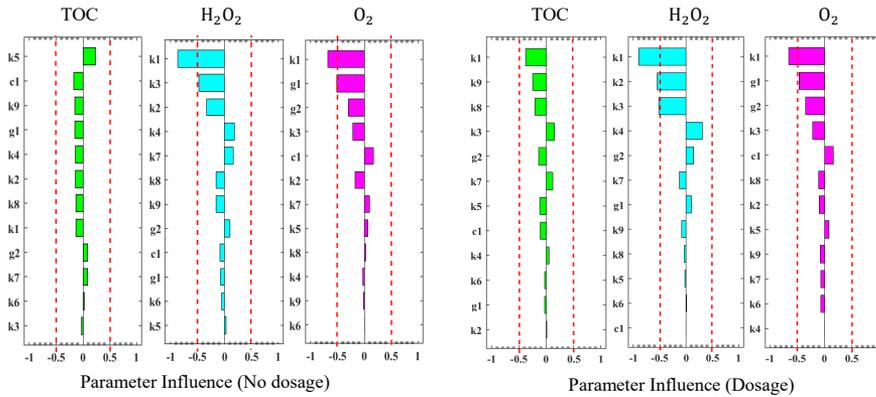


Figure 4. Ranked influence of model parameters on the signal matching with respect to TOC, H_2O_2 , and O_2 : Case without dosage (Left), Case with flexible dosage (Right).

The different ranked series of kinetic parameters were obtained for the measured variables TOC, H_2O_2 , and O_2 considering no dosage and dosage cases. For a better interpretation of the sensitivity analysis plots, a threshold line (red dash line) was defined at $R=\pm 0.5$. k_1 with respect to the criterion $|R|\geq 0.5$ indicated the most significant parameter in the ranking plots to H_2O_2 , and O_2 . In the case of TOC, the order of parameters showed an almost similar magnitude of $|R|$ value lower than 0.5, but again k_1 was in the place of a significant parameter.

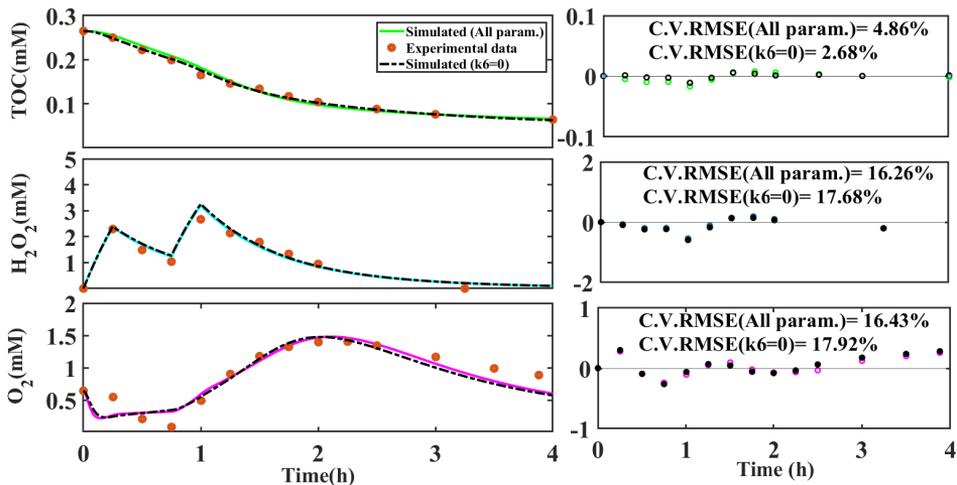


Figure 5. Experimental data and simulated profiles in dosage mode: Considering all parameters for the estimation (solid line); excluding k_6 from model ($k_6 = 0$) (dash line)

Figure 5 shows the model profiles fitting the experimental data with and without considering k_6 . Likewise, similar fitting results were obtained with and without considering k_6 through parameter estimation of experimental data by very close ($\pm 2\%$) C.V.(RMSE) values for TOC, H_2O_2 and O_2 . Whilst this real experimental insight adequately represented the unimportant effect of k_6 on the model fitting, further studies using diverse H_2O_2 dosage profiles would be required to validate the efficient fitting of the model for flexible H_2O_2 profile.

4. Conclusions

This work has developed a kinetic model for the dosage of hydrogen peroxide in photo-Fenton processes to be fit to experimental data obtained from different dosage profiles. The model focuses on the process variables to be easily measured (TOC, H_2O_2 , and O_2). After verifying the model with the adjustment to simulated data a sensitivity analysis of the model fitting has been performed, which provided a valuable insight into the nature of the model. One of the reactions proposed in the literature has been found to have a scarce influence on the fitting of the model, for which it has been excluded. After checking the insignificant effect of this reaction on the fitting, the model has been shown to be able to adjust to a series of experimental data (C.V.(RMSE) of 2.68%, 17.68%, and 17.92% for TOC, H_2O_2 , and O_2 , respectively). Therefore, the model is ready to address further experimental data (dosage profiles) and hence attempt the model-based optimization of the dosage profile.

Acknowledgments

Financial support was received from the Spanish "Ministerio de Ciencia e Innovación" and the European Regional Development Fund, both funding the research Project CEPI (PID2020-116051RB-I00) and also the Secretariat of Universities and Research of the Department of Business and Knowledge of the Generalitat de Catalunya through the program FI SDUR (Grant ref. BDNS 481561) are fully acknowledged.

References

- F. Audino, G. Companyà, M. Pérez-Moya, A. Espuña, M. Graells, 2019, Systematic Optimization Approach for the Efficient Management of the Photo-Fenton Treatment Process, *Science of The Total Environment*, 646, 902–913.
- A. Cabrera Reina, L. Santos-Juanes Jordá, J. L. García Sánchez, J. L. Casas López, J. A. Sánchez Pérez, 2012, Modeling Photo-Fenton Process for Organic Matter Mineralization, Hydrogen Peroxide Consumption and Dissolved Oxygen Evolution, *Applied Catalysis B: Environmental*, 119-120, 132-138.
- J. C. Helton, J. D. Johnson, C. J. Sallaberry, C. B. Storlie, 2006, Survey of sampling-based methods for uncertainty and sensitivity analysis, *Reliability Engineering & System Safety*, 91(10), 1175–1209.
- E. Ortega-Gómez, J. C. Moreno Úbeda, J. D. Álvarez Hervás, J. L. Casas López, L. Santos-Juanes Jordá, J. A. Sánchez Pérez, 2012, Automatic dosage of hydrogen peroxide in solar photo-Fenton plants: Development of a control strategy for efficiency enhancement, *Journal of Hazardous Materials*, 237–238, 223–230.
- X. Yu, A. Somoza-Tornos, M. Graells, M. Pérez-Moya, 2020, An experimental approach to the optimization of the dosage of hydrogen peroxide for Fenton and photo-Fenton processes, *Science of The Total Environment*, 743, 140402.
- J. A. Zazo, J. A. Casas, A. F. Mohedano, J. J. Rodriguez, 2009, Semicontinuous Fenton oxidation of phenol in aqueous solution. A kinetic study, *Water Research*, 43(16), 4063–4069.

Optimising the performance of the condensate stabilisation unit in LNG Processes

Abdul Aziz Shaikh, Ahmed AlNouss, Tareq Al-Ansari*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation,
Doha, Qatar*

**talansari@hbku.edu.qa*

Abstract

Today, Liquefied Petroleum Gas, Natural Gas Liquids Recovery, and Condensate Stabilisation have gained significant interest, given their increased selling prices and market demands. Nevertheless, many operational units do not meet the desired profits and lack proper optimisation initiatives. In this study, a case study from the operating condensate stabilisation unit within the liquefied natural gas process is considered, where a sensitivity analysis of key process operating parameters is applied to investigate process performance. The process is optimised to achieve the highest efficient performance, maximum productivity rate, and minimum energy consumption. Based on the available production data and operating conditions, an existing condensate stabilisation unit is simulated using the Aspen HYSYS V10 simulator, which is based on the Peng-Robinson equation of state. The most significant process variables and constraints that directly affect the performance, and production of condensation stabilisation units are highlighted to demonstrate the connection among process operating conditions and the influence on process objectives. The optimal process operating conditions are determined to achieve a stable column operation. The most effective process variables based on the sensitivity analysis are stabiliser column pressure, stabiliser feed temperature, the outlet temperature of process gas from the wellhead, and stabiliser bottom temperature. After collecting process data from the sensitivity analysis, MATLAB has been used to formulate and solve the multi-objective optimisation problem. Beginning with a natural gas feed flow of 554 MMSCFD, a condensate with 956.9 barrel/day standard liquid volume flow and a reboiler load of 0.86MW have been achieved. After performing the singular optimisation problem in MATLAB, the condensate standard liquid volume flow is increased to 1273 barrels/day, and the reboiler load has decreased to 0.79MW.

Keywords: Condensate stabilisation, optimisation, sensitivity analysis, LNG, simulation.

1. Introduction

Natural gas (NG) is one of the world's most common, promising and affordable fuels, with a wide range of uses. However, in the gaseous state, transportation presents a number of challenges. As a result, the method of transforming NG into intermediate liquid streams has been recommended and used for years to alleviate the difficulties posed by its gaseous composition (Bahmani et al., 2017). Liquefied petroleum gas (LPG) and natural gas liquids (NGLs) are the two main products sourced from all commercial gas refineries' NG stabilisation systems. Natural gas products to be fed into the pipeline for the gas distribution system or to be liquefied must meet the applicable minimum requirement for the pipeline's network or liquefaction unit to work effectively. As a result, natural gas extracted at the wellhead should be refined before it can be liquefied and securely

transported to the long-distance, high-pressure pipelines that transport the fuel to the consumer, as it mainly consists of contamination (non-hydrocarbon gases) and natural gas liquids. Gas condensate is a valuable liquid hydrocarbon mixture that is generated while delivery in raw gas collecting pipelines or naturally from a gas well, can be used as a dilution in oil products or turned to various oil products such as diesel and aviation fuel, amongst others (Uwitonze et al., 2017). To meet the appropriate requirements for transportation and storage, the raw condensate must be reduced in water, salt, and acid content (i.e., H_2S , CO_2 , mercaptans, etc.) (Moradi Kazerooni et al., 2015). The aim of this condensate treatment process is to remove lighter hydrocarbons (methane, ethane, propane, etc.), including acidic components from liquid hydrocarbons in order to make them commercially viable (Zhu et al., 2016). Whereas distillation is the primary method that is widely applied to separate a mixture of components based on their boiling points by boiling the more volatile components out of the mixture preferentially. The degree of separation of a multi-component system depends on the properties of the feed mixture, operating conditions, and other process-imposed restriction. The aim of the unit is to reduce liquid condensate vapor pressure. In order to avoid vapor phase generation while flashes to the storage tank. This method, is structured to distinguish very light hydrocarbon gases, such as methane and ethane, from heavier hydrocarbon components (Adib et al., 2015). Light ends are processed for more commercialisation purposes such as liquefied natural gas (LNG) processing and natural gas liquid extraction while stable condensate is stored (Uwitonze et al., 2017). In the field of condensate stabilisation for natural gas, significant research has been conducted, and numerous papers have been developed. Tavan et al. (2019) investigated the energy and exergy of the condensate stabilisation system and discovered the best location for the water withdraw tray within the condensate stabiliser column. Furthermore, Moghadam et al. (2012) optimised the condensate stabilisation unit's process design and suggested that the design with the minimum operating pressure had the lowest fixed capital cost, and also required the least power due to the high separation performance and low energy demand. In this study, an LNG operating condensate stabilisation is has been selected, where sensitivity analyses are carried out by Aspen HYSYS for the key process parameters to determine their effect on the condensate product and stabiliser energy consumption. After which, the optimal process parameters are determined using the MATLAB software by solving the singular and multi objectives optimisation problems related to the minimisation of stabiliser reboiler duty and the maximisation of condensate flow.

2. Methodology

The sequential structure of methodology adapted in this study is depicted in Figure 1. Firstly, the condensate stabilisation process is developed in the steady-state mode using Aspen HYSYS V10. Subsequently, sensitivity analyses are carried for the key process parameters to allocate their effect on stabiliser reboiler duty and condensate flow. In this, one parameter is subjected to change at a time, while the others are kept constant in the simulation model so that it should not deviate from the actual operating situation. Eventually, these process parameters are optimised using MATLAB to maximise condensate production and minimise the sabiliser energy consumption.

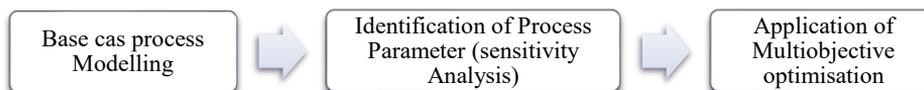


Figure 1: Schematic of methodological steps undertaken in this study.

3. Condensate stabilisation process

The pre-separation units consists of three units, which are a three-phase separation, dehydration unit, and condensate stabilisation unit (Bahmani et al., 2017). Natural gas from different wells is gathered and enters the three-phase separator at nearly 80 bar pressure. In the three-phase separator condensate, water and gases are separated. After the three-phase separation process, the outlet condensate is sent to the condensate stabilisation facility, and water is sent for additional processing and gas as one mainstream sent to the gas dehydration. Raw condensate from the reservoirs has light hydrocarbons, which flash at high temperature and low pressure causing loss of valuable products and environmental pollution (Rahmanian et al., 2016). The separated natural gas condensate requires reduction in acid, water, and salt contents to meet expected standards. Subsequently, it is transformed into a transportable and distributable format that is commercially acceptable (Moradi Kazerooni et al., 2015). While the extracted condensate can be transported without further refining, it must be stabilised often in order to be blended into the crude oil. The aim of the condensate stabilisation process is to lower the vapor pressure of the condensate liquids such that no vapor phase is created when the liquid is flashed to atmospheric storage tanks (Uwitonze et al., 2017). It entails the number of intermediates (C_3 to C_5) and (C_{6+}) components in condensate, while removing light hydrocarbons from liquid hydrocarbon. There are no precise product standards other than meeting process criteria or particular product specifications (Zhu et al., 2016). The primary objective towards stabilising natural gas condensate is to extract as much liquids condensate as possible while maintaining its Reid vapor pressure (RVP) within 10.5 psia. The condensate leaves the three-phase separator V101 where its pressure is dropped by a control valve at 27.5 bar, and the gases are recovered in a flash drum and sent for recompression. The condensate pressure is further dropped by vale V102 at 10 bar. Subsequently, the condensate is sent to the stripper where it is heated, and the light gases are recovered and sent for the recompression and recycle. The stabilised condensate which is obtained from the bottom of stripper is cooled down and sent to the storage tank. The process flow diagram is illustrated in Figure 2.

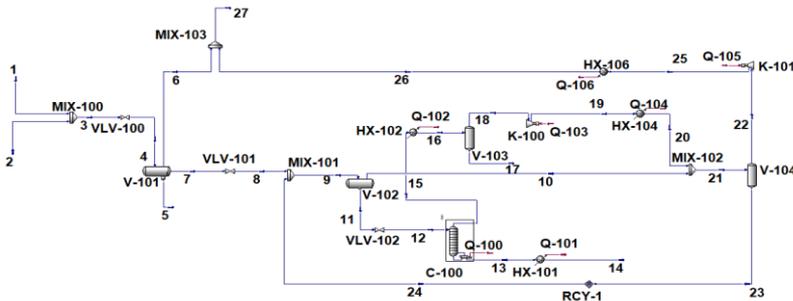


Figure 2: Process Flow Diagram of Condensate stabilisation process.

4. Sensitivity analyses

The most critical process variables and constraints that directly affect the production and the performance of the plant are discussed in the following sections.

4.1 Relationship of feed gas flow rate and stabiliser duty

The feed gas flow rate that is used in this study is 554 MMSCFD, while the feed gas composition is illustrated in Table 1.

Table 1. Feed gas composition.

Name	Mole fraction %
N ₂	0.0500
H ₂ S	0.0150
CO ₂	0.0280
C ₁	0.8093
C ₂	0.0425
C ₃	0.0179
n-C ₄	0.0103
n-C ₅	0.0257
M-Mercaptan	0.0002
Benzene	0.0005
Toluene	0.004
P-Xylene	0.0004
Temperature	27 C
Pressure	80 bar
Mass flow Rate	554 MMSCFD

The sensitivity analysis is carried out while RVP is kept constant. Figure 3a demonstrates the effect of gas flow rate change on the stabiliser duty. As the gas flowrate increases, the duty of the stabiliser increases steadily due to the increase energy requirement to remove light gases from the condensate.

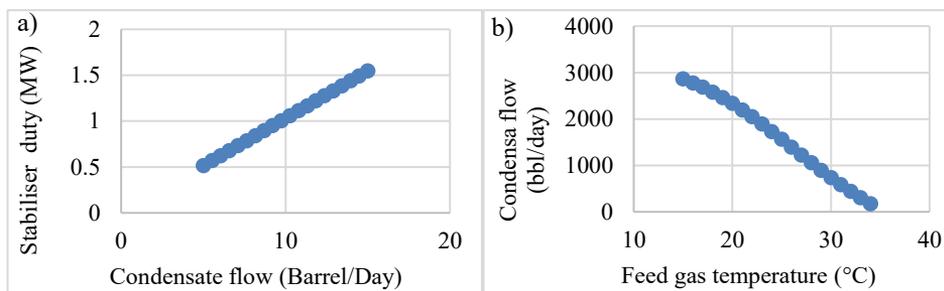


Figure 3: Trends of condensate flow and (a) stabiliser duty and (b) feed gas temperature.

4.2 Relationship of feed gas temperature and condensate flow

The normal temperature of the feed gas is 27 °C. This temperature is varied between 15 and 35 °C while maintain the RVP constant. As illustrated in Figure 3b, condensate production gradually reduced from 2865 bbl/day to 176.1 bbl/day as the temperature of the feed gas increased. It implies that at low condensate temperature at the inlet of the stabiliser plant more condensate is produced. By increasing the temperature, light gases are flashed off which reduces condensate production.

4.3 Relationship of feed gas temperature and energy consumption in the reboiler

As illustrated in Figure 4a, as the temperature of feed gas decreases in the outlet of the wellhead, stabiliser duty increases because condensate productivity rate increase as a result of heavier hydrocarbon condensation. In addition, the total consumed energy increases considerably as a result of increasing reboiler heating duty.

4.4 Relationship of stabiliser pressure and reboiler duty

The condensate stabiliser column pressure is directly controlled by the overhead column vapours. Figure 4b illustrates the relation between condensate stabiliser column pressure and the stabiliser duty. As the pressure of the column increases, the duty of the stabiliser also increases. Hence, the stabiliser feed pressure is directly proportional to the total energy consumed in the stabilisation column.

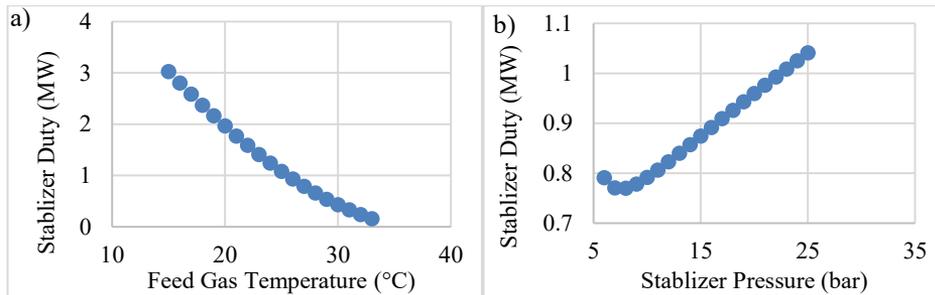


Figure 4. Trends of stabiliser duty and (a) feed gas temperature and (b) stabiliser pressure.

5. Optimisation problem

The optimisation problem is considered for singular and multi objectives, where condensate flow is maximised and energy consumption in the stabiliser is minimised. For optimisation formulation, MATLAB’s genetic algorithm is used as it works for discrete as well as continuous functions. Base case and optimised case results for the singular optimisation problems are given in Table 2 and Table 3, while Figure 5 demonstrates the Pareto front of condensate flow and the stabiliser duty for the multi objective problem.

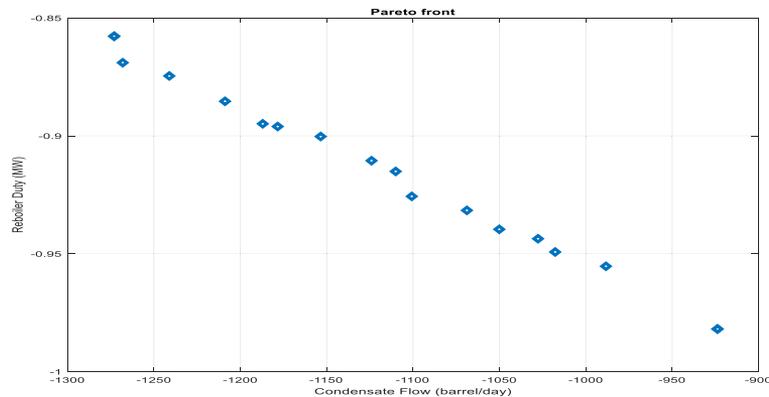


Figure 5. Pareto front of condensate flow and stabiliser Duty.

Table 2: Base case and optimised case results for first objective function.

Optimisation problem \ Cases	Base case	Optimised case
Objective function: Maximisation of condensate flow (bbl/day)	957	1273
Function constrains: RVP (psia)	9.4	10.15
Function Primary variables:		
Stabiliser column pressure bar	10	10
Feed gas Temperature °C	27	25
Stabiliser bottom temperature °C	137.4	133.1

Table 3: Base case and Optimized case results for second objective function.

Optimisation problem \ Cases	Base case	Optimised case
Objective Function: Minimisation of stabilizer column energy (MW)	0.86	0.799
Function constrains: RVP (psai)	9.4	10
Function primary variables:		
Stabiliser column pressure bar	10	10
Feed gas Temperature °C	27	27
Stabiliser bottom temperature °C	137.4	133.1

6. Conclusion

This study has provided a framework for investigating and improving the performance of an existing condensate stabilisation unit recovery plant. The condensate recovery plants in operation require continuous innovation and selection of suitable operating conditions such that it can increase their returns. The outcomes of this study demonstrate the importance of the optimal process operating variables as an essential role in improving condensate production, power consumption, and separation performance of the condensate column. This is demonstrated through the increase in the objective function of condensate production by 33% (316 bbl/d) from the original case with 4.78% increase in pentanes plus mole fraction in the condensate product. Moreover, regarding the second objective for the reduction of reboiler duty in the stabiliser column, the respective energy consumption is reduced by 6.1% (0.061 MW).

References

- H. Adib, A. Sabet, A. Naderifar, M. Adib, and M. Ebrahimzadeh, 2015, Evolving a prediction model based on machine learning approach for Hydrogen Sulfide Removal from Sour Condensate of South Pars Natural Gas Processing plant, *Journal of Natural Gas Science and Engineering*, 27.
- M. Bahmani, J. Shariati, and A.N. Rouzbahani, 2017, Simulation and optimization of an industrial gas condensate stabilization unit to modify LPG and NGL production with minimizing CO2 emission to the environment, *Chinese Journal of Chemical Engineering*, 25, 3, 338-46.
- N.K. Moghadam, M. Samadi, and Z. Hosseini, 2012, Simulation of Gas Condensate Stabilization Unit Aiming at Selecting the Right Technique and Assessing the Optimized Operational Parameters,
- N. Moradi Kazerooni, H. Adib, A. Sabet, M. Adhami, and M. Adib, 2015, Toward an intelligent Approach for H2S Content and Vapor Pressure of Sour Condensate of South Pars Natural Gas Processing Plant, *Journal of Natural Gas Science and Engineering*, 28.
- N. Rahmanian, L.S. Bt Jusoh, M. Homayoonfard, K. Nasrifar, and M. Moshfeghian, 2016, Simulation and optimization of a condensate stabilisation process, *Journal of Natural Gas Science and Engineering*, 32, 453-64.
- Y. Tavan, S.H. Hosseini, and G. Ahmadi, 2019, Energy and exergy analysis of intensified condensate stabilization unit with water draw pan, *Applied Thermal Engineering*, 155, 49-58.
- H. Uwitonze, K.S. Hwang, and I. Lee, 2017, Modelling and improving natural gas condensate process with stripping and heat integration, *Chemical Engineering and Processing: Process Intensification*, 118, 71-77.
- L. Zhu, H. Liu, Z. Zhang, and Y. Pu, 2016, Simulation analysis of stripping fractionation process of gas condensate treatment and practical application, *Journal of Natural Gas Science and Engineering*, 34.

Data-driven modelling of full batch distillation cycles based on recurrent neuronal networks

Gerardo Brand-Rihm^{a,*}, Erik Esche^a and Jens-Uwe Repke^a

^a*Technische Universität Berlin, Process Dynamics and Operations Group, Sekr. KWT 9, Straße des 17. Juni 135, Berlin 10623, Germany*
g.brandrihm@tu-berlin.de

Abstract

Online and offline optimization of dynamic operation trajectories applied to real processes remains a challenging task when models are hard to evaluate, or when their plant-model-mismatch is unsatisfactory. These issues can be tackled by training data-driven models when sufficient measurement data is available. Time dependencies in operation trajectories can be exploited by recurrent neuronal networks (RNN). In this contribution, RNNs with LSTM cells without and with series-parallel (*teacher forcing*) output regressor configurations were trained for their use as surrogate models of an example first-principles batch distillation model simulated from start-up to shutdown. For the offline optimization of a reference batch trajectory, a parallel configuration with predicted output feedback is discussed as a better performing alternative to the classical simulation approach. The prediction accuracies achieved by these configurations is compared and their applicability to optimization of chemical processes is discussed.

Keywords: dynamic data-driven model, recurrent neuronal network, batch distillation

1. Introduction

In the context of process optimization for the reduction of costs and environmental impact in the chemical industry, prediction of the process behaviour under different operating conditions is required. Specifically, online and offline optimization of dynamic operation trajectories applied to real processes remains a challenging task when models are hard to evaluate, or when their plant-model-mismatch is unsatisfactory. These issues can be tackled by training data-driven models when sufficient information through measurement data is available.

Processes with complex dynamics such as discrete changes in the manipulated variables and vanishing phases are no exception, rendering batch distillation a good example. There, simplifications such as neglecting the filling, preheating and shutdown batch phases are often made to obtain simple, first-principles models (Wang et al. (2003)). The flexibility achieved by machine learning models opened new possibilities to avoid such simplifications (Venkatasubramanian (2019)). Furthermore, time dependencies in operation trajectories can be exploited by recurrent neuronal networks (RNN), with long short-term memory (LSTM) cells to tackle long-term dependencies (Hochreiter and Schmidhuber (1997)). Successfully applied to fields such as natural language processing and time series forecasting, these models seem promising for predicting operation trajectories of complex chemical processes (Esche et al. (2022)).

RNNs with LSTM cells may be trained as surrogate models in three different configurations presented in section 2. When the goal of the trained models is to be used for offline optimization (as in

the optimization of a reference batch trajectory to be tracked online), only the manipulated variables are known *a priori*. Typically, a configuration with only control regressors is used for this purpose. Series-parallel configurations trained with *teacher forcing* (Williams and Zipser (1989)), where past measurements are known, have the ability to model the prediction error while advancing in a moving horizon, but cannot be used in offline optimization tasks. A parallel configuration might tackle both restrictions while sacrificing computation time.

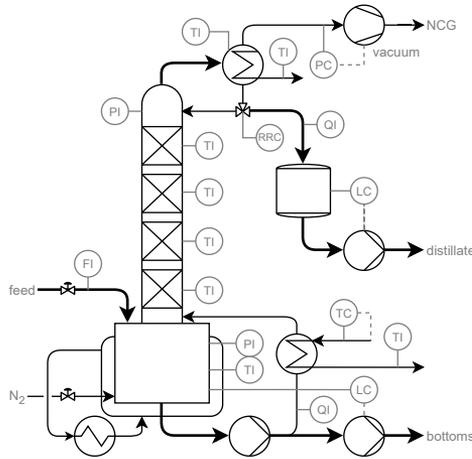


Figure 1: Flowsheet of the case study process: a batch distillation column and its instrumentation

As an example for a chemical process with complex dynamics, the proposed configurations are applied to an example batch distillation column for the separation of water-methanol mixtures shown in Fig. 1, simulated based on a pressure-driven model in Aspen Plus Dynamics (see Brand-Rihm et al. (2021)) with the scope of validation through real plant experiments.

Complexities include vanishing phases and discrete changes in some of the 12 manipulated variables detailed in Tab. 1. As outputs, any variable can be chosen from the flowsheet simulator. Here, $n_y = 26$ outputs are chosen as realistic online measurements, including temperature profiles, pressures, liquid levels, mass flows, and mass fractions. For better visibility, only the 15 main outputs are listed in Tab. 1.

Table 1: Plant controls and main outputs

control	description	output	description
x_{VN_2}	nitrogen valve position	$T_{med,out}^{Cond}$	cooling medium outlet temperature
PC^{SP}	pressure controller setpoint	T_1	temperature at the column top
x_{VF}	feed valve position	T_{10}	temperature at the bottom vessel
RRC^{SP}	reflux ratio controller setpoint	$T_{med,out}^{Reb}$	heating medium outlet temperature
$T_{Reb}^{med,in}$	heating medium inlet temperature	\dot{m}_D	distillate mass flow
LC_D^{MODE}	level controller mode*	\dot{m}_B	bottoms mass flow
LC_D^{SP}	condenser level controller setpoint	\dot{m}_{NCG}	gas flow of non-condensable gases
LC_D^{OP}	condenser level controller output	w_D^{MeOH}	methanol mass fraction in distillate
PC^{MODE}	pressure controller mode*	w_D^W	water mass fraction in distillate
PC^{OP}	pressure controller output	$w_D^{N_2}$	nitrogen mass fraction in distillate
LC_B^{MODE}	sump level controller mode*	p_D	condenser pressure
LC_B^{OP}	sump level controller output	p_B	bottom vessel pressure
		\dot{Q}^{Cond}	condenser cooling duty
		\dot{Q}^{Reb}	reboiler heat duty
		L_B	liquid level in bottom vessel

* (\cdot)^{MODE} $\in \{0 = auto, 1 = manual\}$

The achieved prediction accuracies of the proposed models are discussed in section 3.

2. Methods

2.1. Prediction, simulation, and forecasting

Different definitions exist for prediction and forecasting. Nonetheless, their differences seem to be clear in the field of data science when time plays a decisive role. Since this contribution considers datasets provided by dynamic multiple input-multiple output (MIMO) processes in the form of time series, the specific task of predicting future states is called forecasting.

In the field of system identification, a difference is made between a task focus where regressors are constructed only for the controls \mathbf{x}_k at a discrete timestep k (*simulation*) in contrast to configurations that also include an autoregressive part with output regressors \mathbf{y}_k , called *predictions*.

Following notation can be used for regressor sets, where l and m are the past horizons for controls and outputs, respectively: $x_k^{(l)} = (x_{k-l}, x_{k-l+1}, \dots, x_k)$, $y_k^{(m)} = (y_{k-m}, y_{k-m+1}, \dots, y_k)$.

When the goal is offline optimization based on data-driven models, an alternative to the *simulation* configuration (Fig. 2 a) **Sim** based on a *series-parallel* architecture trained with teacher forcing (Fig. 2 b) **S-P** can be used, which exploits the modelling of prediction errors while maintaining a simulation focus due to its parallel predicted output regressor loop (Fig. 2 c) **Par**.

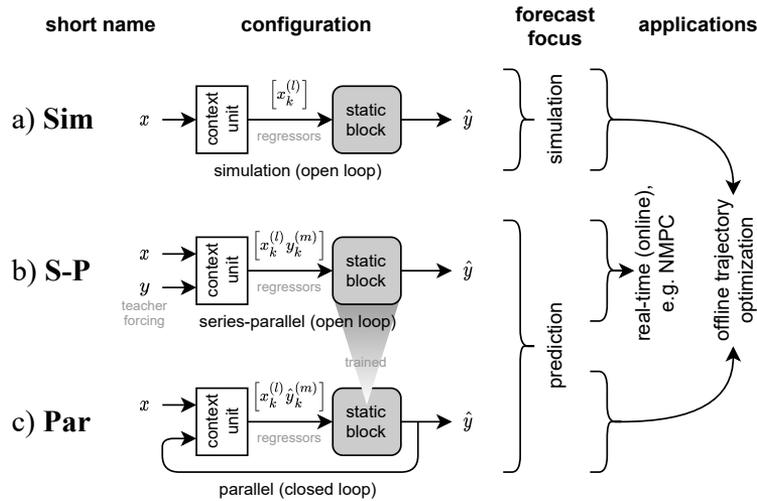


Figure 2: Configurations adopted in this contribution

The **Par** configuration predicts future states using the trained **S-P** model. Since here the regressor set includes predicted outputs instead of real measured outputs, an accumulating prediction deviation is expected. The prediction accuracy of **Par** might be less than in the **S-P** configuration, but only controls are needed.

2.2. Sampling for simulated dataset generation

A simulated dataset consisting of batch runs sampled over an operation recipe to enhance the feasibility ratio of the simulations was generated to train and test different data-driven models according to Brand-Rihm et al. (2021). Here, $n_s = 645$ of 1000 samples converged. A train/test ratio of 3/4 is chosen, leaving $n_s^{est} = 162$ samples for testing. The generated dataset consists of an array with size (samples n_s , timesteps n_t , channels $n_x + n_y$) = (645, 401, 12 + 26).

3. Results and discussion

The prediction accuracies can be analysed comparing the mean squared errors (MSE) between min-max scaled testing (y_{stc}^{test}) and predicted (\hat{y}_{stc}) data for each sample s separately (MSE_s) and for each model (MSE), according to:

$$MSE_s = (n_t \cdot n_y)^{-1} \sum_{t=1}^{n_t} \sum_{c=1}^{n_y} (y_{stc}^{test} - \hat{y}_{stc})^2 \quad MSE = (n_s^{test})^{-1} \sum_{s=1}^{n_s^{test}} MSE_s$$

Since only configurations **Sim** and **Par** support offline optimization, their predictions can be compared qualitatively. Fig. 3 shows the worst test trajectories (highest MSE_s) predicted by **Sim** (top) and **Par** (bottom). It is shown that even in the worst **Par** prediction the test trajectories show correct trends. On the contrary, the worst **Sim** prediction misses the trends while the parallel configuration performs to a satisfactory degree.

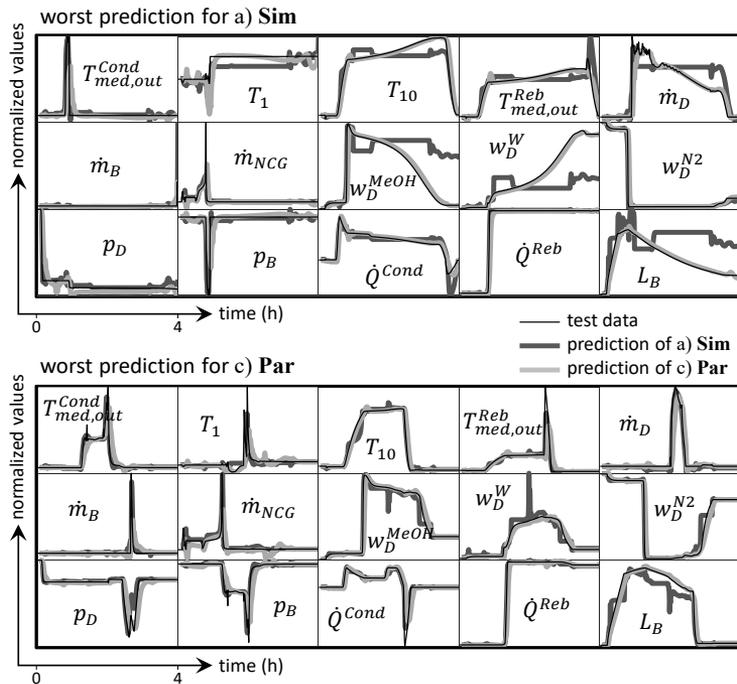


Figure 3: Worst predicted samples (in terms of MSE) for configurations a) **Sim** and c) **Par**

Table 2: Statistics of prediction accuracies for the three configurations

configuration	$MSE \cdot 10^3$				elapsed times (s)	
	mean	P_{25}	P_{50}	P_{75}	training	prediction
a) Sim	4.06	1.91	3.01	5.12	$1.47 \cdot 10^3$	5.14
b) S-P	1.66	1.43	1.67	1.86	$1.54 \cdot 10^3$	6.02
c) Par	3.46	2.70	3.64	4.04	$1.54 \cdot 10^3$	$2.26 \cdot 10^3$

The prediction accuracies in terms of overall MSE among all samples, as well as elapsed training and prediction times are shown in Tab. 2. The prediction error distributions with their mean, median (P_{50}), 25% and 75% percentiles (P_{25} and P_{75}) are presented in Fig 4.

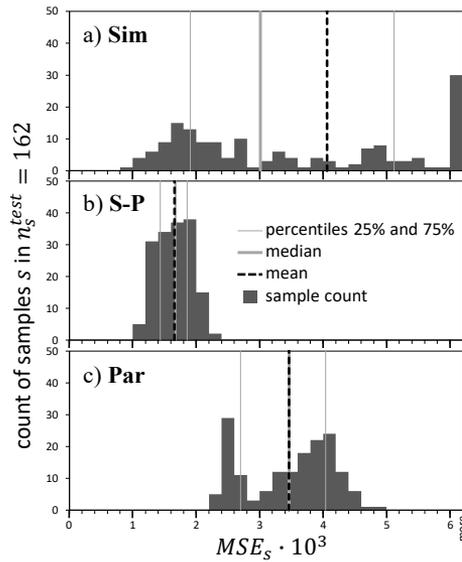


Figure 4: Sample prediction error distributions of the presented methods

It is shown that for the performed sampling and its test data, among the configurations capable of predicting operation trajectories based solely on given controls (**Sim** and **Par**), the parallel architecture outperforms qualitatively and quantitatively (in terms of prediction error) the simulation architecture, although the computation time of the prediction is larger by three orders of magnitude. This might be enhanced by optimizing the code for the feedback loop. The percentile statistics show a lower median (P_{50}) for **Sim**, but having a stronger variance with many test samples with insufficient prediction accuracy. Therefore, the **Par** configuration might be more robust.

Absolute prediction errors for the samples with the best prediction accuracies are shown in Fig. 5. Although no clear advantage is visible for the **Par** over the **Sim** configuration during the first half of the time period, in almost all cases **Par** outperforms **Sim** in the second half.

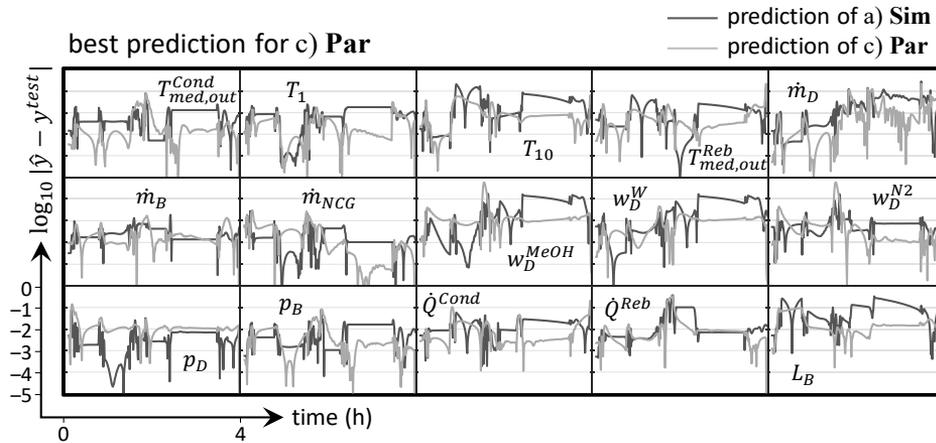


Figure 5: Errors of best predicted samples (in terms of MSE) for configuration c) **Par**

In the context of batch distillation, the second half of the predicted time period is often described by distillate depletion, inertization and shutdown. The conditions in the column states that must be met to proceed with inertization and shutdown strongly depend on the actions taken from the beginning of the operation. These actions in the manipulated variables are bound to past measurements according to an operation recipe. The **Par** configuration can be trained on those conditions using past states, explaining its advantages over **Sim**.

4. Conclusions

For the presented use case, RNN configurations with and without external output feedback loops were tested. A proposed parallel configuration based on a model trained in series-parallel form (*teacher forcing*) showed better overall prediction accuracy and lower variance in the mean squared errors than a simulation configuration with only control regressors. This is due to the modelling of the prediction error considering past testing timesteps. The qualitative behaviour of predicted trends supports the choice of this architecture, although the computation time of the prediction is three orders of magnitude higher. Training of the presented RNNs on large simulated datasets provides data-driven models for trajectory optimization in chemical processes with complex dynamics. The trained models will be used for the optimization of batch distillation cycles in a real batch distillation column.

5. Acknowledgements

This research has been supported by the project "KI-Inkubator-Labore in der Prozessindustrie - KEEN", funded by the Bundesministerium für Wirtschaft und Energie (BMWi) under grant number 01MK20014T.

References

- G. Brand-Rihm, E. Esche, J.-U. Repke, 2021. Sampling space reduction for data-driven modelling of batch distillation - introducing expert process knowledge through operation recipes. In: M. Türkay, R. Gani (Eds.), 31st European Symposium on Computer Aided Process Engineering. Vol. 50 of Computer Aided Chemical Engineering. Elsevier, pp. 611–616.
URL <https://www.sciencedirect.com/science/article/pii/B9780323885065500978>
- E. Esche, J. Weigert, G. Brand Rihm, J. Göbel, J.-U. Repke, 2022. Architectures for neural networks as surrogates for dynamic systems in chemical engineering. *Chemical Engineering Research and Design* 177, 184–199.
URL <https://www.sciencedirect.com/science/article/pii/S0263876221004524>
- S. Hochreiter, J. Schmidhuber, 1997. Long Short-Term Memory. *Neural Computation* 9 (8), 1735–1780.
URL <https://doi.org/10.1162/neco.1997.9.8.1735>
- V. Venkatasubramanian, 2019. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal* 65 (2), 466–478.
URL <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.16489>
- L. Wang, P. Li, G. Wozny, S. Wang, 2003. A startup model for simulation of batch distillation starting from a cold state. *Computers & Chemical Engineering* 27 (10), 1485–1497.
URL <https://www.sciencedirect.com/science/article/pii/S0098135403000942>
- R. J. Williams, D. Zipser, 1989. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation* 1 (2), 270–280.
URL <https://doi.org/10.1162/neco.1989.1.2.270>

Stackelberg Game Design and Operation of a Non-Cooperative Bi-Level H_2 Supply Chain Under Cournot Equilibrium

Jose M. FLORES-PEREZ^a, Catherine AZZARO-PANTEL^a, Antonin PONSICH^b
and Alberto A. AGUILAR LASSERRE^c

^a*Laboratoire de Génie Chimique, Université Toulouse, CNRS, INPT, UPS, Toulouse France.*

^b*IOC, Universitat Politècnica de Catalunya, Barcelona, Spain.*

^c*Instituto Tecnológico de Orizaba, Orizaba, Veracruz, Mexico.*

jose.floresperez@inp-toulouse.fr

Abstract

In the pathway toward decarbonization, hydrogen is presented as a great option as it can be produced from a wide variety of feedstock using multiple production, storage and transportation options for diverse applications. Although H_2 could provide flexibility and sector coupling to energy systems, the design and management of the hydrogen supply chain (HSC) is also identified as a challenging task. Until now, most of the HSC designs are treated as problems with single or multiple objectives without any hierarchical conflict. This paper proposes a mixed-integer bi-level programming (MIBLP) approach as a mathematical model of the Stackelberg game. The solution strategy considers the MIBLP as a multi-parametric problem: if the feasible set of the lower level optimization problem (LLP) of the bi-level programming problem (BLPP) is parametric in terms of the optimization variables of the upper level problem (ULP), each level can be solved with a different approach. Consequently, to handle continuous and discrete variables at both levels, we propose a hybrid method involving Differential Evolution (DE) for the ULP and an Integer Linear Programming Solver (ILPS) for the LLP. The developed hybrid evolutionary-deterministic strategy evaluates the performance of two HSC study cases combining Steam Methane Reforming and Electrolysis processes for H_2 production: a classical Stackelberg game design vs. a Stackelberg one leader - multi followers under Cournot competition. In both scenarios, the ULP objective is to minimize the distribution cost while the LLP objective tries to minimize the production cost of a given producer. The experimental results obtained show that the solution method is efficient and promising for dealing with one-leader multi-objective / multi-follower single objective optimization cases.

Keywords: bi-level programming, hydrogen supply chain, Stackelberg Games, Cournot Equilibrium, MIBLPP.

1. Introduction

Hydrogen (H_2) can be obtained from a wide variety of renewable and non-renewable sources with many different production, storage and transport modes, so that the development of an optimal Hydrogen Supply Chain (HSC) arises as a complex task. With the objective of a low or zero carbon energy future, and with the faster growing research interest in this topic, as proposed in Lindorfer et al. (2019); Borschette (2019), hydrogen can offer new pathways for the energy transition. Until now, most of the HSC designs are treated as single level programming problems (SLPP) with

single or multiple objectives without any hierarchical conflict. SLPPs can be seen as centralized decision-making systems, where the highest level in the hierarchy has enough power to dictate the decisions and have them executed by the lower levels without any chance of opinion. However, a more realistic formulation would recognize the role of the lower levels as part of the decision process. Therefore, the optimization of non-cooperative supply chains using bi-level programming models has received increased scientific attention due to the natural representation of this kind of hierarchical situations.

Bi-level programming problems (BLPP) constitute a class of nested optimization problems where an optimization problem contains another optimization problem as part of the constraint definition. Main approaches used to tackle BLPP are commonly divided into two classes: deterministic / classical approaches and evolutionary approaches, for further detail refer to Sinha *et al.* (2018). Mostly of the classical optimization literature is related with simple, single-objective problems with linear constraints when it is possible to replace the lower level optimization problem with its Karush-Kuhn-Tucker (KKT) conditions. Existing models usually can only handle linear or quadratic problems in the follower level using classical approaches, like the transformation of the bi-level problem into a single level problem using the KKT conditions, e.g. Dempe and Franke (2019); Li *et al.* (2016); Yue and You (2014). On the evolutionary approach branch, studies on computational techniques like evolutionary algorithms had proved the efficiency solving from small to big size problem instances, e.g. Sinha *et al.* (2017); Bylling (2018).

This work attempts to solve a non-cooperative HSC in the presence of multiple followers (producers) that compete in a shared market. In a previous work, we proposed a more realistic HSC formulation that recognize the role and influence of each stakeholder as part of the whole decision process, based on Flores-Perez *et al.* (2020) and represented as a non-cooperative supply chain in the form of a single leader - single follower BLPP.

In Game Theory, the Cournot Competition Game is an economic model to describe how n rival companies compete when they offer the same product in the same market and collusion is not possible or is forbidden. Ledvina and Sircar (2012) define that the Cournot Oligopoly Model occurs when more than two firms (players) choose their individual production level q_i and the market price is determined by the total sum of outputs while each firm wants to maximize their profit. Recent studies, e.g. Lu *et al.* (2021); Zhang *et al.* (2020), demonstrate the usefulness of this approach applied to energy markets, where the authors use a Cournot Game to model the interactions and competition among strategic generators.

In this paper, we formulate the HSC as a BLPP where the leader is a monopolistic distribution company and the Cournot Oligopoly Game model is adopted to formulate production competition between different producers (followers). The rest of this paper is organized as follows. Section 2 defines the HSC model represented as a Mixed Integer Bi-level Programming Problem (MIBLPP), including the shared market competition. Section 3 explains the functionality of the hybrid algorithm to solve MIBLPP, the Cournot model definition and the expected interaction between levels. Section 4 and 5 highlight the results obtained and conclusions, respectively.

2. Problem Definition

Due to the complex and non-general structure of the HSC, and to test the viability of the developed algorithm, we propose a modified numerical example of a Production-Distribution problem adapted from literature, see Avraamidou and Pistikopoulos (2019). There are two main ways to produce hydrogen. Today, around 99% of hydrogen comes from fossil fuels, primarily by reforming natural gas into hydrogen and CO_2 . This hydrogen is often referred to as "grey hydrogen" if greenhouse gases are emitted into the atmosphere. The "green" alternative refers to hydrogen gas that is generated via electrolysis powered by renewable energy such as hydro, solar or wind energy.

The resulting HSC planning activity can be modeled as MIBLPP with continuous and discrete variables at both levels to represent production and distribution activities of a HSC, with the par-

Sets and index		Parameters			
P, p	Set of H_2 producing plants, plant index		$p = 1$	$p = 2$	
S, s	Set of storage units, unit index		13,500	10,000	
R, r	Set of refueling stations, station index		$r = 1$	$r = 2$	$r = 3$
Constants					
$PCap_p$	Max production capacity of plant p (kg)				
$TCap^{prod}$	Connection capacity at production level (kg)				
$TCap^{dist}$	Connection capacity at distribution level(kg)				
CC_{sr}^{trans}	Transmission cost from storage s to station r (\$/kg)	$s = 1$	1	1	1
CC_{ps}^{conv}	Conversion cost at plant p for storage s (\$/kg)	$s = 2$	2	2	2
UTC_{ps}	Transp. cost from plant p to storage s (\$/kg)		$s = 1$	$s = 2$	
UPC_{ps}	Prod. cost of H_2 at plant p for storage s (\$/kg)	$p = 1$	0.7	0.7	
D_r	Demand for refueling station r (kg)	$p = 2$	1	1	
c_p	Marginal production cost for producer p (\$/kg)		$s = 1$	$s = 2$	
a, b_p	Coefficients for the inverse demand function	$p = 1$	2.3	2.3	
		$p = 2$	1.6	1.6	
			$s = 1$	$s = 2$	
		$p = 1$	2.3	2.3	
		$p = 2$	1.6	1.6	
			$r = 1$	$r = 2$	$r = 3$
			5,500	6,500	1,500
Decision variables					
Q_{sr}	H_2 sent from storage s to station r (kg)				
y_{sr}	Existence of route connecting storage s to station r		$p = 1$	$p = 2$	
Q_{ps}	H_2 produced in plant p for storage s (kg)		$1.21e^{-4}$	$3.2e^{-4}$	
x_{ps}	Existence of route connecting plant p to storage s	$p = 1$			
q_t^*	Optimal quantity produced by producer t (kg)	$p = 2$	2.3	1.6	
Π_t^*	Total profit for producer t (\$)	$r = 1$			
Q^*	Optimal quantity produced (kg)	$r = 2$			
P^*	Market price (\$/kg)	$r = 3$			

Table 1: Sets, index and parameters for HSC case study

ticularity of multiple followers competing for the production market. For the case study considered here (see Figure 1), the HSC for mobility application is operated by two different companies, production and distribution, described as follows:

The upper level problem (ULP) represents the distribution company, the objective function consists in minimizing the costs that arise of getting the hydrogen from plants and then to distribute it to the refueling stations, including the cost associated to hydrogen storage in the dedicated facilities. The lower level problem (LLP) represents the decisions of the multiple production companies, which have the interest in minimizing the total operating cost associated to production activities while maximizing their individual profit via a Cournot model, more detail in Osborne (2002). The fixed cost related to the decision of establishing a connection between a production plant and storage units corresponds to the cost of compression/liquefaction.

The resulting mathematical model is presented by equations (1-17) using the nomenclature and parameters of Table 1. The parameters associated with production, storage, distribution

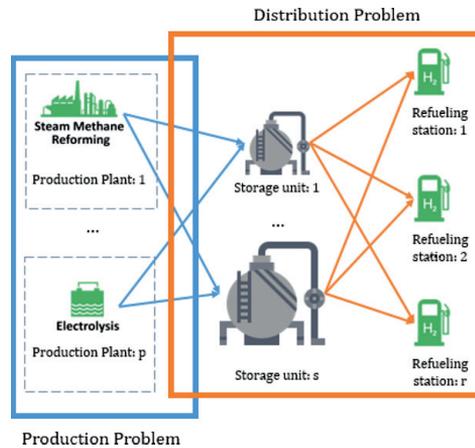


Figure 1: Non-cooperative Hydrogen Supply Chain.

and conversion/transmission for the H_2 supply chain are taken from Hydrogen Council and McKinsey & Company (2021).

3. Hybrid MIBLPP Algorithm. The multi-follower case

The solution algorithm consists of two sections: the Upper Level (UL) problem and the Lower Level (LL) problem. The UL problem solved by a Differential Algorithm approach describes the general properties of the leader problem, including the Upper Level Variables (ULV), each iteration includes a call to the LL problem to optimize the LL objective function, more detail of the algorithm can be found in Flores-Perez et al. (2020). In other words, for each UL population member with no constraint violation, the © IBM CPLEX Solver evaluates the LL problem taking into account the production levels proposed by the general Cournot model. The result of the evaluation returns to the UL as fixed values (parameters) so the algorithm can evaluate the UL objective function. Each UL iteration ends up with the Differential Evolution engine executed. It is possible to define two tolerances for the stopping criteria: the number of generations or a predefined feasible solution; a tolerance is a threshold which if crossed stops the iterations of a solver.

$$\begin{aligned} \min_{y_{sr}, Q_{ps}} \sum_{s \in S} \sum_{r \in R} CC_{sr}^{trans} y_{sr} + \sum_{s \in S} \sum_{r \in R} UTC_{ps} Q_{ps} \quad (1) \\ \sum_{s \in S} Q_{sr} \geq D_r, \quad \forall r \in R \quad (2) \\ Q_{sr} \leq TC_{ap}^{dist} y_{sr}, \quad \forall s \in S, \forall r \in R \quad (3) \\ Q_{sr} \geq 0, \quad \forall s \in S, \forall r \in R \quad (4) \\ C = \sum_{p \in P} c_p \quad (5) \\ q_p = \frac{1}{b_p} \left(\frac{a+C}{n+1} - c_p \right) \quad (6) \\ Q^* = \frac{1}{b_p} \left(\frac{an-C}{n+1} \right) \quad (7) \\ P^* = \frac{a+C}{n+1} \quad (8) \end{aligned}$$

$$\begin{aligned} \Pi_p^* = \frac{1}{b_p} \left(\frac{a+C}{n+1} - c_p \right)^2 \quad (9) \\ \min_{x_{ps}, Q_{ps}} \sum_{p \in P} \sum_{s \in S} UPC_{ps} Q_{ps} + CC_{ps}^{conv} x_{ps} \quad (10) \\ s.t. \quad \sum_{s \in S} Q_{ps} \leq PCap_p, \quad \forall p \in P \quad (11) \\ \sum_{p \in P} Q_{ps} \geq \sum_{r \in R} Q_{sr}, \quad \forall s \in S \quad (12) \\ \sum_{s \in S} Q_{ps} = q_p, \quad \forall p \in P \quad (13) \\ Q_{ps} \leq TC_{ap}^{prod} x_{ps}, \quad \forall p \in P, \forall s \in S \quad (14) \\ Q_{ps} \geq 0, \quad \forall p \in P, \forall s \in S \quad (15) \\ Q_{sr} \in \mathbb{R}, y_{sr} \in 0, 1 \quad (16) \\ Q_{ps} \in \mathbb{R}, x_{ps} \in 0, 1 \quad (17) \end{aligned}$$

The aforementioned Cournot model is a general static linear oligopoly with multiple producers with asymmetric production cost. Market price (Eq. 8) is the same for everyone in the market. We assume there are n groups of H_2 producers grouped by production technology p ($1, 2, \dots, n$), each one with profit function (Eq. 9) where the total production (Eq. 7) is obtained from the optimal individual production quantities (Eq. 6). Equations (5-9) are obtained via the diagonalization solution method, for details see Ioan and Ioan (2015). For this specific example, arbitrary values are assigned to a , that represents the intercept for the price equals to 0, and b , the slope of demand curve, in order to obtain a similar demand as defined in the problem definition (13,500 kg). The total marginal cost C correspond to the summation of unit production cost c_p established for each hydrogen producer as parameters (see Table 1).

4. Results

This section compares the results obtained via our classical MIBLPP approach and the proposed MIBLPP+Cournot Algorithm. Identical algorithmic parameters were chosen for both examples, e.g. a population of 400 individuals and 2,500 generations.

In the classical version, the algorithm converges to a unique solution (see Figure 2) where the leveled cost of H_2 (do not confuse with market price from Cournot Model) is 4.21 USD/kg, there is no overproduction but also there is no shared market since the production is managed entirely by a unique company (monopoly) (see Table 2). Not surprisingly, the production company decides to produce the entire demand in the SMR plant since it implies the lowest connection cost (1 USD/kg) and thus a lower total cost.

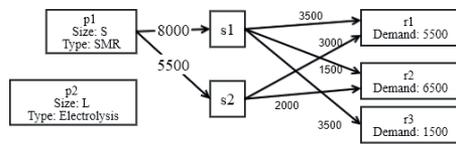


Figure 2: Classic MIBLPP Solution

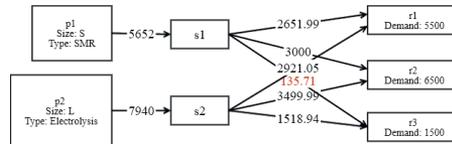


Figure 3: MIBLPP + Cournot Solution

In the case of MIBLPP+Cournot, two different companies in competition, each one with different production technologies and therefore different production costs, must supply the total production of H_2 . In this case, there is a wide variety of feasible solutions, the solution with the lowest total cost of H_2 (Solution 1) is 4.091 USD/kg but with a large overproduction of 2,532 kg of H_2 . As shown in Table 2, a better feasible solution is obtained corresponding to a total cost of 4.30 USD/kg with a low surplus of 135.71 kg of H_2 (Solution 2). The higher total cost, compared with the classical approach, is the result of producing a large amount of H_2 using a large size production plant driven by Electrolysis even if that represents a higher connection cost (2 USD/kg), see Figure 3.

The result obtained using the hybrid algorithm with producers under competition demonstrates the impact of production levels for each producer in order to maximize the individual profit. The results for the Cournot Model are: $q_1 = 8,264.46$ kg, $q_2 = 5,312.50$ kg, $Q = 13,576$ kg, $P = 3.3$ USD, $\pi_1 = 8,264.46$ USD, $\pi_2 = 9,031.25$ USD*

It must be emphasized that even if the cost data used exhibit the order of magnitude observed in the hydrogen roadmaps, the values obtained do not represent a real profit value.

Finally, Table 3 presents both optimal solutions found for comparison purpose between the classical MILPP and the MILPP + Cournot approach.

Approach	Production cost (USD/kg)	Distribution cost (USD/kg)	Total cost (USD/kg)	H_2 produced (kg)	H_2 distributed (kg)	Overproduction (kg)	UL objective function (USD)	LL objective function (USD)
Classic MIBLPP	2.0148	2.1926	4.2074	13,500	13,500.00	0.00	29,600.00	27,200.20
MIBLPP + Cournot (Sol. 1)	1.8911	2.1996	4.0907	13,592	16,123.99	2,531.99	25,703.80	35,466.06
MIBLPP + Cournot (Sol. 2)	1.8911	2.4086	4.2997	13,592	13,727.71	135.71	25,703.80	33,064.12

Table 2: General comparison of results

5. Conclusions

This paper proposes a Cournot Oligopoly Model in the new hydrogen energy market to formulate the competition among non-cooperative suppliers using SMR or electrolysis for hydrogen production. A hybrid strategy to solve mixed-integer BLPP models including multiple players in

Approach	Q_{ps}				Y_{ps}				Q_{sr}			x_{sr}								
	p1		p2		p1		p2		s1	Q_{sr}		s2	r3	s1	x_{sr}		s2	r3		
	s1	s2	s1	s2	s1	s2	s1	s2	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3
Classic MIBLPP	8,000	5,500	0.00	0.00	1	1	0	0	3,500	3,000	1,500	2,000	3,500	1,186.61	1	1	1	1	1	0
MIBLPP+Cournot (Sol. 2)	5,652	0.00	0.00	7,940	1	0	0	1	2,652	3,000	135.71	2,921.06	3,500	1,518.94	1	1	1	1	1	1

Table 3: Optimal solution comparison

competition for the lower level problem is developed for the optimal management of a hydrogen supply chain. This approach has the advantage to couple engineering and economic approaches. The resulting algorithm was tested with a small HSC instance. Further work will now be devoted to a robustness analysis with larger size instances and with the consideration of price elasticity of demand.

6. Acknowledgements

Funding acknowledgement. The National Council of Science and Technology of Mexico (Consejo Nacional de Ciencia y Tecnología, CONACYT), bourse reference: 2018-000003-01EXTF-00046, supported this work.

References

- S. Avraamidou, E. N. Pistikopoulos, 2019. B-POP Bi-level parametric optimization toolbox. *Computers and Chemical Engineering* 122, 193–202.
- A. Borschette, 2019. Green hydrogen opportunities in selected industrial processes Workshop summary report (June).
- H. C. Bylling, 2018. Bilevel Optimization with Applications in Energy.
- S. Dempe, S. Franke, 2019. Solution of bilevel optimization problems using the KKT approach. *Optimization* 68 (8), 1471–1489.
- J. M. Flores-Perez, C. Azzaro-Pantel, A. Ponsich, A. A. Aguilar Lasserre, 2020. A hybrid strategy for mixed integer bi-level optimization applied to hydrogen energy supply chain management. *Modelling and Simulation 2020 - The European Simulation and Modelling Conference, ESM 2020*, 277–281.
- Hydrogen Council, McKinsey & Company, 2021. Hydrogen Insights. A perspective on hydrogen investment, market development and cost competitiveness. Tech. Rep. February.
- G. Ioan, C. Ioan, 2015. The Cournot Equilibrium for n Firms. *EIRP Proceedings* 10 (0).
- A. Ledvina, R. Sircar, 2012. Oligopoly games under asymmetric costs and an application to energy production. *Mathematics and Financial Economics* 6 (4), 261–293.
- H. Li, Q. Zhang, Q. Chen, L. Zhang, Y. C. Jiao, 2016. Multiobjective differential evolution algorithm based on decomposition for a type of multiobjective bilevel programming problems. *Knowledge-Based Systems* 107, 271–288.
- J. Lindorfer, G. Reiter, R. Tichler, H. Steinm, 2019. Hydrogen fuel, fuel cells, and methane.
- W. Lu, Q. Liu, M. Liu, K. Yuan, 2021. Cournot-model-based coordinated secondary voltage control considering reactive power coupling among multiple zones. *International Journal of Electrical Power and Energy Systems* 125 (November 2019), 106470.
- M. J. Osborne, 2002. Nash equilibrium: illustrations. *An Introduction to Game Theory* (1838), 53–95.
- A. Sinha, P. Malo, K. Deb, 2017. Evolutionary algorithm for bilevel optimization using approximations of the lower level optimal solution mapping. *European Journal of Operational Research* 257 (2), 395–411.
- A. Sinha, P. Malo, K. Deb, 2018. A Review on Bilevel Optimization: From Classical to Evolutionary Approaches and Applications. *IEEE Transactions on Evolutionary Computation* 22 (2), 276–295.
- D. Yue, F. You, 2014. Game-theoretic modeling and optimization of multi-echelon supply chain design and operation under Stackelberg game and market equilibrium. *Computers and Chemical Engineering* 71, 347–361.
- Y. Zhang, C. Gu, X. Yan, F. Li, oct 2020. Cournot oligopoly game-based local energy trading considering renewable energy uncertainty costs. *Renewable Energy* 159, 1117–1127.

Modelling of Heat-Driven Water Treatment Systems: Multi-Effect Distillation (MED) model in Modelica

Miguel Castro Oliveira ^{a,b*}, Pedro Coelho ^c, Muriel Iten ^a, Henrique A. Matos ^b

^a *Low Carbon and Resource Efficiency, R&Di, Instituto de Soldadura e Qualidade, 4415 491 Grijó, Portugal*

^b *Centro Recursos Naturais e Ambiente (CERENA), Chemical Engineering Department, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*

^c *Instituto de Engenharia Mecânica (IDMEC), Mechanical Engineering Department, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*
dmoliveira@isq.pt

Abstract

The improvement of water and energy use in the industrial sector is an important concern for decarbonisation of industries, as well as to improve the overall techno-economic performance of single plants. The improvement of water efficiency, in particular, is a potential approach in the context of its interdependencies to energy use, comprehended by the *water-energy nexus*. A practical application corresponds to the use of heat-driven water treatment units such as multi-effect distillation (MED). Such technology is commonly used for the desalination of seawater, as well as industrial waste saline water streams. This work presents a customised model for an industrial MED unit, namely its assembling using the object-oriented Modelica language, and also a brief economic assessment for the viability of such unit for industrial application. The model was validated with literature data by comparing the theoretical temperature profiles with the simulation results. A payback time of 7.1 years was estimated by excluding negligible parcels of operating expenses (OPEX) for the MED project, which may be considered reasonable in terms of economic feasibility.

Keywords: water and energy integration; Multi-effect distillation; waste heat recovery; water efficiency; Modelica.

1. Introduction

Water and energy integration involve a set of methods, measures and practices to improve their usage in industry (Alnouri et al., 2014), corresponding to two of the most extensively used resources in an industrial plant. The interdependencies of water and energy resources, namely the understanding of the improvement of water use through the understanding of the use of energy, are studied in the scope of the *water-energy nexus* (Oliveira et al., 2019). Moreover, practices of *Combined Water and Energy Integration* corresponds to the simultaneous application of water recirculation and heat recovery principles. In practice, these may be implemented through the use of a water stream as a waste heat stream (Savulescu and Alva-Argaez, 2013). An alternative approach which makes use of the principle of heat recovery, relies on the waste heat potential from a waste stream to set the operation of a water treatment and recirculation system.

In the alignment of the latter approach, this work studies a Multi-Effect Distillation (MED) technology, which may be used (among other applications) for industrial wastewater treatment, namely to desalinate saline water streams (Rahimi and Chua,

2017). This technology makes part of a set of heat-driven water treatment technologies, with the operation of such system needing the supply of a sensible heat source. The approach considered in this work, explores the MED technology, namely the development and validation of a numerical model using the Modelica language, followed by the analysis of its performance and a brief economic assessment in an industrial plant.

2. Modelling of a Multi-Effect Distillation (MED) Unit

The modelling and simulation procedure have been performed using the Modelica object-oriented modelling language, namely using the open-source software OpenModelica 1.18.1. The models for each one of the operational units as well as the whole model have been developed through the use and adaptation of the existing code presented in the ThermoPower Modelica Library (Politecnico di Milano, 2021).

2.1. Theoretical Model

The study performed by Rahimi and Chua (Rahimi and Chua, 2017) presents an overall model that describes the physical phenomena occurring in a MED unit. Following this work, a conventional MED configuration model is presented. The assembling of the MED model considered 3 main components: i) first effect, ii) second-to-last effects and iii) condenser sections. The physical phenomena model described by Rahimi and Chua (2017) is based on mass and energy balances, including an analysis of the heat transfer and fluid phenomena. The model enables a division of the temperature profiles of the effects by zones, which divide specific different sections with different trends for the respective temperature profiles of the hot and cold streams. The mass and energy balance equations for each section are presented in Table 1, while the temperature profiles for each section are represented in Fig. 1.

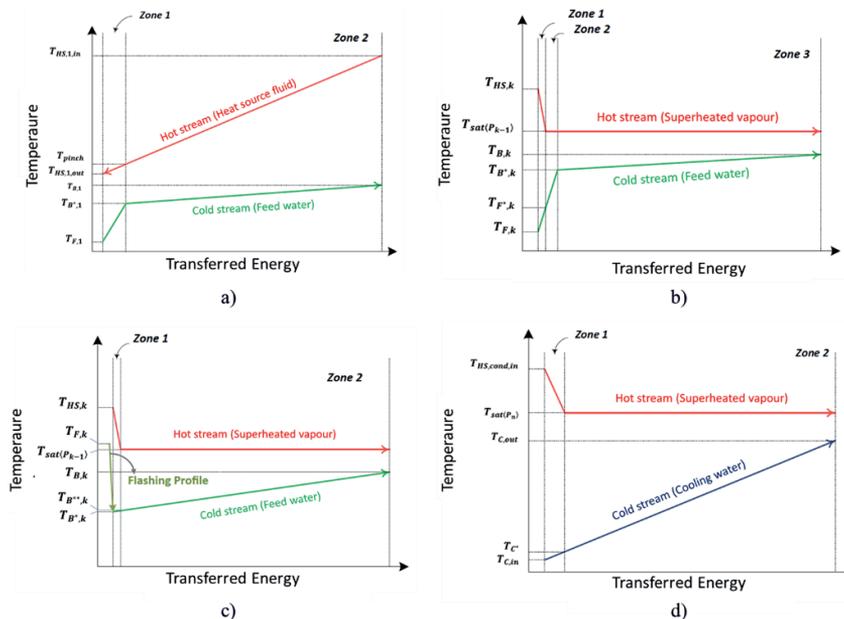


Fig. 1. Temperature profiles for a) first effect section, b) second-to-last effects section without flashed feed stream, c) second to last effect section with flashed feed stream and d) condenser section (adapted from (Rahimi and Chua, 2017))

Table 1. Mass and Energy balances for a conventional MED system

Balance	Zone	Equation
Mass Balance	First effect	$M_{F,1} = M_{B,1} + M_{V,1}$ (1)
		$M_{F,1} \cdot X_{F,1} = M_{B,1} \cdot X_{B,1}$ (2)
	Second-to-last effects	$M_{F,k} = M_{B,k} + M_{V,k}$ (3)
		$M_{HS,k} = M_{V,k-1}$ (4)
		$M_{F,k} \cdot X_{F,k} = M_{B,k} \cdot X_{B,k}$ (5)
	Condenser	$M_{HS,cond} = M_{V,n}$ (6)
Enthalpy Balance	First effect	$M_{F,1} \cdot h_{F,1} + M_{HS,1} \cdot h_{HS,1,in} = M_{B,1} \cdot h_{B,1} + M_{V,1} \cdot h_{V,1} + M_{HS,1} \cdot h_{HS,1,out}$ (7)
	Second-to-last effects	$M_{F,k} \cdot h_{F,k} + M_{HS,k} \cdot h_{HS,k,in} = M_{B,k} \cdot h_{B,k} + M_{V,k} \cdot h_{V,k} + M_{HS,k} \cdot h_{HS,k,out}$ (8)
		$h_{HS,k,out} = h_{f,sat}(P_k)$ (9)
		$M_{HS,cond} \cdot h_{HS,cond,in} + M_C \cdot h_{C,in} = M_{HS,cond} \cdot h_{HS,cond,out} + M_C \cdot h_{C,out}$ (10)
	Condenser	$h_{HS,cond,in} = h_{v,nn}$ (11)
		$h_{HS,cond,out} = h_{f,sat}(P_n)$ (12)

2.2. Development of a MED Unit model in Modelica

The procedure of this work involved the development of a model for each component of a MED unit: i) First effect, ii) Second-to-last effects and iii) Condenser. Each one of these models was assembled considering the mass and energy balances and the heat transfer along the effects and condenser. In order to build the computational models, it was necessary to adapt such equations, in order to simplify and enable the convergence of the simulation. For instance, the heat transfer equations were discretized using the finite volume method (FVM). The assembled model (considering a total of four MED effects) is represented in Fig. 2 (in which Effects 3 and 4 are omitted). The equations used to describe the heat transfer phenomena in a MED unit are presented in Table 2 and Table 3. It is to note that, in the sets of equations, i designates a node of the control volume of an effect or the condenser.

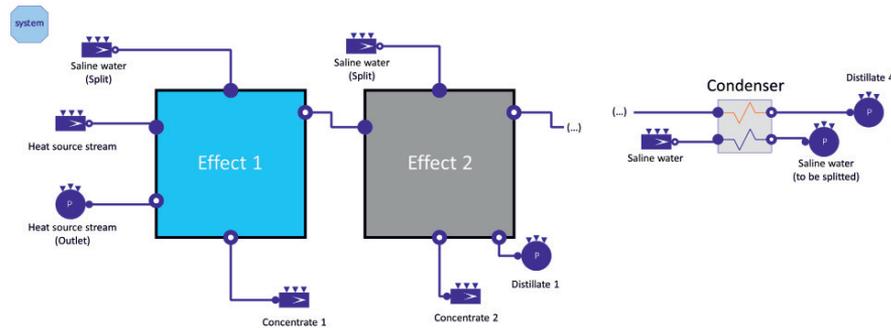

Fig. 2. Assembling of the MED unit model in OpenModelica

Table 2. Model equations for heat transfer associated to Effect 1 and its zones

Section	Zone	Heat Transfer Equations
Effect 1	Zone 1	$q(i) = M_{HS,1} \cdot C_{P,HS,1} \cdot (T_{HS,1}(i) - T_{HS,1}(i+1))$ (13)
		$q(i) = M_{F,1} \cdot C_{P,W} \cdot (T_{F,1}(i+1) - T_{F,1}(i))$ (14)
	$dT(i) = ((T_{HS,1}(i) - T_{F,1}(i)) + (T_{HS,1}(i+1) - T_{F,1}(i+1))) \cdot 0.5$ (15)	
	$q(i) = U_{1,1} \cdot (A(i+1) - A(i)) \cdot dT(i)$ (16)	
	Zone 2	$q(i) = M_{HS,1} \cdot C_{P,HS,1} \cdot (T_{HS,1}(i) - T_{HS,1}(i+1))$ (17)
	$dT(i) = ((T_{HS,1}(i) - T_{F,1,sat}) + (T_{HS,1}(i+1) - T_{F,1,sat})) \cdot 0.5$ (18)	
	$q(i) = U_{1,2} \cdot (A(i+1) - A(i)) \cdot dT$ (19)	

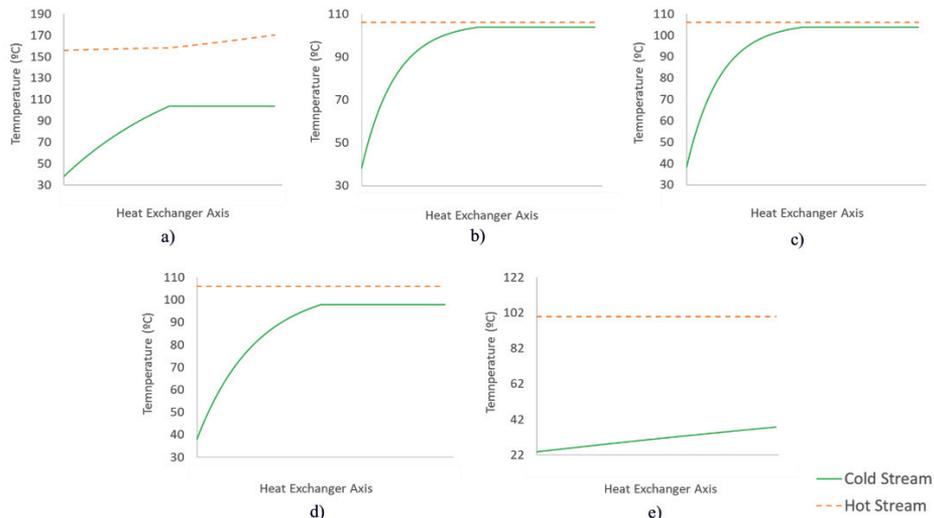
Table 3. Model equations for heat transfer associated to Effect k, Condenser and its respective zones

Section	Zone	Heat Transfer Equations	Section
Effect k	Zone 2	$q(i) = M_{HS,k} \cdot (h_{HS,k}(i) - h_{HS,k}(i+1))$	(20)
		$q(i) = M_{F,k} \cdot C_{P_w} \cdot (T_{F,k}(i) - T_{F,k}(i+1))$	(21)
		$dT(i) = \left((T_{HS,k_{sat}} - T_{F,k}(i)) + (T_{HS,k_{sat}} - T_{F,k}(i+1)) \right) \cdot 0.5$	(22)
	Zone 3	$q(i) = U_{k,2} \cdot (A(i+1) - A(i)) \cdot dT(i)$	(23)
		$q(i) = M_{HS,k} \cdot (h_{HS,k}(i) - h_{HS,k}(i+1))$	(24)
		$dT(i) = \left((T_{HS,k_{sat}} - T_{F,k_{sat}}) + (T_{HS,k_{sat}} - T_{F,k_{sat}}) \right) \cdot 0.5$	(25)
Condenser	Zone 2	$q(i) = U_{k,3} \cdot (A(i+1) - A(i)) \cdot dT(i)$	(26)
		$q = M_{F,cond} \cdot C_{P_w} \cdot (T_{F,cond}(i) - T_{F,cond}(i+1))$	(27)
		$dT(i) = \left((T_{HS,n_{sat}} - T_{F,cond}(i)) + (T_{HS,n_{sat}} - T_{F,cond}(i+1)) \right) \cdot 0.5$	(28)
		$q(i) = U_{cond,2} \cdot (A(i+1) - A(i)) \cdot dT(i)$	(29)

The following modelling assumptions have been considered: steady-state simulation; constant and equal recovery ratio of the primary MED effects; negligible heat losses and pressure losses; constant temperature and salinity of the inlet feed water stream; null salinity of the outlet treated water stream; a temperature difference of 2.5 °C between the condensed vapour and the outlet concentrate streams for primary MED effects. The heat source stream, which is the hot stream in Effect 1, is a liquid with a specific heat capacity of 6643 J/(°C.kg), entering in Effect 1 at a temperature of 170.0 °C and at pressure of 1.5 bar.

2.3. Simulation Results and Model Validation

The model validation has been performed in order to assess its reliability and accuracy in comparison with the theoretical data (Rahimi and Chua, 2017). In this work, model validation has been performed by comparing the heat transfer phenomena occurring in each one of the sections, namely the temperature profiles. The temperature profiles achieved by the simulation of the model for each one of the components of the MED unit are represented in Fig. 3.

**Fig. 3.** Obtained temperature profiles for a) Effect 1, b) Effect 2, c) Effect 3, d) Effect 4, e) Condenser

By comparing the theoretical temperature profiles reported in the literature with the simulated ones (respectively Fig. 1 and 3), it is observed that the simulated temperature profiles are overall consistent with the theoretical ones. The temperature profile of the cold fluid for the zone 2 of effects 2, 3 and 4 (second-to-last effects) presents a sharper variation compared to the linear variation shown in Fig. 1(c), which may be attributed to the use of FVM method for the computation of mean temperature difference, which produces relatively more accurate results. Considering these aspects, it is possible to conclude that the developed model is overall consistent on the prospect of the performed simulation being able to reproduce the real phenomena.

3. Economic Assessment

In addition to the development of a model for a MED unit, Rahimi and Chua (2017) expanded their analysis to a generalized method for the economic assessment of such desalination installations. The simplistic economic model included the capital cost and payback period associated to a MED project, as detailed in Table 4, while the economic savings and payback period estimations considering the water savings results obtained by the simulation are presented in Table 5. A water price of 1.1952 €/m³ was considered for the conversion of water production (assumed to be equal to potential freshwater savings) related to the economic savings.

Table 4. Capital costs and associated payback period determination for a MED unit (Rahimi and Chua, 2017)

Capital Cost	$TCC_{MED} = 2535 \cdot D_t^{0.9751}$	(30)
Payback Period	$PB = TCC_{MED}/(Sav - AEC)$	(31)

Table 5. Determination of water production, economic savings and payback period

Water Production (D_t) (m³/day)	843.28
Water Production (D_t) (dam³/year)	292.41
Investment costs (TCC_{MED}) (k€)	1807.57
Electricity costs (AEC) (k€/year)	95.21
Economic savings (Sav) (k€/year)	349.49
Payback Period (PB) (years)	7.1

Considering the parameters listed in Table 5 and applying equations (30) and (31), a payback period of about 7.1 is obtained. Note that this analysis has not considered certain parcels of the OPEX, for instance, costs associated to chemicals (which includes the costs associated to the acquisition and use of the heat source stream), maintenance, spares, incomes and labour. These are assumed to be common to both MED and conventional wastewater treatment. In this case, these parcels are considered to have already been taken into account in the project of the industrial site in which the MED unit is installed. The payback lies within the acceptable economic viability interval 4 – 16 years of payback time (Baniasad Askari and Ameri, 2021), and therefore it is possible to say that this project is economically feasible. However, considering different studies for the techno-economic assessment of thermal vapour compression MED units, in which payback times correspond to less than 5 years, it is concluded that the MED unit project considered in this work still requires further analysis in terms of potential improvements.

4. Conclusions

This work presents the study of a heat-driven water treatment system, in particular through the exploitation of the multi-effect distillation (MED) technology. This has been achieved by the development and assembling of numerical models describing a conventional MED unit, using the Modelica language. Literature data has been

considered as case-study enabling the modelling of MED unit. The modelling results were validated by comparing the temperature profiles with the theoretical ones. The simulated temperature profiles were, in general, consistent with the theoretical ones, thus ensuring the validity of the model for the simulation of the real occurring phenomena. Furthermore, a brief economic assessment was performed to assess the viability of the MED plant project in an industrial plant. A payback period of 7.1 years was determined, suggesting that the desalination project may be considered economically viable (considering a reasonable payback time of 4 – 16 years), although this is still high when compared to payback periods of other MED unit projects. Further improvement measures (for instance, improved configurations of multi-effect distillation such as preheated MED, boosted MED and flash boosted MED units) will be analysed in the following studies.

Nomenclature

<i>A</i>	Heat Transfer Area (m ²)	Subscripts	
<i>AEC</i>	Cost of electricity (€/year)	<i>B</i>	Concentrate/ Brine Stream
<i>D_t</i>	Daily quantity of water production (m ³ /day)	<i>C</i>	Cooling Water Stream
<i>h</i>	Specific Enthalpy (J/kg)	<i>Cold</i>	Cold Stream
<i>M</i>	Mass flow rate (kg/s)	<i>cond</i>	Condenser
<i>P</i>	Pressure (Pa)	<i>F</i>	Feed Water Stream
<i>PB</i>	Payback Period (years)	<i>Hot</i>	Hot stream
<i>TCC</i>	Capital Cost (€)	<i>HS</i>	Heat Source Stream
<i>Sav</i>	Economic Savings (€/year)	<i>in</i>	Inlet
<i>U</i>	Overall Heat Transfer Coefficient (W/(m ² .°C))	<i>k</i>	Effect number designation
<i>X</i>	Salt concentration (ppm)	<i>MED</i>	MED unit
		<i>n</i>	Total number of effects in a MED unit
		<i>out</i>	Outlet
		<i>total</i>	Total
		<i>sat</i>	Saturation
		<i>V</i>	Vapour Stream

References

- Alnouri, S.Y., Linke, P., El-Halwagi, M., 2014, Water integration in industrial zones: A spatial representation with direct recycle applications, *Clean Technol. Environ. Policy*, 16.
- Baniasad Askari, I., Ameri, M., 2021, A techno-economic review of multi effect desalination systems integrated with different solar thermal sources. *Appl. Therm. Eng.* 185.
- Coelho, P.J., 2011, Discrete ordinates and finite volume methods, *Thermopedia*.
- Oliveira, M.C., Iten, M., Matos, H.A., Michels, J., 2019, Water-energy nexus in typical industrial water circuits. *Water (Switzerland)* 11.
- Panagopoulos, A., 2020, Process simulation and techno-economic assessment of a zero liquid discharge/multi-effect desalination/thermal vapor compression (ZLD/MED/TVC) system. *Int. J. Energy Res.*, 44.
- Politecnico di Milano, 2021, ThermoPower Home Page [WWW Document]. URL <https://build.openmodelica.org/Documentation/ThermoPower.html>
- Rahimi, B., Chua, H.T., 2017, Low grade heat driven multi-effect distillation and desalination, *Low Grade Heat Driven Multi-Effect Distillation and Desalination*.
- Savulescu, L.E., Alva-Argaez, A., 2013, Process Integration Concepts for Combined Energy and Water Integration. *Handb. Process Integr. Minimisation Energy Water Use, Waste Emiss.*

Acknowledgements

The publishing procedure and conference participation have received funding by the European Union's Horizon 2020 research and innovation programmes under grant agreement "No. 810764".

Continuous operation of a solar photobioreactor with linearizing control – A simulation study

Joris Sébile-Meilleroux^a, Mariana Titica^a, Jérémy Pruvost^a

^a *Nantes Université, Oniris, CNRS, GEPEA, UMR 6144, F-44600 Saint-Nazaire, France*

Abstract

This paper focuses on continuous operation of a closed solar photobioreactor using a feedback linearizing control strategy, maximizing productivity, by acting on dilution rate. Lumostatic and turbidostatic operations were investigated in simulation, using a dynamic model from the literature. This model couples a radiative model describing light gradient into the culture bulk and a kinetic growth model for *Chlorella vulgaris* strain. Results on performances and system operation under different solar light profiles have been discussed in a comparative study considering practical implementation. The result is that the use of solar irradiance prediction led to significant productivity improvement over a year scenario, under turbidostatic operation.

Keywords: Linearizing control, turbidostat, solar photobioreactor, microalgae, biomass production.

1. Introduction

Seeking an economic sustainability and reducing carbon emissions are two of the main current scientific challenges. Use of microalgae (unicellular photosynthetic microorganisms, living in saline or freshwater environments), can be an answer to both these challenges. Indeed, by performing a photosynthetic activity, these organisms convert CO₂ into O₂ when exposed to light. The resulting microalgae biomass represent a source of renewable biofuels (hydrogen, biodiesel, bioethanol). Their operating at optimum productivity, in solar conditions (the sine qua non condition to make them profitable), is a real challenge, which process automation can meet. This is the subject of the work presented here, which aims to implement a feedback control strategy keeping biomass concentration into the culture medium at a predefined set point by acting on the dilution rate, in such a way to reject main disturbances in solar conditions, i.e. uncontrolled various light conditions. Linearizing control has been already studied on microalgae cultures under controlled lighted conditions (Tebbani et al., 2016). In this study, main idea is to go a step further investigating continuous operation with linearizing controller under solar conditions; constant and varying set points, optimizing growth have been investigated. The paper is organized as follows: process description, together with modelling and control problem, is presented in Section 2. The results of the simulation study, investigating control performances under various set points are presented in Section 3. Determination of set points on a daily or annual basis, with different optimization criteria, is discussed and the resulting dilution rates and resulting productivities are compared over a “virtual plant” application. Conclusions and Perspectives are presented in last section.

2. Process description and modelling

2.1. Growth modelling

A flat panel photobioreactor, exposed to solar light has been considered. The reactor is operated in continuous mode, the dilution rate being the manipulated variable. Algal concentration within the culture system, called X ($g_X \cdot L^{-1}$), has been controlled with a feedback linearizing controller. Its temporal evolution, in a homogenous culture medium, in a bioreactor operated at constant volume, is given by mass balance equation (Pruvost et al., 2011):

$$\frac{dX(t)}{dt} = \langle r_X \rangle - D(t) \cdot X(t) \quad (1)$$

$\langle r_X \rangle$ represents the biomass growth rate at volume scale and $D(t) \cdot X(t)$ term represents the biomass flowrate removed from the reactor. This term defines system productivity ($g_X \cdot L^{-1} \cdot s^{-1}$). $D(t)$, the dilution rate (s^{-1}), represents the ratio of feeding flowrate over reactor volume, which effects culture density. Microalgal concentration can then be controlled by dilution, thanks to a pumping system.

In this study, non-limiting growth conditions except light have been considered. In these conditions, biomass growth rate depends on available light into the culture bulk. By coupling a radiative model to a photosynthetic growth model for eukaryotic microalgae, the kinetic model links solar radiation to biomass population expansion inside PBR. The model used in this study has been extensively presented (Sebile-Meilleroux et al., 2020) for *Chlorella vulgaris* strain.

Under high lights conditions, biological phenomena leading to decreasing of biomass growth rate were considered by coupling a mathematical penalty to the kinetic model. Based on experimental results (Artu, 2016), $\langle r_X \rangle$ decreases from a MRPA (Mean Rate of Photonic Absorption) value of $30 \text{ mmol}_{hv} \cdot kg^{-1} \cdot s^{-1}$, until reaching a null value above $40 \text{ mmol}_{hv} \cdot kg^{-1} \cdot s^{-1}$.

2.2. Simulated light profiles

In this study, solar radiation has been described as the total flux density q_0 ($\mu\text{mol}_{hv} \cdot m^{-2} \cdot s^{-1}$) with collimated and diffuse components q_{col} and q_{dif} :

$$q_0 = q_{col} + q_{dif} \quad (2)$$

Four light flux density typical profiles have been defined in order to evaluate the impact of sky coverage (cloudy and sunny situations) and season (winter and summer) on the system productivity. For each profile, three parameters have been considered: maximum total flux density $q_{0,max}$, sky coverage q_{col}/q_{dif} , and sunshine duration t_{day} (h). Parameters values, synthetized in Table I, were taken from meteorological data of Nantes city, France, on the whole year of 2015, given by Meteonorm measurements. With ideal simulated light profiles, parameters were kept constant for a given situation, leading to a sinusoidal light profile.

Table I: Parameters of simulated light profiles

Day	Cloudy winter	Sunny winter	Cloudy summer	Sunny summer
$q_{0,max}$ ($\mu\text{mol}_{\text{hv}}\text{m}^{-2}\text{s}^{-1}$)	150	450	450	1350
$q_{\text{col}}/q_{\text{dif}}$	0.1	1.4	0.1	1.4
$t_{\text{day}}(h)$	10	10	15	15

2.3. Control strategy

A feedback linearizing control law has been implemented to control the biomass concentration at a set point maximizing productivity. This supposes the biomass concentration can be measured online. In these conditions, the control law calculates the dilution (at each control step of 10 minutes), in such a way to keep biomass concentration as close as possible to the set point. In the case of linearizing controller, the control law is defined as follows (Tebbani et al., 2016)

$$D(t) = \frac{1}{X(t)} \cdot \langle r_X \rangle - \frac{\lambda}{X(t)} \cdot [X_{sp} - X(t)] \quad (3)$$

where λ is a constant positive tuning parameter and X_{sp} ($g_X \cdot L^{-1}$) is the biomass concentration set point. In this study λ was set at $6.0 h^{-1}$ thanks to preliminary studies which looked for a successful set point tracking at different constant biomass concentration values, with *Chlorella vulgaris* and *Chlamydomonas reinhardtii* strains.

3. Results

Simulations have been performed using the dynamic model presented in Section 2 as a virtual plant, under simulated and real light profiles, in open and closed loop. The linearizing controller has been used for tracking constant and various set points, chosen with purpose to increase system productivity. Simulations were achieved on Matlab software (R2021a). Typical light profiles were repeated every 24 hours until the appearance of a steady-state on culture behaviour, defined by biomass concentration evolution.

3.1. Study on typical daily light profiles

3.1.1. Lumostatic operation

A way for optimizing process operation is to impose light gradient into the bulk, to ensure optimization of light energy absorption. For microalgae, the maximum volumetric productivity was obtained when full-light attenuation occurs without a dark zone, meaning PBR illuminated fraction equals 1 (Takache et al., 2009). This condition defines the lumostatic mode, in which light absorption is maximized. To do so, optimal concentration X_{sp} has been updated every 10 minutes based on incident irradiance measurements, in such a way to keep lumostatic operating conditions

Applying the linearizing controller (equation 3) to track this variable set point leads however to culture wash-out. This phenomenon was due to low incident lights occurring at the start and the end of the day, to which corresponds low optimal biomass concentration. Indeed, to reach such low concentrations, high dilution rates have to be applied. Then, due to low growth kinetics of microalgae species when compared with the incident irradiance increase over the day, it is not possible to reach next values of biomass concentration set point. In other words, the imposed set point is not achievable. Indeed, further investigations on several lumostatic regimes have shown that microalgae growth

kinetics should be increased by a factor of 20 to be able to track biomass concentration set points ensuring full-light attenuation without a dark zone. So, following a biomass concentration related to lumostatic mode in solar conditions was not realistic; this set point tracking strategy has been consequently rejected.

3.1.2. Turbidostatic operation

In the following, biomass concentration set point has been determined in such a way to optimize productivity on a daily or season basis. Constant biomass concentration set point maximizing daily productivity, for each light scenario (as reported in Table I), has been considered. To do so, 100 values between 0 and $2 \text{ g}_X \cdot \text{L}^{-1}$ for X_{sp} were simulated.

Unlike setting an instantaneous light-dependent concentration set point, setting a constant value over the day is efficient to keep a stable culture, without any wash-out (Figure 1). As it can be seen, the control law applied dilution rates in such a way to track the defined set point. Indeed, due to lack of light during night, biomass concentration decreases even when no dilution is applied. The major effect is the absence of dilution operation in the morning, in order to allow biomass concentration increase. Thanks to photosynthesis achieved under light conditions of first hours of the day, culture densified until reaching the defined set point. Once set point attained, dilution rate given by linearizing controller followed a bell shape (with a flattened part around midday). This shape is similar to the light profile, in order to continuously keep the biomass concentration as long as received light amount is sufficient. In the following, a constant set point giving the highest daily productivity have been selected. Optimal set points and associated productivities applied on repeated daily light profiles were summarized in Table II. Optimal set point values appear to vary with light scenarios, meaning that they are highly light-dependent.

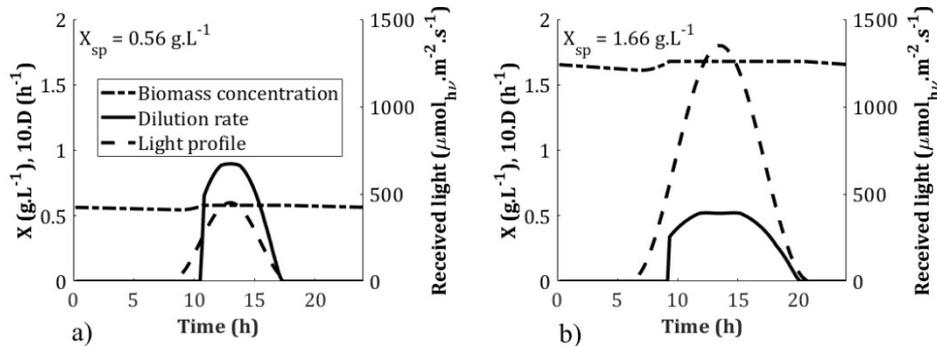


Figure 1: Turbidostatic operation over 24 hours on a) sunny winter and b) sunny summer light

Table II: Daily optimal set points and productivities

Day	Cloudy winter	Sunny winter	Cloudy summer	Sunny summer
Solar flux ($\text{mol}_{hv} \cdot \text{m}^{-2} \cdot \text{day}^{-1}$)	2.70	8.10	12.2	36.4
$X_{sp,opt}$ ($\text{g}_X \cdot \text{L}^{-1}$)	0.26	0.56	0.58	1.66
Productivity ($\text{g}_X \cdot \text{m}^{-2} \cdot \text{day}^{-1}$)	2.22	6.02	8.47	17.6

3.2. Study on year scenario

3.2.1. Closed-loop operation at constant set point over the year

In the following, constant biomass concentration over a year, maximizing annual productivity has been investigated. In order to simulate light decrease from an unperfect light transmission of cover and presence of biofilm on irradiated surface, received light by microalgae was considered half of solar irradiance. A reference scenario has been used to define the constant set point, maximizing the overall productivity. This typical scenario profile represented the yearly light evolution averaged from 10 consecutive years of real data (2005 to 2014), for a given location. Maximal production on this typical profile is attained for X_{sp} at $0.50 \text{ g}_X \cdot \text{L}^{-1}$ for Nantes, France. When applying this set point on year 2015 scenario for the same location, associated productivity hits $5.67 \text{ g}_X \cdot \text{m}^{-2} \cdot \text{day}^{-1}$. For a large-scale “virtual plant”, associated biomass production is $20.7 \text{ t}_X \cdot \text{ha}^{-1}$, meaning turbidostatic operation in solar conditions can satisfy industrial production expectations.

3.2.2. Closed-loop operation at constant set point over the day

As illustrated in previous sub-section, set point value on biomass concentration is highly light-dependent. In the following, a strategy where daily set point value has been adapted to daily light amount, has been investigated. This strategy requires perfect weather predictions over 24 hours to determine the next day total light amount $q_{0,tot}$ ($\text{mol}_{hv} \cdot \text{m}^{-2}$). Thanks to results on typical daily light profiles from Table II, daily optimal set point value appears to follow equation 4 (linear regression: $R^2 = 0.99$).

$$X_{sp} = 0.042 \cdot q_{0,tot} + 0.16 \quad (4)$$

The linearizing controller performed satisfactorily for tracking the biomass concentration set point given by equation 4, as illustrated in Figure 2. Figure 2(a) represents biomass concentration evolution and Figure 2(b) represents the applied dilution rates, over the second week of July of 2015. Adapted set point kept constant over each day (“daily set point” strategy), and set point kept constant over the year (“yearly set point” strategy) were studied. When applying “daily set point strategy”, associated productivity goes up to $6.58 \text{ g}_X \cdot \text{m}^{-2} \cdot \text{day}^{-1}$, meaning a total biomass production of $24.0 \text{ t}_X \cdot \text{ha}^{-1}$. Consequently, daily adaptation of biomass concentration set point improves production by 16 % when compared to an optimal set point kept constant over the year.

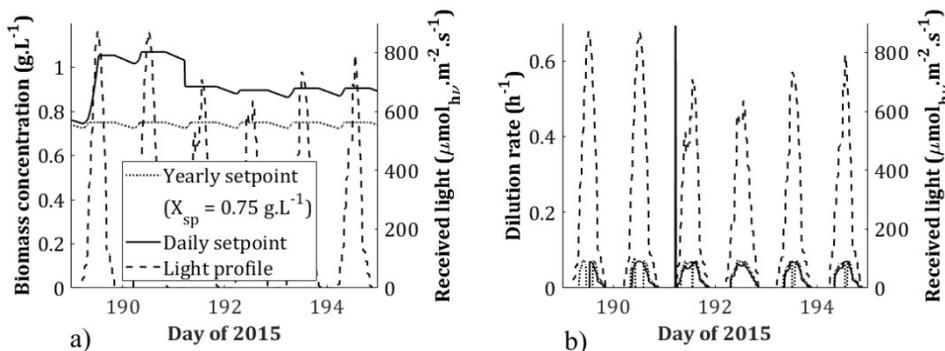


Figure 2 : Daily and yearly strategies effects on a) biomass concentration and b) command law

It can be noticed the efficiency of turbidostat operation using a linearizing control strategy in non-ideal solar conditions, either with a daily or a yearly set point strategy. Nonetheless, using a daily optimal set point can induce a sharp dilution peak, as seen on third day in Figure 2(b). Indeed, if daily set point value drops from previous day

(passing from a high to a lower sunny day), short dilution appears at the beginning of the day in order to quickly attain the new set point value. When diluting a real solar photobioreactor with this daily optimal strategy, this could cause issues with expected productivity, due to pumping limitations. Regarding yearly set point, it has been observed that dilution can be stopped around midday, contrary to daily set point which diluted continuously enduring the sunny hours of the day. Therefore, such a strategy working with a set point optimized at year-scale is not well-suited for high light summer light profiles, because it can induce light stress with negative impact on the productivity. These issues are under study in our ongoing research considering uncertainties on model parameters and saturation on applied dilution rates.

4. Conclusions

When cultivating microalgae in solar conditions, a way to optimize system productivity is to operate in closed-loop continuous mode, at predefined set point for biomass concentration control. The issue in solar condition is to choose a feasible set point, in such a way to optimize system productivity under uncontrolled, varying incident light conditions. In this study, different set point trajectories have been investigated in simulation, using a linearizing controller applied to a “virtual plant”. Working with such strategies requires a dynamic model describing main phenomena related to light influences on the microalgae growth. It has been noticed that the optimum set point ensuring lumostatic conditions wasn’t achievable, as demonstrated in simulation with the linearizing controller applied every 10 minutes. The feedback linearizing controller was able to maintain the biomass concentration at desired set point (constant over the year and piecewise constant over the day), determined in such a way to maximize productivity (turbidostatic operation). Optimizing biomass concentration set point on a daily basis, thanks to global solar irradiance perfect prediction over 24 hours, led to a 16 % improvement of microalgae yearly production, when compared to an optimal set point predefined on a yearly basis. More realistic simulations in the presence of uncertainties in the weather data are currently being studied. On the other hand, practical implementation is envisaged on a solar-scale experimental pilot.

References

- Artu, A., 2016. Etude et optimisation de la culture de microalgues en photobioréacteurs solaires. <https://doi.org/10.13140/RG.2.2.23498.41924>
- Pruvost, J., Cornet, J.F., Goetz, V., Legrand, J., 2011. Modeling dynamic functioning of rectangular photobioreactors in solar conditions. *AIChE J.* 57, 1947–1960. <https://doi.org/10.1002/aic.12389>
- Sebile-Meilleroux, J., Titica, M., Pruvost, J., 2020. Dynamic optimization of closed photobioreactors in solar conditions. A simulation study, in: 2020 24th International Conference on System Theory, Control and Computing (ICSTCC). Presented at the 2020 24th International Conference on System Theory, Control and Computing (ICSTCC), IEEE, Sinaia, Romania, pp. 582–587. <https://doi.org/10.1109/ICSTCC50638.2020.9259653>
- Takache, H., Christophe, G., Cornet, J.-F., Pruvost, J., 2009. Experimental and theoretical assessment of maximum productivities for the microalgae *Chlamydomonas reinhardtii* in two different geometries of photobioreactors. *Biotechnol. Prog.* NA-NA. <https://doi.org/10.1002/btpr.356>
- Tebbani, S., Titica, M., Join, C., Fliess, M., Dumur, D., 2016. Model-based versus model-free control designs for improving microalgae growth in a closed photobioreactor: Some preliminary comparisons, in: 2016 24th Mediterranean Conference on Control and Automation (MED). Presented at the 2016 24th Mediterranean Conference on Control and Automation (MED), IEEE, Athens, Greece, pp. 683–688. <https://doi.org/10.1109/MED.2016.7535870>

Dynamic Inherent Safety Analysis of a Distillation Column under Simultaneous Design and Control

Denis Su-Feher,^{a,b,c} Efstratios N. Pistikopoulos^{a,b}

^a*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA*

^b*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA*

^c*Mary Kay O'Connor Process Safety Center, Texas A&M University, College Station, TX 77843, USA*

Abstract

Inherently safer design (ISD), which focuses on reducing the inherent hazards of a design before applying any safety controls, is one of the most effective and reliable tools for improving the safety of a process. However, while an intensified process may be inherently less hazardous, if the design restricts the controllability of the process, then the design may have a higher risk and be less safe overall. Therefore, considering both the inherent hazard contained within the process and the ease by which these hazards can be controlled is necessary for a more complete evaluation of the inherent safety of a system.

The objective of this research is to implement a strategy to simultaneously design and control an inherently safer distillation column. The PARAmetric Optimization and Control (PAROC) framework is used as a basis for the simultaneous design and control of a distillation column. The Safety WEighted Hazard Index (SWEHI) is incorporated into the PAROC framework, and the distillation column is optimised for cost, while receding control horizon policies are implemented to ensure that the column is capable of controlling disturbances. The dynamic effects of different operating variables on safety are analysed and discussed. The integration of ISD with simultaneous design and control allows for a greater understanding of inherent safety during process design and substantially reduces operability issues that result from an uncontrollable process design and allow for greater tolerance and ease of control.

Keywords: Dynamic Optimization, Distillation Column Design, Inherently Safer Design, Design and Control Integration.

1. Introduction

The design of industrial processes requires the consideration of a number of different characteristics, including profitability, safety, and sustainability. Traditionally, the focus of both academia and industry has been to optimize the early design of process systems primarily for profitability. Other considerations, such as safety and sustainability, are postponed until later stages of engineering design. In such an approach, the process of risk management typically only involves implementing prevention, mitigation, and response methods after the design is nearly finalized. However, such approaches are typically inefficient and result in unsafe designs [1]. Therefore, there has been a movement towards the use of inherent safety principles to improve the safety and profitability of industrial process designs [2].

1.1. Inherently Safer Design

Inherently Safer Design (ISD) is one of the most effective methods for assessing safety in the early stages of design [3]. Inherent Safety uses four main principles to determine the safety of a process.

Intensification (or Minimization) involves reducing hazardous inventories in both process equipment and storage. Substitution means replacing hazardous chemicals with less hazardous chemicals. Attenuation (or Moderation) is using hazardous chemicals in conditions that would make them less hazardous (i.e. reducing the pressure or temperature). Limitation of Effects is to reduce consequences by using designs or reaction conditions rather than by adding protective equipment that may fail or be neglected

Using these principles, inherent safety addresses potential hazards and risks that can potentially lead to the need for costly add-on safety later in the process lifecycle. Therefore, the earlier in the design that inherent safety is considered, the more opportunities there are for the application of inherently safety in the design [4]. One of the limitations of current inherent safety metrics is that they are typically do not consider dynamic disturbances that may impact the safety of the process. During start-up, shutdown, and even during regular operations, the safety of the process will fluctuate with changes in process parameters and variables. However, the fluctuation of safety is not accounted for in the early stages of design by current inherent safety metrics.

It is difficult to determine the dynamics of a system in the early stages of design because most process systems are quite complex. The design of an industrial process influences the ability of its control systems to reject disturbances [5]. Due to the complexity of these processes, the control response is difficult to assess without designing the controller, a development that typically occurs at a later phase. However, without knowing the control response the dynamic inherent safety cannot be assessed.

Recently, the PARAmetric Optimization and Control (PAROC) framework was proposed at a method to allow for the representation, modelling, and solution of simultaneous design and control problems [6]. PAROC allows for the simultaneous simulation of design and operating variables in the early design phase under optimal multi-parametric Model Predictive Control (mp-MPC). This allows for the measurement of the dynamic behavior of industrial processes in the early design phase. Thus, the incorporation of PAROC with inherent safety allows for a dynamic understanding of the interaction between safety and operability in the early stages of design.

The aim of the present work is to demonstrate how PAROC can be used to dynamically measure and optimize the design of an industrial process with respect to its profitability, safety, and controllability. A distillation column is used as a case study. The Safety Weighted Hazard Index (SWeHI) is input as an extra variable into the high-fidelity model to quantify the inherent safety of the distillation column [7]. Disturbances are introduced into the process, and the dynamic change in the safety index is compared to the steady-state index. Finally, the safety index is added as a constraint to the dynamic optimization of the process, allowing for a process that is optimized for cost, operability, and safety.

2. Case Study – Binary Distillation

2.1. Benzene-Toluene-Xylene Production

In 2018, the US produced 11.3 billion lbs of benzene, 6.1 billion lbs of toluene, and 10.0 billion lbs of mixed xylenes [1-3]. Benzene is used to make nylon, styrene, polycarbonate, and epoxy resins [4]. Toluene and xylene are used as solvents in paints, thinners, glues, and other household items [5]. Approximately 45% of the United States' benzene is

produced by conversion of naphtha to benzene, toluene, and xylene via catalytic reforming and separations [6].

The Benzene-Toluene-Xylene (BTX) production process converts naphtha to benzene, toluene, and xylene by catalytically cracking naphtha and then using a series of separations processes to separate BTX from the resulting pyrolysis gas. The benzene-toluene distillation unit of the BTX production process will be optimized in this work.

2.2. Dynamic Challenges in Inherent Safety

Luyben and Hendershot examined the BTX aromatics fractionation process from a control and safety perspective. In this study, they examined the control effects of decreasing molar holdups in the reboilers and the reflux drums and found that making the process inherently safer via intensification may make the column less operable, and thus dynamically less safe [8]. However, no attempt at optimizing the process was performed, and so the problem of determining an inherently safer, controllable, and cost-optimal design remains unsolved.

3. Methodology

3.1. Incorporating safety into the PAROC framework

The basic principles of the PAROC framework in the context of simultaneous design, control, and safety optimisation are presented in Figure 1.

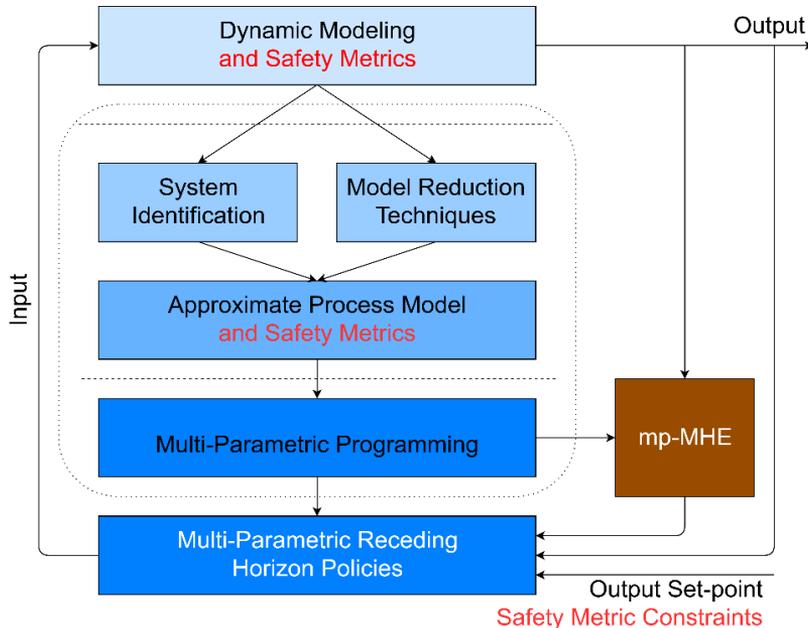


Figure 1. The PAROC framework for simultaneous design, control, and safety.

The first step of the PAROC framework is to create a “high fidelity” model to capture the dynamic behavior of the process, based on first principles modelling and data driven approaches. This includes mass/energy balances, driving force constraints, thermodynamic relations, sizing constraints, and numerous other empirical correlations. The primary basis of the model was that of Bansal et al [9]. The modelling was performed in gPROMS® and steady state accuracy was confirmed in ASPEN Plus [10, 11]. Since

safety may depend on the states, controlled variables, and design variables, the overall safety of the system is defined as a separate variable that is a function of each of these process variables.

The next step is model approximation. While the high-fidelity model presents a highly accurate representation of the process, designing a controller for such a complex system is a difficult task. The MATLAB® System Identification toolbox was used to derive a set of linear state-space models that include design variables in the model expressions.

Multiparametric programming is then performed using the Parametric Optimisation (POP) toolbox to define a set of critical regions where a multiparametric controller can act in order to keep the process within acceptable control limits [12].

Next, the developed receding horizon control policy is tested against the high-fidelity model in-silico. Since the controller was created via a model approximation, this step is necessary to ensure the applicability of the controller to the original high-fidelity system.

Finally, the controller is placed into the original model, and the entire model is dynamically optimised using gPROMS Modelbuilder as a Mixed Integer Dynamic Optimization (MIDO) problem to solve for the design variables.

Continuous design variables that were considered were tray volume and reboiler and reflux drum size. Mixed integer design variables considered were number of trays, feed tray location. Safety was assessed via the Safety Weighted Hazard Index. SWeHI examines material properties such as flammability, reactivity, toxicity, and vapor pressure, as well as operating properties such as operating pressure, tray temperature, and molar flowrate. Molar composition of the distillate and the bottoms is controlled by changing the steam flowrate in the reboiler, and the reflux flowrate into the condenser.

4. Results and Discussion

A 10% sinusoidal disturbance was introduced into the feed composition with a period of 80 seconds. A high-fidelity model and mpMPC was developed and optimized. The high-fidelity model contained 1948 algebraic equations, 81 differential equations, and 2029 unknown variables.

Steam flowrate varied sinusoidally between 6.0 and 7.3 kmol/min, and reflux flowrate varied between 3.5 and 5.0 kmol/min. Figure 2 shows the corresponding output purity. The control response is within acceptable limits, and the output purity is constrained within the bounds.

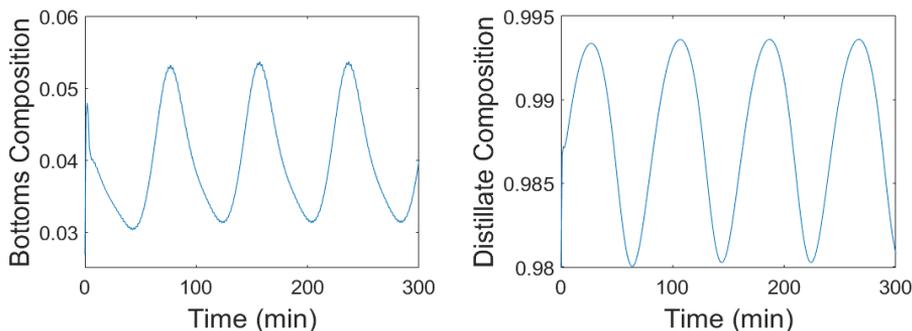


Figure 2. The output purity of the bottoms (left) and distillate (right) streams are constrained above 0.98 and below 0.06 respectively.

Table 1 compares the dynamic optimization performed in this study to the previous dynamically optimized model in literature. The previous model considered a feed flow disturbance that was not examined in this work, which is why its cost is higher.

Table 1. A comparison of the current model optimization with optimized models in literature

	Current Approach	Bansal et al.[9]
Condenser Area (m ²)	121	83.33
Reboiler Area (m ²)	63.3	131.6
Diameter of Column (m)	2.11	1.91
# Trays	22	23
Feed Tray	10	12
Total Cost (\$1k)	7.494	7.908

A dynamic safety assessment was performed. Figure 3 shows the overall safety rating of the column. The total steady state value of the column is 1900. This rating varies from 1888 to 1909, a negligible (1%) difference compared to the steady state value.

Figure 4 shows the safety rating of the reboiler and the reflux drum, which varies from 100 to 110 in the reboiler, and from 50.2 to 51 in the reflux drum. This equates to a 5 % difference in the reboiler, and a 1 % difference in the reflux drum. One important observation to note is that the oscillation in the SWeHI hazard rating in the lower trays and the reboiler are 180 degrees out of phase with those of the upper trays and the reflux drum. This produces destructive interference that muffles the overall change in safety index throughout the column. This oscillation also indicates a dynamic vertical transfer of hazard from one side of the column to the other as a response to a sinusoidal feed composition change. Although the distillation tower as a whole remains at a near constant hazard, the inherent hazard is being transferred up and down the column periodically with the disturbance.

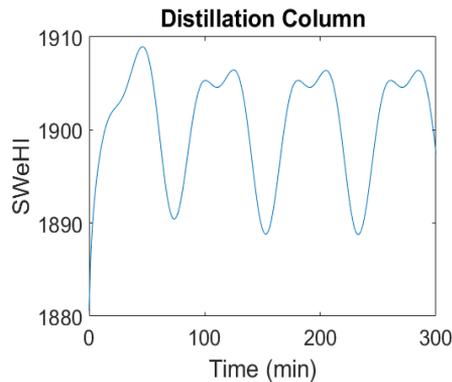


Figure 3: Dynamic SWeHI hazard rating of the overall column.

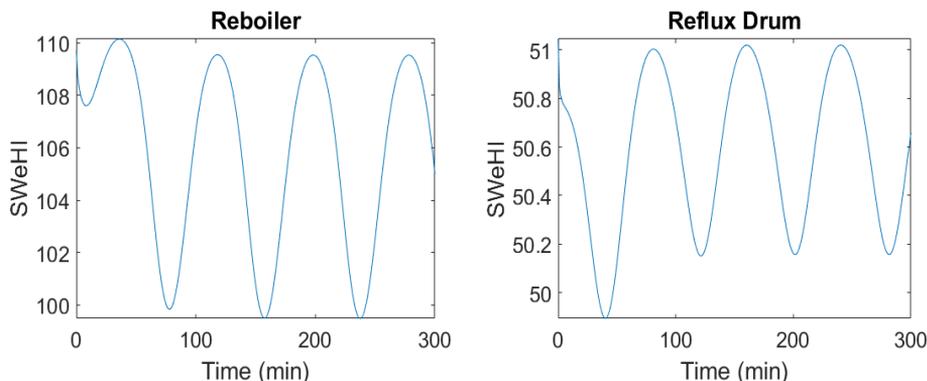


Figure 4: The SWeHI hazard ranking of the reboiler (left) and condenser (right).

5. Conclusions

A methodology for the dynamic optimization of design and safety with guaranteed control was presented. The methodology was applied to a distillation column case study. A dynamic safety assessment was conducted on the distillation column.

An overall 1% difference between steady-state and maximum dynamic damage radius was observed within the process as a result of feed composition disturbance. A 5% difference was observed within the reboiler and a 1% difference was observed in the reflux drum, along with oscillations in safety along vertical sections of the column. A similar assessment and optimization could be performed on any process unit.

References

- [1] T. Kletz, "Making safety second nature," *Process Safety Process*, vol. 17, pp. 196-199, 1988.
- [2] P. Amyotte, F. Khan, and T. Kletz, "Inherently Safer Design Activities Over the Past Decade," presented at the Hazards 11, 2009.
- [3] T. Kletz, "Inherently Safer Plants," *Plant/Operations Progress*, vol. 4, no. 3, pp. 164-167, 1985.
- [4] M. Hurme and M. Rahman, "Implementing inherent safety throughout the process lifecycle," *Journal of Loss Prevention*, vol. 18, pp. 238-244, 2005.
- [5] J. G. Ziegler and N. B. Nichols, "Process Lags in Automatic Control Circuits," *Trans. ASME*, vol. 65, pp. 433-444, 1943.
- [6] E. N. Pistikopoulos and N. A. Diangelakis, "Towards the integration of process design, control and scheduling: Are we getting closer?," *Computers and Chemical Engineering*, vol. 91, no. 1st, pp. 85-92, 2015.
- [7] F. Khan, T. Husain, and S. A. Abbasi, "Safety Weighted Hazard Index (SWeHI): A New, User-friendly Tool for Swift yet Comprehensive Hazard Identification and Safety Evaluation in Chemical Process Industries," *Trans IChemE*, vol. 79, no. 2, pp. 65-80, 2001.
- [8] W. Luyben, "Dynamic Disadvantages of Intensification in Inherently Safer Process Design," *Industrial & Engineering Chemistry Research*, vol. 43, pp. 384-396, 2003.
- [9] V. Bansal, J. D. Perkins, and E. N. Pistikopoulos, "A Case Study in Simultaneous Design and Control Using Rigorous, Mixed-Integer Dynamic Optimization Models," *Industrial & Engineering Chemistry Research*, vol. 41, no. 4, pp. 760-778, 2002/02/01 2002, doi: 10.1021/ie010156n.
- [10] Process Systems Enterprise. "gPROMS." www.psenterprise.com/products/gproms (accessed 1997-2020).
- [11] Aspen Technology Inc. "Aspen Plus." (accessed 2021).
- [12] R. Oberdieck, N. A. Diangelakis, M. M. Papathanasiou, I. Nascu, and E. N. Pistikopoulos, "POP - Parametric Optimization Toolbox," *Industrial & Engineering Chemistry Research*, vol. 55, 33, pp. 8979-8991, 2016.

Modeling of Phosphates Slurry Pipelines Through Dynamic Non-Newtonian Fluid Model with Modelica

Fatima Ez-Zahra El Hamra^{a,*}, Radouan Boukharfane^a, Saad Benjelloun^a, Ahmed Ja^a and Jean-Michel Ghidaglia^{a,b}

^a*Mohammed VI Polytechnic University MSDA Group, Benguerir, Morocco*

^b*ENS Paris-Saclay, Universite Paris-Saclay, CNRS, Centre Borelli, Gif-sur-Yvette, France*
fatimaezzahra.elhamra@um6p.ma

Abstract

The main objective of this work is to develop a one dimensional model to simulate the flow of Non-Newtonian fluids under Modelica software. We built a new library of components that provide the simulation of Newtonian and Non-Newtonian fluids flows. Within the pipe model, a new first order scheme of finite volume method to resolve Navier-Stokes equations for incompressible fluids with varying density is implemented. To introduce pressure drops, a Bingham plastic model is considered to represent the rheological behavior of the Non-Newtonian fluids. With the new library components, we perform a first numerical simulations of water flow in a pipe network of extended length. The computational cost of the performed simulation with the the Modelica Standard Library components reduced by a significant orders of magnitude compared to the previous implementation, while the dynamic results are not affected. Then we simulate the flow of the phosphate slurry through the real Pipeline network design and compare the results with industrial data.

Keywords: Modelica, Bingham, finite volume, Non-Newtonian

1. Introduction

OCP Group a worldwide leader operating in the mining and marketing of the phosphate ore and its derivatives. The industrial process involves transporting the ore from the deposits to the processing plants. Since 2014, as part of the optimization strategy, it was decided to replace the rail transport mode between the city of Khouribga and the Jorf Lasfar plant by a flow of phosphate slurry through a network of pipes (Rusconi et al., 2016). the pipeline extends over 187 km of terrain with varied topographies. However, the control parameters available for this system are the density of the slurry, the prescribed inlet flow rate and the pressure at the network's downstream. Therefore, flow models are required to predict system behavior and ensure optimal control by avoiding ore deposition, pressure drops, and the damages that can result from such events. Fully three-dimensional flow models with water-particles fluid models are too complex to build and are CPU time consuming to be exploitable in any real process control, optimization or design (Ghodke and Apte, 2016). Alternatively, one may consider an averaged rheology for the mixture and in this case the slurry can be characterized by a Non-Newtonian rheology. Given the complexity of the non-Newtonian rheological behavior of the phosphates slurry, besides the length of the network, a phenomenological model to simulate this flow is requested (Lima, 2020).

In the present work we developed a new fluid library under Modelica software, to simulate flows through pipelines for both Newtonian and Non-Newtonian fluids. Indeed, we have developed

a new numerical scheme to solve the Navier-stokes equations for incompressible fluids drained into a circular pipe in one dimension. The new scheme is implemented within the pipe model. Then, we integrated the frictional pressure drop calculation models for both types of fluid. For the Non-Newtonian fluid, we adopt the Bingham model as it give a fair description of the rheological behavior (Maazioui et al., 2021). To validate the new scheme model, we perform numerical simulations through different sizes of pipeline network (10, 100, and 200 pipes) using the new fluid library and the Modelica Standard Library (MSL) components. We compare their results and the computation time. Moreover, we simulate the flow of the phosphate slurry through the first part of the pipeline network formed by 1293 pipes, and we compare the results with industrial data.

2. Methodology

The purpose of this work is to develop a one-dimensional model to simulate the flow of Non-Newtonian fluids withing a pipe of length L , a circular section of diameter D , and inclined to the horizontal by the angle θ . We define T the internal volume of the pipe and C its contour, S_{in} and S_{out} are respectively the inlet and outlet sections of the pipe.

2.1. Mathematical model

Consider the conservation of mass and momentum equations for incompressible flow with suitable boundary conditions

$$\begin{cases} \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{u}) = 0, \\ \frac{\partial (\rho \mathbf{u})}{\partial t} + \text{div}(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla P = \rho \mathbf{g} + \mu \Delta \mathbf{u} \end{cases}, \quad (1)$$

where ρ is the density, \mathbf{u} the velocity vector, P the pressure, \mathbf{g} the gravity acceleration vector, and μ the fluid dynamic viscosity. The hypothesis of flow incompressibility rewrites the continuity equation as

$$\text{div}(\mathbf{u}) = 0. \quad (2)$$

The boundary conditions are specified for the present model as follows. For the inlet, we consider that the average density of the fluid and the mass flow through the inlet section ($\mathcal{A} = L_y L_z$ and $d\sigma = dydz$) are given in input following

$$\begin{cases} \rho_{in}(t) = \frac{1}{\mathcal{A}} \int_{S_{in}} \rho(0, \sigma, t) d\sigma, \\ q_{in}(t) = \int_{S_{in}} \rho(0, \sigma, t) \mathbf{u}(0, \sigma, t) \cdot \mathbf{e}_x d\sigma. \end{cases} \quad (3)$$

Then, we define the average input velocity

$$v_{in}(t) = q_{in}(t) / \mathcal{A} \rho_{in}(t) \quad (4)$$

At the pipe outlet, the pressure P_{out} is determined as $\frac{1}{\mathcal{A}} \int_{S_{out}} P(L_x, \sigma, t) d\sigma = P_{out}(t)$. The boundary condition at the wall is specified according to the no-slip conditions as $u = 0$ By integrating $\text{div}(\mathbf{u})$ we get

$$\int_{S_{in}} (\mathbf{u} \cdot \mathbf{e}_x)(0, y, z, t) dydz = \int_{S_{out}} (\mathbf{u} \cdot \mathbf{e}_x)(L_x, y, z, t) dydz \quad (5)$$

2.2. Finite volume discretization

We discretize the pipe into N segments of equal length L_x/N where $x_i = \frac{iL_x}{N}$, where $i = 0, \dots, N$. Integrating the mass conservation constraint in Eq. (2) over the volume of control $V_{i+\frac{1}{2}}$ to get

$$\begin{cases} m_{i+1/2} \equiv \int_{V_{i+1/2}} \rho d\tau, 0 \leq i \leq N-1, \\ \varphi_{m_i} \equiv \int_{S_i} \rho v d\sigma, 0 \leq i \leq N \end{cases} \quad (6)$$

The equation of mass conservation is therefore expressed as follows

$$\frac{dm_{i+1/2}}{dt} = \varphi_{m_i} - \varphi_{m_{i+1}}, \quad 0 \leq i \leq N-1, \quad (7)$$

Similarly, integrating Eq. (4) over a control volume $V_{i+1/2}$ we find that

$$(\rho v)_{S_i} \simeq \rho_{S_i} v_{in} \quad (8)$$

A first order approximation is hereafter retained, which allowed to obtain

$$m_{j+1/2} = A\Delta x \rho_{j+1/2} \quad (9)$$

from which, at first order, we get

$$\varphi_{m_i}^{(1)} = \frac{m_{i-1/2}}{A\Delta x}, \quad 1 \leq i \leq N \quad (10)$$

then, Eq. (7) is expressed as follows

$$\frac{dm_{i+1/2}}{dt} = \frac{v_{in}}{\Delta x} (m_{i-1/2} - m_{i+1/2}), \quad 0 \leq i \leq N-1, \quad \text{with } m_{-1/2} = A\Delta x \rho_{in} \quad (11)$$

2.2.1. Momentum conservation integration

We integrate (4) over $V_{i+\frac{1}{2}}$. First, we define $I_{i+1/2} \equiv \int_{V_{i+1/2}} \rho v d\tau$, we obtain

$$\frac{dI_{i+1/2}}{dt} = \varphi I_i - \varphi I_{i+1} + A(\Pi_i - \Pi_{i+1}) + \int_{V_{i+1/2}} \rho \mathbf{g} \cdot \mathbf{e}_x d\tau + \int_{V_{i+1/2}} \mu \Delta v d\tau, \quad (12)$$

where

$$\varphi I_i \equiv \int_{S_i} \rho v^2 d\sigma, \quad 0 \leq i \leq N, \quad \Pi_i \equiv \frac{1}{\mathcal{A}} \int_{S_i} P d\sigma, \quad 0 \leq i \leq N. \quad (13)$$

We note $(\Delta P)_{i+1/2}$ the pressure loss due to friction over the control volume $V_{i+\frac{1}{2}}$, therefore

$$(\Delta P)_{i+1/2} = \frac{1}{\mathcal{A}} \int_{V_{i+1/2}} \mu \Delta v d\tau, \quad 0 \leq i \leq N-1 \quad (14)$$

Besides, since $\mathbf{g} \cdot \mathbf{e}_x$ is constant (equal to $-|g| \sin(\theta)$)

$$\int_{V_{i+1/2}} \rho \mathbf{g} \cdot \mathbf{e}_x d\tau = (\mathbf{g} \cdot \mathbf{e}_x) m_{i+1/2} \quad (15)$$

From (8) we get

$$I_{i+1/2} \equiv \int_{V_{i+1/2}} \rho v d\tau = \bar{v}_{V_{i+1/2}} m_{i+1/2} \quad (16)$$

where $\bar{v}_{V_{i+1/2}}$ is the density weighted average velocity over the volume $V_{i+1/2}$. Thus, we replace $\bar{v}_{V_{i+1/2}}$ with v_{in} in (16) to get

$$I_{i+1/2} = m_{i+1/2} v_{in} \quad \text{and} \quad \varphi I_i = v_{in} \varphi m_i, \quad (17)$$

and eventually returning to (12) we get

$$(\Pi_i - \Pi_{i+1}) = \frac{1}{\mathcal{A}} \left(\frac{dv_{in}}{dt} + |g| \sin(\theta) \right) m_{i+1/2} - (\Delta P)_{i+1/2}, \quad 1 \leq i \leq N-1, \quad (18)$$

where $\Pi_N = P_{out}$

2.3. Pressure Loss determination

The general Darcy-Weisbach equation expresses the pressure loss for an incompressible fluid, due to friction along a given length of pipe at the average velocity of the fluid flow

$$\Delta P = \frac{\rho L v^2 f_D}{2D}, \quad (19)$$

where f_D is the Darcy friction factor, which is a dimensionless quantity given for each fluid behavior and flow regime (laminar/turbulent) by specific formulas.

2.3.1. Darcy friction factor for Newtonian fluids

For a **laminar** flow, the friction factor value is given by Poiseuille's law related to the Reynolds number Re by $f_D = 64/Re$, where $Re = \frac{\rho v D}{\mu}$. In the **turbulent** regime ($Re > 4000$), the friction factor is not only dependent on Reynolds number, but also on the relative roughness of the pipe wall k/D . Haaland (1983) expressed this coefficient explicitly as

$$\frac{1}{\sqrt{f_D}} = -1.8 \log \left[\frac{6.9}{Re} + \left(\frac{k}{3.7D} \right)^{1.11} \right] \quad (20)$$

in transient regime, the friction factor is expressed by linear interpolation between the limit values of the two regimes.

2.3.2. Darcy friction factor for Non-Newtonian fluids

Hedstrom (1952) introduces the following dimensionless numbers, to express the friction factor for the **laminar** region for Bingham flows $Re_B = \frac{DV\rho}{\mu_B}$, with $He = \frac{\tau_0 D^2 \rho}{\mu_B^2}$ where Re_B the Bingham Reynolds number, and He the Hedstrom number. The laminar Darcy friction factor for Non-Newtonian fluids was first published by Buckingham-Reiner in an implicit formulation. Later this formula was approached by the explicit equation of Swamee and Aggarwal (2011)

$$f_L = \frac{64}{Re_B} + \frac{64}{Re_B} \left(\frac{He}{6.2218 \times Re_B} \right)^{0.958} \quad (21)$$

Darby et al. (1992) proposed the semi-empirical equation presented above to give a friction factor for turbulent flow $f_T = 10^a Re_B^b$, where $a = -1.47(1 + 0.146 \times e^{-2.9 \times 10^{-5} He})$ and $b = -0.193$. Since the transition from laminar to turbulent is not sudden for Bingham slurries. Darby et al. (1992) reviewed the previous work and proposed to combine the laminar and turbulent Darcy friction factors into the following equation, which is valid for all flow regimes

$$f_D = (f_L^m + f_T^m)^{1/m}, \quad (22)$$

where $m = 1.7 + 40000/Re_B$. Within the pipe model, we integrate the friction calculation equations for Newtonian and non-Newtonian fluids into a package called `Wallfriction`. We extend them into the flow model based on the equations Eq. (11) and Eq. (18). The choice of the friction model used is set according to the value of a Boolean parameter set by the user.

3. Results and Discussions

3.1. Comparison with Modelica Dynamic Pipe

In this section, we will compare the results of our model implemented on Modelica, to the classical pipe model available in the Fluid library of the MSL (`Modelica.Fluid.Pipes.DynamicPipe`),

we compare therefore the computation time between both models. Beside the pipe model, we have also developed two models for pressure and mass flow sources. We attached all these components to the FluidPort connector that we developed in this way, to ensure the connections. We gather them all in one new library.

Network description: To validate our model of dynamic pipe flow, we create a network consisting of an upstream mass flow source (1106.5 kg/s), a pipe series, and a downstream pressure source (10 bar) from the two libraries. We take the pipe geometry data from the OCP pipeline design data. We use the Newtonian fluid model considering the carried fluid is water of density $\rho = 995.5 \text{ Kg/m}^3$ and dynamic viscosity of $\mu = 10^{-3} \text{ Pa}\cdot\text{s}$. The models predict the pressure drops and then the pressure at the inlet of the pipe. We calculate and plot the resulting piezometric lines from each case to compare their values.

Results and discussion:

As shown in Fig. 1 and the RMSE column of Table 1, the results given by the new model fit quite closely to those given by the components provided in the Fluid library of the Modelica MSL.

Moreover, by increasing the number of pipes, we notice that the difference between the number of solved equations and the simulation time becomes more significant. With a network of two hundred pipes, the new model is 7 times faster than the MSL model. Which is more favorable in our case of study where we will use a network containing more than 5000 pipes. It is noteworthy to mention that increasing further the maximum number of pipes by two leads to an RMSE almost constant, which indicates that the solution achieves solution convergence.

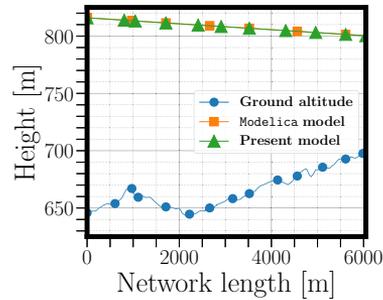


Figure 1: Piezometric lines versus Network length

Pipes Number	Number of equations		Simulation time [s]		RMSE
	MSL Library	New Library	MSL Library	New Library	
10	1277	276	0.0909168	0.0213864	0.0054
100	12437	2706	1.12199	0.083706	0.038
200	24837	5406	2.27997	0.30165	0.082

Table 1: comparison of simulation time and piezometric results between MSL model and the New model

3.2. Simulation of the Non-Newtonian slurry flow

As industrial data, we dispose of the phosphate slurry flow data during a period of 20 minutes along the pipeline. The inlet flow rate as well as the outlet pressure are set by the operator. The fluid density $\rho = 1680.0 \text{ kg/m}^3$. Besides the upstream and the downstream of the network, pressures are captured at 5 points distributed along the pipeline: PMS₁, Valve Station, PMS₂, PMS₃, and PMS₄. We calculate the friction pressure losses from the industrial data set and fit these values in terms of the input velocities to determine the Bingham rheological parameters (μ_B & τ_0) of the transported fluid. We use the `scipy.optimize.curve_fit` function included in the Scipy python library to fit the function ΔP , given by Eq. (19) where we replace the Darcy friction factor f_D by Eq. (22) into industrial data. We obtain a plastic viscosity $\mu_B = 0.012 \text{ Pa}\cdot\text{s}$, and a yield stress of $\tau_0 = 6.1 \text{ Pa}$. Exploiting the fitting parameters, we implement the slurry network between the inlet and PSM₁ with its real geometry, the network consists of 1293 pipes of internal diameter 0.85 m, its longitudinal profile is presented in Fig. 2a. We perform numerical simulations using the inlet

mass flow rates and the PMS_1 pressures from the industrial data set of OCP Slurry Pipeline. We compare the inlet pressures obtained from these simulations with the measured values; see Fig. 2b. The plot of the inlet pressures estimated by our model almost matches the curve of the pressures measured in reality. The mean square error is 218.4, corresponding to a relative mean error of 6%. This error can occur due to different phenomena. The fluid characteristics can be changed during the flow through a long distance. Indeed, it is often noticed at the OCP that the density downstream of the network varies from that of entry. Also, the deposition of the slurry sediment on the inner wall of the pipe can affect the flow. Further, the parameters of Bingham's law that we have used are not derived from a rheological study applied on the slurry mixture conveyed by the pipeline. The total computation time is 3.6 minutes, which is quite satisfying considering the network size considering the network size of 46.43 Km.

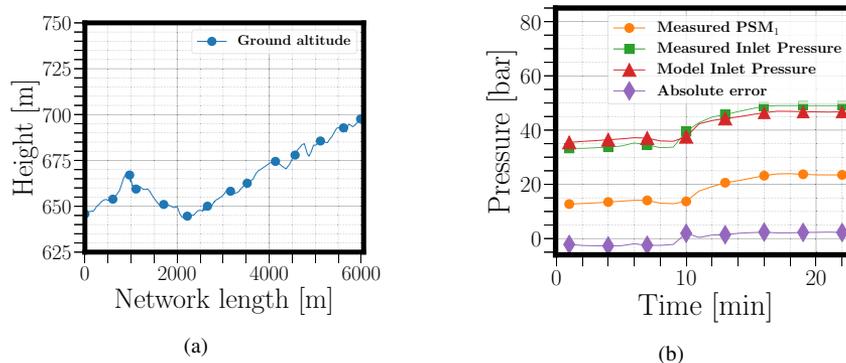


Figure 2: (a) Network altitude versus length and (b) Inlet pressure Comparison

4. Conclusion

The objective of the present work is to build a model able to simulate the flow of phosphate slurry through a pipeline of significant length. The model that has been developed presents a significant gain in computational time compared to the model already available under the Modelica Standard Library, while the flow results are not affected. Using the Bingham parameters from the industrial data fitting, we obtained estimated pipe inlet pressures close to the real values, although a mean relative error of 6% remains. This ongoing work intends to overcome this limitation by taking into account the variation of the fluid characteristics along the network and using more accurate rheology parameters to be obtained from lab measurement.

Acknowledgement: This work was supported by a research fund from OCP group.

References

- R. Darby, R. Mun, D. V. Boger, 1992. Predict friction loss in slurry pipes. *Chemical Engineering* 99 (9), 116–119.
- C. D. Ghodke, S. V. Apte, 2016. Dns study of particle-bed–turbulence interactions in an oscillatory wall-bounded flow. *Journal of Fluid Mechanics* 792, 232–251.
- S. E. Haaland, 1983. Simple and explicit formulas for the friction factor in turbulent pipe flow.
- B. O. Hedstrom, 1952. Flow of plastic materials in pipes. *Industrial & Engineering Chemistry* 44 (3), 651–656.
- L. E. M. Lima, 2020. Application of the one-dimensional drift-flux model for numerical simulation of gas–liquid isothermal flows in vertical pipes: a mechanistic approach based on the flow pattern. *SN Applied Sciences* 2 (4), 1–13.
- S. Maazioui, A. Maazouz, F. Benkhaldoun, D. Ouazar, K. Lamnawar, 2021. Rheological characterization of a concentrated phosphate slurry. *Fluids* 6 (5), 178.
- J. Rusconi, A. Lakhouaja, M. Kopuz, 2016. The design and engineering of the 187 km Khouribga to Jorf Lasfar phosphate slurry pipeline. *Procedia Engineering* 138, 142–150.
- P. K. Swamee, N. Aggarwal, 2011. Explicit equations for laminar flow of Bingham plastic fluids. *Journal of Petroleum Science and Engineering* 76 (3-4), 178–184.

Techno-Economic Analysis of the Conversion of Waste Plastics to Hydrogen Fuel

Ali A. Al-Qadri ^a, and Usama Ahmed ^{a,b*}

^a*Department of Chemical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, 31261. Saudi Arabia*

^b*Interdisciplinary Research Center for Hydrogen and Energy Storage, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia*

usama.ahmed@kfupm.edu.sa

Abstract

The need of hydrogen as a clean fuel has grabbed a lot of global attention. Therefore, various processes have been developed for hydrogen production. Conversion of plastic wastes is one of the attractive option to produce H₂ because of its hydrocarbon structure, higher heating value and availability. Polyethylene (PE) and polypropylene (PP) are considered in this study because of their massive worldwide availability in the category of waste plastics. In this study, the simulation of conversion of plastic wastes (PE and PP) to hydrogen fuel via steam gasification process is developed to perform the techno economic analysis. The composition of produced syngas from steam gasification unit was validated with the literature results followed by sensitivity studies to see the impact of various process parameters on the syngas composition. The syngas produced in the steam gasification unit is then treated in the water gas shift units followed by acid gas (H₂S and CO₂) removal to produce pure H₂. To enhance the hydrogen production capacity and overall process efficiency, plastic gasification is further integrated with the steam methane reforming coincided with gasification to utilize the energy from the gasifier to provide heat to the natural gas reformer. It has been seen from results that the new integrated design containing both gasifier and reformer enhances the hydrogen mass production rate per mass feed rate by 5.6%. Furthermore, the process performance analysis showed that the efficiency of the new process is increased up to 1.82%, where the hydrogen production cost showed the reduction of 29% compared to the standalone gasification cases. In terms of carbon dioxide specific emission, the new design showed the reduction in CO₂ emissions by 4.0%. Overall, the technical and economic analysis favored the new design over the standalone plastic gasification case.

Keywords: Gasification; Reforming; Waste Plastic; H₂ production; GHG Emissions.

1. Introduction

Globally from 1950-2018, 6.3 billion tons of plastics produced; only 9% had been recycled, where 12% had been burnt (Okunola A et al., 2019). Incinerating the plastics produces sever harmful pollutants, such as Dioxins, Furans, and Mercury, which are highly hazardous that negatively impact the creatures health, and cause climate changes . An essential, clean fuel can be produced from plastic wastes (i.e. hydrogen fuel) via gasification process. Hydrogen is more efficient than traditional fuel, because it burns easily with zero carbon emissions. Hence, a strategic way to produce hydrogen from non-considered feedstock is to convert biomass, and plastic wastes into hydrogen. A study conducted by (AlNouss et al., 2020), on techno economic evaluation of biomass into

hydrogen. They used the thermal approach in converting the biomass to hydrogen, investigating the effect of the type of feedstock and the gasifier agent. The study demonstrated the dominance of pure steam as a gasification agent than other agent. A study performed by (Chai et al., 2020), on the production of hydrogen from plastic wastes using co-pyrolysis/gasification process. They tested Ni-CaO-C as a novel catalyst to facilitate H₂ production. Their results revealed a high hydrogen production with a composition of 87.7 mole% with a minimal CO₂ production. Therefore, the conversion of plastics into hydrogen is a feasible process proved experimentally and theoretically. Although, the catalytic gasification process is energetically preferable, however, it is more complex and needs more development. Hence, we will perform our work on purely thermal gasification using steam as a gasification agent as it was recommended by (Namioka et al., 2011).

Our constituted objectives are to produce hydrogen with high purity from polyethylene and polypropylene via steam gasification controlling the greenhouse gases. We will design an alternative model that is anticipated to reduce the energy consumption and production cost. Hence, we will execute energy and economic analysis for the two models likewise to precisely compare the two models.

2. Modelling and Simulation

Aspen Plus (V11) had been used to develop the simulation models to produce hydrogen. The Peng Robinson was used as the effective thermodynamic package, where, polyethylene and polypropylene were defined based on proximate and ultimate analysis. To keep the consistent analysis, the RYield, RGibbs and REquil reactor models are used to match the results with the literature and to achieve the required compositions. Table 1 demonstrates the process conditions of the main units. The steam gasification process conditions were according to (Saebea et al., 2020)(Ahmed, 2021)(Khalafalla et al., 2020).

Table 1: Design Assumptions taken for case 1 and case 2

Equipment	Aspen Model	Assumption
Plastic flowrate	RYield	Plastics = 100 kg/h H ₂ O: Plastic = 1.25
Gasifier	RYield, RGibbs reactors	Entrained flow gasifier; temperature = 650 °C and 900 °C respectively; P = 1 atm
Pre-reformer	RStoic (Reactor)	Heavier hydrocarbon hydrocracking
Reformer	RGibbs (Reactor)	Temperature = 894.3 °C, pressure = 1 bar, H ₂ O: NG = 1.6; nickel-based catalyst
Water Gas Shift (WGS)	REquil (Reactor)	Two Equil reactors H ₂ O: Syngas = 1.09 and 0.73 respectively
Acid Gas Removal (AGR)	RadFrac and flash drums	Rectisol process; temperature = - 30 °C, P = 1 bar CO ₂ Removal = 99%; Remaining H ₂ S = 10 ppm

3. Process Description

Two cases had been designed in Aspen Plus (V11). Figure 1 shows the case 1(i.e. base case design), where, plastics were used to produce synthesis gas, and then WGS was modeled to maximize the hydrogen production. Finally, we removed the acid gases (i.e. H₂S and CO₂) in acid gas removal unit (AGR).

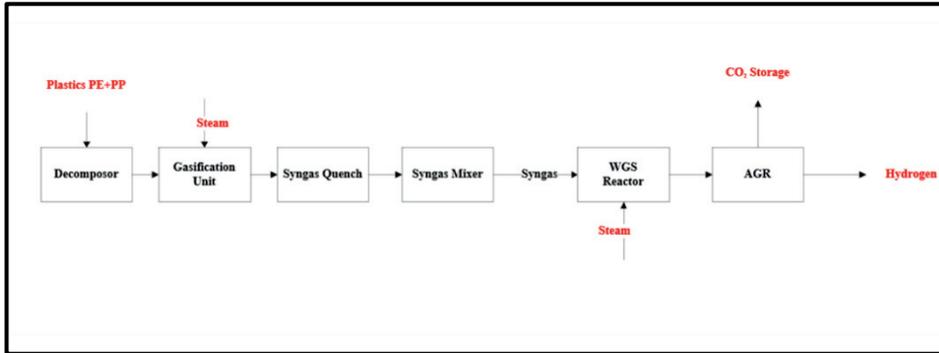


Figure 1: Production of hydrogen from PE and PP through conventional process: Case 1

Case 2 (the integrated design) is given in Figure 2. The SMR produced additional syngas with more hydrogen rich than the plastic gasification. The mixing of syngas from gasification and reforming units will yield higher hydrogen to CO ratio that can enhance the normalized hydrogen production.

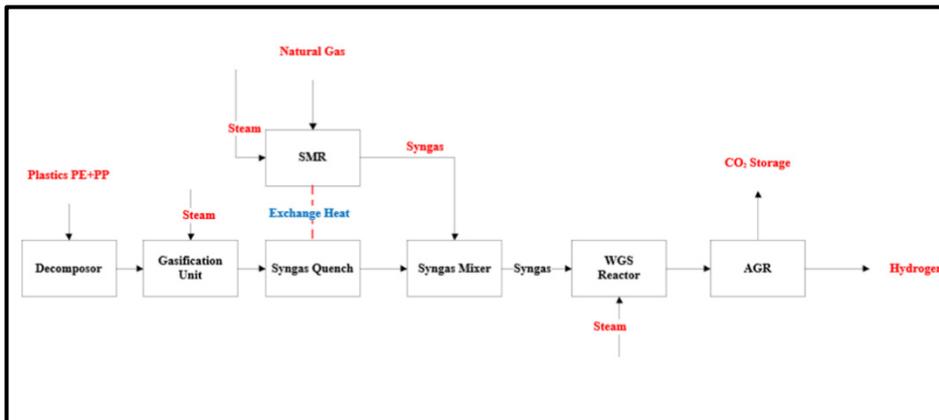


Figure 2: Production of H₂ from PE and PP via an integrated process: Case 2

4. Results and Discussion

The steam gasification results in producing syngas from polyethylene and polypropylene was validated with experimental results (Wu & Williams, 2009) (Erkiaga et al., 2013). Additionally, the SMR syngas composition was validated with (Ghoneim et al., 2016). Our results and the referenced data for syngas composition were in good agreement.

4.1 Syngas Composition at the Outlet of WGS and Overall Production Rates

Case 1 and case 2 results were analyzed in terms of synthesis gas production, heating value, CO₂ specific emissions, energy and economic analysis. The results revealed that, case 2 is superior to case 1 in production syngas with higher heating value. The LHV for case 1 and case 2 were 23.55 and 24.73 MJ/Kg, respectively. On the other hand, the synthesis gas HHV were 26.18 and 27.67 MJ/kg for the two cases correspondingly, which was higher in case 2 compared to case 1 by 5.7%. Additionally, the normalized hydrogen production per feed stock in mass basis was calculated as 50% and 52.8% for case 1 and

case 2, respectively. The integrated case showed higher production rate than the base case by 2.8%.

4.2 Process performance analysis

The overall process efficiency for both cases is calculated using the equation (Eq 1).

$$\text{Process Efficiency} = \frac{\text{H}_2 \text{ thermal energy [MWth]}}{\text{Feed stock thermal energy [MWth] + Energy consumed [MWth]}} \times 100\% \quad (1)$$

The feed stock energy for case 1 included only waste plastics (i.e. PE and PP), however, for case 2 it was combined feed of natural gas and plastics. The feedstock thermal energy for the two cases were 1198.61, and 1757.07 kW respectively. The total consumed energy in the whole design was calculated and deduced to be 3162.28 and 4641.99 kW respectively. Hence, the process efficiency was found to be 31.32% and 31.90% respectively. Figure 3 represents the overall process efficiency and hydrogen to carbon monoxide ratio. It was found that, the H₂/CO ratio was 1.86 and 2.2 for case 1 and case 2 respectively. Those results revealed the superiority of the integrated case than the base case. We also considered the CO₂ specific emissions for the two models which were 0.26 and 0.257. Hence, the carbon dioxide emissions was reduced in case 2 by 1.2%.

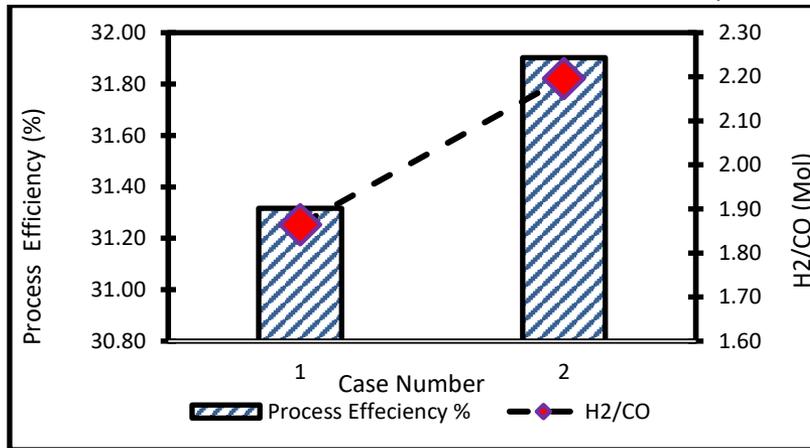


Figure 3: Process efficiency and H₂/CO ratio for case 1 and case 2

5. Economic Analysis and Project Feasibility

The economic analysis was performed for the two cases evaluating the capital expenditures and operating expenditures. The CAPEX mainly encompasses the equipment cost, piping and installation costs, civil work, instrumentation etc. On the other hand, the operating expenditures includes the utilities, maintenance, catalyst replacement, administration, and labor costs. The calculation for CAPEX was based on process costs from previous studies considering the capacity and CEPCI index. The value of x is taken as 0.6 in the Equation 2, where the current CEPCI is taken as 620.

$$\text{Cost}_{\text{New}} = \text{Cost}_{\text{Old}} \times \left(\frac{\text{Capacity}_{\text{New}}}{\text{Capacity}_{\text{Old}}} \right)^x \times \frac{\text{CEPCI}_{\text{New}}}{\text{CEPCI}_{\text{Old}}} \quad (2)$$

For the comparative analysis, the equation 2 is well used in the literature as it can predict the class 3-4 type economic analysis. The economic analysis results revealed that, the total investment cost for case 1 and case 2 in terms M€ were 3.795 and 4.457 respectively. Case 2 required higher investment cost because it involved the SMR process. However,

the total investment cost per hydrogen production in tons (TIC) were 76.53 and 59.25 M€/ton correspondingly. Case 2 reduced the TIC with respect to case 1 by 22.6%. We additionally, calculated the total OPEX per year, and found that, case 2 consumed higher operating expenditures than case 1 by 79.3 in €(10³) / year. However, the operating expenditures per hydrogen production rate for case was lower than case 1 by 32.35%. The normalized hydrogen production rate; which represents the conversion of feed stocks to produce H₂, was higher in case 1 which is attributed in reducing the total OPEX per hydrogen ton. Table 2 displays the main cost analysis parameters for case 1 and case 2.

Table 2: Economic Analysis

	Units	Case 1	Case2
Total Investment Cost	€(10 ³)	3795	4457
TIC per ton of H₂	M€/ton	76.53	59.25
Total OPEX/Year	€(10 ³) / Year	1390.0	1469.3
Total OPEX/ton H₂	€(10 ³) / ton	3.4	2.3
Revenue	M€/year	4.804	7.289
Hydrogen Cost	€/kg	3.675	2.585

The two cases were also evaluated in terms of cash flow diagram. The cash flow revealed that, case 2 is superior to case 1 and the profit is higher. The cash flow return on investment was higher in the integrated case than the conventional case by 52%. Additionally, NPV (the net present value) was larger than the base case by 78%, where PVR (the present value ratio) was higher by 45% compared to the base design. Therefore, the economic analysis proved that, case 2 is much feasible than case 1. It reduced the production cost and it is a promising design to produce hydrogen fuel from waste plastics.

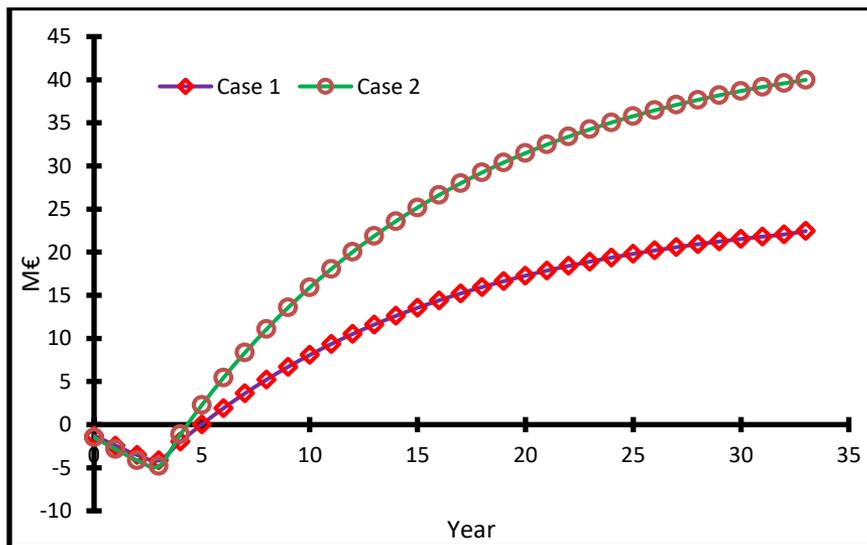


Figure 4: The cash flow diagram for case 1 and case 2

Conclusions

The modeling of production of H₂ from plastic wastes (polyethylene and polypropylene) has been done by developing two models. The energy and economic analysis for the two models conducted. The normalized hydrogen production per feedstock for case 2 was higher than case 1 by 5.6%. Additionally, the heating value, process efficiency, and H₂/CO ratio at the inlet of WGS reactors were estimated. It has been analyzed that case 2 outperformed the case 1 in all the technical and economic evaluations. Moreover, the comparison in terms of CO₂ specific emissions has been conducted, and the results revealed an emission reduction by 1.0%. Additionally economic analysis revealed that the alternative case (2) has lower TIC and fuel production cost per kg of H₂ produced compared to the case 1 by 23%, and 30% respectively. Overall, case 2 showed more promising results than case 1, and it may be experimentally tested in future.

Acknowledgements

The authors would like to acknowledge the support provided by the Deanship of Scientific Research (DSR) at King Fahd University of Petroleum & Minerals (KFUPM) for funding this work through project No. DF201017.

References

- Ahmed, U. (2021). Techno-economic analysis of dual methanol and hydrogen production using energy mix systems with CO₂ capture. *Energy Conversion and Management*, 228. <https://doi.org/10.1016/J.ENCONMAN.2020.113663>
- AlNouss, A., McKay, G., & Al-Ansari, T. (2020). Enhancing waste to hydrogen production through biomass feedstock blending: A techno-economic-environmental evaluation. *Applied Energy*, 266. <https://doi.org/10.1016/j.apenergy.2020.114885>
- Chai, Y., Gao, N., Wang, M., & Wu, C. (2020). H₂ production from co-pyrolysis/gasification of waste plastics and biomass under novel catalyst Ni-CaO-C. *Chemical Engineering Journal*, 382, 122947. <https://doi.org/10.1016/J.CEJ.2019.122947>
- Erkiaga, A., Lopez, G., Amutio, M., Bilbao, J., & Olazar, M. (2013). Syngas from steam gasification of polyethylene in a conical spouted bed reactor. *Fuel*, 109, 461–469. <https://doi.org/10.1016/j.fuel.2013.03.022>
- Ghoneim, S. A., El-Salamony, R. A., & El-Temtamy, S. A. (2016). Review on Innovative Catalytic Reforming of Natural Gas to Syngas. *World Journal of Engineering and Technology*, 04(01), 116–139. <https://doi.org/10.4236/wjet.2016.41011>
- Khalafalla, S. S., Zahid, U., Abdul Jameel, A. G., Ahmed, U., Alenazey, F. S., & Lee, C. J. (2020). Conceptual Design Development of Coal-to-Methanol Process with Carbon Capture and Utilization. *Energies* 2020, Vol. 13, Page 6421, 13(23), 6421. <https://doi.org/10.3390/EN13236421>
- Namioka, T., Saito, A., Inoue, Y., Park, Y., Min, T. J., Roh, S. A., & Yoshikawa, K. (2011). Hydrogen-rich gas production from waste plastics by pyrolysis and low-temperature steam reforming over a ruthenium catalyst. *Applied Energy*, 88(6), 2019–2026. <https://doi.org/10.1016/J.APENERGY.2010.12.053>
- Okunola A, A., Kehinde I, O., Oluwaseun, A., & Olufiropo E, A. (2019). Public and Environmental Health Effects of Plastic Wastes Disposal: A Review. *Journal of Toxicology and Risk Assessment*, 5(2). <https://doi.org/10.23937/2572-4061.1510021>
- Saebea, D., Ruengrit, P., Arpornwichanop, A., & Patcharavorachot, Y. (2020). Gasification of plastic waste for synthesis gas production. *Energy Reports*, 6, 202–207. <https://doi.org/10.1016/j.egy.2019.08.043>
- Wu, C., & Williams, P. T. (2009). Hydrogen production by steam gasification of polypropylene with various nickel catalysts. *Applied Catalysis B: Environmental*, 87(3–4), 152–161. <https://doi.org/10.1016/j.apcatb.2008.09.003>

Evaluating the flexible operation of vacuum-pressure swing adsorption for CO₂ capture from modern gas turbines

Mathew Dennis Wilkes ^a, Solomon Brown ^{a*}.

^a*Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom, s.f.brown@sheffield.ac.uk*

Abstract

Flexible post-combustion capture (PCC) of CO₂ is crucial for reaching net-zero emissions targets. Within this paper, a process model for CO₂ capture via vacuum-pressure swing adsorption (VPSA) is developed, validated, and scaled to process flue gas from a small-scale modern gas turbine. The realistic load-following ‘Flexible’ scenario shows small differences in CO₂ recovery rate, purity, and specific energy demand, compared to the ‘Baseline’ scenario. This highlights VPSA is a suitable technology for Flexible- PCC.

Keywords: CO₂ Adsorption, VPSA, PCC, Flexible Operation, Dynamic Modelling

1. Introduction

Electricity systems require balancing capacity to ensure security of supply and counter-act imbalances due to intermittent renewables and inflexible nuclear (Rai, et al., 2022). Open-cycle gas turbines (OCGT) are quick-response dispatchable generators, that are expected to increase in capacity between 2020-2050 (Heuberger & Mac Dowell, 2018). Therefore, OCGTs that incorporate Carbon Capture, Utilization, and Storage (CCUS) may play a vital role in achieving net-zero by 2050.

Dispatchable power generation by nature is highly transient, and any post-combustion capture (PCC) technology will also need to operate flexibly. The majority of Flexible-PCC revolves around the benchmark MEA CO₂ capture (Wilkes, et al., 2021a). However, adsorption technologies are attractive due to high CO₂ recovery rates and purity, as well as lower specific energy demands compared to MEA (Ben-Mansour, et al., 2016).

1.1. Aims and Objectives

In this study we show how CO₂ can be captured through vacuum-pressure swing adsorption (VPSA) in order to clean highly transient gas turbine exhaust. This study does not investigate the optimized design of the VPSA CO₂ capture system. The aim of this paper is, instead, to highlight the effects transient flue gas production has on the process’s key performance indicators. The OCGT assumed flue gas flowrate is based on our previous paper (Wilkes, et al., 2021a). Hence, within this flexible-VPSA study:

- A process model for VPSA is developed in gPROMS Process.
- Model validation is carried out against pilot-scale data.
- The process model scaled to handle exhaust flow from a small OCGT.
- Baseline and Flexible results are shown for a typical 5-hour OCGT operation.

This study only considers the operation of the capture plant, and not the design of the gas turbine or the CO₂ conditioning system. For information on the effect of transient operation on the conditioning system see (Wilkes, et al., 2021b).

2. Adsorption modelling

Carbon capture by adsorption (CCA) is an inherently dynamic process considered as being in cyclic steady state. Each adsorption column requires a set of partial differential and algebraic equations (PDAEs) to calculate the quantity of CO₂ adsorbed, whilst considering the conservation of mass, energy, and momentum (Kikkinides, et al., 2010). Li et al. (2018) provides an extensive review of CCA modelling. Within this study the following assumptions are made to model Flexible-VPSA:

- One-dimensional axially dispersed plug flow regime.
- No radial variation in temperature, pressure, or concentration.
- Mass transfer described through the linear driving force model.
- Adsorption amount described through the dual-site Langmuir isotherm model.
- Pressure drop along the column is calculated through the Ergun equation.
- Thermal-equilibrium heat transfer model.
- Power demand for the pumps is described in (Nikolaidis, et al., 2017).

As there are multiple CO₂ capture technologies and various sorbents for CCA, it is important to quantify certain key performance indicators (KPI) to enable comparison to alternative processes. Important KPI's for CCA technologies that are commonly used in the literature include CO₂ recovery rate (η), specific energy demand (E_T), and productivity (P_r), calculated using Equation 1, 2, and 3, respectively.

$$\eta [\%] = \frac{M_{CO_2}^{in} - M_{CO_2}^{out}}{M_{CO_2}^{in}} \times 100 \quad 1$$

$$E_T [kWh/tCO_2] = \frac{\text{Total Energy Demand [kWh]}}{M_{CO_2}^{in} - M_{CO_2}^{out} [\text{ton}]} \quad 2$$

$$P_r [tCO_2/m^3 \text{ Adsorbent/day}] = \frac{M_{CO_2}^{in} - M_{CO_2}^{out} [\text{ton}]}{V_A [m^3] \times t_c [s] \times 3600} \quad 3$$

Where M_{CO_2} is the mass of CO₂ either entering (superscript 'in') or exiting (superscript 'out') the process. The total energy demand is the sum of the power supplied to each of the blowers and vacuum pumps divided by the total amount of CO₂ captured. The productivity is the quantity of CO₂ captured by the adsorbent during a given time frame, where V_A is the volume of adsorbent and t_c is the total cycle operating time.

2.1. Model validation

Pilot-scale results from Krishnamurthy et al. (2014) are used to highlight the fidelity of the VPSA process model. The pilot experiment investigated CO₂/N₂ VPSA using Zeolite 13X in a single adsorption column. The Skarstrom style four-step cycle includes: pressurization using the flue gas feed, adsorption of CO₂ onto Zeolite 13x, forward blowdown to de-pressurize the column, and reverse evacuation to desorb and recover CO₂. The step timings and pressure levels are highlighted in Figure 1. The inlet feed flowrate was 1000 SLPM and contained 85 mol.% N₂ and 15 mol.% CO₂. The column geometries, particle characteristics, isotherm data and heat transfer parameters can be found in (Krishnamurthy, et al., 2014) and (Haghpahan, et al., 2013).

The KPI's for the pilot experiment and simulation results are shown in Table 1. The simulation showed almost identical column pressure increase and flowrate profiles (not shown in this paper) compared to the pilot results. As the model is pressure driven, the

accuracy of the pressure profile is important in calculating the quantity of CO₂ adsorbed. The simulation's recovery rate (84.30%) is within the error of the pilot results (85.40% ± 4.52). The purity (89.56%) is lower than expected (94.70% ± 1.05). It can be manipulated by the evacuation valve flow coefficient, but this is a set input parameter that cannot be altered specifically during the evacuation step. The lower purity also leads to an increased specific energy demand, as less CO₂ is captured during the operation. Overall, the model is in good agreement with pilot results and the isotherm model accurately predicted the amount of CO₂ adsorbed onto the zeolite surface.

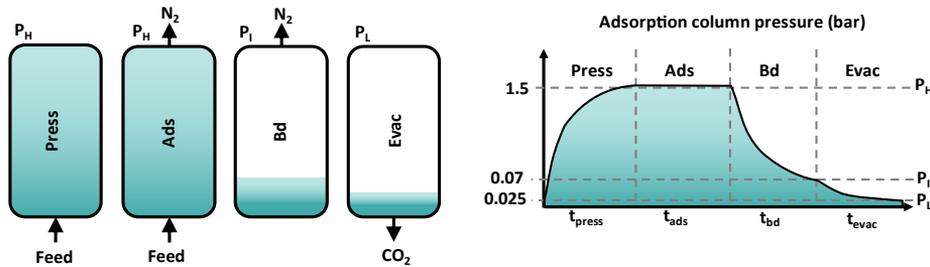


Figure 1: Single column VPSA operation, step notation: "Press" is pressurization to 1.5 bar, "Ads" is adsorption at 1.5 bar, "Bd" is blowdown to 0.07 bar, and "Evac" is evacuation to 0.025 bar (Krishnamurthy, et al., 2014).

Table 1: Key performance indicators for the single column VPSA model and pilot experiment

KPI	Pilot	Simulation	Deviation (%)
η (%)	85.40 ± 4.52	84.30	- 1.29
Purity (%)	94.70 ± 1.05	89.56	- 5.14
E_T (kWh/tCO ₂)	510.50 ± 25.5	569.68	+ 11.70
P_r (tCO ₂ /m ³ ads/day)	1.40 ± 0.07	1.18	- 15.00

2.2. Scaled-VPSA design

The single column model is scaled up to handle the flue gas flow from a small 10.4 MWe open-cycle gas turbine (OCGT). The flexible OCGT operation is described and analyzed in our previous study (Wilkes, et al., 2021a). The scaled system design is based on (Luberti, et al., 2017), where the flue gas source (FGS) is split (FGS1 and FGS2) and processed in two parallel VPSA units, to handle the large flue gas flowrate. The model assumes the flue gas is cooled and treated prior to CO₂ capture. Each VPSA unit consists of two beds operating simultaneously for continuous flue gas processing. Table 2 shows the adsorption steps and timings used in the scaled-VPSA process model.

Table 2: Scaled-VPSA adsorption steps and timings

Step	Time (Sec)	Pressure (bar)
Pressurization	80	1.5
Adsorption	220	1.5
Blowdown	80	0.07
Evacuation	220	0.025

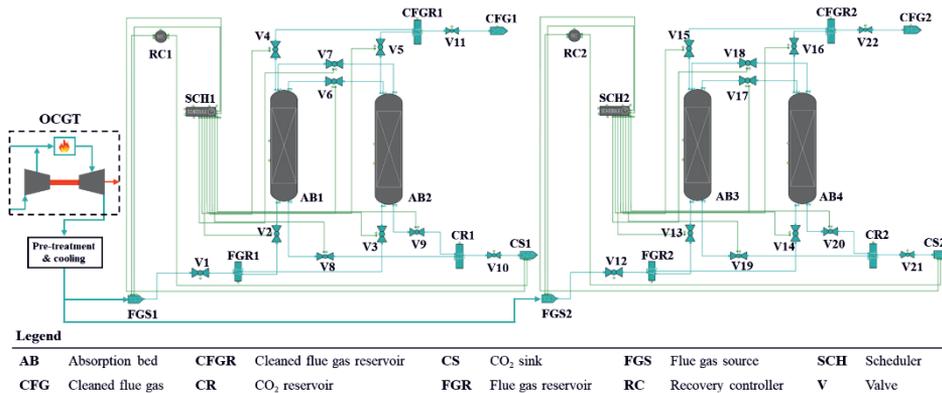


Figure 2: Scaled-VPSA model topology and legend, including cycle steps and durations.

3. Results and discussion

For the Baseline and Flexible scenarios, the flue gas input (FGS1 and FGS2) into each identical VPSA unit is 16.9 kg/s, shown in Figure 3A and Figure 3B. As the total cycle time is 600 seconds and the total operating period is 5 hours (18,000 seconds), the operation requires 30 complete cycles. During the Flexible scenario the start-up and shutdown operations are included, extending the operating period from 18,000 to 18,480 seconds. Also included in the Flexible scenario is ramping to different OCGT power outputs. At 3,600 seconds the flue gas ramps to 70% load. At 10,800 seconds the flue gas ramps to 50% load. This assumes the flue gas and power output are directly proportional. The flowrate profiles for the CFG and CS streams are similar to the pilot results shown in (Krishnamurthy, et al., 2014), where there is an initial spike in flowrate due to the pressure draw, which then equalizes during each operating step. When one bed is pressurizing and adsorbing, the other bed is evacuating N₂ and CO₂ to allow for continuous operation. Due to computational limitations the simulations cannot calculate adsorption characteristics with zero flow. Therefore, during the start-up operation in the Flexible scenario the FGS1 and FGS2 flowrates ramp from 0.9 to 16.9 kg/s in 8 minutes (480 seconds), the inverse occurs during the shutdown operation.

The adsorption bed pressure profiles (see Figure 4A and Figure 4B) also exhibit similar results as the pilot study. During the low-load points in the Flexible scenario, as less flow enters the columns, the bed pressure does not increase to the level shown in the Baseline scenario. As the process is pressure driven, the lower bed pressure should affect the quantity of CO₂ adsorbed. However, the decrease in recovery rate between the Flexible and Baseline scenario (see Table 3) is only 0.04%. It is worth noting that, the scaled-VPSA Baseline recovery rate is 97.07% and 13.66% higher than the pilot study. This increase in efficiency is related to the adsorption step timings, column sizing, valve operation and parameter scaling. This study does not focus on optimized process scaling, the aim is to assess the performance under transient flue gas conditions.

At low load periods (70% and 50%) the blower and vacuum pumps require less energy (processing less fluid), however, the quantity of CO₂ captured is smaller. Therefore, during the Flexible scenario the specific energy demand is 1274.04 kWh/tCO₂, 6.91% higher than the Baseline scenario at 1191.72 kWh/tCO₂. A possible solution is modifying the adsorption step timings and pressure levels specifically during low load operation, to minimize fluctuations in specific energy demand.

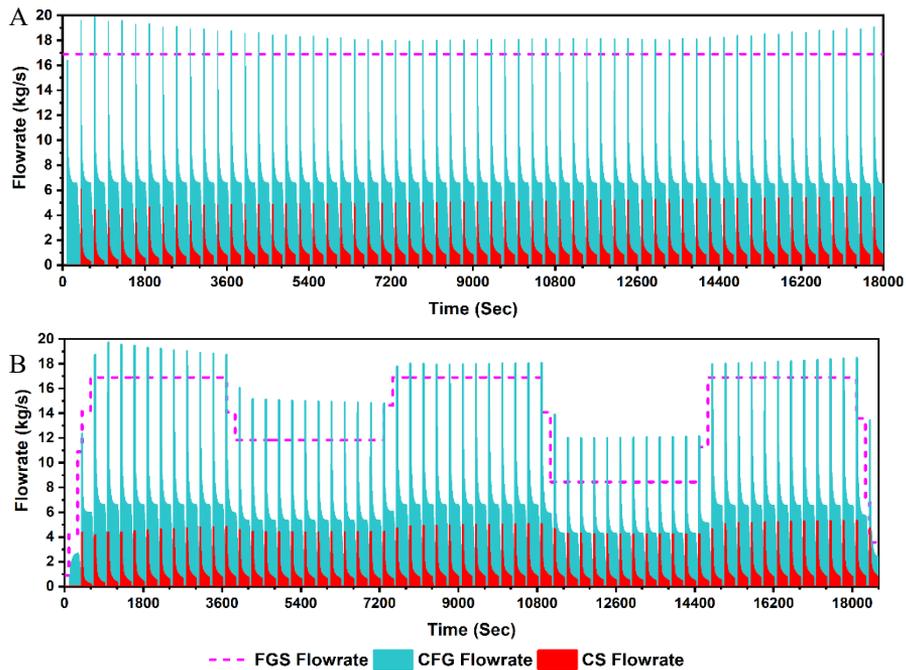


Figure 3: Flue gas source (FGS), cleaned flue gas (CFG) and CO₂ sink (CS) flowrate profiles during the Baseline (A) and Flexible (B) scenarios. for one of the parallel VPSA units.

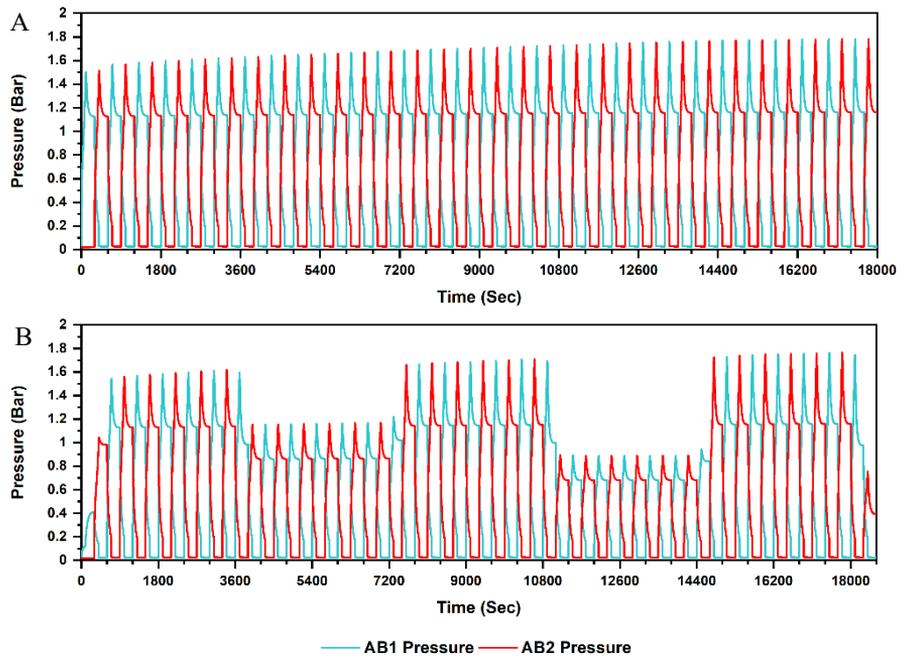


Figure 4: Column pressure for AB1 and AB2 during the Baseline (A) and Flexible (B) scenarios.

Table 3: Scaled-VPSA key performance indicators

KPI	Baseline	Deviation ^a (%)	Flexible	Deviation ^b (%)
η (%)	97.07	+13.66	97.03	-0.04
Purity (%)	80.74	-14.74	79.54	-1.49
E_T (kWh/tCO ₂)	1191.72	+133.67	1274.04	+6.91
P_r (tCO ₂ /m ³ ads/day)	1.28	-9.02	1.07	-15.85

^a deviation between the baseline and pilot results, ^b deviation between Baseline and Flexible scaled-VPSA scenarios

4. Conclusion

To attain net-zero by 2050, small-scale dispatchable gas power generation will require CCUS. Current research in Flexible-PCC revolves around amine-based CO₂ capture. This study considered VPSA as an alternative CO₂ capture technology for modern and highly transient OCGT's. The process model is found to have a high fidelity, with the validation study showing the predicted capture rate is within the error bounds stated in the pilot results. During Flexible-VPSA, the recovery rate and purity deviate <2% compared to the Baseline (no change in flue gas) scenario. The specific energy demand increased during periods at 70% and 50% load (+6.91% overall), as less CO₂ is captured during the total operation. This is lower than Flexible-PCC using MEA (+18.53%) shown in our previous work (Wilkes, et al., 2021a), highlighting VPSA is a viable alternative technology for Flexible-PCC.

5. References

- Ben-Mansour, R. et al., 2016. Carbon capture by physical adsorption: Materials, experimental investigations and numerical modeling and simulations – A review. *Applied Energy*, Volume 161, pp. 225-255.
- Haghpahan, R. et al., 2013. Multiobjective Optimization of a Four-Step Adsorption Process for Postcombustion CO₂ Capture Via Finite Volume Simulation. *Industrial & Engineering Chemistry Research*, 52(11), p. 4249–4265.
- Heuberger, C. F. & Mac Dowell, N., 2018. Real-World Challenges with a Rapid Transition to 100% Renewable Power System. *Joule*, 2(3), pp. 367-370.
- Kikkinides, E. S., Nikolic, D. & Georgiadis, M. C., 2010. Modeling of Pressure Swing Adsorption Processes. In: M. C. Georgiadis, J. R. Banga & E. N. Pistikopoulos, eds. *Dynamic Process Modelling*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, pp. 137-172.
- Krishnamurthy, S. et al., 2014. CO₂ Capture from Dry Flue Gas by Vacuum Swing Adsorption: A Pilot Plant Study. *AIChE Journal*, 60(5), pp. 1830-1842.
- Li, S. et al., 2018. Mathematical modeling and numerical investigation of carbon capture by adsorption: Literature review and case study. *Applied Energy*, Volume 221, pp. 437-449.
- Luberti, M., Oreggioni, G. D. & Ahn, H., 2017. Design of a rapid vacuum pressure swing adsorption (RVPSA) process for post-combustion CO₂ capture from a biomass-fuelled CHP plant. *Journal of Environmental Chemical Engineering*, 5(4), pp. 3973-3982.
- Nikolaidis, G. N., Kikkinides, E. S. & Georgiadis, M. C., 2017. Modelling and Optimization of Pressure Swing Adsorption (PSA) Processes for Post-combustion CO₂ Capture from Flue Gas. In: A. I. Papadopoulos & P. Seferlis, eds. *Process Systems and Materials for CO₂ Capture*. s.l.:John Wiley & Sons Ltd, pp. 343-369.
- Rai, U., Oluleye, G. & Hawkes, A., 2022. An optimisation model to determine the capacity of a distributed energy resource to contract with a balancing services aggregator. *Applied Energy*, 306(Part A), p. 117984.
- Wilkes, M. D., Mukherjee, S. & Brown, S., 2021a. Transient CO₂ capture for open-cycle gas turbines in future energy systems. *Energy*, p. 119258.
- Wilkes, M. D., Mukherjee, S. & Brown, S., 2021b. Linking CO₂ capture and pipeline transportation: sensitivity analysis and dynamic study of the compression train. *International Journal of Greenhouse Gas Control*, Volume 111, p. 103449.

Exergoeconomic assessment of the optimised vapour-recompression assisted column for palm-based fatty acid fractionation

Norul M. Sidek,^{a,b*} Mohamad R. Othman,^{a,b}

^a*Process Systems Engineering & Safety Research Group, Faculty of Chemical & Process Engineering Technology, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.*

^b*Department of Chemical Engineering, College of Engineering, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.*

Abstract

In our preliminary study (Sidek and Othman, 2020), vapour recompression (VRC) proved to have a promising energy performance in the case of palm kernel oil (PKO) fatty acid fractionation. This paper extends the research by applying an exergy analysis as well as incorporating four vacuum fractionation columns to represent a real industry application. The study also included a two-step optimization approach based on response surface methodology. All configurations were evaluated for their exergetic and economic performances. Due to the high costs and low efficiency, the optimised VRC-assisted columns are not in favour of this study. The standard VRC-assisted columns, though exhibit a higher CAPEX than the conventional columns, are more thermodynamically efficient and demonstrate a substantial reduction of 84-88% in OPEX. A thorough optimisation study is needed for future works.

Keywords: vapour recompression, fatty acid fractionation, optimization, exergy.

1. Introduction

Since the pandemic, the oleochemical industry has witnessed greater product demand due to increased hygiene needs. Specific measures must be implemented to ensure a continuous supply of oleochemical products to consumers. The industry's primary separation units are distillation columns, which are used to separate oleochemical cuts. In practice, distillation columns are well-known as major energy consumers. In recent years, research on heat pump-assisted distillation has recently piqued the interest of many academics to reduce the associated energy consumption. However, no research has been devoted to heat pumping in oleochemical distillation. Besides, the majority of established oleochemical refineries in Malaysia still employ mature distillation technologies. In light of these circumstances, oleochemical distillation appears to be a prime focus for improvement. In a prior study, we discovered that vapour recompression (VRC), a popular heat pump scheme, showed promising energy performance in the case of palm kernel oil (PKO) fatty acid fractionation (Sidek and Othman, 2020). The research, although attempting to move towards a more sustainable approach, is still preliminary. The reason is that to facilitate convergence, only one fractionation column was considered to evaluate the feasibility of VRC. In an industrial operation, the PKO fatty acid fractionation involves four operational columns, which include a pre-fractionator for recovering C8-C10 products, a light-cut column for recovering C12, a middle-cut column

for separating C14 as a top product, and a still column for heavy-cut separation (C16-C18) (Othman and Rangaiah, 2020). Another aspect that is yet unknown from the study is the effect of thermodynamic conditions on the simulated VRC-assisted column. Using an exergy analysis, Mangili (2020) presented that a VRC-assisted pressure swing column for di-*n*-propyl-ether and *n*-propanol processes has outperformed another direct heat-integrated column in terms of exergetic efficiency. Intriguingly, a comprehensive study of a VRC-assisted vacuum membrane distillation for sulfuric acid waste treatment by Si et al. (2020) has shown a very low exergetic efficiency of only about 4%. Since the thermodynamic performance of each VRC-assisted column varies depending on the process and operating conditions, a thermodynamic analysis is, therefore, essential to evaluate how much of recovered heat is transformed into useable work. In this paper, the study continues to analyse all operational columns, exactly replicating the industrial operation, from the standpoints of exergetic and economic performance. All process simulations were carried out in Aspen PLUS V12.0 using UNIQUAC thermodynamic model, as recommended by Sidek and Othman (2019) for PKO fatty acid. In Design Expert, the simulated columns were optimised using the response surface methodology.

2. Methodology

2.1. Process Simulation

The PKO fatty acid compositions were adopted from a Malaysian industrial oleochemical refinery, as shown in Othman and Rangaiah (2020). Unlike our previous studies (Sidek and Othman, 2020), this one utilised four (4) major units of fractionation column at a high vacuum pressure to represent the actual process. A process with a crude PKO fatty acid feed flow rate of 9167 kg/h was operated at 80 mbar in a pre-fractionator (V5001), 25 mbar in a light-cut column (V5002), 13 mbar in a middle-cut column (V5003), and 10 mbar in a heavy-cut column (V5004). To avoid chemical degradation, the vacuum application helped to regulate vapour pressure and ensure the working temperature was between 180 and 250 °C. To ensure a fair comparison, a VRC-assisted column configuration was simulated based on the same column conditions and desired product specifications as the CCs. The VRC-assisted column works by compressing the top vapour stream to make it hotter than the reboiler. The elevated vapour temperature allows the utilisation of recovered latent heat for bottom liquid reboiling (Parhi et al., 2019). To achieve that, the vapour stream was first superheated before being introduced into a compressor to avoid partial condensation. A maximum practical compression ratio was taken to be 3.0. The compressed vapour will heat exchange with the bottom liquid stream through a heat exchanger. A minimum temperature approach was kept between 5-10 °C.

2.2. Process Optimisation

In this study, the experimental design based on response surface methodology (RSM) was carried out, which included two-level factorial analysis (TLFA) and central composite design (CCD). RSM develops both statistical and mathematical models to provide adequate predictions of a defined response. Prior to running the TLFA, the variables were first obtained from a sensitivity analysis in Aspen PLUS. For the optimisation part, CCD was primarily chosen to allow the estimation of all regression variables that were required to fit a second-order model to the response. An adequacy test for the developed model was observed in the analysis of variance (ANOVA), whereby the *p*-value must be less than 0.05 and the adequate precision value must be larger than 4. Meanwhile, for chemical processes, the *R*² value must be larger than 0.9. The soundness of the model was determined through a surface contour plot. The optimal conditions must be validated in Aspen Plus to assure convergence.

2.3. Exergy Analysis

Exergy analysis measures how much valuable energy is being neglected in a certain system. A more detailed explanation of exergy computation in the distillation column has been demonstrated by Nguyen and Demirel (2010). According to them, the exergy difference between the product and the feed stream indicates the minimum isothermal separation work which corresponds to Eq. (1).

$$\dot{E}_{x,min} = \sum_{out} \dot{n}e_x - \sum_{in} \dot{n}e_x \quad Eq. (1)$$

where \dot{n} is the mole flow and e_x is the specific exergy of the streams. An exergy loss, $\dot{E}_{x,loss}$, or irreversibility can be computed from Eq. (2).

$$\dot{E}_{x,loss} = \sum_{in} \left(\dot{n}S + \frac{Q_{sink}}{T_{sink}} \right) - \sum_{out} \left(\dot{n}S + \frac{Q_{source}}{T_{source}} \right) \quad Eq. (2)$$

Here, S represents the entropy of a stream and Q is the heat involved with the process. Meanwhile, the conversion of the supplied heat into separation work can be calculated by a thermodynamic efficiency, η using Eq. (3).

$$\eta = \frac{E_{x,min}}{E_{x,loss} + E_{x,min}} \quad Eq. (3)$$

2.4. Economic evaluation

The economic performance of the VRC-assisted columns was evaluated using capital expenses (CAPEX) and operational expenses (OPEX). The plant was assumed to be in operation for approximately 10 years, with 8,000 hours of operation per year. The electricity tariff for the industry was \$0.084/kWh, high-pressure (HP) steam was \$17/t and cooling water was \$0.06/t (Parhi et al., 2019).

3. Results and Discussions

3.1. VRC-assisted fractionation columns

A schematic representation of the proposed VRC-assisted fractionation process for PKO fatty acid is depicted in Fig. 1. In the studied configuration, two columns namely V5002 and V5003 were determined to be the ideal candidates for integration with the VRC system. The V5001 and V5004 were unable to deliver at least 10% of heat duty for the evaporation of the bottom liquid streams even when a maximum practicable compression ratio (CR) of 3.0 was used in a two-stage compression. The proposed VRC system for both V5002 and V5003 exhibits two compression stages to a) comply with the maximum CR, b) fulfil the reboiler's entire heat requirement and c) ensure the outlet temperature is

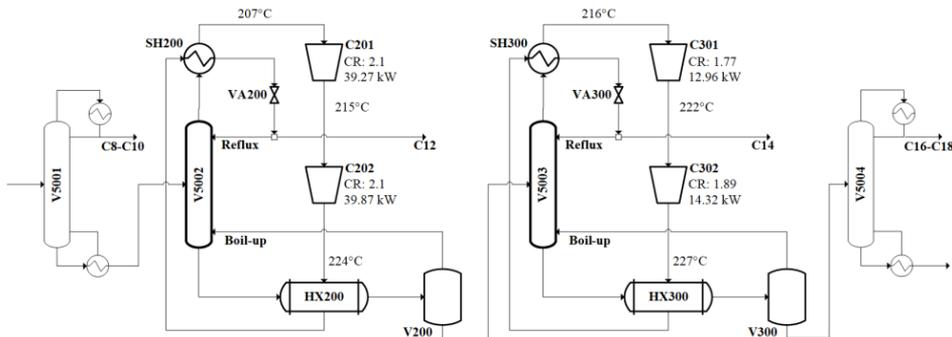


Fig. 1. VRC-assisted fractionation process for PKO fatty acid

sufficiently high for heat exchange. The overhead vapour stream of V5002 was superheated prior to being fed into the first compressor, C201. This procedure is crucial not only for preventing liquid rupture in the compressors but also for avoiding the use of higher CRs. Condensation of the C202 outlet vapour stream was accomplished by transferring heat to the heat exchanger, HX200. In contrast to our earlier work, the compressed vapour of the V5002 column effectively delivered the necessary heat to evaporate the bottom liquid stream, obviating the need for a trim reboiler. Furthermore, the resulting condensate from HX200 was used to superheat the overhead vapour in the SH200, thus, no auxiliary heating was required in this column. Following that, the stream was pressure reduced by the throttle valve, VA200, to the column operating pressure and separated into reflux and distillate sections. Meanwhile, the bottom liquid was routed to the flash drum, V200, whereby the vapour was returned to the column as a boil-up and the liquid was introduced to the V5003. The identical VRC configuration was simulated in the V5003, except that the two compressors, C301 and C302, had lower CRs and hence performed less mechanical work.

3.2. Optimisation using sensitivity analysis and RSM

Although the compressed overhead vapour could satisfy the heating requirements of the V5002 and V5003, there were some deviations in the heat exchanger's minimum temperature approach, ΔT_{\min} . Technically, a small ΔT_{\min} requires a larger heat exchanger to transfer heat. In the sensitivity analysis, the superheated temperature, CR, and exchanger duty were identified as the key factors in the VRC performance. These factors were then investigated as independent variables for their impacts on the ΔT_{\min} in RSM optimisation alongside mechanical compression work. A 2^4 fractional factorial design with 16 runs was used to analyse all variables. However, only 12 of the 16 proposed runs showed convergence in the simulation. This resulted in a slight effect in the interaction prediction between the variables. The results for the V5002 and V5003 optimisations agree with all of the TLFA outputs necessary to demonstrate that the models were a good fit and significant. One solution for optimal conditions from the TLFA was selected for the centre points in the CCD. The two most contributing factors, namely superheated temperature and second stage CR, were used, while other variables were held fixed. Fig. 2a and Fig. 2b illustrate the 3D surface contour plots for the V5002 and V5003.

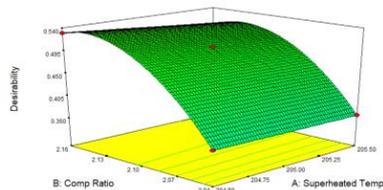


Fig. 2a. 3D surface plot for V5002

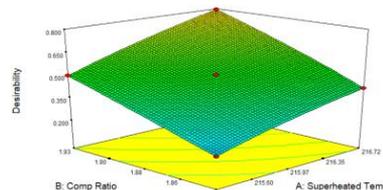


Fig. 2b. 3D surface plot for V5003

The form of the plot for V5002 shows some curvature, suggesting that the factors had a quadratic effect on the ΔT_{\min} , thus the optimum conditions could be determined from the model. The V5003 plot has a relatively flat surface but is still acceptable. For the V5002, the proposed optimal points are the superheated temperature of 204.5 °C, first stage CR of 2.29, second stage CR of 2.14 and exchanger duty of 690 kW. Whereas, the proposed optimal points for the V5003 are the superheated temperature of 216.72 °C, first stage CR of 2.10, second stage CR of 1.93 and exchanger duty of 340 kW.

3.3. Energy and exergy performances

The V5002 and V5003 utility requirements, as well as the findings of the exergy analysis of conventional columns, VRC-assisted columns, and optimised VRC-assisted columns,

are summarised in Table 1. The table clearly indicates that the CC's energy requirements were excessive in contrast to the VRC-assisted columns.

Table 1. Summary of utility requirements and exergy analysis.

Configuration	CC		VRC		Optimised-VRC	
	V5002	V5003	V5002	V5003	V5002	V5003
Cold utility (kW)	924.87	413.53				
Hot utility (kW)	766.58	378.31			76.58	38.31
Mech. work (kW)			79.14	27.27	84.20	31.66
Total (kW)	1691.45	791.84	79.14	27.27	160.78	69.97
Compression ratio 1			2.10	1.89	2.29	2.10
Compression ratio 2			2.10	1.77	2.14	1.93
Thermodynamic efficiency (%)	78.55	97.66	93.93	98.34	92.90	97.01

The VRC needed no cooling since the vapour condensed during heat exchange with the bottom liquid and in the superheater. Moreover, the compressed vapour successfully fulfilled the reboiler's total heating requirement. Because the optimized-VRC assisted columns were largely focused on the heat exchanger's ΔT_{\min} , nearly 10% of external heating was necessary. The exchanger sizes were reduced at the expense of the overall energy required. This is also seen by the higher compression ratios of the optimised VRC columns. As mentioned before, only 12 TLFA proposed runs were converged, which consequently affected the optimisation results. In view of this, the variables selection, defined ranges and response objectives should be comprehensively reassessed in order to generate more sensible findings. Meanwhile, the exergy analysis shows that the VRC-assisted V5002 and V5003 columns were more thermodynamically efficient. Remarkably, the VRC-assisted V5002 column outperforms the CC by around 20%. This indicates that more heat was converted into usable work. On the other hand, the optimised-VRC columns were somewhat less effective than the original VRC columns for V5002, and surprisingly, the optimised-VRC V5003 column has an even lower thermodynamic efficiency than the CC.

3.4. Economic evaluations

Table 2 summarises the economic evaluation results for conventional columns, VRC-assisted columns, and optimized VRC-assisted columns. The Aspen Process Economic Analyzer was used to compute the CAPEX and OPEX. As opposed to the CC, both the VRC and optimized-VRC systems show greater CAPEX. This is to be anticipated since the VRC system for the studied process needs two stages of compression, despite the removal of the CC's condenser and reboiler. However, as the data shows, the optimised VRC columns exhibit a higher CAPEX than the ordinary VRC-assisted columns. While the cost of the exchanger has been greatly lowered, the V5002's first compressor cost has risen by 11% after the optimisation. Meanwhile, the average 15% increment of compressor costs was seen in V5003. The additional trim reboilers has also impacted the overall CAPEX for the optimised-VRC columns. Attention must be paid to the VRC's excellent performance in OPEX, whereby the only source of energy needed is electricity to drive the compressors. As can be seen in the table, approximately 84% and 88% of energy savings can be achieved by integrating the V5002 and V5003 columns, respectively, with the VRC system. Taking a wider perspective, there is a good chance that over the course of ten years of the VRC-assisted column operation for the fractionation of PKO fatty acid, plant revenues may be boosted by the amount saved on utility usage.

Table 2. Summary of CAPEX and OPEX.

Configuration	CC		VRC		Optimised-VRC	
	V5002	V5003	V5002	V5003	V5002	V5003
CAPEX (x10⁵ \$)						
Column tower	2.616	2.367	2.616	2.367	2.616	2.367
Column condenser	0.113	0.096				
Column reboiler	0.397	0.255				
Superheater			0.107	0.111	0.096	0.106
Heat exchanger			0.313	0.300	0.236	0.179
Compressor 1			0.523	0.191	0.581	0.221
Compressor 2			0.530	0.172	0.539	0.199
Flash drum			0.334	0.324	0.334	0.324
Trim reboiler					0.097	0.094
Total	3.126	2.718	4.423	3.465	4.500	3.491
OPEX (x10⁵ \$/y)						
Electricity			0.583	0.200	0.620	0.233
High-pressure steam	2.637	1.301			0.263	0.132
Cooling water	0.924	0.413				
Total	3.561	1.714	0.583	0.200	0.883	0.365

4. Conclusions

The present study was designed to investigate the exergetic performance of the VRC-assisted columns to fractionate out PKO fatty acid cuts in four operating columns. The optimized version of the VRC configurations showed a less promising performance than the ordinary one. Without a doubt, despite increased CAPEX at the start of the business, a considerable rise in profitability might be realised in the coming years. Although, a comprehensive optimisation is to be done for future works for more sensible data. Despite the PKO fatty acid feed's limited flexibility and familiarity, the positive results from this research serve as a baseline for future environmentally friendly and cost-effective oleochemical separation processes.

References

- M. R. Othman and G. P. Rangaiah, 2020, Performance analysis of dividing wall column model for fractionation of oleochemical fatty acid, *J. of Oil Palm Res.*, 32, 2, 326-340.
- N. M. Sidek and M. R. Othman, 2019, Assessment of heat pumping technology in oleochemical fatty acid fractionation, *Mat. Sci. Eng.*, 702, 012012.
- N. M. Sidek and M. R. Othman, 2020, Techno-economic analysis of heat pumping technology for oleochemical fatty acid fractionation, *Comp. Aided Chem. Eng.*, 48, 1075-1080.
- N. Nguyen and Y. Demirel, 2010, Retrofit of distillation columns in biodiesel production plants, *Energy*, 35, 4, 1625-1632.
- P. V. Mangili, 2020, Thermoeconomic and environmental assessment of pressure-swing distillation schemes for the separation of di-n-propyl ether and n-propyl alcohol, *Chem. Eng. Proc.*, 148, 107816.
- S. S. Parhi, G. P. Rangaiah and A. K. Jana, 2019, Optimizing reboiler duty and reflux ratio profiles of vapor recompressed batch distillation, *Sep. Purif. Technol.*, 213, 553-570.
- Z. Si, D. Han, J. Gu, Y. Song and Y. Liu, 2020, Exergy analysis of a vacuum membrane distillation system integrated with mechanical vapor recompression for sulfuric acid waste treatment, *Appl. Therm. Eng.*, 178, 115516.

A Reduced Population Balance Model for Coupled Hydrodynamics and Mass Transfer in Shallow Bubble Column Reactors

Menwer Attarakih^{a,b}, Ferdaous Al-Slaihat^a, Armin Fricke^{b,c}, Hans-Jörg Bart^{b,d}

^aUniversity of Jordan, Chem. Eng. Dept., Amman, 11942, Jordan

^bPopulation Balance Modeling GmbH, Berlin, 10178, Germany

^cCGC Capital-Gain Consultants GmbH, Berlin, 10178, Germany

^dUniversity of Kaiserslautern, TVT, Kaiserslautern, 67653, Germany
m.attarakih@ju.edu.jo

Abstract

Shallow bubble column reactors are gas-liquid reactors which are intensively used in chemical industries because of their simple design and efficiency. They are characterized by a low height-to-diameter ratio when compared to traditional columns. In such equipment, the bubble size distribution is an important design parameter that defines the interfacial area which plays an important role in designing and scaling-up these reactors. In this contribution a new reduced population balance model is developed based on OPOSPM (One Primary One Secondary Particle Method) framework to predict the coupled bubble hydrodynamics and mass transfer in shallow bubble column reactors. The model is composed of four transport equations where two of which are used to describe the total number and volume bubble concentrations and the other two transport equations are devoted for chemical species balances. Other consecutive equations are used to model the bubble relative velocity, breakage, coalescence, interphase mass transfer, and reconstruction of bubble probability density. The model capability to predict the experimental data for the coupled hydrodynamics and mass transfer in the shallow bubble column reactor of Lau et al. (2012) is successfully demonstrated. In conclusion, the model is simple and still efficient for modelling such reactors with the essential phenomena from the detailed population balance equation (PBE) is captured.

Keywords: OPOSPM, Population balance, Bubble columns, Mass transfer.

1. Introduction

Bubble column reactors are widely used in the chemical, petrochemical and biochemical industries (Jakobsen, 2008). The design of these columns is characterized by a high L/D (height-to-diameter) ratio which is required to allow the bubble hydrodynamics to approach a fully developed state. On the contrary, shallow bubble columns are characterized by a low L/D ratio to make use of the high gas holdup due to small size population of bubbles produced above the distributor. This is required to estimate the specific interfacial area to provide a closure for the momentum and mass transport equations which are solved for each phase (Jakobsen, 2008). However, with limited available experimental studies on shallow bubble columns (Lau et al., 2012), the design and scale up of this equipment is still dependent on pilot scale experiments. As an approach for modelling of the bubble column reactors, the mixture model does not take into account the instantaneous discrete events of bubble breakup and coalescence. On the

other hand, multi-fluid models which include bubble-bubble interactions are expensive from computational point of view if detailed column geometry is needed. Along these lines, the emergence of the reduced population balance models as fast solvers of the PBE have gained considerable interest in reflecting the discrete flow and accommodate particle-particle interactions (Wu et al., 1998, Drumm et al., 2010, Shonibare and Wardle, 2015, Attarakih and Bart, 2020, Santis et al., 2021). In this regard, Wu et al. (1998) derived the interfacial area transport equation from the particle number density transport equation followed by the derivation of bubble breakage and coalescence kernels which were improved by Ishii and Hibiki (2011). Drumm et al. (2010) introduced the coupled OPOSPM-CFD as a reduced population balance model, which is derived from the discrete PBE based on the Sectional Quadrature Method Of Moments (SQMOM) (Attarakih et al., 2009). Shonibare and Wardle (2015) used OPOSPM for the numerical investigation of vertical plunging jet using a hybrid multifluid CFD Solver which is implemented in OpenFoam version 2.2.1. Santis et al. (2021) used OPOSPM in the GEneralized Multifluid Modelling Approach (GEMMA) which is used by the well-known open-source CFD code OpenFOAM version 7.0. This modelling technique based on reduced population balances presents a valuable tool for understanding the hydrodynamics of annular centrifugal contactors. Motivated by these findings, we used the OPOSPM framework to model the coupled hydrodynamics and mass transfer of the bubbly flow in shallow bubble column reactors as a building block for detailed CFD models.

2. Coupled bubble hydrodynamics and mass transfer using OPOSPM

The model for the coupled bubble hydrodynamics and mass transfer is derived from the PBE. This equation transports the number concentration function in space-time coordinates and the particle property space through the number concentration function $f(z,d,t)$ where z is the special domain, d and t are bubble size and time respectively. This function is assumed continuous, satisfies regulatory conditions and describes the number concentration of bubbles. In this work we used the reduced population balance model based on the OPOSPM framework to describes the coupled hydrodynamics and mass transfer which is given by:

$$\frac{\partial(N)}{\partial t} + \nabla \cdot (N \langle \langle v_g \rangle \rangle) = S_B - S_C + N^{in} \delta(z - z^{in}) \quad (1)$$

where N is the total bubble concentration, $\langle \langle v_g \rangle \rangle$ is the mean bubble velocity, S_B and S_C are the instantaneous rate of change of bubble number concentration due to bubble breakup and coalescence respectively which are described by the consecutive equations of Hibiki and Ishii (2000) in terms of mean bubble diameter (d).

$$S_B = C_b \frac{\varepsilon^{1/3} (1 - \alpha)}{d^{2/3} (\alpha_{max} - \alpha)} \exp \left(-K_b \frac{\sigma}{\rho_f d^{5/3} \varepsilon^{2/3}} \right) N \quad (2)$$

$$S_C = \frac{C_2 d^{7/3} \varepsilon^{1/3}}{(\alpha_{max} - \alpha)} \exp \left(-K \sqrt[6]{\frac{d^5 \rho_L^3 \varepsilon^2}{\sigma^3}} \right) N^2 \quad (3)$$

The values of the constants in Eqs. (2) and (3) have the same values that were given by Hibiki and Ishi (2000): $C_b = 0.0037$, $C_2 = 0.0014$, $K_b = 1.37$, $K = 1.29$ and $\alpha_{max} = 0.52$.

The source terms are function of the physiochemical properties, gas phase holdup, mean bubble diameter and energy dissipation. The transport of the gas phase holdup (α_G) presents the second OPOSPM transport equation which is given by:

$$\frac{\partial(\alpha_G \rho_G)}{\partial t} + \frac{\partial(\alpha_G \rho_G \langle \langle v_g \rangle \rangle)}{\partial z} = -k_L a \left(\frac{C_G}{H} - C_L \right) M_G + \rho_G^{in} U_G^{in} \delta(z - z^{in}) \quad (4)$$

where C refers to oxygen concentration, H is the dimensionless Henry's constant, $k_L a$ is the volumetric liquid phase mass transfer coefficient and M_G is the solute (oxygen) molecular weight. The energy dissipation ε (m^2/s^3) is calculated using the correlations given by WU et al. (1998). Similarly, the solute transport equation in the gas phase reads

$$\frac{\partial(C_G \alpha_G)}{\partial t} + \frac{\partial(C_G \alpha_G \langle \langle v_g \rangle \rangle)}{\partial z} = -k_L a \left(\frac{C_G}{H} - C_L \right) + C_G^{in} U_G^{in} \delta(z - z^{in}) \quad (5)$$

while the solute transport equation in the liquid phase is given by:

$$\frac{\partial(C_L \alpha_L)}{\partial t} + \frac{\partial(C_L \alpha_L \langle v_L \rangle)}{\partial z} = +k_L a \left(\frac{C_G}{H} - C_L \right) + C_L^{in} U_L^{in} \delta(z - z^{in}) \quad (6)$$

where the subscript L refers to the liquid phase and the mean gas velocity is given by:

$$\langle \langle v_g \rangle \rangle = \langle v_L \rangle / (1 - \alpha_G) + v_{gr} \quad (7)$$

In the above equation the relative bubble velocity is calculated from the equation:

$$v_{gr} = \sqrt{g d_{30} \Delta \rho / (3 C_D \rho_L)} \quad (8)$$

where C_D is the drag coefficient which is function of the mixture Reynolds number (Wu et al., 1998) and d_{30} is the bubble mean volume diameter which couples the bubble number and volume concentrations:

$$d_{30} = \sqrt[3]{\frac{6 \alpha_G}{\pi N}} \quad (9)$$

In the OPOSPM framework, two low-order moments are conserved (α_G, N), and the secondary particle position (d_{30}) represents a Lagrangian fluid particle carrying information about the mean bubble size of the whole population through its low-order moments. On the other hand, the interfacial area concentration is calculated from an algebraic model based on the ratio between gas holdup and the bubble Sauter mean diameter:

$$a = \frac{6 \alpha_G}{d_{32}} \quad (10)$$

The Sauter mean diameter is related to the mean volume diameter (d_{30}) by the empirical relation $d_{30} = (0.75 - 0.80) d_{32}$ as recommended by Shonibare and Wardle (2015) and Santis et al. (2021). As the full bubble size distribution is concerned, the OPOSPM-Weibull distribution, as a least biased probability number density estimator, can be used as derived and explained in our previous work (Attarakih and Bart, 2020).

3. Numerical solution

The model equations (1-9) represent a set of hyperbolic conservation laws which requires special attention during space-time discretization phase (Attarakih et al., 2013). This

system can be casted for a given spatial cell of index j and width Δz using the first order upwind scheme as follows:

$$\left(\frac{dU}{dt}\right)_j + \frac{1}{\Delta z}(F_j - F_{j-1}) = S_j, \quad j = 1, 2, 3 \dots n \quad (11)$$

where $U = [N \ \alpha_G \rho_G \ C_G \alpha_G \ C_L \alpha_L]^T$, $F = U \bullet [\langle v_g \rangle \ \langle v_g \rangle \ \langle v_g \rangle \ \langle v_L \rangle]^T$ are the conserved variables and numerical flux vectors respectively and S is the source term of Eqs. (1-6). For the time discretization, the MATLAB ODE23 solver is used with the built-in default settings. The relative velocity which is given by Eq.(8) is solved iteratively and found to converge with 5 to 10 iterations. By assuming the liquid phase is stagnant and perfectly mixed and there are negligible changes in the oxygen concentration in the rising bubbles, Eq.(6) simplifies into an ODE which has an analytical solution (Lau et al., 2012).

4. Results and discussion

In this section, the numerical results from solving the semi-discrete OPOSPM model (Eqs.11) are compared to the experimental data of Lau's et al. (2012) measured in a pilot scale shallow bubble column reactor of 14 cm diameter.

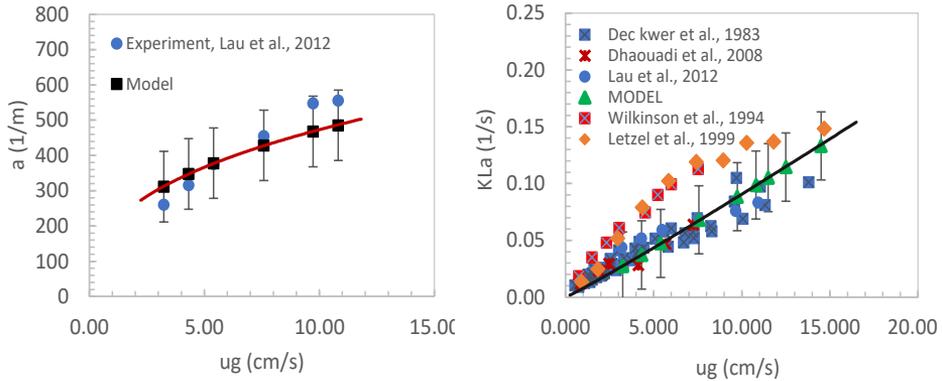


Figure (1): Predicted steady state mean interfacial area concentration and volumetric mass transfer coefficient as compared to published experimental data at $L/D = 7.2$.

The column is equipped with a perforated plate distributor having 211 orifices that are 3 mm in diameter and configured in triangular pitch. Tap water is used as a stagnant phase, while air is used at 1.0 bar and 25 °C as the gas phase and the gas velocity is varied from 3.2 to 10.8 cm/s. The number of spatial cells (n) is set to 25 with further increase is found to have no effect on the numerical results. The initial conditions for the gas phase are set to small positive numbers close to the square root of the machine epsilon since the stagnant water phase is purged by nitrogen before presetting the air flow. In Lau's et al. (2012) experiment, the steady state gas holdup and the interfacial area were measured, while in the mass transfer experiments, the dynamic oxygen concentration at $L/D = 7.2$ of the column height is measured using an optical fiber oxygen probe. In addition to this, the mass volumetric mass transfer coefficient is calculated from the measured data. In our present model, the inlet mean bubble diameter (d_{32}) is calculated using Wilkinson et al. (1994) correlation based on perforated gas distributor and corrected such that $d_{30} = (0.8)d_{32}$ as recommended by Shonibare and Wardle (2015) and Santis et al. (2021). The inlet area concentration is calculated from Hikita correlation, while the liquid side mass

transfer coefficient is calculated by the correlation of Akita and Yoshida which are reported in Shah et al. (1982). Note that the liquid side mass transfer coefficient is independent of the (u_g) which is consistent with the experiments of Lau et al. (2012).

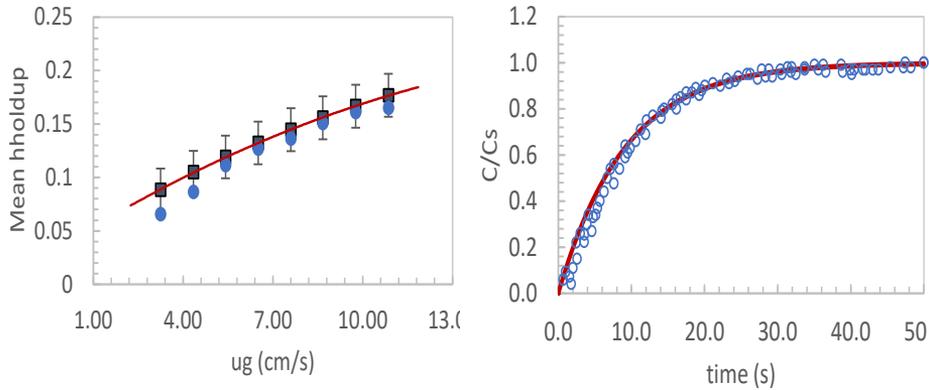


Figure (2): Steady state mean gas holdup at $L/D = 7.2$ and the dynamic evolution of the oxygen concentration in a stagnant tap water column with filled and open circles are the Lau's et al. (2012) experimental data.

The bubble column hydrodynamics coupled with mass transfer is simulated using the present OPOSPM model. A sample of results is shown in Fig.(1) which compares the predicted mean interfacial area concentration as function of the superficial gas velocity (u_g) to the experimental data of Lau et al. (2012) in a shallow bubble column. The model equations were integrated with respect to time until steady state is achieved at 50 s. As can be noticed, the interfacial area concentration increases due to the shear-induced bubble breakup at the nozzle as u_g increases. We used ± 100 (m^{-1}) as fixed bandwidth for the error bar which covers the experimental fluctuations around the predicted values, where rigorous parameter estimation with confidence intervals is beyond the scope of this work. On the right hand side of Fig.(1) is the predicted liquid side volumetric mass transfer coefficient as compared to the published experimental data including those of Lau et al. (2012) with error bandwidth ± 0.03 (s^{-1}). According to Lau et al. (2012), the large dispersion in the measured data is due to different designs of gas distributors, presence of impurities and the sensitivity of mass transfer to system temperature. Nonetheless, the model predicts most of the experimental data except those of Letzel et al. (1999) where slight under predictions is present in the gas velocity middle range (5 - 13 cm/s). On the left hand side of Fig.(2) is the predicted mean gas holdup (filled square) as compared to the experimental data of Lau et al. (2012) with error bandwidth equals to ± 0.02 . Compared to the interfacial area predictions (Fig.(1), Left), the error bandwidth is narrower which may be attributed to the higher measurement accuracy of the gas holdup compared to the interfacial area. By assuming that the liquid phase is perfectly mixed, Eq.(6) can be reduced to a first-order ODE and is solved at $L/D = 7.2$, while the gas phase is modelled as plug flow from the gas distributor. In Fig.(2, right), the dynamic evolution of oxygen absorption in the stagnant water phase is compared to model prediction. It is obvious that the most accurate result of the model is that of mass transfer profile despite the complex coupling with the uncertainties induced by the bubble hydrodynamics. While bubble breakup and coalescence are calibrated based on the work of Hibiki and Ishii (2000), we did not adjust any model or correlation parameters to improve the model prediction. Therefore, we believe that building a database on experimental data and

calibrating individual population balance sub-models, will certainly improve and widen the window of model predictions.

5. Summary and conclusions

In this work, we used the OPOSPM framework as a reduced population balance model to derive the transport equations of coupled hydrodynamics and mass transfer in shallow bubble columns. The model consists of four transport equations with sub-population balance models and consecutive relations. No further sub-model calibration is used other than that in the published literature for those of breakup and coalescence for bubbly flow in vertical tubes. The predicted results by the present model are compared to the published experimental data with good and high accuracy for bubble hydrodynamics and mass transfer respectively. The 1D semidiscrete hyperbolic system is solved using MATLAB ODE23 solver with CPU time as fraction of second. We believe that the relative simplicity of the model offers the bubble column designers one-, two- and three-dimensional models that can be implemented in complex CFD codes.

References

- M. Attarakih and H.-J. Bart, 2020, Beyond OPOSPM: A corrected maximum entropy weibull distribution for Solving population balances, *Computer Aided Chemical Engineering*, 48, 691-696.
- M. Attarakih, M. Abu-Khader and H.-J. Bart, 2013, Modelling and dynamic analysis of an RDC extraction column using OPOSPM, *Chem. Eng. Sci.*, 91, 180-196.
- M. Attarakih, C. Drumm and H.-J. Bart, 2009, Solution of the population balance equation using the sectional quadrature method of moments (SQMOM), *Chem. Eng. Sci.*, 64, 742-752.
- W.D. Deckwer, K. Nguyen-Tien, B.G. Kelkar, Y.T. Shah, 1983, Applicability of axial dispersion model to analyze mass transfer measurements in bubble columns, *AIChE J.*, 29, 915-922.
- A. De Santis, B. C. Hanson and M. Fairweather, 2021, Hydrodynamics of annular centrifugal contactors: A CFD analysis using a novel multiphase flow modelling approach, *Chem. Eng. Sci.*, 242, 116729.
- H. Dhaouadi, S. Poncin, J. M. Hornut, N. Midoux, 2008, Gas-liquid mass transfer in bubble column reactor: Analytical solution and experimental confirmation, *Chem. Eng. Proce.: Process Intensification*, 47,548-556.
- C. Drumm, M. Attarakih, M. W. Hlawitschka and H.-J. Bart, 2010), One-group reduced population balance model for CFD simulation of a pilot-plant extraction column, *Ind. Eng. Chem. Res.*, 49, 3442-3451.
- T. Hibiki and M. Ishii, 2000, One-group interfacial area transport of bubbly flows in vertical round tubes, *Int. J. Heat Mass Transfer.*, 43, 2711 - 2726.
- M. Ishii and T. Hibiki, 2011, *Thermo-fluid dynamics of two-phase flow*, Springer, New York.
- H. A. Jakobsen, 2008, *Chemical reactor modeling: Multiphase reactive flows*, Springer-Verlag, Heidelberg.
- R. Lau, P. H. V. Lee and T. Chen, 2012, Mass transfer studies in shallow bubble column reactors. *Chem. Eng. Proc.: Process Intensification*, 62, 18-25.
- H. M. Letzel, J. C. Schouten, R. Krishna and C. M. van den Bleek, 1999, Gas holdup and mass transfer in bubble column reactors operated at elevated pressure. *Chem. Eng. Sci.*, 54, 2237-2246.
- Y. T. Shah, B. G. Kelkar, S. P. Godbole and W. D. Deckwer, 1982, Design parameters estimations for bubble column reactors, *AIChE Journal*, 28, 353-379.
- O. Y. Shonibare and K. E. Wardle, 2015, Numerical Investigation of vertical plunging jet using a hybrid multifluid-VOF multiphase CFD solver. *Int. J. Chem. Eng.*, 2015, 1-14.
- P. M. Wilkinson, H. Haringa and L. L. Van Dierendonck, 1994, Mass transfer and bubble size in a bubble column under pressure, *Chem. Eng. Sci.*, 49, 1417-1427.
- Q. Wu, S. Kim, M. Ishii and S.G. Beus, 1998, One-group interfacial transport in vertical bubbly flow, *Int. J. Heat. Mass Transfer*, 41, 1103-1112.

Modeling of the crystallization of gypsum produced in the digestion tank of an industrial phosphoric acid manufacturing process

Ilias Bouchkira^{a,b}, Abderrazak M. Latifi^{a,b,*}, Lhachmi Khamar^b and Saad Benjelloun^b

^a*LRGP, Université de Lorraine, 1 Rue Grandville, 54000, Nancy, France*

^b*Université Mohammed VI Polytechnique, 43150, Benguerir, Morocco*

Abstract

The present paper deals with the development of a model of gypsum crystallization during the production of phosphoric acid. The objective is to predict the kinetics of formation of products and consumption of reagents, as well as the time-varying crystal size distribution. The population balance model involving primary and secondary nucleation and crystal growth rates is thus investigated using the method of moments, along with the Pitzer thermodynamic model. Both models involve unknown parameters to be identified from experimental data. Two databases are therefore used, (i) the first one which consists of equilibrium measurements is used to identify the unknown parameters of the Pitzer model, (ii) and the second one which contains temporal profiles of supersaturation ratio, is used to determine the nucleation and growth unknown parameters. The identified parameters are then used to compare the predictions of the developed models with the experimental measurements. The results are very promising, and the models can be used in the design and optimization of gypsum crystallization processes.

Keywords: Crystallization, Population balance model, Pitzer model, Method of the moments.

1. Introduction

In the phosphate industry, phosphoric acid is a major component in the manufacture of fertilizers. It is mainly produced by the digestion of phosphate ore by a concentrated sulfuric acid solution. The products are phosphoric acid and solid gypsum, which are separated in a vacuum filtration unit downstream of the digestion tank (Becker, 1983). The optimization of the process performance is very challenging and aims mainly to: (i) extract the maximum amount of phosphate from the ore during the digestion, (ii) crystallize gypsum with optimum filterability and washing characteristics, (iii) and produce the most concentrated and purest phosphoric acid possible, at a lower cost and with little impact on the environment. In this work, we focus mainly on the modeling of the crystallization of gypsum which strongly influences the quality of the produced phosphoric acid, and the performance of the units downstream of the process, namely, the filtration units.

The objective of this paper is to develop a crystallization model based on the population balance equations (PBEs) along with the Pitzer thermodynamic model (Pitzer, 2018). The latter enables to predict the supersaturation ratio which is involved in the driving force of the crystallization process.

The Pitzer model is first calibrated using equilibrium measurements, and the unknown parameters of the PBEs are identified from the kinetic measurements. The PBEs are then solved using the method of moments.

2. Modeling framework

2.1. Population balance model

For a population of crystals whose size distribution evolves over time, the population balance allows to express its number density function (n) as a function of time (t) and of a characteristic size (x). Assuming that the crystallization is carried out in an isothermal perfectly mixed batch reactor with a volume of 50 cm^3 , and that the growth rate of crystals is size independent, the PBE writes (Hanhoun et al., 2013):

$$\frac{\partial n(t, x)}{\partial t} = -G(t) \frac{\partial n(t, x)}{\partial x} + R_N(t) \delta(x - x_0) \quad (1)$$

where R_N is the sum of primary and secondary nucleation rates, G is the crystal growth rate, and δ is the Dirac function. The nucleation and crystal growth rates are defined by (Hanhoun et al., 2013; Zhu et al. 2016):

$$R_N(t) = A \times \exp\left(\frac{-B}{\ln S(t)^2}\right) + k_s \times [\ln S(t)]^s \times \mu_3(t) \quad (2)$$

$$G(t) = k_g [\ln S(t)]^g \quad (3)$$

where A, B, k_s, s, k_g and g are unknown parameters to be identified from experiments, S is the supersaturation ratio which depends mainly on the speciation of the reacting medium, and μ_3 is the third order moment of the number density function (n).

2.2. Method of moments

The method of moments is a model reduction technique used mainly to solve the population balance equation. It transforms the partial differential equation (1) into a dynamic system of ordinary differential equations (ODEs) with unknowns independent of the spatial variable (x). It is based on the calculation of the k^{th} -order moment μ_k which is defined as follows:

$$\mu_k(t) = \int_0^{\infty} x^k n(t, x) dx \quad (4)$$

Multiplying Eq. (1) by x^k and integrating from zero to infinity, the following system of ODEs is obtained:

$$\frac{d\mu_0(t)}{dt} = R_N(t) \quad (5)$$

$$\frac{d\mu_k(t)}{dt} = k G(t) \mu_{k-1}(t), \quad k \geq 1 \quad (6)$$

$$\mu_k(0) = \int_0^{\infty} x^k n_0(x) dx \quad (7)$$

where n_0 is the crystal size distribution (CSD) at $t=0$.

2.3. Mass balance equations and supersaturation ratio

In this work, we assume that gypsum crystals are formed as a result of direct attack of pure tricalcium phosphate with a concentrated sulfuric acid solution according to the following reaction:



The use of the third moment of the CSD allows to calculate the amount of gypsum produced as follows (Hanhoun et al., 2013; Zhu et al., 2016):

$$C_{\Delta\text{gyp}}(t) = \frac{k_v \rho_s}{M_c} (\mu_3(t) - \mu_3(0)) \quad (9)$$

where k_v is a volume shape factor, ρ_s is the density and M_c is the molecular weight of the formed crystals, respectively. The material balance in the batch reactor can be written as:

$$C_i(t) = C_i^0 + v_i C_{\Delta\text{gyp}}(t) \quad (10)$$

where C_i^0 and v_i are the initial concentration and the stoichiometry coefficient of the component i involved in the reaction (8), respectively. The resolution of the material balance allows then to calculate the supersaturation ratio expressed as:

$$S = \left(\frac{a_c a^2 + a_{\text{SO}_4^2} - a_w^2}{K_{sp}} \right); a_i = m_i \cdot \gamma_i \quad (11)$$

where a_i, γ_i, m_i denote the activity, the activity coefficient, and the molality of the component i , respectively, K_{sp} is the solubility product of gypsum. The activity coefficients are predicted using the following Pitzer model:

$$\ln(\gamma_i) = \frac{z_i^2}{2} f + 2 \sum_j m_j \lambda_{i,j}(I) + z_i^2 \sum_j m_j m_k \lambda'_{j,k}(I) + 3 \sum_{j,k} m_i m_j m_k \psi_{i,j,k} \quad (12)$$

where z_i are the charges of component i , f is the Debye-Huckel function, I is the ionic strength, $\lambda_{i,j}$ and $\psi_{i,j,k}$ represent the short-distance binary and ternary interaction parameters, respectively. Their values are to be identified from experimental equilibrium measurements.

The resulting crystallization model equations (2-3, 5-7, 9-12) are implemented and solved within MATLAB environment.

3. Experimental measurements

Several experimental measurements are carried out in our laboratory. They mainly consist of equilibrium measurements, namely the speciation of sulfuric and phosphoric acids, for a wide range of temperature and concentration conditions. In addition, measurement data of gypsum solubility and water activity are collected from the literature (Bouchkira et al. 2021d). They are used for the identification of the unknown parameters of the Pitzer model. Moreover, experimental measurements of the kinetics of gypsum crystallization during the production of phosphoric acid are taken from the literature (Zhu et al., 2015). They consist of supersaturation ratio of gypsum and are used to identify the kinetic parameters of the PBE.

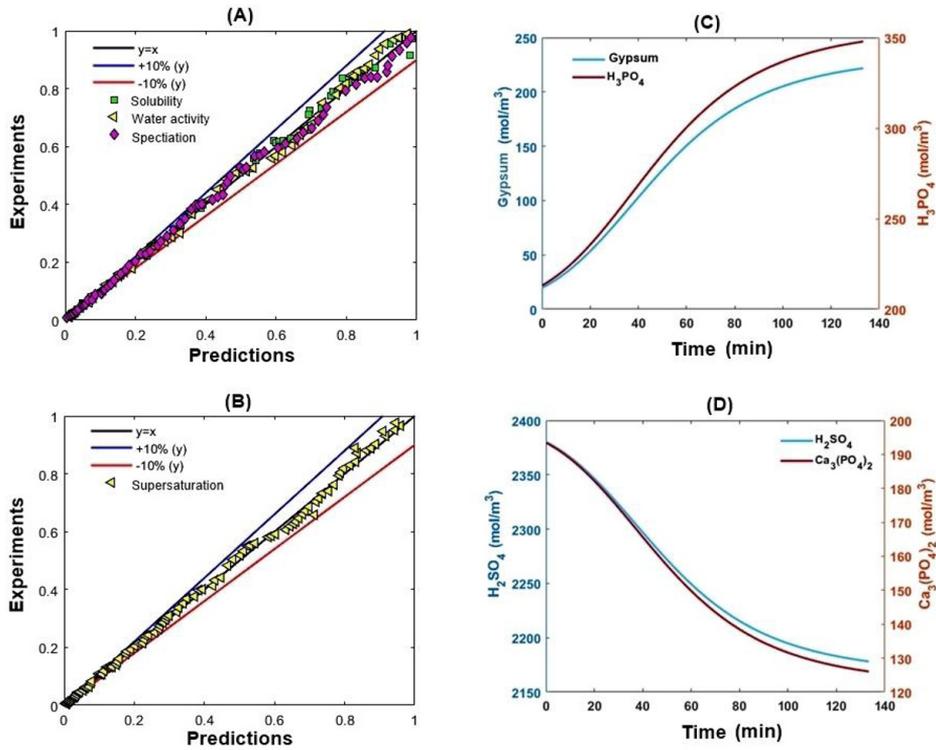


Figure 1: Modeling results. (A): Parity diagram for the identification of the Pitzer unknown parameters. (B): Parity diagram for the identification of the PBE unknown parameters from the supersaturation data. (C): Gypsum and phosphoric acid production profiles. (D): Tricalcium phosphate and sulfuric acid consumption profiles.

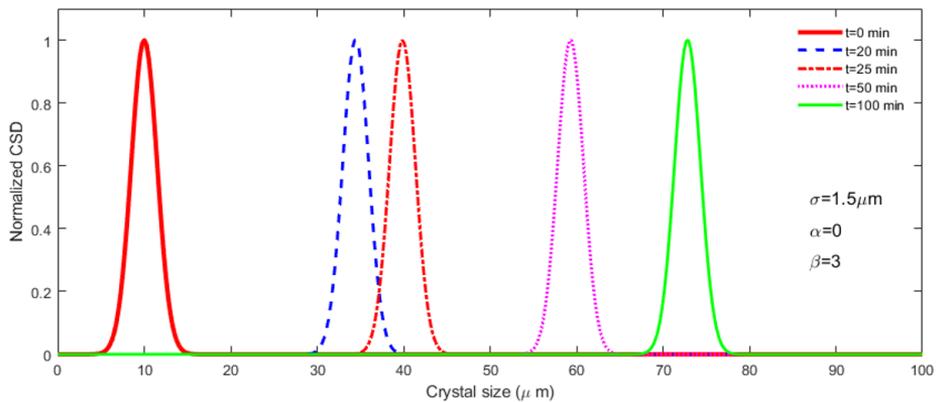


Figure 2: Normalized CSD at different times.

Table 1: Variations of the average size of gypsum crystals with time.

Time (min)	0	20	25	50	100
\bar{x} (μm)	10	34.46	39.87	59.31	72.84

4. Results and discussions

Equilibrium measurements (i.e., gypsum solubility, water activity and speciation) are used to determine the optimal values of the unknown parameters of the Pitzer model, i.e., $\lambda_{i,j}$ and $\psi_{i,j,k}$. Indeed, the optimization of these parameters has been widely investigated in our recent works (Bouchkira et al., 2021a, b, c, d, e). The comparison of the predictions of the identified Pitzer model with the experimental measurements (Fig.1(A)) shows a very good agreement. On the other hand, the data on supersaturation ratio are used to determine the unknown parameters of the PBE, i.e., A, B, k_s, s, K_g and g . The optimal values of these parameters are then used to compare the model predictions with the experiments. It can be seen that the model predictions are very accurate. It is noteworthy that the estimability of the unknown parameters from the available data was carried out using an estimability analysis method based on global sensitivities (Bouchkira et al., 2021c). The most estimable thermodynamic parameters were identified by means of a branch-and-bound method (Sahinidis, 1996) using the Neos server (Czyzyk et al., 1998) hosted by machines at Arizona State University, whereas the kinetic parameters were identified based on a derivative-free method using a Dell Precision T7810 Bi-Xeon 12x Core 64 GB workstation.

Fig.1(C) shows the predictions of phosphoric acid and gypsum production, and Fig.1(D) presents those of sulfuric acid and tricalcium phosphate consumption. The results are coherent since the amounts of phosphoric acid and gypsum increase over time as the reaction proceeds, while those of sulfuric acid and tricalcium phosphate decrease. Furthermore, as expected, Table 1 shows that the average size of gypsum crystals increases with time, meaning that the growth of the particles formed by nucleation takes place uniformly in the reactor. On the other hand, the parameters of the CSDs are computed using different centred moments of order k ($\mu_k^*(t) = \mu_k(t) - \mu_k(0)$) at different crystallization times. Thus, the standard deviation is deduced from the centred moment of order 2 as $\sigma = \sqrt{\mu_2^*} = 1.5 \mu\text{m}$, the skewness is computed using the centred moment of order 3 as $\alpha = \mu_3^*/\mu_2^{3/2*} = 0$, and the kurtosis is calculated from the centred moment of order 4 as $\beta = \mu_4^*/\mu_2^{2*} = 3$. It is worth noticing that the values of these parameters do not change over time and the last two values are typical of a normal distribution. Moreover, since the initial CSD is assumed to be normal, and only nucleation and growth kinetics are involved in the reactor, the parameters σ, α and β remain unchanged as illustrated in Fig.2.

The developed model shows that it is able to predict relevant variables to the design, optimization and control of gypsum crystallization processes. Moreover, it can be used to determine the optimal operating conditions that allow to extract the maximum amount of phosphate from the ore, and to better monetize the gypsum.

5. Conclusions

In this work, we considered the modeling of gypsum crystallization that occurs during the production of phosphoric acid. A PBE is developed and solved by means of the method of moments. The required gypsum supersaturation ratio and the activity coefficients of the components involved in the process are estimated by means of the Pitzer model. Although the results obtained are interesting and promising, the crystallization model developed is a one-dimensional model and does not allow to account for the influence of impurities present in the phosphate ore. It is a very important issue since the impurities have a strong influence on the type and shape of gypsum crystals and therefore on the performance of the digestion tank and of the units downstream. Further work will deal with a multi-dimensional PBE in order to quantify the effects of impurities on gypsum crystal growth.

References

- P. Becker, 1983. Phosphates and phosphoric acid. Marcel Dekker, Inc..
- I. Bouchkira, S. Benjelloun, L. Khamar, A. M. Latifi, 2021a. Thermodynamic-based model for the prediction of the fouling phenomena in a wet phosphoric acid process. *Chemical Engineering Transactions* 86, 1273–1278.
- I. Bouchkira, A. Latifi, L. Khamar, S. Benjelloun, 2021b. Process modeling and multi-criteria optimization of an industrial phosphoric acid wet-process. In: *Computer Aided Chemical Engineering*. Vol. 50. Elsevier, pp. 499–504.
- I. Bouchkira, A. M. Latifi, L. Khamar, S. Benjelloun, 2021c. Global sensitivity based estimability analysis for the parameter identification of pitzer's thermodynamic model. *Reliability Engineering & System Safety* 207, 107263.
- I. Bouchkira, A. M. Latifi, L. Khamar, S. Benjelloun, 2021d. Modeling and multi-objective optimization of the digestion tank of an industrial process for manufacturing phosphoric acid by wet process. *Computers & Chemical Engineering*, 107536.
- I. Bouchkira, A. M. Latifi, L. Khamar, S. Benjelloun, 2021e. Multi-objective optimization of the digestion tank of an industrial phosphoric acid manufacturing process. In: *2020 6th IEEE Congress on Information Science and Technology (CiSt)*. IEEE, pp. 389–394.
- J. Czyzyk, M. P. Mesnier, J. J. Moré, 1998. The neos server. *IEEE Computational Science and Engineering* 5 (3), 68–75.
- M. Hanhoun, L. Montastruc, C. Azzaro-Pantel, B. Biscans, M. Frèche, L. Pibouleau, 2013. Simultaneous determination of nucleation and crystal growth kinetics of struvite using a thermodynamic modeling approach. *Chemical Engineering Journal* 215, 903–912.
- V. John, I. Angelov, A. Öncül, D. Thévenin, 2007. Techniques for the reconstruction of a distribution from a finite number of its moments. *Chemical Engineering Science* 62 (11), 2890–2904.
- K. S. Pitzer, 2018. Activity coefficients in electrolyte solutions, CRC press.
- N. V. Sahinidis, 1996. Baron: A general purpose global optimization software package. *Journal of Global Optimization* 8 (2), 201–205.
- Z. Zhu, Y. Peng, T. A. Hatton, K. Samrane, A. S. Myerson, R. D. Braatz, 2016. Crystallization of calcium sulphate during phosphoric acid production: modeling particle shape and size distribution. *Procedia Engineering* 138, 390–402.

Development of a whole-body physiologically-based pharmacokinetic model for high-dose methotrexate

Giuseppe Pesenti,^a Dario Massari,^b Marco Foppoli,^c Davide Manca^a

^a*PSE-Lab, Process Systems Engineering Laboratory, Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

^b*Anesthesia and Intensive Care, Ospedale Fatebenefratelli e Oftalmico, ASST Fatebenefratelli Sacco, Piazzale Principessa Clotilde 3, 20121 Milano, Italy*

^c*Unit of Lymphoid Malignancies, Division of Onco-Hematological Medicine, Department of Onco-Hematology, IRCCS San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milano, Italy*

davide.manca@polimi.it

Abstract

Whole-body physiologically-based pharmacokinetic (PBPK) models are complex mathematical models that consist of a large number of compartments and describe the anatomical structure of the organism. We present the development of a whole-body PBPK model for high-dose methotrexate by expanding a previously developed minimal PBPK model. We increase the number of body compartments described and introduce an *a priori* description of the fluid volumes, blood flows, and capillary exchanges in each tissue and organ, which do not increase the complexity and dimensionality of the identification procedure. The model is identified and validated with an experimental dataset and produces detailed simulations of concentration within the plasma, ISF, and ICF of each modeled tissue and organ. The original modeling approach demonstrated can be a reference for the development of other whole-body PBPK models.

Keywords: Pharmacokinetics, PBPK, Whole-body, Methotrexate, HDMTX.

1. Introduction

Pharmacokinetic (PK) models describe the relationship between the dose administered to an individual and the resulting concentrations inside various body compartments. Several forms of PK models have been described, with profound differences in terms of structure, complexity, and effort required for the model development and application. PK models usually employ compartments (*i.e.* a defined volume within which concentration is assumed to be homogeneous) as their building blocks. Physiologically-based pharmacokinetic (PBPK) models are a class of compartmental PK models whose compartments are based on the anatomical structure of the organism (Nestorov, 2003).

Whole-body PBPK models generally consist of a large number of compartments, representing the fluid volumes within several organs and tissues, and emulate the actual human anatomy and physiology. Each compartment is usually described by a differential equation representing the drug material balance. The equations feature drug-specific parameters that characterize the interactions and the physicochemical properties of the specific compound and system-specific parameters that describe physiological processes and human anatomy. While many parameters can be estimated and assigned *a priori* (*e.g.*, drug-specific parameters that have been experimentally measured, or physiological

parameters that have been estimated in previous studies), there is usually a significant number of parameters that cannot be assigned *a priori*.

These adaptive parameters are fitted through a nonlinear regression procedure by minimizing the distance (described by a suitable objective function) between model simulations and experimental concentration-time profiles. This approach may lead to numerical and identifiability issues. In this context, minimal PBPK models lump together similar compartments, thus reducing the complexity of the model and the identification procedure, while retaining their physiological foundations (Cao & Jusko, 2012).

We present the development of a whole-body PBPK model for high-dose methotrexate (HDMTX, an intravenously administered antitumor drug), that expands a previously developed minimal PBPK model (Pesenti et al., 2021a) by increasing the number of described body compartments representing different organs and tissues. The additional parameters introduced to expand the previous model are assigned *a priori*, by combining a detailed characterization of anatomical fluid volumes, physiological blood flows, and a comprehensive description of the capillary exchange in each tissue.

2. Methods

The structure of the whole-body PBPK model is shown in Figure 1. The model describes 14 tissues and organs and employs for each one a dedicated compartment for the plasma, interstitial fluid (ISF), and intracellular fluid (ICF). Four additional plasma compartments (*i.e.* large veins, large arteries, left and right heart chambers) represent the general blood circulation that connects all the plasma compartments.

For each i -th tissue and organ, we estimate the weight using a recently developed repository of equations (Stader et al., 2019), and compute the overall volume (V_{tot}) with the specific tissue density (Valentin, 2002). The ISF and ICF volumes (V_{ISF} and V_{ICF}) are evaluated as fractions of V_{tot} (Schmitt, 2008), whereas the estimation of plasma volumes (V_p) follows the approach presented in Pesenti et al. (2021b), based on the blood amount within each tissue (Valentin, 2002) and the patient's hematocrit.

Mathematically, the whole-body PBPK model describes each compartment with an ordinary differential equation, representing the drug material balance. While each material balance depends on the specific flows and connections that involve a compartment, consistently with Fig. 1, the following equations present the formulation of these balances in plasma, ISF, and ICF, for a representative tissue (*e.g.*, adipose tissue, brain, muscle).

$$V_{p,i} \frac{dc_{p,i}}{dt} = Q_{p,i} c_{p,large\ arteries} - Q_{p,i} c_{p,i} - k_{p \rightarrow ISF} A_{cap,i} c_{p,i}^u + k_{ISF \rightarrow p} A_{cap,i} c_{ISF,i}^u \quad (1)$$

$$V_{ISF,i} \frac{dc_{ISF,i}}{dt} = k_{p \rightarrow ISF} A_{cap,i} c_{p,i}^u - k_{ISF \rightarrow p} A_{cap,i} c_{ISF,i}^u + \\ - k_{ISF \rightarrow ICF} K_{50}^u \frac{c_{ISF,i}^u}{K_{50}^u + c_{ISF,i}^u} V_{ICF,i} + k_{ICF \rightarrow ISF} K_{50}^u \frac{c_{ICF,i}^u}{K_{50}^u + c_{ICF,i}^u} V_{ICF,i} \quad (2)$$

$$V_{ICF,i} \frac{dc_{ICF,i}}{dt} = k_{ISF \rightarrow ICF} K_{50}^u \frac{c_{ISF,i}^u}{K_{50}^u + c_{ISF,i}^u} V_{ICF,i} - k_{ICF \rightarrow ISF} K_{50}^u \frac{c_{ICF,i}^u}{K_{50}^u + c_{ICF,i}^u} V_{ICF,i} \quad (3)$$

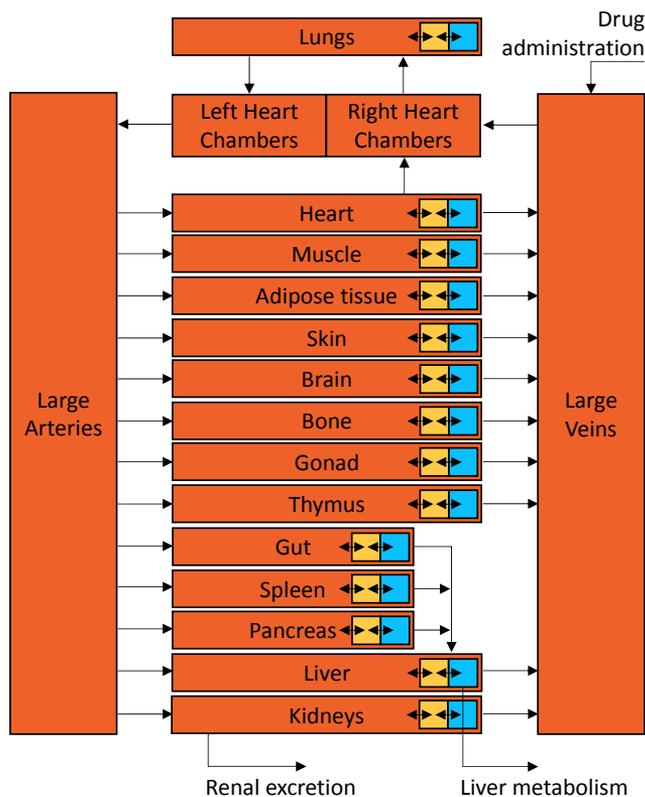


Figure 1 – Structure of the whole-body PBPK model. The rectangles represent the compartments of plasma (dark orange), ISF (mustard), and ICF (cyan) volumes. Black arrows are the drug flows, *i.e.* administration, blood circulation, capillary and cellular exchanges, metabolism, and excretion.

The drug is administered intravenously and enters the large veins plasma compartment, where it reversibly binds to plasma proteins. Plasma compartments are subject to the exchanges due to blood circulation, represented by the plasma flows Q_p , which are estimated as described in Stader et al. (2019). Renal excretion is the main elimination pathway for HDMTX and is estimated as a function of the patient's body weight, height, gender, age, and serum creatinine, as previously described in Pesenti et al. (2021a). Drug clearance due to hepatic metabolism is estimated as approximately 4% of renal excretion. The cellular exchanges, which affect the ISF and ICF compartments, are described with Michaelis-Menten kinetics (Pesenti et al. (2021a)).

The capillary exchange takes place between plasma and ISF of tissues and organs, allowing the unbound drug fraction (denoted by superscript u) to be exchanged with ISF through the gaps and openings of the capillary wall. We compile information from the literature describing capillary density n_{cap} (Freitas, 1999) and the type of capillary in each tissue and organ, as reported in Table 1. For the geometrical modeling of capillaries, we assume average values of capillary length (0.75 mm) and diameter (6 μm) (Freitas, 1999), attaining a geometrical surface a_{cap} of about 0.014 mm^2 . For each capillary type, we obtain rough estimates of the average surface fraction f_{open} that is open to capillary

exchange according to available data (Bundgaard, 1988; Levick, 2004). Eq. 4 estimates the capillary surface A_{cap} that is available for capillary exchange in each tissue.

$$A_{cap,i} = V_{tot,i} n_{cap,i} a_{cap} f_{open,i} \quad (4)$$

Table 1 – Characteristics of the capillaries in each tissue and organ described.

Organ/tissue	Capillary density $n_{cap,i}$ (mm ⁻³)	Capillary type	Open surface fraction $f_{open,i}$ (-)
Muscle	300	Continuous	0.0005
Adipose	100	Continuous	0.0005
Skin	68	Continuous	0.0005
Brain	2750	Continuous (tight junctions)	0 - 0.0005*
Bone	100	Continuous	0.0005
Gonad	300	Fenestrated	0.0062
Thymus	300	Fenestrated	0.0062
Heart	2750	Continuous	0.0005
Lungs	2750	Continuous	0.0005
Gut	300	Fenestrated	0.0062
Spleen	2750	Discontinuous	0.07
Pancreas	300	Fenestrated	0.0062
Liver	2750	Discontinuous	0.07
Kidneys	2750	Fenestrated	0.019

The number of unassigned remaining parameters is four, *i.e.* the same as the minimal PBPK model. These four parameters, describing capillary exchange ($k_{p \rightarrow ISF}$ and $k_{ISF \rightarrow p}$ in cm/min) and cellular exchange ($k_{ISF \rightarrow ICF}$ and $k_{ICF \rightarrow ISF}$ in min⁻¹), are identified with a nonlinear regression using the `fminsearch` routine in MATLAB R2020b (The MathWorks, Natick, MA, US). The regression minimizes the objective function, defined as the mean squared logarithmic error, using an experimental dataset of 84 patients and 657 methotrexate plasma concentrations (Pesenti et al., 2021a).

3. Results and discussion

The model is successfully identified and validated with two different subsets of the experimental dataset, which is split randomly, comprising respectively 56 and 28 patients. The objective function obtained with the identification procedure is 0.9831, which is comparable to the value obtained with the original minimal PBPK model (0.8568) (Pesenti et al., 2021a). Consistently, plasma predictions show a satisfactory agreement with experimental data and are similar to the results of the minimal PBPK model. Figure 2 shows the simulated predictions of the whole-body PBPK model for a patient whose objective function (0.8190) represents the median value of the validation dataset. The investigated time frame covers the administration of three consecutive HDMTX infusions (Fig. 2a). Panel 2c presents the generally good agreement between experimental and predicted plasma concentrations (representing large veins, where experimental

* While the brain features continuous capillaries with tight junctions, in the treatment of brain tumor the blood-brain barrier is usually disrupted, and f_{open} is tentatively assumed equal that of continuous capillaries with loose junctions.

concentrations are measured), and includes the average predicted ISF and ICF concentrations.

Panels 2c-f represent the phenomena of accumulation and exchange described by Eqs. 1-3 in different organs and tissues. For the sake of space and clarity, we present the results for only a few relevant tissues and organs: muscle, adipose tissue, kidneys, skin, and brain. As expected, our results show that predicted plasma concentrations in different compartments are comparable, due to the rapid drug distribution determined by the systemic blood circulation.

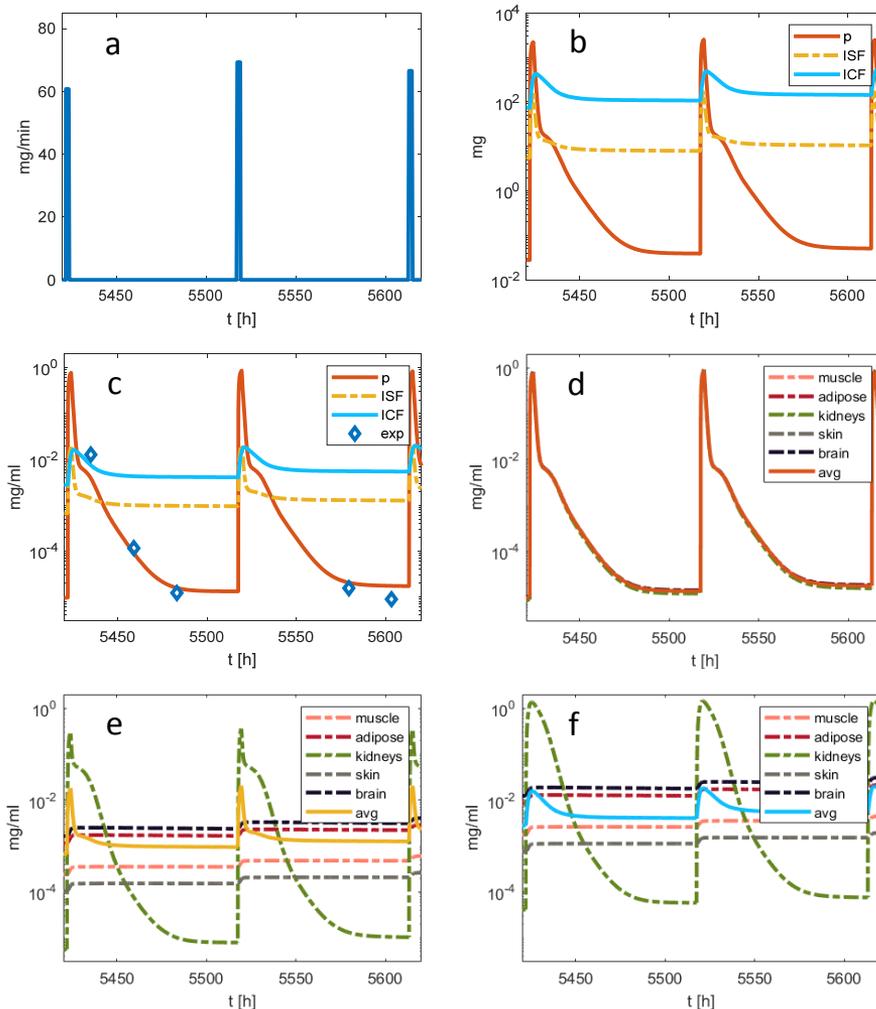


Figure 2 – Simulated curves for “630.3” patient: a) infusion rate, b) drug amounts in plasma, ISF, ICF, c) simulated and experimental concentrations in plasma, compared with the average ISF and ICF concentrations. Panels d-f compare the average concentration with the trends in a few tissues and organs (muscle, adipose tissue, kidneys, skin, and brain) related to d) plasma, e) ISF, f) ICF.

Conversely, the predicted concentrations in ISF and ICF differ significantly from tissue to tissue, consistently with the wide range of capillary exchange features (Table 1). For example, the predicted ISF concentration in kidneys (Fig. 2e) closely resembles the

predicted plasma concentration, thanks to both high capillary density and large open fraction of fenestrated capillaries. On the contrary, muscle and skin reach lower ISF concentration values due to slower capillary exchanges. Interestingly, the predicted ISF concentrations of adipose tissue are slightly higher than the average, despite the capillary features similar to those of the skin (low capillary density and permeability), likely due to the significantly smaller ISF and ICF volumes. The paramount effect of capillary exchange on predicted concentrations is also reflected by ICF concentrations, which present trends that are analogous to those described in ISF.

4. Conclusions

We presented a whole-body PBPK model for HDMTX featuring 46 body compartments, which was successfully identified and validated with an experimental dataset of 657 plasma concentrations. Plasma predictions showed a satisfactory agreement with experimental data and were comparable to the results of the minimal PBPK model. The detailed simulations of concentration within the plasma, ISF, and ICF of each modeled tissue and organ allowed straightforward physiological interpretations and better insights. The results allowed carrying out comprehensive analyses and defining improved therapeutical targets.

This work demonstrates an original approach to the development of a whole-body PBPK based on the extension of a minimal PBPK model. By introducing an *a priori* description of the 46 fluid volumes, 16 blood flows, and 14 capillary exchanges in each tissue and organ, the resulting whole-body PBPK model increased the level of detail of the minimal PBPK model while maintaining equivalent predictive capabilities, without amplifying the complexity and dimensionality of the identification procedure. The proposed modeling approach can serve as a reference for the development of other whole-body PBPK models that can retain their strong physiological consistency while minimizing the number of adaptive parameters, which facilitates the development process.

References

- Bundgaard, M. (1988). The paracellular pathway in capillary endothelia. *Vascular Endothelium in Health and Disease*, 3-8.
- Cao, Y., & Jusko, W. J. (2012). Applications of minimal physiologically-based pharmacokinetic models. *Journal of Pharmacokinetics and Pharmacodynamics*, 39, 711-723.
- Freitas, R. A. (1999). *Nanomedicine, Volume I: basic capabilities (Vol. 1)*: Landes Bioscience Georgetown, TX.
- Levick, J. R. (2004). Revision of the Starling principle: new views of tissue fluid balance. *The Journal of physiology*, 557, 704-704.
- Nestorov, I. (2003). Whole Body Pharmacokinetic Models. *Clinical Pharmacokinetics*, 42, 883-908.
- Pesenti, G., Foppoli, M., & Manca, D. (2021a). A minimal physiologically based pharmacokinetic model for high-dose methotrexate. *Cancer chemotherapy and pharmacology*.
- Pesenti, G., Massari, D., & Manca, D. (2021b). Physiological estimates of human fluid volumes. In preparation.
- Schmitt, W. (2008). General approach for the calculation of tissue to plasma partition coefficients. *Toxicology in Vitro*, 22, 457-467.
- Stader, F., Siccardi, M., Battagay, M., Kinvig, H., Penny, M. A., & Marzolini, C. (2019). Repository Describing an Aging Population to Inform Physiologically Based Pharmacokinetic Models Considering Anatomical, Physiological, and Biological Age-Dependent Changes. *Clinical Pharmacokinetics*, 58, 483-501.
- Valentin, J. (2002). Basic anatomical and physiological data for use in radiological protection: reference values: ICRP Publication 89. *Annals of the ICRP*, 32, 1-277.

Analysis of an industrial adsorption process based on ammonia chemisorption: model validation

Cristian Cardenas^{a,b}, Abderrazak M. Latifi^{a,*}, Cécile Vallières^a and Stéphanie Marsteau^b

^a*LRGP, Université de Lorraine, 1 rue Grandville, Nancy 54000, France*

^b*Institut national de recherche et de sécurité, 1 Rue du Morvan, Vandoeuvre 54519, France
abderrazak.latifi@univ-lorraine.fr*

Abstract

In this study, two one-dimensional first-principles models are developed to simulate an isothermal process where ammonia is adsorbed on a doped activated carbon. A single-scale LDF model and a multi-scale Fickian model. Both models are described by mass balance equations along with thermodynamic and kinetic relations. The modeling methodology is based on experimental measurements of ammonia adsorption isotherms and breakthrough fronts combined with a parameter estimability analysis and an identification method. The models are implemented and solved within COMSOL Multiphysics[®]. The Kolmogorov-Smirnov non-parametric test is used to statistically validate the predictions of the models. The results of the identified and validated multi-scale model using the fitted parameters show very good agreement with the experimental measurements. This is confirmed by the performance indices such as the Pearson correlation coefficient, the mean-square error, the (modified) index of agreement and the chi-square error.

Keywords: Air purification boxes, Ammonia adsorption, Doped activated carbon, Experimental measurements, Model validation.

1. Introduction

In chemical hazards, ammonia is a component that deserves careful study to protect the health of exposed people. Indeed, when used in large quantities, for example in agro-chemistry, in building and civil engineering, or in waste treatment and recycling, ammonia can be dangerous to the respiratory tract. Its concentration in the air should therefore be minimized and controlled. One of the most used processes to reduce the ammonia concentration in the air is adsorption.

In this work, an industrial ammonia adsorption process used in air purification, where ammonia is adsorbed on a doped activated carbon, is considered. The activated carbon is stacked in a parallelepipedic bed and placed in the centre of a box equipped with a feed and an outlet. The box is designed so that air highly concentrated in ammonia is sucked through the fixed-bed of activated carbon, and the purified air is recovered at the outlet. The fixed-bed is replaced once the ammonia starts to breakthrough, and naturally, the idea is to reduce the frequency of the renewal of activated carbon. To this end, different adsorption models are developed ranging from a simple single-scale LDF (linear driving force) model to a complex multi-scale model (Cardenas, 2021; Cardenas et al., 2021). All these models are identified and then tested with additional breakthrough front measurements performed at three gas flow rates and two concentrations, different from those used for parameter identification.

In the present paper, the objective is to statistically validate these models and use them in order to optimize the design and operation of air purification boxes. Many statistical tests are used to validate the predictions of the models. They include the Student t -test, the Fisher-Snedecor F -test, the chi-square test, etc. However, the use of these parametric statistical tests requires the distributions of the model predictions and measurements to be normal which is very often not the case. To overcome this problem, the Kolmogorov-Smirnov non-parametric test is used. In addition, the performance of the models is evaluated using different criteria such as the Pearson correlation coefficient, the root mean-squares error, the (modified) index of agreement, and the chi-square error.

2. Ammonia adsorption process modeling

Different models are developed to describe the behavior of dynamic adsorption of ammonia on doped activated carbon. They range from a single-scale model to more complex models involving multi-scales (Cardenas, 2021). The modeling is based on mass balance, momentum balance, thermodynamic, hydrodynamic, and kinetic equations. These balances are established at both the adsorption box scale and at the particle scale, and are based on the following assumptions: (i) the gas mixture obeys the ideal gas law, (ii) the pressure drop in the box is neglected, (iii) ammonia is the only adsorbed molecule, (iv) the resistance to the mass transfer through the boundary layer surrounding the solid particles is characterized by a mass transfer coefficient, (v) the intra-particle mass transport is due to the Fickian diffusion and is characterized by an effective diffusion coefficient in the macropores, (vi) the mass transport of ammonia in the crystal micropores is due to intracrystalline diffusion, (vii) the Toth equation describes the equilibrium isotherm of the gas phase with the adsorbent, (viii) the adsorbent particles are assumed to be spherical and homogeneous in size and density, (ix) the temperature of the box and the physical properties of the adsorbent are assumed to be constant, (x) only the axial dimension of the box is considered (mono-dimensional model). On the other hand, for the lack of space reasons, only two models will be presented below: the single-scale LDF model and the multi-scale model.

2.1. Single-scale LDF model

In this model, the mass balance of the gas phase ammonia in the box is described by a dynamic equation involving diffusion, convection, and adsorption, given as:

$$\frac{\partial c}{\partial t} - D_{ax} \frac{\partial^2 c}{\partial z^2} + \frac{\partial (vc)}{\partial z} = - \frac{1 - \varepsilon_b}{\varepsilon_b} \frac{\partial \bar{q}}{\partial t} \quad (1)$$

where c and \bar{q} (mol.m^{-3}) are the ammonia concentrations in the gas phase and in the adsorbed phase respectively, v (m.s^{-1}) is the interstitial velocity, D_{ax} ($\text{m}^2.\text{s}^{-1}$) is the axial dispersion coefficient, and ε_b is the bed void fraction.

The kinetics of mass transfer is approximated by means of the LDF equation as:

$$\frac{\partial \bar{q}}{\partial t} = k_{LDF} (q_e^* - \bar{q}) \quad (2)$$

where q_e^* (mol.m^{-3}) is the concentration of ammonia in the solid phase at equilibrium, and k_{LDF} (s^{-1}) is the overall mass transfer coefficient which accounts for the contributions of gas/solid mass transfer, and macropore and intracrystalline diffusion.

The initial conditions of the variables are given as follows:

- For $0 \leq z \leq L$, at $t = 0$: $c = 0$, $\bar{q} = 0$,

and the boundary conditions are expressed as:

- For $t > 0$, at $z = 0$: $-D_{ax} \frac{\partial c}{\partial z} = v(c_0 - c)$, $\bar{q} = 0$ and at $z = L$: $\partial c / \partial z = 0$, where c_0 (mol.m^{-3}) is the concentration of ammonia at the box inlet.

2.2. Multi-scale model

This model takes into account the following three spatial scales: the bed, the particle, and the crystal scales. The particle is divided into two phases: a solid phase for the crystal (micropores) and a gas phase (macropores) and are coupled by continuity at outer surface of the crystal. The model equations are detailed below.

(i) In the box, the ammonia mass balance equation in the gas phase is expressed as follows:

$$\frac{\partial c}{\partial t} - D_{ax} \frac{\partial^2 c}{\partial z^2} + \frac{\partial(vc)}{\partial z} = - \frac{1 - \epsilon_b}{\epsilon_b} \frac{3}{r_{pe}} k_f (c - c_{pe}|_{r_p=r_{pe}}) \quad (3)$$

(ii) In the particle, assuming the absence of lateral and azimuthal variations of the concentration, the mass balance in the radial direction can be written as:

$$\epsilon_{pe} \frac{\partial c_{pe}}{\partial t} + (1 - \epsilon_{pe}) \frac{\partial \bar{q}}{\partial t} = \frac{D_e}{r_p^2} \frac{\partial}{\partial r_p} \left(r_p^2 \frac{\partial c_{pe}}{\partial r_p} \right) \quad (4)$$

(iii) In the crystal, the ammonia mass balance equation is given as:

$$\frac{\partial \bar{q}}{\partial t} = \frac{15D_\mu}{r_c^2} (q_e^* - \bar{q}) = k_1 (q_e^* - \bar{q}) \quad (5)$$

where $c(z, t)$ and $c_{pe}(z, r_p, t)$ (mol.m^{-3}) are the concentrations of ammonia in the gas phase of the box and in the particle respectively, r_{pe} and r_c (m) are the radii of the particle and the crystal respectively, k_f (m.s^{-1}) is the external film mass transfer coefficient, $c_{pe}|_{r_p=r_{pe}}$ (mol.m^{-3}) is the ammonia concentration at the surface of the particle, \bar{q} (mol.m^{-3}) is the average amount of ammonia adsorbed on the crystal, D_e and D_μ ($\text{m}^2.\text{s}^{-1}$) are the effective and intracrystalline diffusion coefficients respectively, and k_1 (s^{-1}) is the internal mass transfer coefficient.

The initial conditions of the variables are given as:

- For $0 \leq z \leq L$, at $t = 0$: $c = 0$, $c_{pe} = 0$, $\bar{q} = 0$,

and the boundary conditions are expressed as:

- For t and $r_p > 0$, at $z = 0$: $-D_{ax} \frac{\partial c}{\partial z} = v(c_0 - c)$, $\bar{q} = 0$ and at $z = L$: $\partial c / \partial z = 0$;

- For t and $z > 0$, at $r_p = 0$: $\frac{\partial c_{pe}}{\partial r_p} = 0$ and at $r_p = r_{pe}$: $-D_e \frac{\partial c_{pe}}{\partial r_p} = k_f (c - c_{pe}|_{r_p=r_{pe}})$

2.3. Implementation and simulation of the models

The external mass transfer coefficient is calculated using the following Wakao and Funazkri (1978) correlation:

$$k_f = \frac{D_m Sh}{2r_{pe}} \quad \text{with} \quad Sh = c_{m1} + c_{m2} Re_p^{0.6} Sc^{1/3} \quad \text{and} \quad c_{m1} = 2 \quad (6)$$

where Sh , Re_p and Sc are the dimensionless Sherwood, Reynolds and Schmidt numbers respectively, D_m ($\text{m}^2.\text{s}^{-1}$) is the molecular diffusion coefficient. For the axial dispersion coefficient, the correlation of Rastegar and Gu (2017) is used:

$$D_{ax} = \frac{2r_{pe}v}{Pe'} \quad \text{with} \quad \frac{1}{Pe'} = \frac{0.7D_m}{2r_{pe}v} + \frac{\epsilon_b}{c_{m3} + c_{m4} (Re_p)^{0.59}} \quad (7)$$

where Pe' is the dimensionless Peclet number. The effective diffusion coefficient is linked to the mixture diffusion by means of the following relation: $D_e = \varepsilon_{pe}^{4/3} D_M$. It is important to point out that D_{ax} and k_f increase with increasing gas flow rate. To keep this property, c_{m2} , c_{m3} and c_{m4} are also considered as unknown parameters to be identified from the measurements. Hence, in the LDF model three unknown parameters are to be estimated, i.e., k_{LDF} , c_{m3} and c_{m4} , whereas in the multi-scale model, five unknown parameters are to be determined, i.e., k_1 , D_M , c_{m2} , c_{m3} and c_{m4} . Comsol Multiphysics[®] version 5.5 is used to implement and solve the set of partial differential equations of the two models.

2.4. Parameter estimability analysis and identification

A preliminary step in the development of a reliable mathematical model, before the problem of parameter identification, is to evaluate the structural identifiability and estimability of the model parameters. The objective is to determine the most estimable parameters from the available measurements. The estimability algorithm used in this work is the same as the one developed by Yao et al. (2003) and used in Cardenas et al. (2020). Moreover, the choice of the estimability threshold value which defines the limit between estimable and non-estimable parameters, is still arbitrary and depends on the studied process. In this work, it is set equal to 0.04 as in Yao et al. (2003) and gives a good idea about the actual sensitivity of the adsorption models to the unknown parameters. However, it should be noted that there are more advanced and sophisticated methods in the literature (Bouchkira et al., 2021).

2.4.1. Parameter estimability analysis results

The estimability analysis algorithm using the orthogonalization algorithm described by Yao et al. (2003) is applied to the ammonia adsorption process models and leads to the following estimability order: $k_{LDF} > c_{m3} > c_{m4}$ for the LDF model, and $k_1 > D_M > c_{m2} > c_{m3} > c_{m4}$ for the multi-scale model. It should be noted that all parameters are estimable on the basis of available experimental measurements.

2.4.2. Parameter identification

The experimental data used consist of breakthrough front measurements carried out at four different concentrations of ammonia (0.066, 0.054, 0.041 and 0.029 mol.m⁻³) and three different gas flow rates (13.8, 9.0 and 4.4 L.min⁻¹), for a total of 12 experiments to identify the parameters, using the Toth isotherm model. The results obtained show that the LDF model fits the adsorption process well with identified parameter values close to those in the literature (see figure 1). However, this model does not allow to understand all the different mechanisms involved in the adsorption box since all the mass transport resistances are lumped into a single global coefficient. The multi-scale model which shows very good agreement between measurements and predictions allows a better understanding the adsorption phenomena by considering the three scales: gas phase in the box, gas phase and solid phase in the particle.

2.4.3. Model prediction performance

The prediction performance of the model is quantitatively assessed by means of criteria which compares model predictions with the experimental measurements. In this study, it is evaluated according to two different criteria: The root mean-square error function (*RMSE*) and the Pearson correlation coefficient (*r*).

3. Model validation

The identified models are then validated with five additional breakthrough front measurements performed at three gas flow rates 17, 15 and 7 L.min⁻¹ with an ammonia concentration of 0.041 mol.m⁻³; and at a gas flow rate of 13.8 L.min⁻¹ with two ammonia concentrations of 0.021 and 0.012 mol.m⁻³, different from those are used for parameter identification. The optimized values of the parameters previously estimated are used to simulate these fronts (Cardenas, 2021). Figure 1 shows the comparison between the time-varying predicted and measured fronts at different gas flow rates and different ammonia concentrations.

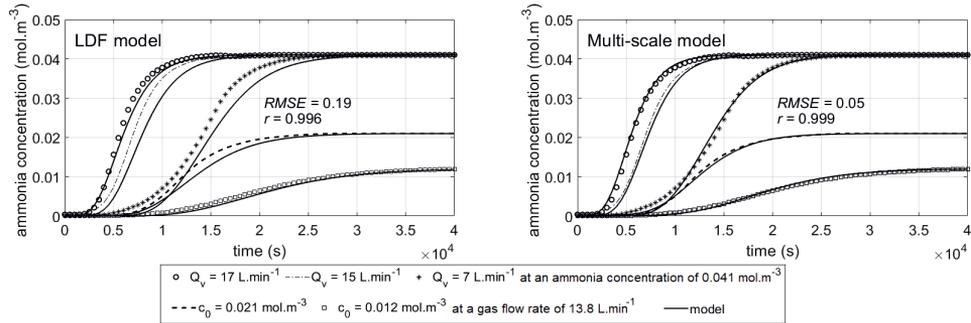


Figure 1: Comparison between predicted and measured breakthrough fronts at three gas flow rates 17, 15 and 7 L.min⁻¹ with an ammonia concentration of 0.041 mol.m⁻³; and at a gas flow rate of 13.8 L.min⁻¹ with two ammonia concentrations of 0.021 and 0.012 mol.m⁻³

The high values obtained for the Pearson coefficient (close to unity, $r \geq 0.996$) and low $RMSE$ values (at most 0.19) show excellent prediction performance of the developed models. Furthermore, to validate these models, many statistical tests may be considered and can be parametric or non-parametric. Parametric tests are based on assumed statistical distributions in the data. Therefore, certain validity conditions must be verified for their results to be reliable. For example, the Student t -test for independent samples is only reliable if the data associated with each sample follow a normal distribution and if the variances of the samples are homogeneous. However, since the validity conditions are not always met, which is the case in our data, the parametric tests can no longer be used. Non-parametric tests which are not based on statistical distributions can therefore be considered. More specifically, the Kolmogorov-Smirnov test is used and is based on the calculation of the divergence (D_n) between two distributions. It is expressed as the maximum value between the predicted (c) and measured (c^{mes}) ammonia concentration values at each instant t , i.e.:

$$D_n = \max_t |c(t) - c^{mes}(t)| \quad (8)$$

The null hypothesis (H0) is defined as: the predicted and measured breakthrough fronts have the same distribution with the probability of 95%. Therefore, the alternative hypothesis (H1) is that the predicted and measured breakthrough fronts do not have the same distribution with the probability of 95%. On the other hand, the critical value of the divergence (D_n^*) corresponding to the number of experimental measurements minus one, and to the significance level of the test (5%) is expressed by $(D_n^* = \frac{1.36}{\sqrt{n}})$ and is equal to 0.06. To pass the validity test, the divergence D_n must be less or equal to the critical value D_n^* . In table 1, bold values correspond to the calculated divergence which is less than the critical value, meaning that the null hypothesis (H0) cannot be rejected. We can therefore affirm that the predicted and measured breakthrough fronts have the same distribution with a probability of 95%. In addition, the performance of the model is evaluated by means of different criteria, i.e., the index of agreement (IA), the modified index of agreement

(IAM), and the nonlinear chi-square error (χ^2). These performance criteria are calculated and their values are reported in table 1. It can be seen that the indices *IA* and *IAM* are close to unity and the low values of chi-square confirm the excellent agreement between the model predictions and measurements. The multi-scale model is therefore the most performing prediction model, as it has successfully passed all the tests. On the other hand, despite the good performance indices of the LDF model, its validation at the 5% significance level is not guaranteed.

Table 1: Performance criteria for adsorption models

	LDF model					Multi-scale model				
Q_v L.min ⁻¹	17	15	7	13.8		17	15	7	13.8	
c_0 mol.m ⁻³	0.041			0.021	0.012	0.041			0.021	0.012
D_n	0.076	0.149	0.128	0.127	0.076	0.048	0.052	0.047	0.038	0.040
χ^2	0.001	0.032	0.009	0.004	0.007	0.001	0.020	0.003	0.002	0.002
<i>IA</i>	0.996	0.995	0.996	0.996	0.998	0.999	0.999	0.999	0.999	0.999
<i>IAM</i>	0.979	0.954	0.955	0.961	0.961	0.986	0.982	0.988	0.991	0.978

4. Conclusions

Among the developed models, the multi-scale model is the most effective since it exhibits higher performance indices and successfully passes the validation tests considered. This model can now be used as a tool to predict the breakthrough time in an ammonia adsorption box with doped activated carbon to protect operators working in an ammonia-contaminated environment. It can also be used to study the effect of different operating conditions such as temperature, concentration, flow rate, bed geometry, porosity, etc., and more generally in optimal design and operation of such adsorption processes.

References

- I. Bouchkira, A. M. Latifi, L. Khmar, S. Benjelloun, 2021. Global sensitivity based estimability analysis for the parameter identification of pitzer's thermodynamic model. *Reliability Engineering & System Safety* 207, 107263.
- C. Cardenas, 2021. Analyse et modélisation du comportement des caissons d'épuration de l'air équipant les engins de chantier pour la protection des opérateurs contre les gaz et vapeurs. Ph.D. thesis, Université de Lorraine. URL <https://www.theses.fr/s283000>
- C. Cardenas, A. M. Latifi, C. Vallières, S. Marsteau, L. Sigot, 2021. Analysis of an industrial adsorption process based on ammonia chemisorption: Modeling and simulation. *Computers & Chemical Engineering* 154, 107474.
- C. Cardenas, S. Marsteau, L. Sigot, C. Vallières, A. M. Latifi, 2020. Analysis of an industrial adsorption process based on ammonia chemisorption: Modeling and simulation. In: *Computer Aided Chemical Engineering*. Elsevier, pp. 625–630.
- S. O. Rastegar, T. Gu, 2017. Empirical correlations for axial dispersion coefficient and pecllet number in fixed-bed columns. *Journal of Chromatography A* 1490, 133–137.
- N. Wakao, T. Funazkri, 1978. Effect of fluid dispersion coefficients on particle-to-fluid mass transfer coefficients in packed beds. *Chemical Engineering Science* 33 (10), 1375–1384.
- K. Z. Yao, B. M. Shaw, B. Kou, K. B. McAuley, D. W. Bacon, 2003. Modeling ethylene/butene copolymerization with multi-site catalysts: Parameter estimability and experimental design. *Polymer Reaction Engineering* 11 (3), 563–588.

CFD modeling and simulation of an ammonia adsorption process

Flora Esposito^{a,b}, Cristian Cardenas^{a,b}, Abderrazak M. Latifi^{a,*} and Stéphanie Marsteau^b

^a*LRGP, Université de Lorraine, 1 rue Grandville, Nancy 54000, France*

^b*Institut National de Recherche et de Sécurité, 1 Rue du Morvan, Vandoeuvre 54519, France
abderrazak.latifi@univ-lorraine.fr*

Abstract

Due to increased ammonia emissions, there has been a growing interest in developing ammonia removal technologies. Adsorption is one of the most widely used separation techniques to reduce the concentration of ammonia in the air. This study aims to use CFD-based modeling to investigate ammonia adsorption on a doped activated carbon in an industrial air purification box. An experimental study is first carried out, then followed by detailed modeling of the processes that occur in the box. The hydrodynamic and adsorption models are implemented and solved using COMSOL Multiphysics[®] software in the 5.6 version. Adsorption is modeled by means of a multi-scale Fickian model, which demonstrated to be very efficient to predict the experimental data. The performance of the predictions of the model is assessed using several statistical criteria such as the mean-square error and the Pearson correlation coefficient.

Keywords: Air purification boxes, Ammonia adsorption, Doped activated carbon, Experimental measurements, Modeling and simulation

1. Introduction

Ammonia is one of the most toxic pollutants, it causes massive burning by inducing an exothermic reaction in the respiratory system. Despite its harmful effects on humans, it is one of the most widely used gases in the manufacture of fertilizer, polymers, and resins. Controlling pollutant emissions at the source is the most effective way to reduce total air pollution and protect human health. However, if this is not possible, pollutant concentrations should be targeted, with preventative and abatement measures. Among different air purification techniques, adsorption of pollutants on activated carbon is one of the most effective. Air purifiers are becoming increasingly popular as a means of removing contaminants from ambient air. CFD is a powerful tool for modeling and investigating mass transfer and diffusion mechanisms in fixed-bed adsorption boxes in less time and at a lower cost than experiments, thereby improving the design and optimization of processing equipment. This paper deals with a CFD modeling and simulation of an industrial fixed-bed adsorption box where ammonia is removed from the air using a doped activated carbon. The model predictions are compared to experimental measurements of breakthrough fronts carried out at different gas flowrates.

2. Mathematical modeling

The air purification box used for the flow modeling is the Honeywell SP Défence Filagric K100. Depending on the manufacturer, the air flowrate supplied can range from 40 to 100 m³.h⁻¹. Thus,

the modeling and simulation of the phenomena will concern the two extreme values of the flow range.

2.1. Hydrodynamic model

The hydrodynamic model developed is based on the following assumptions: (i) negligible gravity force, (ii) the flow is considered as incompressible, (iii) the gas properties are those of air, (iv) the temperature of the box and the physical properties of the adsorbent are constant, (v) the average properties of the porous medium are isotropic. On the other hand, particular attention to the flow regime is required for proper modeling of the problem. As shown in figure 1, the purification box does not have a constant section. Hence, the Reynolds number is calculated in inlet, outlet and bed sections. The values are all exceeding the critical Reynolds number found in the literature (Reynolds, 1883; Hanks and Ruo, 1966), showing that the purification box operates in a turbulent flow regime.

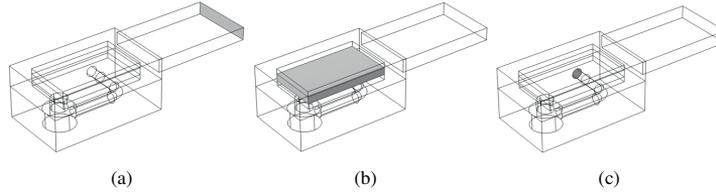


Figure 1: (a) Inlet section (b) Fixed bed section (c) Outlet section

The flow model is then described by the continuity and the Navier-Stokes equations along with the Brinkman-Forchheimer equation for the porous medium domain. The continuity equation is expressed as:

$$\nabla \cdot \mathbf{u} = 0 \quad (1)$$

where \mathbf{u} ($\text{m}\cdot\text{s}^{-1}$) is the average velocity vector. The momentum balance is given by Reynolds averaged Navier-Stokes (RANS) equations as:

$$\rho_f \nabla \cdot \mathbf{u} + \nabla \cdot \overline{\rho_f (\mathbf{u}' \times \mathbf{u}')} = -\nabla p + \nabla \cdot (\mu_f (\nabla \mathbf{u} + (\nabla \mathbf{u})^T)) + \mathbf{F} \quad (2)$$

where p (Pa) is the average pressure, ρ_f ($\text{kg}\cdot\text{m}^{-3}$) is the fluid density, μ_f (Pa.s) is the fluid dynamic viscosity and \mathbf{F} (N) is the vector representing all the external forces applied to the fluid. However, because the activated carbon filter is a porous medium, the RANS alone cannot describe the fluid flow inside the process. The following Brinkman equations are therefore used:

$$\frac{1}{\varepsilon_b} \rho_f \mathbf{u} \cdot \nabla \frac{1}{\varepsilon_b} = -\nabla p + \nabla \cdot \left(\frac{\mu_f}{\varepsilon_b} (\nabla \mathbf{u} + (\nabla \mathbf{u})^T) \right) - \left(\mu \kappa^{-1} + \beta \rho_f |\mathbf{u}| + \frac{Q_m}{\varepsilon_b^2} \right) \mathbf{u} + \mathbf{F} \quad (3)$$

where ε_b ($\text{m}^3\cdot\text{m}_{\text{box}}^{-3}$) is the porosity of the bed, κ (m^2) is the permeability, \mathbf{u} ($\text{m}\cdot\text{s}^{-1}$) is the fluid average velocity over the porous medium, which is related to the interstitial velocity v ($\text{m}\cdot\text{s}^{-1}$). The Brinkman-Forchheimer drag term β (m^{-1}) and the permeability κ (m^2) are modeled using Ergun equation. The corresponding boundary conditions are given as follows. At the inlet, the velocity is determined by the volumetric flow rate Q_v ($\text{m}^3\cdot\text{s}^{-1}$) and the inlet section Ω_m (m^2) as: $\mathbf{u} = -U_0 \mathbf{n}$ and $U_0 = \frac{Q_v}{\Omega_m}$, where \mathbf{n} represents the boundary outward pointing normal. At the outlet, the following pressure condition is defined: $(-p \mathbf{I} + (\mu + \mu_T) (\nabla \mathbf{u} + (\nabla \mathbf{u})^T)) \mathbf{n} = -p_0 \mathbf{n}$, where p_0 (Pa) is the gauge pressure, set equal to 0 kPa, and \mathbf{I} is the identity matrix. A crucial point

is the correct modeling of the fluid flow in the vicinity of the wall, given the formation of a viscous sub-layer and a buffer region. The behavior of the flow near the walls is therefore approximated, since the RANS model is not suitable anymore to properly describe this area at low Reynolds numbers. A non-slip condition is then considered on every wall, and this is written as: $\mathbf{u} = 0$. There is no need for an additional boundary condition since the turbulence model we used is algebraic.

2.2. Adsorption model

The adsorption model developed takes into account three scales: the fixed-bed scale, the intra-particle scale and the crystal scale. It is based on the following assumptions: (i) the gas mixture obeys the ideal gas law, (ii) only ammonia is adsorbed, (iii) resistance to mass transfer across the boundary layer surrounding the solid particles is characterised by an external film mass transfer coefficient, (iv) intra-particle mass transport is due to Fickian diffusion, (v) intra-crystalline mass transport of ammonia in the crystal micropores is due to the intra-crystalline diffusion, (vi) the adsorbent particles are assumed to be spherical and homogeneous in size and density, (vii) only the axial dispersion in the bed is taken into account (one-dimensional model). The amount of ammonia adsorbed on the doped activated carbon is described by the Toth adsorption isotherm, better detailed by Cardenas (2021). The model equations are given below.

(i) The ammonia mass balance in the bed domain is expressed as:

$$\frac{\partial c}{\partial t} + \nabla \cdot (-D_{ax} \nabla c) + v \cdot \nabla c = -\frac{(1 - \varepsilon_b)}{\varepsilon_b} \frac{3}{r_{pe}} k_f (c - c_{pe}|_{r_p=r_{pe}}) \quad (4)$$

where c and $c_{pe}|_{r_p=r_{pe}}$ (mol.m^{-3}) are the concentrations of ammonia in the gas phase and on the particle surface respectively, k_f (m.s^{-1}) is the external mass transfer coefficient, D_{ax} ($\text{m}^2.\text{s}^{-1}$) is the axial dispersion coefficient, and r_{pe} (m) is the mean radius of the particle.

(ii) The intra-particle mass balance in the radial direction can be written as:

$$\varepsilon_{pe} \frac{\partial c_{pe}}{\partial t} + (1 - \varepsilon_{pe}) \frac{\partial q}{\partial t} = \frac{D_e}{r_p^2} \frac{\partial}{\partial r_p} \left(r_p^2 \frac{\partial c_{pe}}{\partial r_p} \right) \quad (5)$$

where ε_{pe} ($\text{m}_g^3.\text{m}_{pe}^{-3}$) is the porosity of the particle, q (mol.m^{-3}) is the amount of ammonia adsorbed on the solid and D_e ($\text{m}^2.\text{s}^{-1}$) is the effective diffusion coefficient. The associated initial and boundary conditions of equations (4-5) are:

— for $t = 0$: $c = 0$, $c_{pe} = 0, \forall r_p$ and z ; — for $z = 0$: $\frac{\partial c}{\partial z} = \frac{v}{D_{ax}} (c - c_0), \forall r_p$ and z

— for $z = H$: $\frac{\partial c}{\partial z} = 0, \forall r_p$ and t ; — for $r_p = 0$: $\left. \frac{\partial c_{pe}}{\partial r_p} \right|_{r_p=0} = 0, \forall z$ and t

— for $r_p = r_{pe}$: $-D_e \frac{\partial c_{pe}}{\partial r_p} = k_f (c - c_{pe}|_{r_p=r_{pe}}), \forall z$ and t

where c_0 (mol.m^{-3}) is the concentration of ammonia at the box inlet.

(iii) The ammonia mass balance in the crystal is given as:

$$\frac{\partial \bar{q}}{\partial t} = k_1 (q_e^* - \bar{q}) \quad (6)$$

where \bar{q} (mol.m^{-3}) is the average amount of ammonia adsorbed on the crystal, q_e^* (mol.m^{-3}) is the amount of ammonia adsorbed at equilibrium, and k_1 (s^{-1}) is the internal mass transfer coefficient in the crystal. The associated initial condition is given as: for $t = 0$: $c = 0, c_{pe} = 0, \bar{q} = 0, \forall r_p$ and z .

3. Results and discussion

It is important to notice that since the geometry shown in figure 1 leads us to very high computational times, we decided to change the ventilator shape in order to take advantage of symmetry

and therefore only implement one half of the domain.

3.1. Mesh quality

Accurate solution of CFD models requires mesh refinement to achieve a grid-independent solution. The first mesh used for the discretization of the domain is a fine physics controlled mesh, and its convergence is evaluated with respect to a user-controlled one. The first has 2.08×10^6 elements and an average quality of 0.7, while the latter has a number of elements of 3.82×10^6 and a mean quality of 0.74. Both meshes predicted the same velocity and pressure profiles; we concluded then that our solution is grid independent.

3.2. Pressure and velocity fields

Figure 2 (a) shows the pressure profile in the activated carbon bed for an inlet flowrate of $100 \text{ m}^3 \cdot \text{h}^{-1}$. The pressure drop computed is about 113 Pa. It decreases to 38 Pa when the overall inlet flowrate decreases to $40 \text{ m}^3 \cdot \text{h}^{-1}$. Figure 2 (b) shows a heterogeneous velocity distribution throughout the bed with an average velocity of about $0.22 \text{ m} \cdot \text{s}^{-1}$. However, this is true only in the gas inlet section. In fact, as soon as the flow encounters the activated carbon particles the velocity distribution becomes more homogeneous.

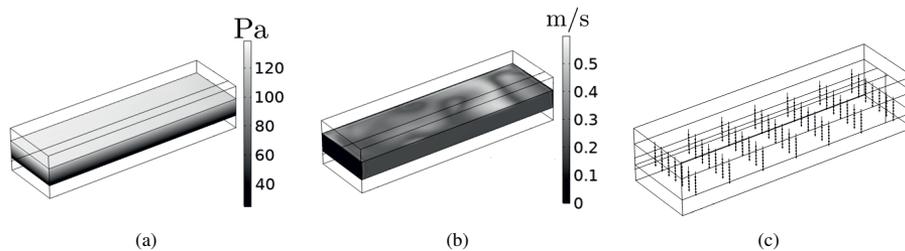


Figure 2: Fixed bed hydrodynamics: (a) Pressure drop (b) Velocity distribution (c) Arrow plot

Figure 2 (c) shows that the velocity field in the fixed bed domain is homogeneous. It is therefore clear that the z component is the only significant one, but in order to prove it, a volume average evaluation group is used to determine the values of the three components of the velocity. The evaluation showed us that the absolute value of the z component is very close to the velocity magnitude in the fixed bed domain, while the x and y components have a magnitude of $10^{-6} \text{ m} \cdot \text{s}^{-1}$.

3.3. Residence time distribution

Numerical residence time distributions (RTDs) are computed in order to determine the time available for adsorption. Figure 3 shows a normalized RTD and the resulting mean residence time. The corresponding space-time is equal to 3.2 s. It can be seen that, the average residence time is higher than the space-time, meaning that the molecules of ammonia remain on average longer than would remain if the velocity distribution was homogeneous.

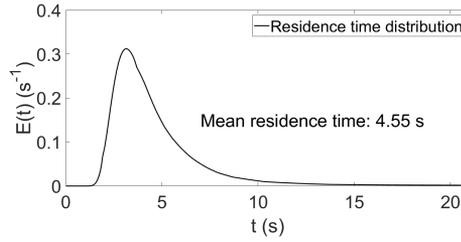


Figure 3: Residence time distribution in the purification box

On the other hand, the RTD curve exhibits a long tail indicating the presence of dead zones which degrade the air purification box performance.

3.4. Adsorption results

In order to solve the model equations, it is necessary to determine the unknown parameters such as the internal mass transfer coefficient (k_1), the axial dispersion coefficient (D_{ax}), the effective diffusion coefficient (D_e), and the external mass transfer coefficient (k_f). The latter is calculated by means of the Wakao and Funazkri (1978) correlation as:

$$k_f = \frac{D_m Sh}{2r_{pe}} \quad \text{with} \quad Sh = c_{m1} + c_{m2} Re_p^{0.6} Sc^{1/3} \quad \text{and} \quad c_{m1} = 2 \quad (7)$$

where D_m ($\text{m}^2 \cdot \text{s}^{-1}$) is the molecular diffusion coefficient. The axial dispersion coefficient is evaluated using the Rastegar and Gu (2017) correlation as:

$$D_{ax} = \frac{2r_{pe} v}{Pe'} \quad \text{with} \quad \frac{1}{Pe'} = \frac{0.7 D_m}{2r_{pe} v} + \frac{\epsilon_b}{c_{m3} + c_{m4} (Re_p)^{0.59}} \quad (8)$$

The effective diffusion coefficient is linked to the mixture diffusivity by means of the following relation : $D_e = \epsilon_{pe}^{4/3} D_M$. It is noteworthy that D_{ax} and k_f increase with increasing gas flowrate. To keep this property, c_{m2} , c_{m3} and c_{m4} are also considered as unknown parameters to be identified from the experimental data. The most estimable parameters are determined from experimental measurements of the breakthrough fronts at different gas flowrates and the non-estimable parameters have their values fixed from the literature or from previous studies (see Cardenas et al. (2021)). A total number of five estimable parameters is therefore identified, i.e. k_1 , D_M , c_{m2} , c_{m3} and c_{m4} .

3.4.1. Parameter estimation

The COMSOL optimization module is used to estimate the optimal values for the model parameters. More specifically, the Monte Carlo solver is employed. Table 1 shows the values as well as their 95% confidence intervals (CI).

Table 1: Identified values of the Fickian Multi-scale model parameters.

Parameter	Value	CI (95 %)
k_1 (s^{-1})	5.79×10^{-4}	$\pm 1.26 \times 10^{-5}$
c_{m2}	0.10	$\pm 6.52 \times 10^{-3}$
c_{m3}	2.18×10^{-2}	$\pm 4.97 \times 10^{-3}$
c_{m4}	7.75×10^{-3}	$\pm 1.82 \times 10^{-3}$
D_M ($\text{m}^2 \cdot \text{s}^{-1}$)	6.53×10^{-5}	$\pm 2.54 \times 10^{-5}$

As the *CI* values are lower than the parameter values, we can conclude that the parameters are determined with good precision. The optimal values of the parameters are then used to simulate ammonia breakthrough fronts at different gas flowrates and at a concentration of 0.041 mol.m^{-3} , shown in figure 4. It can be seen that the breakthrough occurs more rapidly as the flow rate increases. A higher flowrate leads to a faster movement of the mass transfer zone. On the other hand, a lower flowrate results in slower transport of ammonia molecules, which increases the breakthrough time.

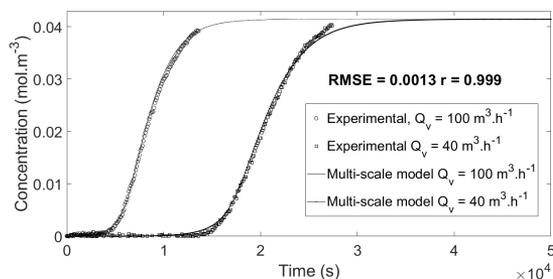


Figure 4: Comparison between predicted and measured breakthrough fronts

It should be noted that, based on the overall calculated *RMSE* (0.0013), the multi-scale model fits the adsorption process very well. Moreover, the Pearson coefficient value (very close to 1) shows that the model predictions are in excellent agreement with the experimental measurements.

4. Conclusions

After performing a hydrodynamic analysis of the air purification box, the CFD model was expanded to include the ammonia adsorption breakthrough profile on the activated carbon filter. Then, the COMSOL optimization tool allowed to determine the optimal values of the estimable parameters and their confidence intervals. For various inlet gas flowrates, the Fickian multi-scale model has proven to be the best suitable fit with experimental breakthrough profiles. However, by modeling adsorption in 3D, there is still the possibility for model improvement. Furthermore, the performance of the model predictions could be improved by working with different flowrates and inlet ammonia concentrations. Despite this, the numerical model developed could be used to further analyze and optimize the design and operation of industrial air purification boxes.

References

- C. Cardenas, 2021. Analyse et modélisation du comportement des caissons d'épuration de l'air équipant les engins de chantier pour la protection des opérateurs contre les gaz et vapeurs.
- C. Cardenas, A. M. Latifi, C. Vallières, S. Marsteau, L. Sigot, 2021. Analysis of an industrial adsorption process based on ammonia chemisorption: Modeling and simulation. *Computers and Chemical Engineering* 154, 107474.
- R. W. Hanks, H. C. Ruo, 1966. Laminar-turbulent transition in ducts of rectangular cross section. *Industrial and Engineering Chemistry Fundamentals* 5, 558–561.
- S. O. Rastegar, T. Gu, 2017. Empirical correlations for axial dispersion coefficient and pecllet number in fixed-bed columns. *Journal of Chromatography A* 1490, 133–137.
- O. Reynolds, 1883. III. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. Vol. 35. *Proceedings of the Royal Society of London*.
- N. Wakao, T. Funazkri, 1978. Effect of fluid dispersion coefficients on particle-to-fluid mass transfer coefficients in packed beds. correlation of sherwood numbers. *Chemical Engineering Science* 33, 1375–1384.

Polygeneration from sugarcane industries enhanced by functionalizing novel cultivars and excess thermal energy

Shoma Fujii^a, Yuichiro Kanematsu^b, Yasunori Kikuchi^{abc}

^a*Institute for Future Initiatives, The University of Tokyo, Tokyo, 113-8654, Japan*

^b*Presidential Endowed Chair for “Platinum Society”, The University of Tokyo, Tokyo 113-8656, Japan*

^c*Department of Chemical System Engineering, The University of Tokyo, Tokyo 113-8656, Japan*

Abstract

Changes in the generation of crop residues are expected due to the introduction of cultivars adapted to climate change, and there is a need to promote polygeneration of crop-derived production processes in order to effectively utilize crop residues for sustainable development. A process flow model of a Japanese sugar mill was developed and was incorporated into the integrated modeling of agricultural-industrial process. A boiler designed to maximize the power sold after 10 years from now cannot be operated in its current state because the heat transfer area of the bagasse boiler is too large to maintain the temperature of the final flue gas. A boiler designed to produce only the steam needed for sugar production after 10 years from now can make effective use of excess bagasse and can still be operated today.

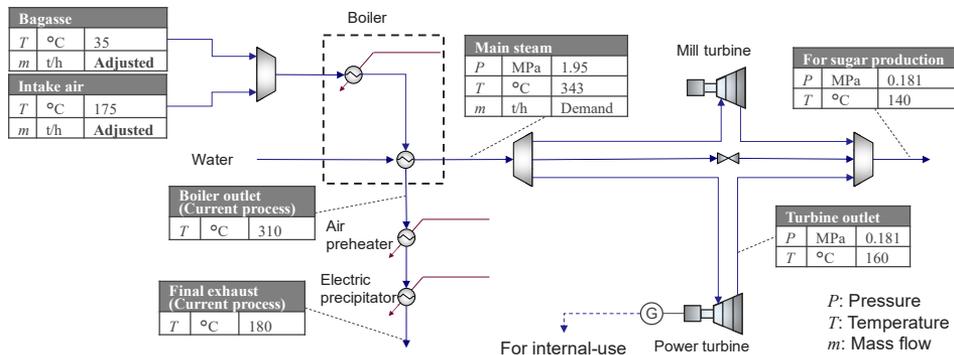
Keywords: Sugarcane bagasse, Bagasse boiler, Process modeling, Integrated modeling

1. Introduction

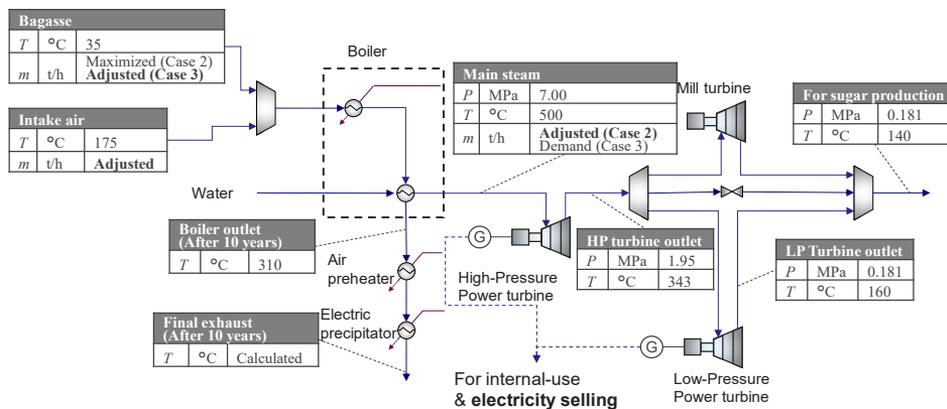
In the context of sustainable developments, polygeneration of crop-derived production processes will have a significant role. For many crops, agricultural operations are followed by industrial processing, and the residues may be used as fuel for operations. Currently, as the polygeneration of sugarcane industry, technological options for energy systems such as selling surplus electricity (Palacio et al., 2018; Birru et al., 2019), producing ethanol (Ohara et al., 2019) and producing unused heat to vicinal area (Kikuchi et al., 2016; Fujii et al., 2019) are being explored to effectively use the residue, bagasse. However, many crops need to be adapted to local environmental changes (Ma et al., 2017), and as new cultivars such as high-yielding cultivars, are introduced and gradually diffused, the balance between product and residue yields will change dynamically. Therefore, factories that are considering expanding their functions hereafter need to plan their facilities replacements carefully, taking into account the balance between maintaining production yield and efficient production of by-products. This study focuses on polygeneration in the sugarcane industry and aims to support the planning of equipment replacement to functionalize new high-yielding cultivars and excess thermal energy using computer-aided process engineering.

2. Material and methodology

Figure 1(a) shows a schematic of a conventional sugar mill that is not connected to the regional power grid. As an example, the flow of a sugar mill in Japan (Kikuchi et al., 2016) is shown. Sugarcane is supplied to the mill turbine first, then bagasse is transported to the bagasse boiler to be burned as fuel. Steam is generated by bagasse combustion, and 1.95 MPa of steam is split to the mill and power turbine, respectively, then depressurized steam is provided to evaporating and crystalizing process of sugar. Thus, typical sugar mills can provide its own steam and electricity demand utilizing bagasse. A process flow model was developed to simulate the mass and heat balance around a bagasse boiler in the sugar mill. This process flow model was incorporated into the already developed integrated modeling of agricultural and industrial processes for sugarcane-derived products (Ouchida et al., 2017). The expected increase of sucrose and fiber calculated by the expected composition of the cultivars for the next decade are shown in Figure 2. The cultivars considered were NiF8, NiTn18, Ni22, and a new cultivar (KY10-1380) with high-yielding and high percentage of fiber content (Hattori et al., 2019). The introduction



(a) Current process



(b) Selling surplus electricity

Figure 1 Schematic of sugar mill around bagasse boiler

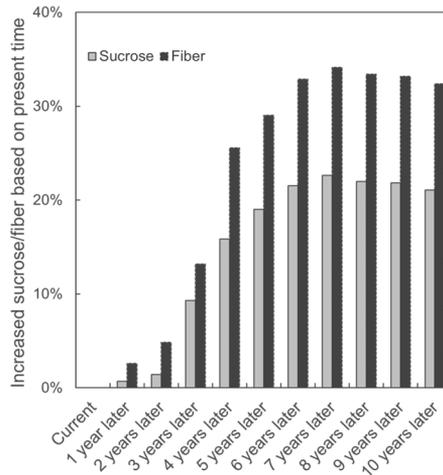


Figure 2 Increased sucrose and fiber in sugarcane based on present time

and expansion of new cultivar with high fiber content leads to further increase bagasse as biomass energy in the future. In order to make effective use of these surplus resources, this study examines the functionalization of new cultivars, using electricity sales as a case study.

The mass and heat balance of the plant for each year were calculated for the three cases. Case 1 is the case where the existing boiler is used without replacement. In the current energy flow of the sugar mill, the flow rates of bagasse and intake air are unknown. The UA value of the existing bagasse boiler was calculated by adjusting these unquantified parameters until the target values of 310 °C for the boiler outlet temperature and 10% for the oxygen concentration in the final exhaust flue gas were reached. The energy balance in the heat exchanger was divided into the phase change part of the steam and the other part, and the phase change part was divided into 10 intervals, and UA was calculated in each interval by the logarithmic mean temperature difference method, and the UA value of the entire heat exchanger was calculated by summing these values according to the temperature and flow rate of each fluid. The mass flow rate of bagasse supplied to the bagasse boiler in each year was adjusted to maintain this quantified UA value. In the cases where surplus electricity is sold to the local power grid (Cases 2 and 3), as shown in Figure 1(b), the bagasse boiler generates steam at 7 MPa, a high-pressure power turbine is installed, and exhaust steam from the high-pressure power turbine is introduced into the mill turbine and the low-pressure power turbine. In Case 2, the UA value of the bagasse boiler was designed to maximize the amount of electricity sold to the local power grid after 10 years from now by consuming excess bagasse as much as possible, except for exporting excess bagasse as bedding for livestock. The mass flow rate of the generated steam in each year was adjusted to maintain the UA value. In Case 3, the UA value of the bagasse boiler was designed to generate high-pressure steam required only for sugar production after 10 years from now. The generated steam is assumed to be introduced into a high-pressure power turbine to generate surplus electricity that is sold. The UA value was maintained by adjusting the mass flow rate of bagasse in each year.

3. Result and discussion

3.1. Amount of excess bagasse after 10 years with the current boiler (Case 1)

The UA value of the existing bagasse boiler was calculated by adjusting the volumetric flow of intake air and bagasse input to match the current operating data (temperature of the boiler outlet and oxygen concentration of the final exhaust flue gas). The UA value of the boiler was maintained by adjusting the amount of bagasse input to the bagasse boiler in each year, assuming that the existing boiler would not be replaced when a new high-yielding cultivar was introduced. Figure 3 shows the results of the bagasse production estimated by the integrated modeling of agricultural and industrial processes for sugarcane derived products, which incorporates the heat exchange process around the bagasse boiler in the sugar mill, and the change in bagasse input to the boiler adjusted by the energy flow model. About three years after the introduction of the new high-yielding cultivar, total sugar production increases with the spread of the new cultivar, and bagasse input to the bagasse boiler and required steam flow also increase. The new cultivar (KY10-1380) has a higher stem height and higher yield than other typical cultivar, but because of the high proportion of fibrous material within the cane, the amount of bagasse generated per unit of sugar produced increases as the new cultivar spreads, and the excess amount of bagasse increases. The current excess bagasse is exported as much as possible as bedding for livestock, but with the increase in bagasse production, it was expected that 7500 tons/year of excess bagasse would be generated which would be difficult to store as shown in Figure 3. Since this excess bagasse could be used to supply additional biomass energy to the surrounding area, the introduction of the electricity selling process is discussed in the next section.

3.2. Selling surplus electricity (Case 2 and 3)

The process selling surplus electricity by increasing the main steam pressure of the bagasse boiler and by utilizing excess bagasse that is expected to increase with the introduction of new high-yielding cultivar was simulated. The results of the ratio of electricity sold to the local power grid and self-consumption in the sugar mill are shown in Figure 4(a) when the UA of the bagasse boiler is designed to maximize the amount of electricity sold after 10 years. The UA value of the bagasse boiler in this case was 1.3 times higher than the current process. Although the internal power consumption of the sugar mill increases in accordance with the increase in total sugar production due to the spread of new high-yielding cultivar, the results show that 9.8 GWh of surplus power can be sold to the local power grid, mainly due to the increase in total power generation by the introduction of the high-pressure steam turbine. Figure 4(a) shows the temperature change of the final exhaust flue gas. Since the heat transfer area of the bagasse boiler designed to maximize the amount of electricity sold after 10 years is too large for the total heating value of bagasse that can be supplied to the bagasse boiler at present, the temperature of the final exhaust flue gas in this case at the present stage is below 140 °C, as shown in Figure 4(a). In order to avoid the temperature of the final flue gas falling below the acid dew point, sugar mills empirically maintain the temperature above 150 °C. Therefore, in this case, the current process should be maintained until enough excess bagasse is generated to introduce the process of selling surplus electricity, and the boiler should be replaced after 3 to 4 years when sufficient bagasse amount is obtained and the final exhaust gas temperature exceeds 150 °C.

Next, the case where the bagasse boiler generates only the steam required for sugar production in each year at high pressure was considered. The mass flow rate of bagasse was adjusted to generate the mass flow rate of steam required for sugar production after

10 years, and the UA value of bagasse was quantified. The quantified UA value was fixed, and the bagasse mass flow rate was adjusted annually to meet the sugar production demand. The results of the simulation of the ratio of electricity sold to the local power grid and self-consumption at the sugar mill and the temperature change of the final exhaust flue gas are shown in Figure 4(b). In this case, the UA value of the bagasse boiler was 1.2 times higher than that of the current process. The total amount of electricity sold after 10 years resulted in 8.0 GWh, which was about 20% lower than Case 2, but the temperature of the final exhaust flue gas was found to be able to be maintained at around 180 °C, the same level as the current process, because the mass flow rate of bagasse was optimized in each year to generate only the steam required for sugar production. Although not all of the bagasse generated in this case can be effectively used after 10 years, the excess amount can be reduced to less than half of that in Case 1, and furthermore, it is found that operation is possible even if the boiler is currently replaced.

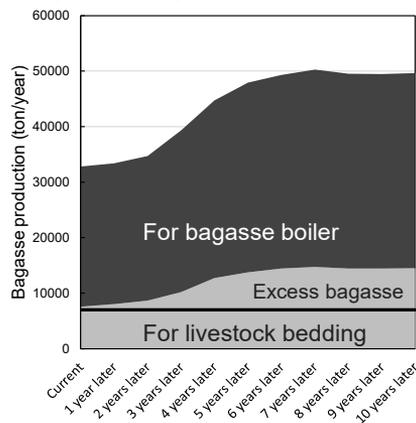


Figure 3 Change in bagasse amount in 10 years

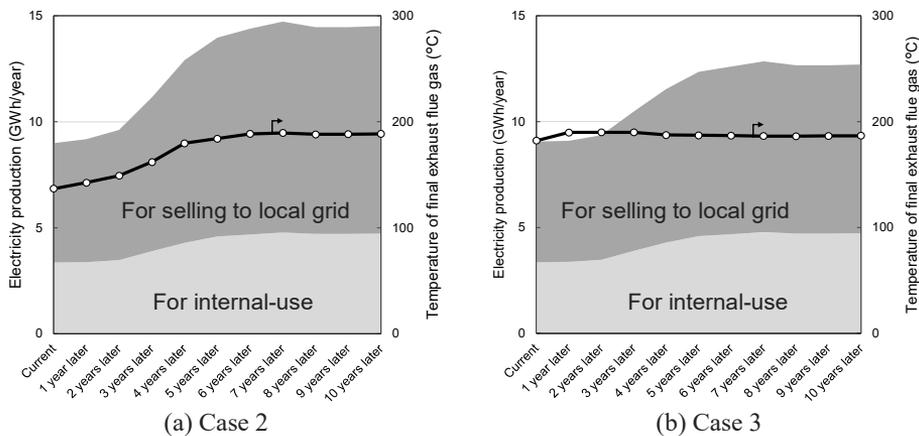


Figure 4 Change in balance between internal-use and selling electricity

4. Conclusion

In this study, a process flow model of a sugar mill was incorporated into an integrating model of agricultural-industrial process, and an electricity selling process using the excess thermal energy of sugarcane bagasse combustion associated with the introduction and diffusion of a new high-yielding cultivar was investigated. It was found that the bagasse, which is estimated to be unprocessable with the current facilities after 10 years from now, can be effectively utilized by introducing a high-pressure steam turbine and a process for selling electricity to the local community. In addition, by constructing a process flow model that takes into account the size of the heat transfer area of the boiler, it is possible to construct a strategy for the timing and scale of the replacement of the bagasse boiler based on the constraints of the flue gas temperature.

Acknowledgement

This work was supported by Shinko Sugar Mill Co., Ltd., JSPS Grant-in-Aid for Research Activity Start-up (Grant Number 20K23360), JST COI-NEXT (Grant Number JPMJPF2003), and the Environment Research and Technology Development Fund (Grant Number JPMEERF20213R01 and JPMEERF20192010) of the Environmental Restoration and Conservation Agency of Japan. The activities of the Presidential Endowed Chair for “Platinum Society” at the University of Tokyo are supported by the KAITEKI Institute Incorporated, Mitsui Fudosan Corporation, Shin-Etsu Chemical Co., ORIX Corporation, Sekisui House, Ltd., East Japan Railway Company, and Toyota Tsusho Corporation.

References

- E. Birru, C. Erlich, A. Martin, 2019, Energy performance comparisons and enhancements in the sugar cane industry, *Biomass Conversion and Biorefinery*, 9, 267–282
- S. Fujii, N. Horie, K. Nakaibayashi, Y. Kanematsu, Y. Kikuchi, T. Nakagaki, 2019, Design of zeolite boiler in thermochemical energy storage and transport system utilizing unused heat from sugar mill, *Applied Energy*, 238, 561-571
- T. Hattori, Y. Terajima, T. Sakaigaichi, T. Terauchi, Y. Tarumoto, K. Adachi, M. Hayano, M. Tanaka, S. Ishikawa, M. Umeda, M. Matsuoka, A. Sugimoto, 2019, High ratoon yield sugarcane cultivar “Harunoogi” developed for Kumage region by using an interspecific hybrid between a commercial cultivar and *Saccharum spontaneum* L. (in Japanese), *Journal of the NARO Research and Development*, 2, 21-44
- Y. Kikuchi, Y. Kanematsu, M. Ugo, Y. Hamada, T. Okubo, 2016, Industrial symbiosis centered on a regional cogeneration power plant utilizing available local resources: A case study of Tanegashima, *Journal of Industrial Ecology*, 20, 276-288
- L. Ma, L.R. Ahuja, A. Islam, T.J. Trout, S.A. Saseendran, R.W. Malone, 2017, Modeling yield and biomass responses of maize cultivars to climate change under full and deficit irrigation, *Agricultural Waste Management*, 180, 88-98
- S. Ohara, Y. Kikuchi, K. Ouchida, A. Sugimoto, T. Hattori, T. Yasuhara, Y. Fukushima, 2019, Reduction of greenhouse gas emissions in the introduction of inversion system to produce sugar and ethanol from sugarcane (in Japanese), *Journal of Life Cycle Assessment, Japan*, 15, 86-100
- K. Ouchida, Y. Fukushima, S. Ohara, A. Sugimoto, M. Hirao, Y. Kikuchi, 2017, Integrated design of agricultural and industrial processes: A case study of combined sugar and ethanol production, *AIChE Journal*, 63, 560-581
- JCE. Palacio, MLG. Renó, AM. Martínez Reyes, GF. de Souza, EAO. Batlle, OAA. del Olmo, EES. Lora, 2018, Exergy and Environmental Analysis of a Polygeneration System of Alcohol Industry. *Waste and Biomass Valorization*, 11, 1539-1554

Economic optimization of a reactive distillation column with multiple reactive sections for silane production

Alcántara-Maciel, Francisco D., ^a Victor E. Casillas-Céspedes, ^a J. Armando López-García, ^a Julián Cabrera-Ruiz, ^{a*} César Ramírez Marquez, ^a J. Rafael Alcántara-Ávila ^b

^a *Departamento de Ingeniería Química, Universidad de Guanajuato. Noria Alta S/N, Guanajuato, 36050, Mexico*

^b *Department of Chemical Engineering, Kyoto University, Kyoto, 615-8510*

Abstract

Solar cell manufacturing is based on solar grade silicon which can be obtained using Silane as precursor instead of Trichlorosilane in the well-known Siemen's process. The silane production can be done in a reactive distillation (RD) column in which reaction and separation take place simultaneously. The typical RD column has one reactive section, but in a recent work (Zang et al., 2017) it has been shown that an RD column with multiple reactive sections can enhance the internal mass integration and/or energy interaction between the reactions and the separation of Silane. Therefore, this work aims to optimize an RD column for silane production with up to three reaction sections. The number and location of the reactive sections are found after solving a proposed global stochastic optimization framework based on a modified Simulated Annealing Algorithm. The results show that the increase in the number of reactive sections does not necessarily mean and reduction in the cost of the process.

Keywords: Silane, Global stochastic optimization, Reactive Distillation, Process Intensification

1. Introduction

Silane is produced by three simultaneous disproportionation reactions of Trichlorosilane (SiHCl_3). These reactions have a rather unfavorable reaction kinetics and a thermodynamic conversion close to zero. Therefore, this process can be intensified by means of a reactive distillation (RD) column where the reactions and the separation of Silane are carried out simultaneously.

The most important process to obtain Solar-grade Silicon (SoG-Si) is the Siemens process, accounting for approximately 90% of worldwide polysilicon production (Bye and Ceccaroli, 2014) while the second most widely used commercial process, accounting for almost all non-Siemens technology-based production, utilizes the thermal decomposition of Silane (SiH_4) through a pyrolysis reaction in fluidized bed reactors (FBR) (Yadav et al., 2017).

The FBR method has the following advantages: 1) low production cost, 2) low energy consumption, and 3) high productivity. This method can reduce the energy consumption per unit mass of Silicon from the 40–60 kW h/kg (Mongstad et al., 2016) of the Siemens

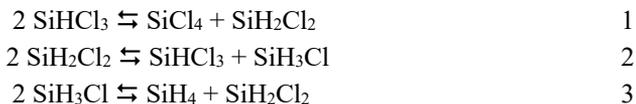
process to 5–8 kW h/kg (Tejero-Ezpeleta et al., 2004) for the silane pyrolysis step. Thus, the energy-saving and pollutants discharge reduction may generate a tremendous social and economic benefit (Zhang et al., 2018).

RD can overcome the chemical equilibrium limitation and separate Silane continuously. Thus, RD and intensified process for Silane production. The typical RD column feeds SiHCl_3 as raw material and generates Silane as the top product and Tetrachlorosilane (SiCl_4) as the bottom byproduct (Huang et al., 2013). However, Zang et al. (2017) studied an RD column with multiple reaction sections. Their results showed that three reaction sections attained higher condenser and reboiler duty savings. Thus, they claimed that the use of multiple reaction sections is a better intensification alternative for the RD column. This work proposes a rigorous global stochastic optimization framework that uses a modified Simulated Annealing Algorithm (m-SAA) in which integer optimization variables (e.g., number of stages, number and location of reactive stages) are treated more effectively to avoid unnecessary continuous relaxations of variables. Moreover, the proposed framework can find the number and locations of reactive sections without any prespecified structural constraint.

2. Methodology

2.1. Case study

An RD column for the production of Silane and Tetrachlorosilane from a feed containing Trichlorosilane was studied as shown in Figure 1a. Intermediate components are Monochlorosilane (SiH_3Cl) a Dichlorosilane (SiH_2Cl_2) respectively. The feed and product specifications are shown in Table 1. The kinetic parameters to model the disproportionation reactions in Eqs. 1 to 3 were taken from Huang et al. (2013). Also, the Peng Robinson model was used to estimate the vapor-liquid equilibrium relationships. The simulations were done in the Aspen Plus V11.0 simulation software, and the built-in block RADFRAC was used. Finally, the number of reactive sections ranges between one and three as previously considered by Zang et al. (2017) for the sake of comparison.



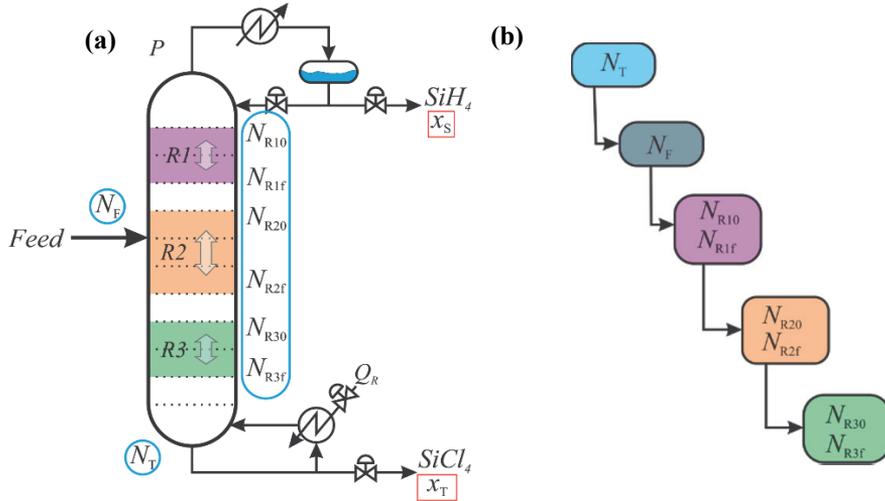


Figure 1. Reactive distillation column: (a) design variables, (b) sequential evaluations

N_T and N_F are the number of total stages and feed stage, respectively. N_{R10} , N_{R20} , and N_{R30} are the top stages for the first, second, and third reactive section, while N_{R1f} , N_{R2f} , and N_{R3f} are those for the bottom stages, respectively. x_S and x_T are the SiH_4 and SiCl_4 purities at the RD column top and bottom, respectively.

Table 1. Feed and product specifications for the RD column

Variable	Value
SiHCl_3 feed flow rate (kmol/h)	10
Pressure (atm)	5.0
Pressure drop (kPa)	0.5
SiH_4 purity (%mol)	99.3
SiCl_4 purity (% mol)	99.0

2.2. Mathematical treatment

The minimization of the total annual cost (TAC) was performed. The optimization variables are those related to N_T and the locations of N_F , N_{R10} , N_{R20} , N_{R30} , N_{R1f} , N_{R2f} , and N_{R3f} as shown in Eq.4. It is worth to mention that the operating variables such as reflux and reboiler duty were calculated by the build-in *design specification* and *vary* features of the module RADFRAC. Therefore, the optimization was solely done in terms of structural, integer variables.

$$\min(TAC) = f(N_T, N_F, N_{R10}, N_{R1f}, N_{R20}, N_{R2f}, N_{R30}, N_{R3f}) \quad 4$$

subject to Eq. 5

$$\vec{y}_i \geq \vec{x}_i \quad 5$$

where \vec{y}_i is the vector for the calculated purities for each component i and \vec{x}_i is the purity assigned target.

The optimization was carried out by a modified Simulated Annealing Algorithm (m-SAA) capable of handling integer variables in a more effective way than their typical continuous relaxation (Cabrera-Ruiz et al., 2021). The m-SAA was programmed in the Spyder platform (Spyder, 2021), which is a free and open-source scientific environment written in Python, mainly because it is free access and open-source software that has been gaining great interest by scientists and engineers in recent years. Also, it offers direct connectivity with the simulation software Aspen Plus. Thus, the optimization-simulation interface was done as shown in Figure 2.

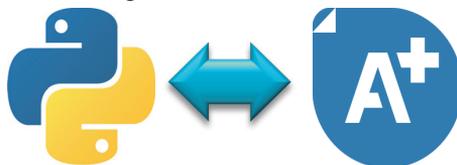


Figure 2. Optimization-simulation software interface

This work considers that the bounds of optimization variables change dynamically as the problem is being solved to handle the number of reactive sections and their locations. Thus, the variable bounds are expressed in terms of optimization variables as shown in Table 2. It was possible to assign these bounds on the random variables in the m-SAA by performing a normalization of variables between 0 and 1 as shown in Eq. 6. The generated values by the m-SAA take a normalized value (\bar{x}_i) and the actual value can then be calculated since the upper bounds (UL_i) and lower bounds (LL_i) are known despite they change dynamically. Furthermore, the values of those bounds depend sequentially on the result of the previous variable as shown in Fig. 1b. This strategy also offers the advantage of avoiding the generation of infeasible design points and the redundancy in terms of iterations for the calculation of new points due to design constraints (Cabrera-Ruiz et al., 2021).

Table 2. Optimization variables

Variable	Optimization bounds
N_T	[20,50]
N_F	[2, $N_T - 1$]
N_{R10}	[2, $N_T - 3$]
N_{R1f}	[N_{R10} , $N_T - 3$]
N_{R20}	[$N_{R1f} + 1$, $N_T - 2$]
N_{R2f}	[N_{R20} , $N_T - 2$]
N_{R30}	[$N_{R2f} + 1$, $N_T - 1$]
N_{R3f}	[N_{R30} , $N_T - 1$]

The solution procedure to find RD columns with multiple reaction sections is shown in Fig. 4a.

$$\bar{x}_i = \frac{UL_i - x_i}{UL_i - LL_i} \quad 6$$

The TAC was calculated using the Guthrie method (Turton et al., 2018) with a payback time of 5 years using the CEPCI value corresponding to the one for 2019. The cost of utilities was taken from Seider et al. (2009).

3. Results and Discussion

The solution of the m-SAA showed short computation time and few iterations compared to a traditional stochastic optimization in a software interface, proving that the normalization of variables offered an advantage not only in computational terms but also in the handling of dynamic bounds on optimization variables.

Figure 4b shows the optimization results of several solutions. Different designs are shown because their TAC values are very similar. Equation 7 shows the TAC average value and its standard deviation. It becomes clear that it is possible to obtain different designs with a similar cost.

$$\overline{TAC} = \$144,481.56 \pm 150.32 \quad 7$$

It also can be observed that in terms of TAC, there is not a significant improvement towards the selection of three reactive zones because the second solution from left to right in Fig. 4b has only one reactive zone, the one in the middle has two reactive zones, while the rest have three reactive sections. Nevertheless, the solutions with three reactive sections have one or two stages between them. Therefore, it is not completely clear if the increase of reactive sections attains better solutions. The existence of one equilibrium stage between reactive sections can be due in part to the rounding of the continuous variables that is used in all stochastic algorithms (Cabrera-Ruiz et al., 2021). In general, the variables that remain almost the same in all solutions are N_T and N_{R10} while N_F tends to be located near the top of the column.

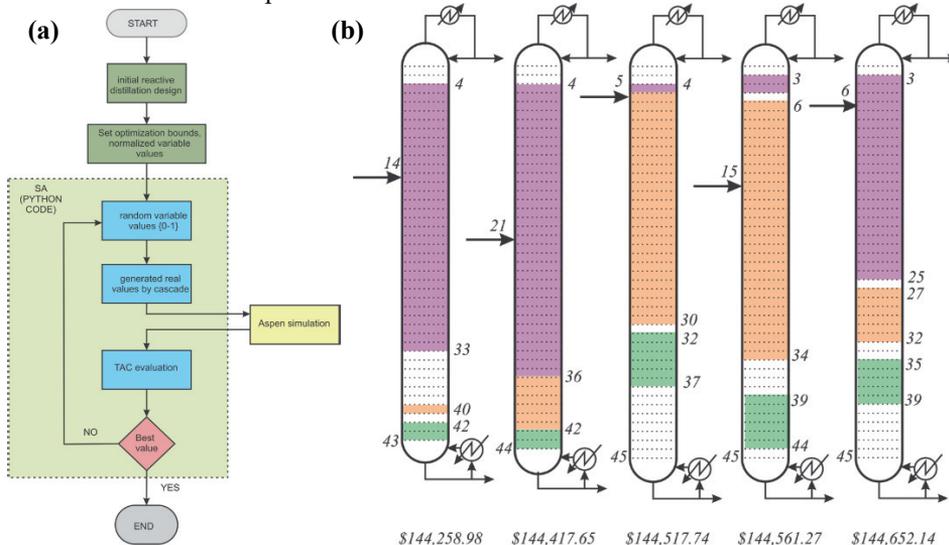


Figure 4. (a) Solution procedure for a general RD column, (b) optimization results

4. Conclusions

In this work, a global stochastic optimization framework was proposed for the economic minimization of an RD column with the possibility of having up to three reactive sections. The selection of lower and upper bounds for the reactive sections was done dynamically through a normalization of the optimization variables. The proposed m-SAA performed satisfactorily showing great flexibility and capability for handling integer variables

dynamically. The results showed that there are multiple solutions with similar cost despite the number of reactive sections. It was observed that one to two reactive zones are favored sometimes. Therefore, the selection of three reactive sections was not remarkably superior. Therefore, it is imperative to perform further optimizations considering alternative objective functions.

References

- G. Bye, B. Ceccaroli, 2014, Solar grade silicon: Technology status and industrial trends, *Sol Energy Mater Sol Cells*, 130, 634–646.
- J. Cabrera-Ruiz, S. Hasebe, J.R. Alcántara-Avila, 2021, Improvement of the Optimal Design Procedure Using Randomized Algorithm and Process Simulators, *MATEC Web Conf* 2021, 333, 06004.
- X. Huang, W.J. Ding, J.M. Yan, W.D. Xiao, 2013, Reactive distillation column for disproportionation of Trichlorosilane to Silane: Reducing refrigeration load with intermediate condensers, *Ind Eng Chem Res*, 52, 6211–6220.
- T. Mongstad, H. Klette, Ø. Nordseth, T.J. Preston, M. Guro, W.O. Filtvedt, 2016, New technologies for silicon production from monosilane; Pilot testing and on-line process monitoring aiding the development, *Silicon Chem Sol Ind*, 12, 1–10.
- W.D. Seider, J.D. Seader, D.R. Lewin, S. Widagdo, 2009, *Product and Process Design Principles: Synthesis, Analysis and Design*, 3rd Edition. John Wiley & Sons Ltd.
- Spyder, 2021, <https://www.spyder-ide.org/> (accessed on 29 November 2021).
- M.P. Tejero-Ezpeleta, S. Buchholz, L. Mleczko, 2004, Optimization of reaction conditions in a fluidized-bed for silane pyrolysis, *Can J Chem Eng*, 82, 520–529.
- R. Turton, J. Shaeiwitz, D. Bhattacharyya, W. Whiting, 2018, *Analysis, Synthesis, and Design of Chemical Processes*, 5th Edition. Boston, USA: Pearson.
- S. Yadav, K. Chattopadhyay, C.V. Singh, 2017, Solar grade silicon production: A review of kinetic, thermodynamic and fluid dynamics based continuum scale modeling, *Renew Sustain Energy Rev*, 78, 1288–1314.
- X. Zang, K. Huang, Y. Yuan, H. Chen, L. Zhang, S. Wang, K. Wang, 2017, Reactive Distillation Columns with Multiple Reactive Sections: A Case Study on the Disproportionation of Trichlorosilane to Silane, *Ind Eng Chem Res*, 56, 717–727.
- P. Zhang, J. Duan, G. Chen, J. Li, W. Wang, 2018, Production of polycrystalline Silicon from silane pyrolysis: A review of fines formation, *Sol Energy*, 75, 44–53.

A multiscale model of proliferating and quiescent cell populations coupled with cell cycle dynamics

Iqra Batool^{a,*} and Naim Bajcinca^a

^a*Mechanical and Process Engineering, Technical University of Kaiserslautern, 67663 Kaiserslautern, Germany*

**iqra.batool@mv.uni-kl.de*

Abstract

This paper proposes a nonlinear multiscale mathematical model for age-structured proliferating and quiescent cell populations (PDEs) coupled with cell cycle protein dynamics (ODEs). The model assumes a bidirectional transition between the proliferating and quiescent subpopulations. The coupling between the two scales is introduced based on biological findings inherited from the literature. Numerical simulations are performed using the finite volume method to examine the impact of parameters on the nonlinear dynamics of the model. The main focus of this paper is to investigate the balance between proliferating and quiescent cell populations, which play a crucial role in maintaining homeostasis in a cell population.

Keywords: Multiscale model, Tumor growth, Population dynamics

1. Introduction

The cells subject to division can be categorized into two compartments, i.e., proliferating and quiescent cells. Proliferating cells undergo division by passing through various stages of cell division cycle (G_1, S, G_2, M). However, quiescent cells neither grow nor divide but either transit to the proliferative compartment or stay in the G_0 phase until differentiation or death. The switching of the cells between proliferating and quiescent phases thus plays a crucial role in maintaining tissue homeostasis. In fact, any disturbance in the transitional balance between proliferating and quiescent phases results in an unconditional growth in tumoral tissue, van Velthoven and Rando (2019). The switching mechanism, however, depends on cell signaling molecules representing growth or anti-growth factors, Heldt et al. (2018). In this paper, we focus on modeling and analyzing the dynamics of cell population distributed in proliferating and quiescent phases coupled with cell cycle protein dynamics. The motivation stems from recent experimental outcomes, see Hartwell and Kastan (1994), which state that the cyclins are the most determinant control molecules in cell cycle phase transitions. Therefore, we couple the cell cycle protein dynamics to predict the transitional balance between proliferating and quiescent populations at macroscale, which is critical to homeostasis. At microscale, various proteins are expressed but for the sake of simplicity, we consider only four proteins (Cyclin D/CDK 4-6, E2F, p21 and Rb) which are mainly contributing in G_1 to S phase transition. We further assume at microscale that the molecular interactions are happening in a fast renewing cell population and not in an individual cell. Therefore, the molecular concentrations of these proteins act only as averaged concentrations in a subpopulation without accounting for cell-cell variability. The overall aim of this paper is to analyse the role of Cyclin D/CDK 4-6 complex in the deregulation of transitional balance between proliferative and quiescent cell compartments, which can eventually lead to an unlimited tumor growth.

2. Mathematical model

We consider two predominant scales comprising of cell populations and subcellular protein dynamics. The macroscale is characterised by age-structured partial differential equations to model cell population in quiescent and proliferating phases. Whereas, sub-cellular protein interactions involved in cell cycle dynamics are modeled using ordinary differential equations. Here age refers to biological clock of the cell in the cell cycle phases. The two scales are coupled via feedback in both directions. The schematics of the model is shown in Figure 1.

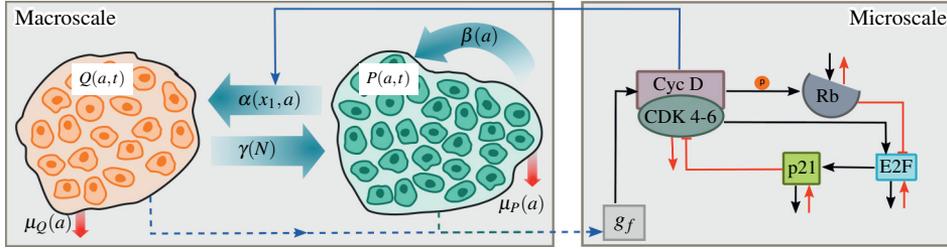


Figure 1: Schematics of the model.

2.1. Physiologically structured model

The macroscale level comprising of quiescent $Q(a,t)$ and proliferating $P(a,t)$ cell populations is described by transport type nonlinear hyperbolic partial differential equations (PDEs) which describe the number densities of the cells with respect to their physiological age a and time t :

$$\begin{cases} \frac{\partial}{\partial t} Q(a,t) = \alpha(x_1,a)P(a,t) - (\gamma(N) + \mu_Q(a))Q(a,t), \\ \frac{\partial}{\partial t} P(a,t) + \frac{\partial}{\partial a} (g(a)P(a,t)) = \gamma(N)Q(a,t) - (\beta(a) + \alpha(x_1,a) + \mu_P(a))P(a,t). \end{cases} \quad (1)$$

The first term on the right side of Eq. (1)a is the influx from the proliferative compartment by the rate function $\alpha(x_1,a)$ which is regulated by a microscale state, i.e., Cyclin D/CDK 4-6 complex x_1 . The second term describes the decrement in cell density due to either transitioning back to proliferating phase with the recruitment function $\gamma(N)$ or by cell death due to the apoptosis (or necrosis) with death rate $\mu_Q(a)$. In Eq. (1)b, $g(a)$ denotes the evolution speed of age a with respect to time t . The first term on the right side is the influx from the quiescent phase. The second term represents the number of cells completing the cell division with rate function $\beta(a)$, whereas the cells transitioning from the proliferating phase to quiescent phase without undergoing division are denoted by the third term. Finally, the death rate $\mu_P(a)$ represents the loss of cells in proliferating phase. The function $N(t)$, represents the total cell number in proliferating and quiescent phases, is defined as $N(t) = \int_0^{a^*} (P(a,t) + Q(a,t)) da$, where a^* represents the maximum age of cells.

To complete the model, we define the initial conditions at $t = 0$ as $Q(a,0) = Q_0(a)$, $P(a,0) = P_0(a)$, $\forall a \geq 0$. The boundary condition for $t > 0$ takes the form: $g(0)P(0,t) = 2 \int_0^{a^*} \beta(a)P(a,t) da$, where the number 2 represents two newborn cells which start with age 0 in the proliferating phase. The recruitment function $\gamma(N)$ is described using a Hill function $\gamma(N) = \frac{\nu \theta^\kappa}{\theta^\kappa + N^\kappa}$, where κ is the Hill coefficient, ν is the maximum transition rate and θ defines the total cell population achieving the half maximum of the transition rate ν . The rate function $\beta(a)$ is given as $\beta(a) = \rho_1 a^{\gamma_1} / (\rho_2^{\gamma_1} + a^{\gamma_1})$, where, ρ_1 represents the maximum proliferation rate, ρ_2 is the value of age achieving the half maximum effect and γ_1 is the Hill coefficient. Finally, the rate function α given as $\alpha(x_1,a) = \sigma_1 \cdot \sigma_2^{\gamma_2} / (\sigma_2^{\gamma_2} + x_1^{\gamma_2}) \cdot \sigma_3^{\gamma_3} / (\sigma_3^{\gamma_3} + a^{\gamma_3})$, determines the number of cells which cannot undergo division due to anti-growth factors. Here, γ_2 and γ_3 are Hill coefficients, σ_2 and σ_3 are the switching values of Cyclin D/CDK 4-6 complex and age a beyond which the rate function α becomes effectively

zero, preventing escape to quiescence. This implies that at age equals to σ_3 , the cell population is irreversibly committed to proceed into the proliferative phase. The value of σ_2 is the threshold of Cyclin D/CDK 4-6 complex determining the restriction point R. In the signaling mechanism among the cells, the growth response is modulated by the cytokine proteins, the details are given in Batool and Bajcinca (2021b,a). Thereby, the growth factors are defined as $g_f = 1/(1 + k_f N)$.

2.2. Cell cycle model

We employ Michaelis-Menten kinetics to explain the series of enzymatic reactions in the form of ODEs. The list of selected proteins and their description is given in the Table 1. Given the evolution speed of age a with respect to time t as $g := da/dt$, the dynamics of the cell cycle with respect to age a is then described by the following system, Gérard and Goldbeter (2012):

$$\begin{cases} \frac{dx_1}{da} = \frac{1}{g(a)} \left(k_{1s} \left(\frac{g_f}{k_{gf} + g_f} \right) - k_{14} x_4 x_1 - k_{1d} \left(\frac{x_1}{k_1 + x_1} \right) \right), \\ \frac{dx_2}{da} = \frac{1}{g(a)} \left(k_{21} \left(\frac{x_{2t} - x_2}{k_2 + (x_{2t} - x_2)} \right) x_1 - k_{32} x_2 x_3 - k_{2d} x_2 \right), \\ \frac{dx_3}{da} = \frac{1}{g(a)} \left(k_{3s} - k_{32} x_2 x_3 - k_{31} \left(\frac{x_3}{k_3 + x_3} \right) x_1 - k_{3d} x_3 \right), \\ \frac{dx_4}{da} = \frac{1}{g(a)} \left(k_{4s} + k_{42} \left(\frac{k_{34}}{k_{34} + x_3} \right) x_2 - k_{41} \left(\frac{x_4}{k_4 + x_4} \right) x_1 - k_{4d} x_4 \right). \end{cases} \quad (2)$$

Description	State
Cyclin D/CDK 4-6	x_1
E2F	x_2
Rb	x_3
p21	x_4

Table 1: Cell cycle states at microscale.

Param.	Description	Value	Unit
k_{1s}	Rate constant for synthesis of x_1 induced by growth factors g_f	0.175	h^{-1}
k_{gf}	Michaelis constant for synthesis of the x_1 induced by g_f	0.1	μM
k_{14}	Bimolecular rate constant for binding of cyclin D/CDK4-6 to p21	0.15	$\mu M^{-1} h^{-1}$
k_{1d}	Maximum degradation rate of cyclin D/CDK4-6	0.245	$\mu M h^{-1}$
k_1	Michaelis constant for the degradation of Cyclin D/CDK4-6	0.1	μM
k_{21}	Rate constant for activation of E2F by cyclin D/CDK4-6	0.805	h^{-1}
k_2	Michaelis constant for E2F activation by cyclin D/CDK4-6	0.01	μM
x_{2t}	Total concentration of the transcription factor E2F	2	μM
k_{32}	Bimolecular rate constant for binding of Rb to E2F	0.01	$\mu M^{-1} h^{-1}$
k_{2d}	Apparent first-order rate constant for non-specific E2F degradation	0.02	h^{-1}
k_{3s}	Basal rate of synthesis of Rb	0.8	h^{-1}
k_{31}	Rate constant for phosphorylation of Rb by cyclin D/CDK4-6	2.2	h^{-1}
k_3	Michaelis constant for Rb phosphorylation by cyclin D/CDK4-6	0.1	μM
k_{3d}	Apparent first-order rate constant for Rb degradation	0.01	h^{-1}
k_{4s}	Basal, E2F-independent rate of synthesis of p21	0.8	$\mu M h^{-1}$
k_{42}	Rate constant for synthesis of p21 induced by E2F	0.1	h^{-1}
k_{34}	Constant of inhibition by Rb of p21 synthesis	0.1	μM
k_{41}	Rate constant for inactivation of p21 through phosphorylation by x_1	50	h^{-1}
k_4	Michaelis constant for p21 phosphorylation by Cyclin D/CDK 4-6	0.5	μM
k_{4d}	Apparent first-order rate constant for non-specific p21 degradation	0.06	h^{-1}

Table 2: Parameters of cell cycle model.

In Eq. (2)a, the first term on right hand side describes the synthesis of Cyclin D/CDK 4-6 complex (x_1) induced by the growth factors g_f . The last two terms describe the binding of x_1 with tumor suppressor protein p21 and the degradation rate induced by other cell cycle proteins. In Eq. (2)b, first term on right side describes the synthesis of transcription factors E2F (x_2) induced by x_1 . The

second term denotes the decrement of x_2 due to inhibition by retinoblastoma protein Rb, while the last term depicts a constant inactivation rate of x_2 . In Eq. (2)c, the first term on right side represents the synthesis of free un-phosphorylated retinoblastoma protein Rb (x_3). The second term denotes the decline in Rb by making a complex with x_2 to inhibit it. The third term refers to deactivation of Rb by phosphorylation from x_1 and the last one to the degradation of Rb. In Eq. (2)d, the first and second term represent the synthesis of p21 (x_4) by ATM/ATR, TGF β pathways and by x_2 , respectively. The third and fourth term represent the decrement in p21 due to inhibition of x_1 and the degradation of p21. The description of the involved parameters is given in Table 2.

3. Numerical solution

In this section, we present the numerical method used to solve the system (1)-(2) in MATLAB. Finite volume method (FVM) is implemented using central upwind discretization scheme, Kuzmin (2010). Hereby, we introduce the following notations: Δa is mesh size and Δt is the time step, $\Delta a = a^*/N_a$, where N_a is a maximum number of age nodes given by $a_i = i\Delta a$, $0 \leq i \leq N_a$ and time is discretized into N_t steps with equidistant interval $\Delta t = t^{k+1} - t^k$. The computational mesh is shown in the Figure 2, where the domain is divided into many control volumes (for instance, the one highlighted in green) and we approximate the integral conservation law on each control volume. The red arrows are pointing the flux through the boundary of the control volume where the flux \mathcal{F} is computed at each grid point using central upwind scheme as schematically depicted with yellow boxes for $k = 2$ at $i = 2, 3$ with the green arrows. In the sequel, we describe the discretized model and fluxes. The discretized cell densities of quiescent cells associated with the i^{th} spatial interval at time k reads

$$P_i^k = \frac{1}{\Delta a} \int_{a_{i-\frac{1}{2}}}^{a_{i+\frac{1}{2}}} P(a, t^k) da, \quad Q_i^k = \frac{1}{\Delta a} \int_{a_{i-\frac{1}{2}}}^{a_{i+\frac{1}{2}}} Q(a, t^k) da.$$

The necessary Courant-Friedrichs-Lewy (CFL) condition for convergence of the solution requires

$$\left. \begin{aligned} & \Delta t [\gamma^k + \max(\mu_Q(a_i))] \\ & \Delta t \left[\frac{\max(g(a_i))}{\Delta a} + \max(\alpha(x_{1,i}^k, a_i) + \beta(a_i)) \right] \end{aligned} \right\} \leq 1.$$

The initial conditions for Q_i^0 and P_i^0 are defined below

$$Q_i^0 = \frac{1}{\Delta a} \int_{a_{i-\frac{1}{2}}}^{a_{i+\frac{1}{2}}} (a, t^0) da, \quad P_i^0 = \frac{1}{\Delta a} \int_{a_{i-\frac{1}{2}}}^{a_{i+\frac{1}{2}}} P(a, t^0) da.$$

Next, the discretized form of the PDEs (1) are given as

$$\begin{aligned} Q_i^{k+1} - Q_i^k &= \Delta t \alpha(x_{1,i}^k, a_i) P_i^k - \Delta t (\gamma^k + \mu_Q(a_i)) Q_i^k, \\ P_i^{k+1} - P_i^k - \frac{\Delta t}{\Delta a} (\mathcal{F}_{i+1/2}^k - \mathcal{F}_{i-1/2}^k) &= \Delta t \gamma^k Q_i^k \\ &\quad - \Delta t (\alpha(x_{1,i}^k, a_i) + \beta(a_i) + \mu_P(a_i)), \end{aligned}$$

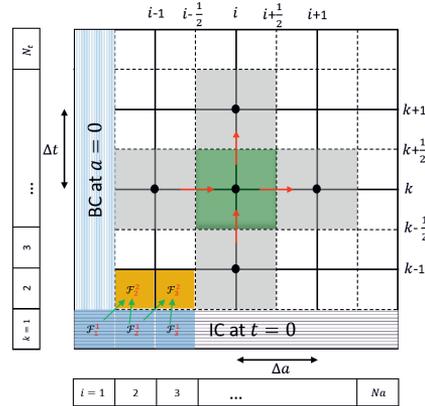


Figure 2: Computational mesh.

Algorithm 1: Numerical method

```

initialization;
forall time step  $k = 1, 2, \dots, N_t$  do
  Compute  $N^k, g_f^k$ ;
  Compute  $x_1^k, x_2^k, x_3^k, x_4^k$  using ode45;
  forall age step  $i = 1, 2, \dots, N_a$  do
    Compute  $\beta(a_i)$ ;
    Compute  $\alpha(a_i, x_{1,i}^k)$ ;
    Compute  $\Gamma(N^k)$ ;
    Compute  $Q_i^{k+1}$  and  $P_i^{k+1}$ ;
  end
  Compute CFL condition;
  if  $CFL < 1$  then
    continue;
  else
    break;
    disp('CFL is not satisfied')
  end
end

```

where $\gamma^k = \gamma(N^k)$ and $N^k = \Delta a \sum_i [Q_i^k + P_i^k]$. The fluxes represented by \mathcal{F} are defined using central upwind scheme as $\mathcal{F}_{i+1/2}^k = g(a_{i+1/2})P_{i-1}^k$. The cell division boundary condition at age $a = 0$ reads $g(a_0)P_0^{k+1} = 2\Delta a \sum_{i=1}^{a^*} \beta(a_i)P_i^k$. Then, we define the growth factors as $g_f = 1/(1 + k_f N^k)$. We have used ode45 function of MATLAB to solve cell cycle model at each time step k . The resulting outcome of Cyclin D/CDK 4-6 complex $x_{1,i}^k$ is then used in the macroscale model.

4. Results and discussion

In this section, we present the numerical results of the model proposed in Section 2. The initial states and the used parameters are given in Table 2 and 3. In all the simulations, we used the spatial step size $\Delta a = 0.5$ with a maximum age of cells $a^* = 50$ and the time step $\Delta t = 0.02$. In the sequel, we will discuss two case studies. First, we investigate the stability of the non-trivial steady-state solutions. Figure 3(a) and (b) represent the cell density distribution of proliferating $P(a,t)$ and quiescent $Q(a,t)$ cells, respectively. Both subpopulations show the trends of achieving a steady-state with time. The growth factors influence the behavior of Cyclin D/CDK 4-6 complex as depicted in Figure 3(c). The total cell count is initially low (see Figure 4(a)) while the growth factors are at their maximum (see Figure 4(b)), which results in the proper activation and degradation of cyclin D/CDK 4-6 complex along with the age of the cell. The latter depicts a complete cell cycle or successful division of cells on average. However, as the growth factors decline to a point where no (or fewer) new cells are required, the average behavior of Cyclin D/CDK 4-6 complex in proliferating cells also exhibit non-oscillatory dynamics, and it remains at lower concentration throughout, which is a depiction of no cell divisions, see Figure 4(c). In Figure 4(a), we plot the total cell population $N(t)$ comprised of proliferating, and quiescent phases exhibit an exponential increase in cell number and ultimately achieve a steady-state. On the other hand, Figure 4(b) shows the growth factors influenced by total cell population are maximum initially due to low cell count and gradually start declining until achieving an equilibrium. The transition rate $\gamma(N)$ from quiescent to proliferating phase is depicted in Figure 4(c). The increase in the total cell population invokes a decrement in the transition from quiescent to proliferating phase.

Param.	Value
ν	0.6 day^{-1}
θ	0.095×10^6
κ	1
ρ_1	0.7
ρ_2	0.35
γ_1	5
σ_1	0.01
σ_2	0.35
σ_3	14 h
γ_2	7
γ_3	7
k_f	1.80×10^{-9}

Table 3: Parameters used in the simulations.

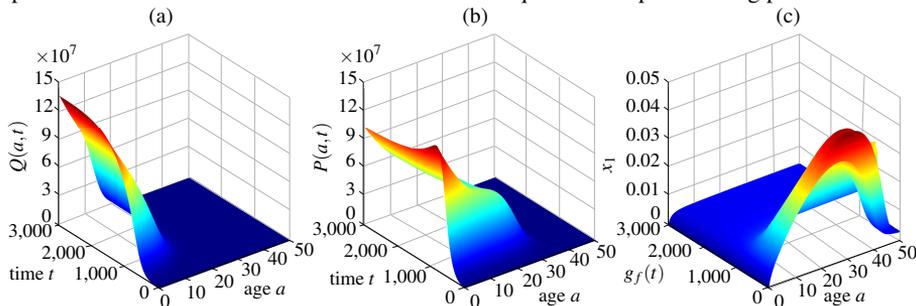


Figure 3: Density distribution of (a) proliferating cell population and (b) quiescent cell population with respect to age a and time t . (c) Concentration of Cyclin D/CDK 4-6 complex x_1 .

Next, we investigate the instability of the solution in Figure 5. The dynamics of the proposed model is very robust in general due to the feedback loops. However, the transition function $\alpha(a, x_1)$ is sensitive concerning noise in the cell cycle states. Here, to analyze a situation in which the cell cycle behaves abnormally, we changed a parameter value, i.e., $k_{gf} = 0.0001$, which depicts that the influence of the growth factors on the production of Cyclin D/CDK 4-6 complex is somehow compromised, or in other words, we are inducing delays in completing a Cyclin D/CDK 4-6

complex oscillation. Resultantly, the cell number grows exponentially due to more cells in proliferative compartment as compared to quiescent phase. All other parameters used in this case study are similar to the first case of non-trivial steady-states. The total cell number is plotted in Figure 5 (a). It grows exponentially in the presence of a larger amount of growth factors, see Figure 5 (b).

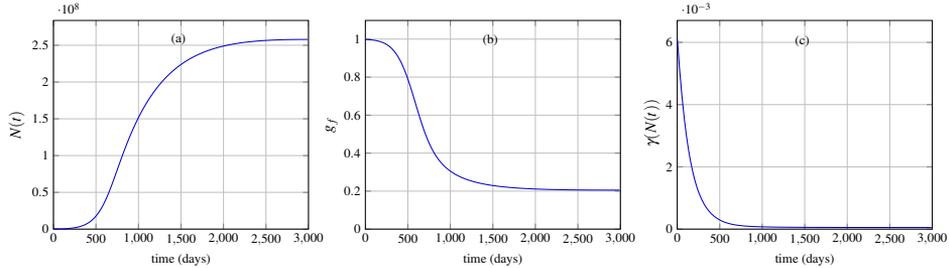


Figure 4: Steady state. (a) Total cell population $N(t)$. (b) Growth factors g_f . (c) Gamma function γ representing the rate at which the cells move back to proliferation phase from quiescent phase.

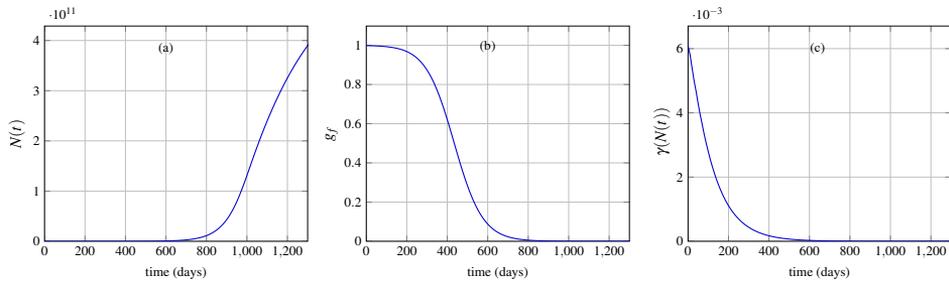


Figure 5: Exponential growth. (a) Total cell population. (b) Growth factors. (c) Gamma function.

5. Conclusion

This paper proposes a nonlinear, multiscale model of an age-structured population of proliferating and quiescent cells coupled with an averaged cell cycle dynamics which plays a crucial role in the cell's commitment for irreversible cell division. There exists a bi-directional transition between the two populations. The closed feedback loop couples the two scales and further aids in keeping the overall growth of the cells in homeostasis. We investigated the role of the complex formed by Cyclin D with its inhibitor CDK 4-6 in the bi-directional transition, and we concluded that an imbalance in this transition could result in tumor initiation and progression.

References

1. Batool, N. Bajcinca, 2021a. Evolution of cancer stem cell lineage involving feedback regulation. *PLoS ONE* 16 (5), e0251481.
2. Batool, N. Bajcinca, 2021b. Well-posedness of a coupled pde-ode model of stem cell lineage involving homeostatic regulation. *Results in Applied Mathematics* 9, 100135.
3. C. Gérard, A. Goldbeter, 2012. The cell cycle is a limit cycle. *Mathematical Modelling of Natural Phenomena* 7 (6), 126–166.
4. L. H. Hartwell, M. B. Kastan, 1994. Cell cycle control and cancer. *Science* 266 (5192), 1821–1828.
5. F. S. Heldt, A. R. Barr, S. Cooper, C. Bakal, B. Novák, 2018. A comprehensive model for the proliferation–quiescence decision in response to endogenous dna damage in human cells. *Proceedings of the National Academy of Sciences* 115 (10), 2532–2537.
6. D. Kuzmin, 2010. A guide to numerical methods for transport equations. Friedrich–Alexander university, Germany.
7. C. T. van Velthoven, T. A. Rando, 2019. Stem cell quiescence: dynamism, restraint, and cellular idling. *Cell stem cell* 24 (2), 213–225.

Multi-objective optimization of the food chain as a support for the development of agricultural sector

Jan Drofenik^a, Klavdija Zirngast^a, Bojan Pahor^{b,a}, Zdravko Kravanja^a, Zorka Novak Pintarič^a

^a*University of Maribor, Faculty of Chemistry and Chemical Engineering, Smetanova 17, 2000 Maribor, Slovenia*

^b*Saubermacher Slovenija d.o.o., Sp. Porčič 4a, 2230 Lenart, Slovenia*
zorka.novak@um.si

Abstract

An optimization model for the food supply chain is presented. This model is formulated as a Mixed Integer Linear Programming (MILP) problem and optimizes the land distribution, the amount and types of food produced, imported and exported, the number of livestock, the greenhouse gas emissions, the nitrogen uptake, the amount of food waste, etc. The production technologies considered are conventional, organic, conservation agriculture and food production in greenhouses. The requirements of the European Green Agreement and the Farm-to-Fork Strategy are considered. The model was applied to optimize different scenarios in the food chain in Slovenia, such as population growth and changing dietary habits, and to optimize different objectives, such as food self-sufficiency, greenhouse gas emissions, nitrogen balance, economic efficiency and a multicriteria objective. The results show that food supply could be improved through optimization even with the expected population growth. Greenhouse gas emissions and synthetic fertilizer use could be reduced. The results of the model can provide a guide for the development of agricultural policies for sustainable supply of locally produced food.

Keywords: food chain, locally produced food, optimization, MILP, multi-objective.

1. Introduction

The food supply chain is a complex system that includes the production of agricultural crops and animals, their processing into food, the distribution and consumption of food, and the collection and processing of waste. The agricultural sector is one of the main sources of environmental impacts and is accompanied by a number of uncertainties (Zirngast et al., 2019). In rapidly developing countries, the negative impact of agriculture on the environment increases as the food security of the population improves (Qi et al., 2018). The objectives of reducing greenhouse gas emissions, the use of artificial fertilizers, nitrogen inputs, the preservation of biodiversity, etc. must be taken into account (European Commission, 2019). Therefore, systematic methods need to be applied to optimize the efficient use of available agricultural land, as demonstrated in the case of Estonia (Põldaru et al., 2018) and in the case of the Milan metropolitan region, where changes in land use and dietary habits were assessed through linear programming (Sali et al., 2016). Optimization has also been applied to accurately determine nutritionally adequate meals, considering environmental impacts and economic aspects (Benvenuti et al., 2021).

This paper presents a MILP model for optimizing food production of plant and animal origin on available arable land in Slovenia. The novelty of the model is that it considers

a balance between production, consumption, import and export of food, considers different technologies of food production and incorporates the requirements of European and national directives. The optimization is performed for several single objectives and as a multi-objective problem by combining the most important indicators.

2. Mathematical model for food supply optimization

The cultivation land for distribution is constrained by Eq. (1):

$$\sum_i \sum_k A_{i,k} \leq A_{TOT} \quad (1)$$

where i is a set of food and feed of plant origin, k set of cultivation technologies (e.g. conventional, organic, conservation, greenhouses), A area (ha), A_{TOT} total available arable land (ha). Production of food and feed of plant origin is defined as:

$$mpp_{i,k} = A_{i,k} \cdot yha_{i,k} \quad (2)$$

where mpp is a mass of plant food produced (t/y), yha hectare yield (t/(ha·y)). Balance for food and feed of plant origin is as follows:

$$\sum_k mpp_{i,k} + mpi_i = mpc_i + mpe_i + \sum_k mpl_{i,k} \quad (3)$$

where mpi and mpe are masses of imported and exported plant food (t/y), mpc consumed plant food (t/y), mpl lost plant food (t/y). According to the statistical definition, the domestic production of food of animal origin is defined as:

$$map_j = NZS_j \cdot mcrcS_j + NZE_j \cdot mcrcE_j - NZI_j \cdot mcrcI_j \quad (4)$$

where j is a set of food of animal origin (e.g. beef, pork, poultry, milk, eggs), map a mass of domestically produced food of animal origin (t/y), NZS , NZE and NZI are numbers of livestock slaughtered domestically, exported or imported (y^{-1}), $mcrcS$, $mcrcE$, $mcrcI$ are carcass masses per slaughtered, exported and imported animal (t). For milk and eggs, NZS represents the number of dairy cows or laying hens, and $mcrcS$ represents yearly milk or egg production per animal. For milk and eggs, $mcrcE$ and $mcrcI$ are equal to 0.

Balance of food of animal origin is given by Eq. (5):

$$mac_j + mae_j + mal_j = \sum_j NZS_j \cdot mcrcS_j + mai_j \quad (5)$$

where mac is a total mass of consumed food of animal origin (t/y), mai , mae and mal imported, exported and lost food products of animal origin (t/y), respectively.

Consumption of plant fodder for animals:

$$mpc_{ia} = \sum_j map_j \cdot fconv_{ia,j} \quad (6)$$

where ia is a subset of plants used for animal fodder, $fconv$ is a specific mass of consumed plant fodder in tons of plant fed per ton of carcass. Greenhouse gas emission is estimated as follows:

$$mghg = \sum_i \sum_k mpp_{i,k} \cdot fghg_{i,k} + \sum_j map_j \cdot fghg_j \quad (7)$$

where $mghg$ mass of greenhouse gas emission (t CO_{2eq}/y), $fghg$ specific emission per mass of food (t CO_{2eq} per ton of food). The amount of mineral fertilizers used is calculated:

$$mmf = \sum_i \sum_k A_{i,k} \cdot mfer_{i,k} \quad (8)$$

where mmf is total mass of mineral fertilizers used (t/y) and $mfer$ mass of mineral fertilizers used (t/(ha·y)). Nitrogen balance is presented in Eq. 9:

$$mnit = \sum_j map_j \cdot fnit_j + mmf \cdot wnit - \sum_i \sum_k A_{i,k} \cdot mnup_{i,k} \quad (9)$$

where $mnit$ is a mass of nitrogen (t N/y), $fnit$ specific nitrogen emission from livestock manure (t of N per ton of carcass), $wnit$ mass fraction of nitrogen in synthetic fertilizers, $mnup$ nitrogen uptake by plants (t N/(ha·y)). Economic added value (€/y) is defined as:

$$\begin{aligned} value = & \sum_i mpc_i \cdot ppd_i - \sum_i \sum_k mpp_{i,k} \cdot cpd_{i,k} + \sum_j mac_j \cdot pad_j - \sum_j map_j \cdot cad_j + \\ & \sum_i mpe_i \cdot ppe_i + \sum_j (NZE_j \cdot mcrcE_j + mae_j) \cdot pae_j - \sum_i mpi_i \cdot ppi_i - \\ & \sum_j (NZI_j \cdot mcrcI_j + mai_j) \cdot pai_j \end{aligned} \quad (10)$$

where ppd and pad are prices of food of plant and animal origin at domestic market (€/t), ppe and pae prices of exported food (€/t), ppi and pai prices of imported food (€/t), cpd production cost of plant food/feed (€/t), cad production cost of food of animal origin (€/t). Self-sufficiency for food for human is calculated by Eq. (11):

$$SS_{hum} = \left(\sum_{ih} \sum_k mpp_{ih,k} + \sum_j map_j \right) / \left(\sum_{ih} mpc_{ih} + \sum_j mac_j \right) \quad (11)$$

where ih is a subset of plants used for human food. A set of logical relations is used for the exclusive choice between food import and export:

$$\begin{aligned} ypi_i + ype_i &\leq 1 & mpi_i &\leq MB \cdot ypi_i & mae_j &\leq MB \cdot yae_j \\ yai_j + yae_j &\leq 1 & mpe_i &\leq MB \cdot ype_i & NZI_j &\leq MB \cdot yNZI_j \\ yNZI_j + yNZE_j &\leq 1 & mai_j &\leq MB \cdot yai_j & NZE_j &\leq MB \cdot yNZE_j \end{aligned} \quad (12)$$

where ypi and ype are binary variables for imported and exported plant food/feed, yai and yae binary variables for imported and exported food of animal origin, $yNZI$ and $yNZE$ for imported and exported livestock. MB is a large positive constant.

The basic objective function is formulated as the total mass of food and fodder produced:

$$Z = \max \left(\sum_i \sum_k mpp_{i,k} + \sum_j map_j \right) \quad (13)$$

where Z (t/y) represents a domestic production of food of plant and animal origin. In addition, other objective functions can be optimized, namely maximizing economic value, minimizing greenhouse gas emissions, and minimizing excess nitrogen from mineral and organic fertilizers. A normalized multi-objective function was also formulated to account for multiple conflicting criteria.

$$Z_{multi} = \max \left(\frac{value}{value^0} + \frac{SS_{hum}}{SS_{hum}^0} - \frac{mghg}{mghg^0} - \frac{mmf}{mmf^0} \right) \quad (14)$$

where subscript 0 represents the current value of specific variable. The list of variables in the model is as follows:

$$A_{i,k}, mpp_{i,k}, mpc_i, mpi_i, mpe_i, mpl_{i,k}, NZD_j, NZI_j, NZE_j, map_j, mac_j, mai_j, mae_j, mal_j, mghg, mnit, SS_{hum}, value \geq 0; \quad ypi_i, ype_i, yai_j, yae_j, yNZI_j, yNZE_j \in \{0,1\}$$

3. Slovenian case study

The developed model was used to optimize the distribution of arable land in the Republic of Slovenia, which has about 2 million inhabitants and 460,000 ha at its disposal.

3.1. Scenarios of various changes in food supply chain

The first part of the experiments dealt with scenarios that considered projected population growth, legal directives, and health recommendations.

Scenario 1. Population growth. It was estimated (Drofenik et al., 2021) that the population could increase by 9 % by 2050, which will affect food demand, so all further scenarios were also optimized for this forecast.

Scenario 2. Reduction of fertilizer use. The European Farm-to-Fork Strategy foresees a reduction in fertilizer use of at least 20 % by 2030. Therefore, the upper limit for fertilizer use has been set at 80 % of current levels.

Scenario 3. Proportion of land devoted to organic farming. The 2030 Agenda for Sustainable Development envisages an increase in the share of agricultural land used for organic farming to 25 %. In Slovenia, this share is currently 7.7 %. Therefore, the requirement to increase this share was included in the model.

Scenario 4: Healthier diet. Per capita consumption of red meat in Slovenia is significantly higher than the European average. The consumption of red meat should be reduced. In this scenario it was assumed that the per capita consumption of beef and pork would decrease by 30 %.

3.2. Optimization of the various objectives

In the second part, optimizations of various single objective functions and a composite criterion for multi-objective optimization were performed.

Scenario 5. Greenhouse Gas Emissions. Agriculture is an important source of GHG emissions. Therefore, the criterion in Eq. (7) was minimized to estimate the target potential for reducing emissions from the agricultural sector.

Scenario 6. Surplus in the nitrogen balance. Nitrogen input from animal manure and synthetic fertilizers often exceeds the nitrogen uptake that plants take from the soil for growth, which is an environmental problem. Therefore, the criterion $mnit$ defined in Eq. (9) was minimized to balance out the surplus of nitrogen.

Scenario 7. Added economic value. Although the main objective is to ensure an adequate supply of quality food, the food supply chain is also an economically important system. The economic value defined in Eq. (10) has therefore been maximized.

Scenario 8. Self-sufficiency. Maximizing food self-sufficiency as defined in Eq. (11) is the most important criterion for formulating national food policy.

Scenario 9. Multi-objective optimization. Many important factors for the agricultural sector have conflicting effects on food supply chain optimization. Therefore, a multi-objective optimization of the composite objective function of the four most important criteria defined in Eq. (14) was performed.

4. Results

The results of the first set of optimizations are shown in Table 1. All values of Z_{multi} are positive, indicating an improvement over the current situation. In scenario 1, self-sufficiency increases despite a projected increase in population. This is the effect of an optimized distribution of land and food produced compared to the current situation, which is not optimized. The area devoted to human food production increases significantly. Vegetable production in greenhouses increases and is even exported.

Table 1: Results of scenarios

	Unit	Current	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Added value	M€/y	443.7	499.8	496.8	451.8	482.6
GHG emission	kt/y	1,464	1,493	1,456	1,180	1,342
Nitrogen balance	t/y	1,565	11,866	11,855	6,159	2,510
Synthetic fertilizers	t/y	87,613	107,349	70,091	67,359	98,570
Total self-sufficiency	%	86.1	101.9	99.8	96.7	103.1
Self-sufficiency human	%	80.9	102.9	96.6	94.8	111.4
Area for human food	ha	60,129	101,116	93,726	90,584	105,516
Area for animal feed	ha	402,877	361,890	369,280	372,422	357,490
Area (conventional)	ha	427,155	456,592	263,681	340,841	456,592
Area (organic)	ha	35,651	0	0	115,752	0
Area (conservation)	ha	0	0	192,911	0	0
Area (greenhouses)	ha	200	6,414	6,414	6,414	6,414
Z _{multi}	%	0.0	15.3	51.8	61.4	42.2

A requirement for a 20 % reduction in fertilizer use (scenario 2) leads to the introduction of conservation agriculture with lower fertilizer use but almost the same yields per hectare as conventional production. The requirement of 25 % of the land dedicated to organic production (scenario 3) leads to a significant reduction in the economic criterion. The level of self-sufficiency decreases compared to the previous scenarios. Imports of plant and animal foods increase. A 30 % reduction in red meat consumption (scenario 4) would have several positive impacts, increasing self-sufficiency while decreasing emissions and excess nitrogen inputs. The production of plant foods for human consumption would increase significantly. Necessary imports of animals and meat products would decrease, and value added would increase.

The results of the optimizations for different objectives are shown in Table 2. Minimizing GHG emissions (scenario 5) provides a solution in which GHG emissions are 33 % lower than at present, but the economy and self-sufficiency deteriorate. The lower emissions are achieved by switching a significant part of production from conventional to organic and conservation agriculture with lower fertilizer use.

Table 2: Results of various objective functions

	Unit	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9
Added value	M€/y	431.2	448.1	511.4	472.5	451.4
GHG emission	kt/y	1,085	1,187	1,495	1,221	1,143
Nitrogen balance	t/y	5,163	0	17,832	6,894	6,587
Synthetic fertilizers	t/y	44,626	92,826	110,155	109,166	46,309
Total self-sufficiency	%	82.6	93.6	90.3	96.8	92.9
Self-sufficiency human	%	81.3	85.8	100.1	120.1	108.7
Area for human food	ha	125,727	88,277	90,584	144,954	131,840
Area for animal feed	ha	337,279	374,729	372,422	318,052	331,166
Area (conventional)	ha	179,285	459,265	456,592	456,003	179,285
Area (organic)	ha	129,942	0	0	0	76,593
Area (conservation)	ha	153,779	0	0	590	200,714
Area (greenhouses)	ha	0	3,741	6,414	6,414	6,414
Z _{multi}	%	72.7	18.9	11.1	46.9	105.2

Minimizing nitrogen surplus (scenario 6) offers a solution with an ideal balance between nitrogen supply and uptake, but also lower economic value added. A solution with maximum economic value (scenario 7) has high self-sufficiency, but also higher GHG emissions and excess nitrogen. Many animals and vegetables are produced, which are also exported. Maximizing self-sufficiency (scenario 8) achieves 120 % food self-sufficiency for people, and excess food could be exported. Optimization with a composite objective function (scenario 9) yields a compromise solution in which 21 % of the land is devoted to organic agriculture. GHG emissions would be 25 % lower and fertilizer use 53 % lower than at present, while food self-sufficiency would be over 100 %.

5. Conclusions

A MILP food supply chain optimization model was developed. The model determines the optimal distribution of arable land, crops, livestock and meat production. Besides, it considers the requirements of various European and national directives regarding the environmental and social impacts of agriculture, while the main motivation is to increase the food security through stable production of safe, quality and accessible food, and to shorten the supply chain taking into consideration CO₂ emissions that could be avoided. The results point out the need to promote a shift from intensive meat production, especially cattle, to grass-based production, and to change people's diets from meat and meat products to a greater proportion of vegetables. Promoting organic farming and other unconventional forms of agriculture is another driver for reducing the environmental impact of agricultural production. The consideration of multiple criteria in the optimization suggests a favorable trade-off solution, confirming the need for multi-objective decision-making in guiding national agricultural policy. The assumptions set in the strategies can have a strong impact on the results in the different scenarios. The sensitivity of these assumptions will be explored in future work, which will include options for food waste reduction and processing, and nutritional information about foods.

Acknowledgment

The authors thank the Ministry of Agriculture, Forestry and Food and the Slovenian Research Agency (Project V2-2009 and Program P2-0414) for support.

References

- L. Benvenuti, A. De Santis, P. Cacchione, 2021, Multi-indicator design and assessment of sustainable diet plans, *Journal of Cleaner Production*, 313, 127699.
- J. Drofenik, B. Pahor, Z. Kravanja, Z. Novak Pintarič, 2021, Scenario Analysis of Changes in the Food Supply Chain, *Chemical Engineering Transactions*, 88.
- European Commission, Farm to Fork Strategy, 2020, ec.europa.eu/food/system/files/2020-05/f2f_action-plan_2020_strategy-info_en.pdf, accessed 18.10. 2021.
- R. Põldaru, A.-H. Viira, J. Roots, 2018, Optimization of arable land use to guarantee food security in Estonia, *Agronomy Research*, 16 (4), 1837-1853.
- G. Sali, F. Monaco, C. Mazzocchi, S. Corsi, 2016, Exploring land use scenarios in metropolitan areas: food balance in a local agricultural system by using a multi-objective optimization model, *Agriculture and Agricultural Science Procedia* 8, 211-221.
- X. Qi, Y. Fu, R. Y. Wang, C. N. Ng, H. Dang, Y. He, 2018, Improving the sustainability of agricultural land use: An integrated framework for the conflict between food security and environmental deterioration, *Applied Geography*, 90, 214-223.
- K. Zirngast, L. Čuček, Ž. Zore, Z. Kravanja, Z. Novak Pintarič, 2019, Synthesis of flexible supply networks under uncertainty applied to biogas production, *Computers & Chemical Engineering*, 129, 106503.

Optimal CO₂ allocation for enhanced oil recovery operations within carbon utilisation networks in Qatar

Razan Sawaly, Ikhlas Ghiat, Abdulkarim Mohamed, Ahmad Abushaikha*, Tareq Al-Ansari*

College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

**Corresponding email: talansari@hbku.edu.qa; aabushaikha@hbku.edu.qa*

Abstract

Carbon capture and storage (CCS) is one of the most efficient techniques for reducing carbon dioxide (CO₂) emissions into the atmosphere. Combining CCS with enhanced oil recovery (EOR) processes is a very attractive method for carbon capture and utilisation (CCU). These operations enable CO₂ emissions to be reduced through geological sequestration, whilst generating additional revenue from enhanced oil production due to CO₂ re-injection via EOR. In practice, mass balance and temporal features of a given location are considered when planning EOR operations. When numerous oil reservoirs are involved, it is vital to allocate available CO₂ supplies and schedule EOR operations for these reservoirs at suitable timings. As a result, CO₂ allocation and scheduling are crucial for maximising the economic benefits of EOR operations. As such, this study introduces a resource trade scheme for CO₂ integration and utilisation within the state Qatar, where a mixed integer linear programming (MILP) model is developed to address CO₂ allocation and scheduling based on environmental and economic objectives. The model considers a single CO₂ source (Qatar Gas) within an multi sink scenario which includes several sinks within an industrial setting (QAFCO, QAFAC, PEARL GTL, ORYX GTL, Dukhan Field Well (EOR)). Two scenarios are considered to allocate CO₂ to different sinks (including EOR) to obtain the optimal solution for each scenario. The outcome of scenario 1 demonstrates that the optimal solution is to utilize 13.5Mt/y of carbon dioxide, which results in an annual profit varying from 14.3 to 42.8 billion US dollars. The maximum CO₂ utilisation occurs at Dukhan Field Well (EOR), which utilises up to 67%. Scenario 2 is implemented based on scenario 1 to further improve the model; where the profit increased annually, and the model became more sustainable.

Keywords: Sustainability, CO₂ utilisation, carbon capture and storage, EOR

1. Introduction

The allocation of carbon dioxide (CO₂) emissions is critical in defining reduction responsibilities at the national level or emission permits at the company level. Different techniques for calculating CO₂ emissions have been developed and applied throughout the last few decades. Various approaches could be used in order to reduce carbon dioxide emissions, among the most notable examples are increased energy efficiency by system integration, the adoption of energy saving technologies, carbon storage and sequestration, as well as fuel switching and renewable energy sources. Many studies have focused on the use of renewable and clean energy sources, carbon capture and storage (CCS), and

energy efficiency, yet minimal research has been published on methods for systematically investigating carbon utilisation approaches. Carbon dioxide can be converted chemically or biologically into fuel, food supplements, polymers, or other value-added products. (Mikkelsen *et al.*, 2010). Furthermore, in the case of enhanced oil recovery (EOR), storing CO₂ underground may provide financial incentives to reduce emissions (Hasan *et al.*, 2014). Injecting CO₂ into reservoirs for EOR is the only commercially recognized carbon usage method that allows permanent large-scale storage for captured CO₂ (Núñez-López *et al.*; Moskal *et al.* 2019). This study aims to design a framework allocation model between a single CO₂ source and several sinks based in the State of Qatar and evaluating a network design for carbon capture, storage, utilisation and transportation. A multi-period mathematical optimization is used that considers varying marketing prices of value-added products to determine the optimal solution in the model's revenue and utilisation target.

2. Literature Review

In the context of increased oil recovery and storage sinks, steps and mechanisms related to the allocation of carbon have been previously investigated. Turk *et al.* (1987) was able to provide a framework of sources particularly improved oil recovery sites, power plants and sinks, while taking into account the injection and transportation costs by the use of a nonlinear mathematical model. In addition, Middleton and Bielicki (2009) studied the supply and allocation of carbon dioxide. The study used the MILP spatial model, which addresses carbon capture and storage facilities. Weihs and Wiley (2012) also investigated the cost-optimal for CO₂. The transport of carbon dioxide in various phases was studied by Knoope *et al.* (2013), while He *et al.* and Tan *et al.* (2013) investigated the storage of multiphase carbon. Tan *et al.* (2013) used the source and sink approach for carbon capture and storage, capacity and injection. He *et al.* (2013) used the addressed the uncertainties of source-to-sink matching, where elements of physical and time limitations were integrated in a MILP to address the allocation of CO₂. The formulation focused mainly on the geological storage of sinks and attempted to optimise CO₂ storage by one source connected to a sink, relying on the flow and storage capacity. Hasan *et al.* (2014) analysed the matching of enhanced oil recovery sinks at a national level considering a wide network of CO₂ supply chains that selected the appropriate capture technology for various CO₂ sources and used geographical sites and transportation of pipelines. Alhajaj *et al.* (2013) applied the same approach at a regional level in the UAE.

3. Model Design and Optimisation

3.1. Available data and assumptions

The framework structure in this study consists of a single CO₂ source, Qatar gas (QG), since it is the world's largest LNG producer, with an annual CO₂ emissions of 35.8 Mt/year approximately (Mohammed, 2016), and six application sinks: (a) Qatar Fertiliser Company (QAFCO) and Qatar Fuel Additives Company (QAFAC) located south of Qatar in Mesaieed industrial zone (approximately 100 km from QG), (b) Dukhan Field Well (EOR) located west of Qatar in Dukhan City (approximately 100 km from QG); and

(c) Pearl GTL and Oryx GTL plants located north of Qatar in the Ras Laffan industrial zone (approximately 5 km and 3 km from QG) (Al-Yaeshi et al., 2020).

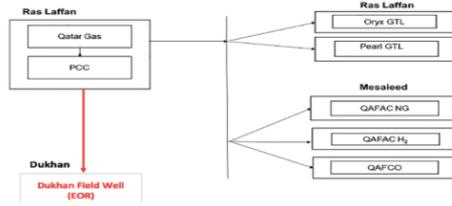


Figure 1: Framework design for CO₂ allocation.

Table 1 presents the products of each sink with their production rates when utilising 1 tonne of CO₂. Chemical absorption with amine solvents was anticipated to capture CO₂ at 100% efficiency, with an OPEX of \$ 40/t CO₂ (Al-Yaeshi et al., 2020), while maintaining its stability as a CO₂ source. In this model, a set amount of 15 Mt CO₂/year is assumed to be captured and sent to sinks reliably. The CCU is estimated to invest \$169 million on CAPEX. Post combustion carbon capture technology (PCC) is utilised due to its high level of commercial availability and technology development. Therefore, in the case of Qatar, PCC technology is considered to be extremely viable for absorbing CO₂ from power plants and other industrial operations. It provides a highly reliable CO₂ source, despite its higher cost compared to alternative capture technologies. Furthermore, the CO₂ composition of flue gas available in Qatar is less than 5%, which meets PCC technology standards (Herzog and Golomb, 2004). In this model, the PCC unit is placed in Qatar Gas LNG plant since it has the highest CO₂ emissions in Qatar (35.8Mt/y) compared to other plants (Mohammed, 2016).

Table 1: Production Rate of each sink for one tonne of CO₂.

Sink	Product	Production rate (tonne of product/tonne of CO ₂)	Reference
QAFAC CH ₄	Methanol	3	(Al-Yaeshi et al., 2020)
QAFAC O	Urea	1.35	(Al-Yaeshi et al., 2020)
QAFAC H ₂	Methanol	5.9	(Al-Yaeshi et al., 2020)
ORYX GTL	Wax, Diesel, and Gasoline	0.884	(Al-Yaeshi et al., 2020)
Pearl GTL	Wax, Diesel, and Gasoline	0.884	(Al-Yaeshi et al., 2020)
EOR	Oil	0.2	(Li and Abushaikha, 2021)

3.2. Model Formulation

The mixed integer linear programming (MILP) model is developed to address the optimal CO₂ allocation from Qatar gas to QAFAC, Pearl GTL, ORYX GTL, QAFAC with H₂, QAFAC with natural gas, and Dukhan Field Well for EOR. This model determines the allocation of CO₂ as a function of commodity price variability which essentially directs the optimal CO₂ allocation from the source to the sinks. The data is collected over a period of 15-year (Al-Yaeshi et al., 2020). This time period is chosen since it is frequently symbolic of the life planning/cycle phase of a CO₂ utilisation project, for which the model tries to identify allocations that optimize the network's long-term economic returns. Table 1 presents the products of each sink with their production rates when utilising 1 tonne of CO₂. The objective is to identify a feasible design solution for allocating CO₂ between the source and sinks that maximises the system's overall cost over multiple periods. Economic returns and technical capacities of sink applications are the main constraints addressed in the model.

Model Formulation

Maximise

$$\sum_{j=1}^n [R_j - C_j] x_j Q_{total}$$

Subject to
$$\sum_{j=1}^n x_j Q_{total} \leq Q_j$$

$$\sum_{j=1}^n x_j \leq 1$$

Where;
$$C_j = C_{cc} + \sum_{j=1}^n (C_{transport} + C_{OPEX j}) \quad n = 1,2 \dots, 6$$

R_j= revenue

C_j= cost for pcc, transport, opex

X_j= fractional amount allocated to sink j

Q_{total}= total amount 15MT/year

Q_j= Sink j capacity

n= number of sinks

4. Results and Discussion

The proposed methodology is used to implement the model's outcomes using the Excel LP optimisation solver. Two scenarios are implemented in this study: (1) a base case scenario for allocating CO₂ from the source to the sinks depending on the sinks capacity; and (2) a scenario that focuses on testing different EOR features by reducing the amount of CO₂ allocated in different years. *Scenario 1* is a base case scenario which aims to determine the optimal solution for allocating 15MT/year of CO₂ to all 6 sinks. The maximum CO₂ utilization, number of sinks, and revenue are included in the solution. According to the findings, the market prices of the products are one of the key drivers of the techno-economic optimization. As a result, the CO₂ utilisation findings in Table 2 remained constant between years 2005 and 2018 for the optimized sinks. This is due to the price linkage between products, which causes them to shift together, except for wax, since the price remained constant from 2011 to 2018 (Al-Yaeshi et al., 2020).

Table 2: The profile for CCU system for period 2005 to 2018.

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Profit \$ (x10 ¹⁰)	1.76	2.43	2.19	3.47	1.87	2.87	3.64	4.15	4.28	4.07	2.08	1.43	2.25	2.47
CO ₂ Utilisation (x10 ⁶)	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35

The CO₂ utilisation of the system is 13.5x10⁶ tonne/year, which implies that the captured amount of CO₂ from Qatar Gas (15x10⁶ tonne/year) is not fully utilised by the sinks. The allocation throughout the years is relatively in the same range. Some changes or fluctuations in the results from one year to another may be due to product prices fluctuations each year. As shown in Figure 2, the optimal allocation solution is to allocate 67% (9Mt/y) of CO₂ to Dukhan Field Well, which is a very high percentage compared to other sinks. This could be due to the high commodity prices, high production rate, low transportation cost, or lower CAPEX and OPEX compared to other sinks which directly affects the economic measure of the system.

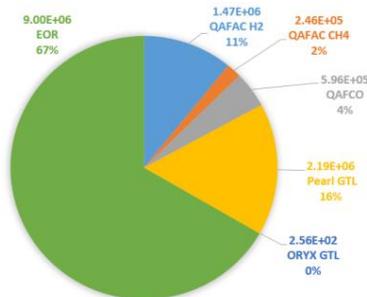


Figure 2: Average CO₂ utilisation by the sinks.

In **Scenario 2**, a simulation using **QASR Reservoir Simulator** (Li and Abushaikha, 2021) is made to examine the natural drive production of oil and CO₂ in Dukhan Field Well (without injecting CO₂ into the reservoir) and the production after CO₂ injection. The natural drive production normally produces a low percentage of the oil inside the reservoir (10 -15%). When CO₂ gas is injected to the reservoir, this can improve the overall oil displacement efficiency it and drives it to the production wellbore, which results in the recovery of a higher percentage of the original oil in place (Massarweh and Abushaikha, 2021). Therefore, different scenarios were conducted in order to examine the reservoirs production in the primary drivee production (no CO₂ injection), when injecting 9Mt of CO₂, and 15Mt of CO₂. Based on this simulation, it is observed that more oil can be produced with CO₂ injection as compared to natural drive production. Figure 3 and 4 illustrates the oil and gas production rate when injecting different CO₂ values. As illustrated in the figures, as the injection rate increases, more production is achieved. This is because the oil greatly inflated as it combined with CO₂ gas, resulting in decreased interfacial tension and viscosity; which results in higher oil and gas recovery.

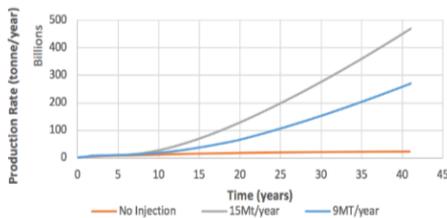


Figure 3: Cumulative Gas Production Rate.

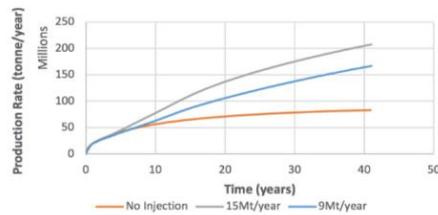


Figure 4: Cumulative Oil Production Rate.

Furthermore, since injecting CO₂ results in producing more CO₂, a sensitivity analysis was made in order to improve the overall system by reinjecting the CO₂ recovered after injection (non-sequestered CO₂) back into the reservoir. By doing so, the required CO₂ amount from Qatar gas each year decreases depending on the production rate in the previous injection (assuming the allocation for all other sinks stay the same). Therefore, in this scenario, it is assumed that Qatargas will allocate 100% of the capacity required for EOR for the first 5 years (9Mt/y) only. Subsequently, this rate will begin to decrease gradually depending on the amount of CO₂ required in this reservoir (see figure 6). Overtime, the allocated CO₂ keeps decreasing until Qatar gas stops allocating CO₂ to EOR since Dukhan Field will have enough “recovered after injection CO₂” from the previous allocations. This scenario improves the sustainability and profit of the system by reducing the amount of CO₂ purchased from Qatar gas annually. As shown in Figure 5, Qatar gas will stop allocating CO₂ to Dukhan Field Well at almost year 21. When comparing the profit from scenario 1 and scenario 2 (Figure 6), it is observed that the profit is constant till year 2010, then, scenario 2’s profit increases further due to the decrease in CO₂ amount allocated.

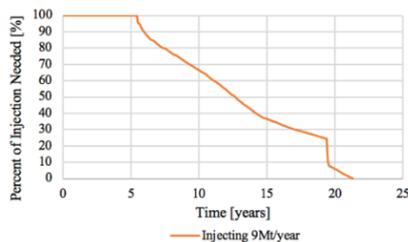


Figure 5: CO₂ utilisation scenario 2.

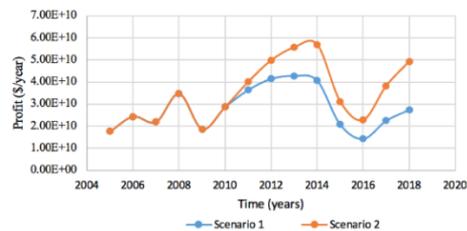


Figure 6: Profit comparison Scenario 1 & 2.

5. Conclusion

The results obtained in this study for both scenarios aims to devise the most effective means for eliminating, reducing, and converting carbon dioxide emissions into a revenue-generating stream. Mixed integer linear programming (MILP) model is used to address CO₂ allocation. The model considers a single CO₂ source (Qatargas) with multiple sinks that use CO₂ as raw material for their production. Two scenarios are considered to allocate CO₂ to different sinks (including EOR) to obtain the optimal solution for each scenario. The outcomes of scenario 1 illustrate that using 13.5 million tonnes of carbon dioxide per year is the best option, with annual profits ranging from 14.3 to 42.8 billion dollars. Dukhan Field Well (EOR) has the highest carbon dioxide utilisation rate, with up to 67 %. Scenario 2 was applied to develop the model further, with the profit increasing annually and the model becoming more sustainable.

References

- Al-Yaeshi, A. A., AlNouss, A., McKay, G., & Al-Ansari, T. (2020). A simulation study on the effect of CO₂ Injection on the performance of the GTL process. *Computers & Chemical Engineering*, 136, 106768. <https://doi.org/10.1016/j.compchemeng.2020.106768>
- Al-Yaeshi, A. A., Govindan, R., & Al-Ansari, T. (2020). Techno-economic-based dynamic network design for optimum large-scale carbon dioxide utilisation in process industries. *Journal of Cleaner Production*, 275, 122974. <https://doi.org/10.1016/j.jclepro.2020.122974>
- Hasan, M.; Boukouvala, F.; First, E.; Floudas, C. Nationwide, Regional, and Statewide CO₂Capture, Utilization, and Sequestration Supply Chain Network Optimization. *Industrial & Engineering Chemistry Research*. 2014, 53, 7489-506.
- He, Y.; Zhang, Y.; Ma, Z.; Sahinidis, N.; Tan, R.; Foo, D. Optimal Source–Sink Matching in Carbon Capture and Storage Systems under Uncertainty. *Industrial & Engineering Chemistry Research*. 2013, 53, 778-85.
- Herzog, H., & Golomb, D. (2004). Carbon capture and storage from fossil fuel use. *Encyclopedia of Energy*, 277–287. <https://doi.org/10.1016/b0-12-176480-x/00422-8>
- Knoope, M.; Ramírez, A.; Faaij, A. Economic Optimization of CO₂ Pipeline Configurations. *Energy Procedia*. 2013, 37, 3105-12.
- lhajaj, A.; MacDowell N.; Shah, N. Multiscale Design and Analysis of CO₂ Capture, Transport and Storage Networks. *Energy Procedia*. 2013, 37, 2552-61.
- Li, L., Abushaikha, A. A fully-implicit parallel framework for complex reservoir simulation with mimetic finite difference discretization and operator-based linearization. *Comput Geosci* (2021). <https://doi.org/10.1007/s10596-021-10096-5>
- Massarweh, O., & Abushaikha, A. S. (2021). A review of recent developments in CO₂ Mobility Control in enhanced oil recovery. *Petroleum*. <https://doi.org/10.1016/j.petlm.2021.05.002>
- Middleton, R.; Bielicki, J. A Comprehensive Carbon Capture and Storage Infrastructure Model. *Energy Procedia*. 2009, 1, 1611-16.
- Mikkelsen, M., Jorgensen, M., Krebs, F., 2010. The teraton challenge. A review of fixation and transformation of carbon dioxide. *Energy Environ. Sci.* 3, 43e81.
- Mohammed, S.S., 2016. Qatar's National Emission Inventory Repor. Qatar Environment and Energy Research Institute , Qatar Foundation/Hamad Bin Khalifa University, Figshare open access book license CC BY-4.0. <http://creativecommons.org/licenses/by/4.0>
- Núñez-López, V., & Moskal, E. (2019). Potential of CO₂-eor for near-term decarbonization. *Frontiers in Climate*, 1. <https://doi.org/10.3389/fclim.2019.00005>
- Weih, F.; Wiley, D. Steady-State Design of CO₂ Pipeline Networks for Minimal Cost per Tonne of CO₂ Avoided. *International Journal of Greenhouse Gas Control*. 2012, 8, 150-68.

Optimal operation of an evaporator for the controlled production of titania nanoparticles

Filippo Tamagnini^a and Sebastian Engell^a

^a*TU Dortmund University, Biochemical and Chemical Engineering, Process Dynamics and operations Group, Emil-Figge-Straße 70, 44227 Dortmund, Germany*
filippo.tamagnini@tu-dortmund.de

Abstract

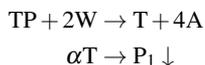
Titania nanoparticles are an important building block for materials with photocatalytic activity or specific optical properties. The production via the hydrolysis of an alcoxide precursor involves a distillation step which is needed to control the final particle size but is time consuming and energy-intensive. The concern of this paper is the optimization of the production process by optimizing the trajectory of a batch evaporator with the goal of maximizing its energy efficiency under predefined production constraints. Depending on the goals, significant reductions in energy consumption are possible, while simultaneously reducing the total processing time. Furthermore, in the situation where electricity from intermittent renewable sources is available, it is possible to tailor the operation of the process to leverage this availability by adapting the input sequence, such that the most energy-intensive phases of the process are performed in the window where energy is available from renewable sources, hence producing a positive economic and environmental impact.

Keywords: trajectory optimization, batch evaporation, titanium dioxide nanoparticles

1. Introduction

Titania nanoparticles have gained attention due to their optical properties, which find application in a very broad range of fields (Ijaz and Zafar (2021)), including photocatalytic materials. The light absorbance properties of nanoparticles depend on the material as well as on their geometrical parameters (Bohren and Huffman (1998)), thus making good control over the production process necessary to fine-tune the photoactivity of the final product.

The synthesis considered here involves the hydrolysis of a titanium precursor (TP) in water (W) to yield dissolved titanium dioxide (T) as well as an alcoholic byproduct (A). Following the hydrolysis, α titanium dioxide units condense to form primary particles (P_1):



The nanoparticles population evolves due to aggregation and breakage towards an equilibrium Particle Size Distribution (PSD). The particle size can be described in terms of the number of primary units in a nanoparticle and in terms of the particle diameter. The first description is useful to model the evolution of the system, the second is used to define measurable goals for the product characteristics. The particle diameter is measured using Dynamic Light Scattering (DLS), which measures the so called Z–Average Particle Size (ZAPS). The typical evolution of the ZAPS during the process is shown in the first plot of figure 1.

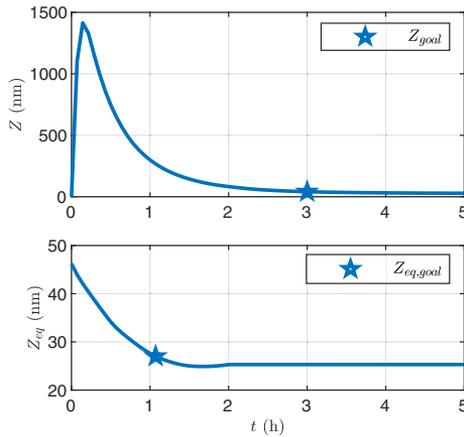


Figure 1: Evolution of the ZAPS and equilibrium ZAPS during the distillation. The times when the product constraints are met are marked with stars in the plot.

find the optimal mode of operation of the batch evaporator that minimizes a measure of cost (processing time, energy consumption or environmental/economic cost) while satisfying the final product constraints. In the optimization, energy can have a time-varying price, e.g. because it is sourced from intermittent renewables, and it is assumed that this availability can be forecasted.

2. Modelling

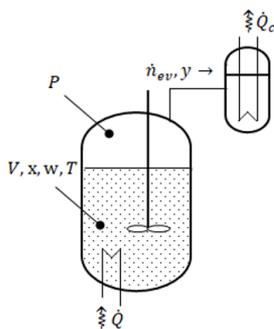


Figure 2: Setup of the plant

Aggregation is a second order process that is limited by the diffusion of the particles through the liquid medium and by the interparticle repulsion (Bal and Bandyopadhyaya (2018)). Alcohols decrease the interparticle repulsion, hence increasing the particle size obtained at equilibrium (Paradisi et al. (2021)). Breakage on the other hand follows first order kinetics and is mainly temperature-driven (Vorkapic and Matsoukas (1999)).

To obtain small nanoparticles sizes, a distillation step is required. The distillation is performed under vacuum, with the power supplied to the system, \dot{Q} and the vacuum pressure, P being the manipulated variables. The setup is shown in figure 2. Heat is supplied through an electric heater, while pressure is controlled through a condenser (also operated electrically). The process is terminated when the particle population has reached the desired size characteristics and the particle size that results at equilibrium after a ripening process is sufficiently small. The goal of this work is to

The state of the system is described by the number of moles of the chemical species n_A, n_W, n_T (alcohol, water and titanium dioxide) and the temperature of the liquid (T). An additional state variable $s \in [0, 1]$ is introduced to describe the normalized fraction of alcohol in the solvation layer of the nanoparticles. Quantities such as the liquid volume $V = (n_A M_A + n_W M_W + n_T M_T) / \rho$, the alcohol fraction in the liquid $x = n_A / (n_A + n_W)$ and the mass fraction of titania $w = \frac{n_T M_T}{\rho V}$ can be computed from the state of the system.

The evolution of the amount of the species is described by their mole balances as shown in eq. 1, where \dot{n}_{ev} is the molar evaporation rate and y the mole fraction of alcohol in the vapor stream in equilibrium with the liquid. The evolution of s , as shown in eq. 2 is a diffusion process driven by the difference between the alcohol fraction in the solvation layer of the particles and the corresponding equilibrium value $K_c x$. The temperature T results from an energy balance (eq. 3) as a function of the heat input \dot{Q} to the system and the heat losses due to evaporation. Constant enthalpy of vaporization H_v and molar heat capacity C_p are used. The evaporation rate is written as a function of the pressure difference between liquid and vapor $\dot{n}_{ev} = k_{ev} \max(P_{ev} - P; 0)$, with k_{ev} a fitting parameter and P_{ev} the equilibrium pressure of the liquid mixture computed using the extended Raoult's law (Wilson model). The heat to be removed by the condenser is $\dot{Q}_c = \dot{n}_{ev} (H_v + C_p (T - T_c))$.

$$\frac{d}{dt} \begin{pmatrix} n_A & n_W & n_T \end{pmatrix}^T = -\dot{n}_{ev} \begin{pmatrix} y & 1-y & 0 \end{pmatrix}^T \quad (1)$$

$$\frac{ds}{dt} = -k_c(s - K_c x) \quad (2)$$

$$\frac{dT}{dt} = \frac{\dot{Q} - \dot{n}_{ev} H_v}{(n_A + n_W + n_T) C_P} \quad (3)$$

The evolution of the population of the nanoparticles is described using a direct moment-based method (Hulburt and Katz (1964)). The k -th moment of the PSD $N_{p,m}$ is defined as $\mu_k = \sum_{m=1}^{\infty} m^k N_{p,m}$. The first moment is directly related to the total number of moles of T as $\mu_1 = \frac{N_A n_T}{\alpha}$, with N_A Avogadro's constant and α the number of T units per primary particle. The dynamics of the k -th moment and the corresponding equilibrium values are given in equations 4-5, where k_a and k_b are the aggregation and breakage rate constants (assumed to be independent of the particle sizes), ϕ_k the k -th moment of the breakage distribution function and, with M_X the molar mass of X and ρ the density of the liquid.

$$\frac{d\mu_k}{dt} = \frac{k_a}{V^2} \left(\frac{1}{2} \sum_{n=0}^k \binom{k}{n} \mu_k \mu_{k-n} - \mu_0 \mu_k \right) + \frac{k_b}{V} \mu_k (\phi_k - 1) \quad (4)$$

$$\frac{\mu_{k,eq}}{V} = \frac{1}{2} \frac{k_a}{k_b} \frac{1}{1 - \phi_k} \sum_{n=1}^{k-1} \binom{k}{n} \frac{\mu_{n,ss}}{V} \frac{\mu_{k-n,ss}}{V} \quad (5)$$

According to the theory of Brownian aggregation, the aggregation process depends linearly on temperature (Lyklema (1991)). Furthermore, it depends on the electrostatic potential between the approaching particles' surfaces, which is strongly influenced by the fraction of alcohol in their solvation layers (Vorkapic and Matsoukas (1998)). A semi-empirical model was used to describe this dependency, with x_a a parameter obtained by fitting steady state particle size-alcohol fraction data. The breakage kinetic constant only depends on temperature with an Arrhenius-like dependency. The choice of the reference temperature T_{ab} in the denominator of (6) is arbitrary, however by using the same temperature for both the aggregation and breakage kinetic constants, T_{ab} is the temperature where the ratio k_a/k_b is at its minimum for a given alcohol fraction. T_{ab} will be used in the optimization as a terminal constraint, as it results in the minimum ZAPS at equilibrium.

$$k_a = k_{a,T_{ab}} \frac{T}{T_{ab}} e^{\left(\frac{k_a s}{k_c x_a}\right)^2} \quad (6)$$

$$k_b = k_{b,T_{ab}} e^{1 - \frac{T}{T_{ab}}} \quad (7)$$

The breakage distribution moments are expressed as a function of the parameter $r \in (0, 1)$ by assuming that a particle of size 1 breaks into fragments of size nr and $n(1-r)$ (Marchisio et al. (2003)):

$$\phi_k = r^k + (1-r)^k. \quad (8)$$

The Z-average diameter of the particles as measured by a DLS instrument is expressed as $Z = Z_0 \left(\frac{\mu_4}{\mu_3}\right)^{\frac{1}{\gamma}}$, with Z_0 the measured diameter of a primary particle and $\gamma \sim 1.9$ the fractal dimension of the particles. The third moment weighted average is used to approximate the weighting behaviour of light scattering instruments (Stetefeld et al. (2016)). $Z_{eq} = Z_0 \left(\frac{\mu_{4,eq}}{\mu_{3,eq}}\right)^{\frac{1}{\gamma}}$ denotes the

measurement obtained from the equilibrium PSD. An important observation is that the moments of the equilibrium PSD in eq. 5 ultimately depend on w , x and T , therefore it is possible to write $Z_{eq} = Z_{eq}(x, w, T)$. During the course of the distillation, the value of Z_{ss} initially decreases due to the depletion of alcohol until a minimum is reached. Past this point, the increase in weight fraction determines an increase in the obtainable particle size, as visible in the second plot of figure 1.

3. Optimization

The optimization problem is formulated in equations 9-18, with respect to the state vector $x(t) = (n_A(t), n_W(t), n_T(t), s(t), T(t), \mu_0(t), \dots, \mu_4(t))^T$ and the input vector $u = (\dot{Q}(t), P(t))^T$:

$$\min_{u(t)} J_n[u(t)] \quad (9)$$

$$s.t. : \quad \frac{dx(t)}{dt} = f(x(t), u(t)) \quad \forall 0 \leq t \leq t_{max} \quad (10)$$

$$x(0) = x_0 \quad (11)$$

$$0 \leq \dot{Q}_c(t) \leq \dot{Q}_{c,max} \quad \forall 0 \leq t \leq t_{max} \quad (12)$$

$$0 \leq \dot{Q}(t) \leq \dot{Q}_{max} \quad \forall 0 \leq t \leq t_{max} \quad (13)$$

$$P_{min} \leq P(t) \leq P_{max} \quad \forall 0 \leq t \leq t_{max} \quad (14)$$

$$Z(t_{max}) \leq Z_{goal} \quad (15)$$

$$Z_{eq}(t_{max}) = Z_{eq,goal} \quad (16)$$

$$T(t_{max}) = T_{ab} \quad (17)$$

$$J_1 \leq t_{max}. \quad (18)$$

Three cost functions are defined: the duration of the process J_1 , the total energy use J_2 and the energy cost J_3 :

$$J_1[u(t)] = \min t \text{ such that: } (Z(t) \leq Z_{goal}) \text{ and } (Z_{eq}(t) \leq Z_{eq}) \text{ and } (T(t) = T_{ab}) \quad (19)$$

$$J_2[u(t)] = \int_0^{t_{max}} \dot{Q}(t) + \frac{\dot{Q}_c(t)}{COP} dt \quad (20)$$

$$J_3[u(t)] = \int_0^{t_{max}} p(t) \left(\dot{Q}(t) + \frac{\dot{Q}_c(t)}{COP} \right) dt. \quad (21)$$

where the Boolean variable $p(t) \in \{0, 1\}$ denotes the energy price (economic or environmental, e.g.: $\text{kg}_{\text{CO}_2}/\text{kW}$) at time t and $\frac{\dot{Q}_c(t)}{COP}$ the electrical power consumed by the condenser. This definition of the energy price is used to simplify the formulation of the problem. In reality, $p(t)$ is a real number varying between a minimum and a maximum. Therefore, the cost function J_3 is more accurately the cost incurred by not buying energy at the lowest price.

The function $f(x(t), u(t))$ is the vector of the right hand sides from equations 1-4. The initial conditions of the system are fixed prior to the optimization from the product recipe and stoichiometry. The particulate system is initialized as a point distribution with $\mu_k(0) = \mu_1 = \frac{N_{AB}V_T}{\alpha}$, $s(0) = 1$ and $T(0) = 30^\circ\text{C}$. The input variables and the heat removed by the condenser have minimum and maximum bounds. The terminal constraints are chosen such that the ZAPS at the end of the ripening step after the process meets the product constraints with the final values of x and w (16), the PSD at the end of the process is sufficiently close to the equilibrium PSD (15), and the temperature of the system is the one for which the minimum ZAPS can be achieved, for given values of x and

w (17). Unless the distillation proceeds significantly past the point where the minimum Z_{ss} is obtained (hence failing to meet the constraint on Z_{ss}) and except for the initial phase of the process where the nanoparticles grow to their maximum size, $Z(t)$ is strictly decreasing and $Z(t) \geq Z_{eq}(t)$ holds. For this reason, once the constraint on Z is met, it is guaranteed to be satisfied until the end of the process. Furthermore, once $T(t) = T_{ab}$ is achieved, any variation would either determine an increase in the duration of the process or the energy expense, making the solution sub-optimal. Thus, during optimization, the terminal constraints can be checked at t_{max} only, while the duration of the process is computed by equation 19. t_{max} can be used to further impose a constraint on the productivity of the process.

To solve the problem, the states and inputs are discretized in time using orthogonal collocation on finite elements (Cuthrell and Biegler (1987)). The time interval from 0 to 6 hours is subdivided in 12 subintervals, within which the inputs are assumed to have a given shape (\dot{Q} piecewise constant, P piecewise linear). Each subinterval is further subdivided in 6 finite elements to control the precision of the solution. The constraints are evaluated at the boundaries of each finite element. $N_c = 3$ collocation points are placed on each finite element, where the states are approximated by Lagrange interpolating polynomials. Continuity is enforced between interpolating polynomials at the boundary of each two finite elements and with the initial conditions. A Radau scheme is chosen for the collocation points for numerical stability. The model and optimization were implemented in python using CasADi (Andersson et al. (2019)), with IPOPT as solver (Wächter and Biegler (2005)).

4. Results

An overview of some results is given in figure 3. It was found that under stringent product constraints and despite being the most energy-intensive mode of operation, the J_1 -optimal input can be up to 10% less energy demanding than the least efficient input policy with the same terminal constraints. The J_2 -optimal operation requires more time to complete the process, but reduces the total energy input by an additional 10%, compared to the J_1 -optimal solution. The trade-off between the duration of the process and the energy expense in the J_2 -optimization can be controlled by choosing t_{max} . For the J_3 -optimization, several scenarios were tested where the energy price (economic or environmental) was set to 0 in a time window between t_1 and t_2 and 1 outside. The J_3 -optimized processes adapts the operation such that the energy input is maximized within the window as shown in figure 4. The total energy use in this scenario can be very close to the optimal energy use if the low price window is not close to the end of the process. The J_3 -optimal operation is less advantageous and less energy-efficient if the window begins close to the end of the process. This is visible in the second J_3 -optimal trajectory in figure 4 and illustrative in figure 3. For this reason it is advisable to schedule the process such that cheap energy is available within the first half of the batch run.

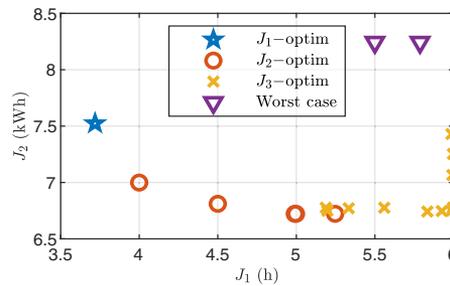


Figure 3: J_1 and J_2 for different optimization scenarios with $Z_{eq,goal} = 25$ nm and $Z_{goal} = 35$ nm. Different J_2 -optimal solutions were obtained by setting $t_{max} \in \{4, 4.5, 5, 5.5\}$ in eq. 18. The J_3 -optimal solutions were obtained by varying the position and size of the cheap energy price window. For these, J_1 reflects the position of the window. The worst case scenarios were computed by maximizing energy for the given product constraints and $t_{max} \in \{5.5, 6\}$.

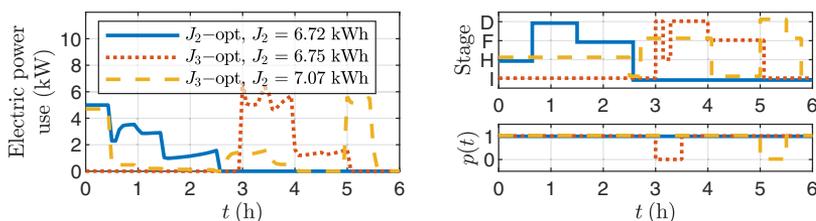


Figure 4: Comparison of the power consumption and process stages in J_2 - and J_3 -optimal solutions for $Z_{goal,eq} = 25$ nm, $Z_{goal,dyn} = 35$ nm for variable energy price $p(t)$. Process stages: Idle (I), Heating (H), Flashing (F), Distillation (D). The total energy use J_2 is given for all scenarios

5. Conclusion and future work

The problem of finding the optimal operation of a batch evaporator was addressed and solved for several objective functions. It was found that improvements in the performance indicators can be obtained, and that using trajectory optimization yields a significant reduction in the total processing time and total energy consumption, compared to the unoptimized operation. Furthermore, the process was found to be flexible enough to efficiently make use of intermittent energy, making it a promising candidate for the application of demand-side response.

6. Acknowledgements

The project leading to this publication has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 820716 (SIMPLIFY).

References

- J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, M. Diehl, 2019. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11 (1), 1–36.
- V. Bal, R. Bandyopadhyaya, 2018. Generalized model for nano- and submicron particle formation in liquid phase, incorporating reaction kinetics and hydrodynamic interaction: Experiment, modeling, and simulation. *The Journal of Physical Chemistry C* 122 (35), 20489–20499.
- C. F. Bohren, D. R. Huffman, Apr. 1998. Absorption and scattering of light by small particles. Wiley.
- J. E. Cuthrell, L. T. Biegler, 1987. On the optimization of differential-algebraic process systems. *AIChE Journal* 33 (8), 1257–1270.
- H. M. Hulburt, S. Katz, 1964. Some problems in particle technology: a statistical mechanical formulation. *Chemical Engineering Science* 19 (8), 555–574.
- M. Ijaz, M. Zafar, 2021. Titanium dioxide nanostructures as efficient photocatalyst: progress, challenges and perspective. *International Journal of Energy Research* 45 (3), 3569–3589.
- J. Lyklema, 1991. Fundamentals of interface and colloid science. Vol. 4. Academic Press, London San Diego.
- D. L. Marchisio, R. D. Vigil, R. O. Fox, 2003. Quadrature method of moments for aggregation-breakage processes. *Journal of Colloid and Interface Science* 258 (2), 322–334.
- E. Paradisi, R. Rosa, G. Baldi, V. Dami, A. Cioni, G. Lorenzi, C. Leonelli, 2021. Effect of isopropanol co-product on the long-term stability of tio2 nanoparticle suspensions produced by microwave-assisted synthesis. *Chemical Engineering and Processing - Process Intensification* 159.
- J. Stetefeld, S. A. McKenna, T. R. Patel, Dec 2016. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews* 8 (4), 409–427.
- D. Vorkapic, T. Matsoukas, 1998. Effect of temperature and alcohols in the preparation of titania nanoparticles from alkoxides. *Journal of the American Ceramic Society* 81 (11), 2815–2820.
- D. Vorkapic, T. Matsoukas, 1999. Reversible agglomeration: a kinetic model for the peptization of titania nanocolloids. *Journal of Colloid and Interface Science* 214 (2), 283–291.
- A. Wächter, L. T. Biegler, apr 2005. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106 (1), 25–57.

A model-based approach for the prediction of banana rust thrips incidence from atmospheric variables

Carlos A. Estrada^a, José Manrique-Silupú^{a,*}, William Ipanaqué^a, Bogdan Dorneanu^{b,c}
and Harvey Arellano-García^{b,c}

^aUniversidad de Piura, Av. Ramón Mujica 131, Piura 20009, Perú

^bUniversity of Surrey, Stag Hill University Campus, GU2 7HX Guildford, UK

^cLS- Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg,
Cottbus D-03044, Germany
jose.manrique@udep.edu.pe

Abstract

This work focuses on the development of a mathematical model for the population growth of banana red rust thrips (*Chaetanaphothrips signipennis*) based on a modified temperature-based growth rate with the addition of climatic variables, such as relative humidity, wind speed and rainfall rate. The aim is to enable better prediction of the pest incidence and improve decision making, productivity, as well as quantifying the influence of these variables on the development of red rust thrips. The developed model is then compared with current solutions for predicting the pest incidence, showing improved accuracy (higher than 67%) versus experimental data, for which the state-of-the-art models indicate extremely poor fits.

Keywords: Mathematical modelling; IoT sensors; Precision agriculture; Pests; Organic banana; Red rust thrips.

1. Introduction

Banana is one of the most important food crops in the world, contributing to the food security of billions of people by providing income and employment to rural communities (MINAGRI, 2014). In the last ten years, Peru has joined the large community of organic banana exporting countries, with the main production located on its northern coast, reaching around 223,298 tons of organic bananas exported in 2019 (FAO, 2021). Since 2010, red rust thrips have become a serious issue on local organic banana plantations, causing yield losses of 30%-40% if effective measures are not applied (Lopez et al., 2020).

Currently there are several (organic) products that help to control this pest, but their efficacy depends on the level of incidence present. If the plot is already infested, corrective measures need to be applied with products that cost 7 to 8 times more than the solutions used preventively. Thus, a total infestation of the banana plots is extremely detrimental to the crop and the farmers because, in addition to losing part or all of the production, the cost of recovering from infestation will also increase significantly.

For this reason, tools that help in the decision-making process and enable farmers to reduce as much as possible the impact of these pests on the production are needed. Precision agriculture offers technologies that can support this development. Among these, intelligent sensing, automated data collection, big data analysis and modeling approaches, have the potential to increase

crop yields, detect diseases, pests or weeds early, as well as improve the control, maximize the economic benefits and ensure resource conservation (Annosi et al., 2019).

This research is part of a project aimed at developing tools based on precision agriculture and digitalisation technologies that can help improve the control of pests during the organic banana production. To achieve this objective, an investigation of the pest behaviour is required, to better understand what factors influence their growth throughout the year. In this contribution, the development of a mathematical model for the prediction of the pest population, with application to the banana red rust thrips (the pest that most affects the organic banana production), as a function of the atmospheric variables recorded by IoT sensors located in an organic banana plot (1 ha area) is presented. These models are based on bacterial growth models. By considering other climate variables, such as environmental humidity, rainfall and wind speed, improvement over existing pest incidence prediction models (e.g., the Campbell model (Donatelli et al., 2017)) is sought.

2. Methodology

2.1. Data acquisition

The input data used in the model is collected by a smart sensor network located in the district of Buenos Aires, province of Morropon, department of Piura (Peru). Placed in an organic banana plot, the network consists of two nodes and a weather station illustrated in Figure 1. The station uploads data to the cloud every 15 minutes, so there are 96 data per day. The daily average values are used further for the model.



Figure 1: Weather station (left) and a node (right).

To enable prediction of the pest incidence (the number of thrips on the plot), this parameter needs to be quantified. This task is performed manually and recorded on a weekly basis. For this assessment, the number of eight different types of pests, namely the red rust thrips - *Chaetanaphothrips signipennis*, the mealybug - *Pseudococcus elisae*, the weevil - *Cosmopolites sordidus*, the red spider mite - *Tetranychus spp*, the ceramid - *Ceramidia sp*, the aphid - *Pentalonia nigronervosa* and the whitefly - *Aleurodicus dispersus*, is counted for 25 plants, and averaged to provide an idea on the condition of the plot.

The behavior of the pests has been monitored for almost two years (from November 2019). During this period, it has been observed that, during the summer season, when the maximum temperature reaches values above 30°C, the thrips population increases significantly compared to the other seasons of the year.

2.2. The Campbell Model (CM) for predicting insect growth

Previous work (Campbell et al., 1974) in the area of parasite incidence prediction shows that the growth rate of insects depends on the temperature to which they are exposed (Figure 2). Over a range of temperatures, the relationship between the growth rate and the temperature can be represented by a straight line which, when extended, cuts the x-axis at the temperature threshold (T). Looking at Figure 2, one conclusion that can be drawn is that it is impractical to analyze the average growth rate in the range A domain, due to the high mortality caused by low temperatures. At high temperatures, characteristic of the range C domain, the average rate of development decreases with respect to the linear growth observed in the range B domain. However, this growth rate reduction observed at high temperatures occurs only if the temperature remains constant or fluctuates within a small range for a long period of time, hence it is not normally experienced. Therefore, in practice the field conditions are almost always in the linear section corresponding to the range B domain.

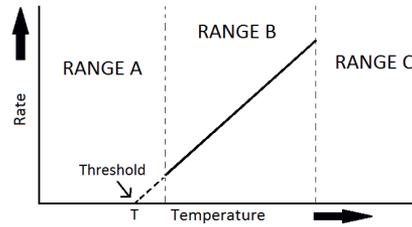


Figure 2: The relationship between the rate of insect development and temperature.

For practical purposes, the growth rate of insects can be approximated by a linear equation (Prasad et al., 2021):

$$\rho = a + b \cdot T \quad (1)$$

The parameters a and b are constants estimated from “experimental” data, with a the intercept of the rate of development, and b the slope relating temperature to the rate of thrips growth, while T is the temperature ($^{\circ}\text{C}$).

2.3. The Modified Campbell Model (MCM)

In the CM, only the temperature is taken into account, while important climatic variables such as rainfall, wind speed and relative humidity, which have an impact on the insect population growth (Elbehri et al., 2015), are neglected. To this end, by considering these variables measured on the plot, an improvement to the prediction of the CM is expected. Thus, the following equation is considered for the rate of growth of insects:

$$\rho = a + b \cdot T + c \cdot H + d \cdot W + e \cdot R \quad (2)$$

Where H is the relative humidity (%), W the wind speed (m/s), and R the rain rate (mm/h), while a ($1/\text{day}$), b ($1/\text{day}\cdot^{\circ}\text{C}$), c ($1/\text{day}\cdot\%$), d ($1/\text{day}\cdot\text{m/s}$), and e ($1/\text{day}\cdot\text{mm/h}$).

2.4. Mathematical Modelling

To describe the population dynamics of insects, Murray’s exponential model of bacterial growth can be used, an implementation considered in (Hernández and Rivera, 2018) to find the parameters and determine the number of microorganisms. This particular case is used to approximate the number of thrips (insects per plant) in an organic banana crop at any instant of time.

$$\frac{dP_T}{dt} = \rho \cdot P_T - \psi \cdot P_T = (\rho - \psi) \cdot P_T = r \cdot P_T \quad (3)$$

Where P_T represents the population of thrips per plant, ρ , the birth rate ($1/\text{day}$), ψ , the death rate ($1/\text{day}$) and r describes the growth rate ($1/\text{day}$).

Two types of pesticides are regularly used on the organic banana plot, and hence considered for this model. The first one is exclusive for rust thrips and is used in a corrective way, being also more aggressive. The second one is the most frequent type of pesticide used, and its application is performed in a preventive fashion for pests such as red spider mites, white flies, thrips, etc. Moreover, its effect is weaker compared to the one used for the rust thrips. For this reason, pesticide spraying terms, F_1 and F_2 , are considered in the model, and expressed based on the degradation of antibiotics in bacteria (Romero Leiton et al., 2011):

$$F_1 = e^{-\varphi_{F_1} \cdot \tau_{F_1}} \quad (4)$$

$$F_2 = e^{-\varphi_{F_2} \cdot \tau_{F_2}} \quad (5)$$

Where F_1 is the concentration of organic pesticide 1, F_2 is the concentration of organic pesticide 2, φ_{F_1} is the degradation rate of pesticide 1, φ_{F_2} is the degradation rate of pesticide 2, τ_{F_1} is the difference between the current day and the day pesticide 1 was sprayed (*day*), τ_{F_2} is the difference between the current day and the day pesticide 2 was sprayed (*day*).

During the third nymphal stage, mature nymphs migrate from the plant to the soil or epidermis and transform into prepupae. At this stage, cleaning of the plot is an influential factor in the thrips population growth, since it directly interferes with their life cycle. If this activity is not carried out, it would result in a significant increase of the number of insects. Therefore, the cleaning term is described in a similar way to the spraying terms, as its impact on the insect population decreases over time:

$$C = e^{-\varphi_c \cdot \tau_c} \quad (6)$$

Where C is the variable representing the cleanliness of the plot, φ_c is the degradation rate of the cleanliness level, τ_c is the difference between the current day and the day on which the cleaning was carried out (*day*).

Unlike the case of Equation 3, where the growth rate is constant, in the following the growth rate is expressed as a function of the crop conditions, such as pesticide spraying and cleaning, which directly affect the thrips population:

$$r = \rho - \psi - \phi_1 \cdot F_1 - \phi_2 \cdot F_2 - \gamma \cdot C \quad (7)$$

Where ρ is the rate of thrips development (*1/day*), ψ is the natural mortality rate (*1/day*), ϕ_1 is the mortality rate from spraying with chemical 1 (*1/day*), ϕ_2 is the mortality rate from spraying with chemical 2 (*1/day*), γ is the mortality rate from the cleaning (*1/day*).

Finally, the rate of change of thrips intensity with respect to time for the modified Cambell model (MCM) is obtained based on the conditions in the crop field:

$$\frac{dP_T}{dt} = (\rho - \psi - \phi_1 \cdot F_1 - \phi_2 \cdot F_2 - \gamma \cdot C) \cdot P_T \quad (8)$$

3. Results

In the following, the predictions for the red rust thrips incidence of the two models described above, the MCM and the CM, are compared. For this purpose, 100 data sets are generated from the experimental data gathered across a period of one and half years. For cross validation, each set has been divided into training (75%) and validation (25%) data. The sets are then used for parameter estimation, and the resulting models' metrics are shown in Table 1.

For the training and validation sets, mean values of 0.78 and 0.66 were obtained for R-squared, while 1.50 and 1.60 were obtained for RMSE, respectively. For the training and validation sets,

average values of 0.80 and 0.70 were obtained for R-squared while for RMSE 1.45 and 1.55 respectively.

Table 1: Model metrics for the CM and MCM: Coefficient of determination (R^2) adjusted and Root mean square error (RMSE)

	R^2 adjusted		RMSE	
	CM	MCM	CM	MCM
Train	0.78	0.80	1.50	1.45
Test	0.66	0.70	1.60	1.55

Figure 3 shows the real data compared to the prediction of the two models. For the MCM, an R^2 of 0.73 and an RMSE of 1.76 were obtained, while for CM these were 0.72 and 1.78, respectively. As the same data is used in the cross-validation, there is not much difference in the results for these metrics, although it seems that the MCM is slightly more accurate than the CM.

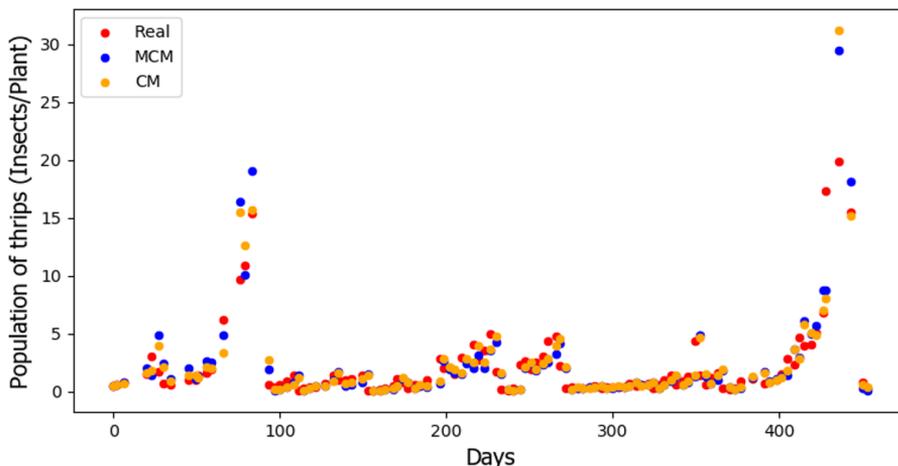


Figure 3: Comparison of model predictions with experimental data used for model development.

When the two models are used for prediction of new data (the validation set), corresponding to the period from July 8 to October 19, 2021, the metrics shown in the Table 2 indicate that the MCM has a better accuracy compared to the CM. Furthermore, it can also be observed that the MCM predictions are not so far away from the real values. At the same time, the CM results indicate a very poor fit.

	CM	MCM
R^2	-1.68	0.67
RMSE	4.00	1.40

Table 2: Metrics obtained for the model predictions for the validation set.

Figure 4 shows the response of the models for the data recorded in the plot during the 103 days corresponding to the validation set. During this time interval, 10 pest evaluations were performed and used to assess the accuracy of the models.

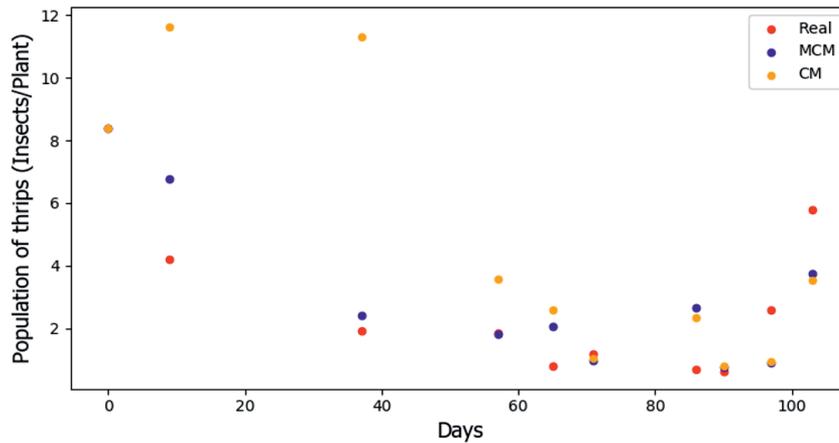


Figure 4: Comparison of model predictions with experimental data used for validation.

4. Conclusions

Understanding the growth behavior of the rust thrips with respect to atmospheric variables enables farmers to take better preventive control measures to avoid infestation as much as possible. In this way, production losses as well as plot maintenance costs can be significantly reduced. This contribution presents the development of a model for the predictions of the thrips incidence in an organic banana plot based on climatic variables. The results show good agreement with the experimental data, providing a solution that can be further implemented in decision-making tools for the organic banana producers.

These can be used in the future work to define and improve the monitoring and control of the pest incidence by defining better fumigation and cleaning schedules that will increase their efficiency and benefit the quality of the final product.

References

- M. C. Annosi, F. Brunetta, A. Monti, *F. Nat.*, 2019. Is the trend your friend? an analysis of technology 4.0 investment decisions in agricultural smes. *Computers in Industry* 109, 59–71.
- A. Campbell, B. D. Frazer, N. Gilbert, A. P. Gutierrez, M. J. P. Mackauer, 1974. Temperature requirements of some aphids and their parasites. *Journal of Applied Ecology* 11, 431–438.
- M. Donatelli, R. Magarey, S. Bregaglio, L. Willocquet, J. Whish, S. Savary, 2017. Modelling the impacts of pests and diseases on agricultural systems. *Agricultural Systems* 155, 213–224.
- A. Elbehri, G. Calberto, C. Staver, A. Hospido, L. Roibas, D. Skully, P. Siles, J. Arguello, I. Sotomayor, A. Bustamante, 2015. Cambio climático y sostenibilidad del banano en el Ecuador: Evaluación de impacto y directrices de política. Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), Roma, Italia.
- FAO, 2021. Banana market review – Preliminary results 2020.
- D. M. Hernández, J. D. Rivera, 2018. Determinación de parámetros del modelo exponencial mediante una aplicación móvil. *CIENCIA ergo-sum* 25 (2).
- M. Lopez, R. Corozo-Ayovi, R. Delgado, B. Osorio, D. Moyón, D. Rengifo, P. Suarez, A. Paulino, S. Medrano, L. Sanchez, J. Rojas Llanque, U. Vegas, D. Alburqueque, C. Staver, R. Van Tol, L. Clercx, 2020. Red rust thrips in smallholder organic export banana in Latin America and the Caribbean: Pathways for control, compatible with organic certification. *Acta Horticulturae*, 153–161.
- MINAGRI, 2014. El banano peruano, producto estrella de exportación. Tech. rep.
- Y. G. Prasad, M. Gayathri, V. Sailaja, M. Prabhakar, G. Ramachandra Rao, G. Rekha, S. Vennila, 2021. Evaluation of linear and nonlinear models for temperature driven development of *spodoptera litura* (fabricius) on soybean crop. *Journal of Agrometeorology* 23 (2), 169–175.
- J. Romero Leiton, E. Ibarguen-Mondragon, L. Esteva, 2011. Un modelo matemático sobre bacterias sensibles y resistentes a antibióticos. *Matemáticas: Enseñanza Universitaria* 20, 55–73.

Comparison between 3D numerical simulations and experimental results of a lab-scale liquid-solid fluidized bed

Almir G. S. L. Ritta^a, Renaud Ansart^a and Olivier Simonin^b

^a*Laboratoire de Génie Chimique (LGC), Université de Toulouse, CNRS, INPT, UPS, 4 allée Emile Monso, Toulouse 31432, France*

^b*Institut de Mécanique des Fluides de Toulouse (IMFT), Université de Toulouse, CNRS, 2 All. du Professeur Camille Soula, Toulouse 31400, France*
almirguilherme.siqueiralopesritta@toulouse-inp.fr

Abstract

The high consumption of fossil fuels is running out the petroleum sources, therefore it is extremely important to improve the efficiency or diversify the industry's main route. To deal with these new processes, the liquid-solid fluidized bed is one of the available technologies. Mastering this complex industrial device requires a better understanding of the intrinsic multi-scale hydrodynamics couplings that play an important role between solid and liquid phases. For that reason, the numerical simulation of multi-phase flow has become a useful tool for examining the particle behavior in fluidized beds. The fluidized bed simulation may be addressed using the Discrete Element Method (DEM-CFD) approach, which is a model that requires a lot of computational effort, because of the large scale of the equipment and the huge number of particles. Hence, for industrial purpose, this approach becomes infeasible, and simulations are carried out in the frame of Euler-Euler model such as implemented in the NEPTUNE_CFD code. But such approach was originally developed for gas-solid flows, then a specific model developments and validation studies are needed for liquid-solid fluidized bed (Gevrin et al., 2008). This is one of the main objectives of the MUSCATS project (N°ANR-19-CE05-0010-02), which is carrying out an innovative multi-scale modelling program for liquid-solid fluidized beds, based on complementary numerical approaches and carefully designed experiments. In a first stage, 3D numerical simulations were performed using the NEPTUNE_CFD code on a laboratory configuration in order to assess the model assumption and to compare the mean and fluctuating fluid and particle velocity predictions with available experiment data from Aguilar-Corona (2008). Thus, some discrepancies on fluid and particle kinetic fluctuating energy and areas of improvement are identified and proposed in order to capture accurately the particle-fluid and inter-particle interactions in the liquid-solid fluidized bed.

Keywords: CFD, Euler-Euler, Fluidized bed, Liquid-solid bed

1. Introduction

The fluidization phenomenon occurs when solid particles are suspended by a fluid, gas or liquid, and they behave like a fluid-like state. The liquid fluidization is widely utilized in several processes such as chemical, petrochemical, hydro-metallurgical, mining, biochemical, pharmaceutical and water treatment (Wang et al., 2014; Nirmala and Muruganandam, 2019). As reported by Kunii and Levenspiel (1994) the bed properties are heavily dependent of the fluid velocity and it plays

an important role over the fluid-dynamics of the equipment, because the flow regime changes in function of the fluidization velocity. For instance, when the fluid acquires enough kinetic energy to support the particles, but without dragging them out, the regime changes from a fixed bed to a fluidized bed with certain intermediate states. As it is a fluid transport equipment, the fluidized bed can be dimensioned based on factors such as pressure loss, the fluid density and viscosity, the particle density, diameter and shape, the fluid velocity and etc (Yates and Lettieri, 2016).

The Computational fluid dynamics (CFD) has been applied to simulate the liquid fluidized bed system and according to Zhang et al. (2013) the models used to simulate the liquid-solid fluidized bed can be divided into two major groups: At a macro-scale it is represented by the Two-Fluid Models (TFM) approach and at meso-scale it is distinguished by the use of CFD coupled with the discrete element method (DEM). The mass and the momentum equations are solved separately for each phase taking into account the momentum transfer between liquid and solid. Hence, the numerical results is strongly dependent on the drag force between the phases and the solid stress (Ren et al., 2021). Many authors has studied the liquid-solid fluidized bed and they noted that the existence of meso-scale structures play an important role over the hydrodynamics and it is not well computed when coarse grid is employed. Beside that, they have investigated the influence of many parameters such as density ratio, fluid viscosity, restitution coefficient, forces acting on particles, turbulence model and etc. They concluded that fluidized bed hydrodynamics results are considerably impacted by turbulence and the drag model (Ren et al., 2021; Khan et al., 2017). In order to investigate the capability and development needs of mathematical modeling and the lack of accurate prediction of fluid and particle agitation for various fluidization velocities, numerical predictions from NEPTUNE_CFD code are compared with laboratory measurements from Aguilar-Corona (2008).

2. Experimental measurement from Aguilar-Corona (2008)

Aguilar-Corona (2008) performed many experiments using a solution of Potassium Thiocyanate (KSCN) 64 % (w/w) and solid particles of Pyrex beads. She investigated many parameters experimentally using a high-speed camera to capture the trajectory of the particles. The experimental results performed by Aguilar-Corona (2008) were used to build the reference case, which is a cylinder consisting of 80 mm in diameter and 500 mm in height. The fluid and particle properties are provided in Tab. 1. For more details see Aguilar-Corona (2008).

Table 1: The fluid & particle properties.

Description	Value
Fluid density	1400 kg/m^3
Fluid viscosity	3.8E-3 Pa.s
Particle density	2230 kg/m^3
Particle diameter	6 mm
Total solid mass	0.624 kg
Initial bed height	9.5 cm
Initial solid packing	0.585
Terminal velocity	0.226 m/s

3. Mathematical modeling and numerical parameters

Three-dimensional numerical simulations are carried out using NEPTUNE_CFD. This Eulerian n-fluid unstructured parallelized multiphase flow software is developed in the framework of the NEPTUNE project financially supported by CEA (Commissariat a l'Énergie Atomique), EDF

(Electricité de France), IRSN (Institut de Radioprotection et de Sureté Nucléaire) and Framatome. The modeling approach for fluid–particle flows is implemented by Institut de Mécanique des Fluides de Toulouse (IMFT) (Neau et al., 2020). The Eulerian n-fluid approach used is a hybrid approach where the transport equations are derived by phase ensemble averaging for the continuous phase and by using kinetic theory of granular flows supplemented by fluid and turbulent effects for the dispersed phase thanks to joint fluid–particle Probability Density Function (PDF) approach. In the proposed modeling approach, transport equations (mass, momentum and fluctuating kinetic energy) are solved for each phase and coupled through inter-phase transfer terms. The momentum transfer between liquid and particle phases is modeled using the drag law of Wen and Yu limited by Ergun equation for the dense flows (Gobin et al., 2003). The collisional particle stress tensor is derived in the frame of the kinetic theory of granular media. The fluid turbulence modeling is achieved by the two equations of $\kappa - \varepsilon$ model extended to particle-laden flows accounting for additional source terms due to the inter-phase interactions (Vermorel et al., 2003). For the dispersed phase a coupled transport equation system is solved on particle fluctuating kinetic energy and fluid–particle fluctuating covariance ($q_p^2 - q_{fp}$). The effects of the particle–particle contact force in the very dense zone of the flow are taken into account in the particle stress tensor by the additional frictional stress tensor. For more details about the model see (Hamidouche et al., 2018; Gevrin et al., 2008).

The constructed mesh is composed of 510.000 cells in size $\Delta x = \Delta y = 0.001$ m and $\Delta z = 0.002$ m. The inlet, outlet and wall boundaries were defined, respectively, as inlet velocity, free pressure outlet and friction for both phases. We ran several simulations using the following fluidization velocities: $U_f = 0.073, 0.09, 0.12, 0.15$ and 0.17 m/s. The inlet velocity was assumed to be homogeneous over the inlet surface. For these velocities tested, the Reynolds Number is ranging from 82 to 378 and the Stokes Number is varying from 0.9 to 6.1. The coefficient of restitution and maximum packing α_p are, respectively, 0.9 and 0.64.

4. Fluidization law

In order to determine the fluidization law, many simulations was performed with different fluidization velocity. Figure 1(a) illustrates, for different U_f , the time average for solid volume fraction, $\langle \alpha_p \rangle$, at the center of the column. As overall trend, the bed height moves upward when fluidization velocity increases, while $\langle \alpha_p \rangle$ drops. In other words, the fluidized bed becomes more diluted when the fluidization velocity is increased. Figure 1(b) presents the mean value of the solid volume fraction averaged over the entire bed height for different velocities. Below you can see the numerical results, experimental value from Aguilar-Corona (2008) and Wen and Yu (1966), Tenneti et al. (2011) & Richardson and Zaki (1954) correlations. The Richardson and Zaki (1954) equation is showed in Eq. 1 and the parameters utilized to plot Fig. 1(b) are $U_0 = 0.226$ m/s and $n = 2.39$.

$$U_t = U_0(1 - \alpha_p)^n \quad (1)$$

Analysing Fig. 1(b) we can see that the numerical results are much more in agreement with Tenneti et al. (2011) and Wen and Yu (1966) correlations, the last one is a correlation that takes into account a homogeneous system without collision. The drag law utilized in our simulations is based on Wen & Yu. Therefore, an investigation using different drag laws seems to be important, as the numerical results were much closer to the correlation than the experimental results. As a result, the findings differ from those obtained in the experiments and Richardson and Zaki (1954) correlation, although for dense regime ($\langle \alpha_p \rangle > 0.3$) the results present a good agreement with the experimental and the correlations.

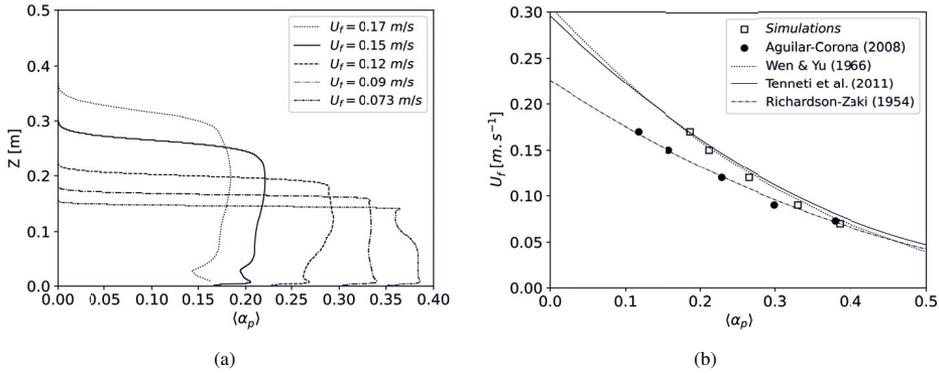


Figure 1: Numerical and experimental results where figure (a) shows $\langle \alpha_p \rangle$ profiles for different fluidization velocities at a central line inside the cylinder and figure (b) shows the spatial average value for $\langle \alpha_p \rangle$ in function of fluidization velocity U_f .

5. Fluid and particle analysis of the total fluctuating kinetic energy

To calculate the fluctuating kinetic energy, some nomenclatures need to be clarified. Where Ψ_k is an ordinary variable and $\widetilde{\Psi}_k$ is the time average taking in account α_k as represented in Eq. 2.

$$\widetilde{\Psi}_{k,i} = \frac{\sum \alpha_k \Psi_{k,i} \Delta t}{\sum \alpha_k \Delta t} \quad (2)$$

In order to calculate the fluctuating kinetic energy in each direction is necessary to compute the 3 terms shown in Eq. 3. Where the first term is the spatial variance of the mean velocity, the second one is the spatial mean of the local temporal variance and the third one is the spatial mean of the random fluctuation of q_k^2 .

$$\alpha_{k,i} = \widetilde{U_{k,i}^2} + \langle \widetilde{u_{k,i}^2} \rangle + \langle 2/3 q_k^2 \rangle \quad (3)$$

For comparisons purposes, the fluid and particle velocity variances are being plotted side by side (see Fig. 2). The $\langle \alpha_p \rangle$ value used to plot the numerical result is the same as the experimental, as the value is being over predicted by the model. As can be seen, the vertical particle variance is being overestimated much more than the others, it is about 3 times larger than the experimental value for $\langle \alpha_p \rangle = 0.12$, however it is well estimated for $\langle \alpha_p \rangle > 0.3$. In contrast, the vertical and the horizontal components for the fluid are being underestimated independently of $\langle \alpha_p \rangle$.

The total fluctuating kinetic energy can be calculated using Eq. 4, where it takes into account all components (x, y, z) to figure out the energy. However, it is known that the horizontal components (x, y) are similar, then, only the x contribution was used to estimate the kinetic energy.

$$E_k = \frac{1}{2} (\alpha_{k,x} + \alpha_{k,y} + \alpha_{k,z}) = \frac{1}{2} (2 * \alpha_{k,x} + \alpha_{k,z}) \quad (4)$$

Using Eq. 4 and results from Aguilar-Corona (2008) we can plot energy for both phases. As it is shown in Fig. 3, the energy prediction value for both phases is decreasing with increasing of $\langle \alpha_p \rangle$. For dispersed phase, the energy is being overestimated until $\langle \alpha_p \rangle = 0.3$, after that, the kinetic energy is being underestimated. The fluid phase is being underestimated independently of $\langle \alpha_p \rangle$. It is worth noticing that the kinetic energy predicted for both phases are approximately of the same order of magnitude, which is not the case for the measured results from Aguilar-Corona (2008).

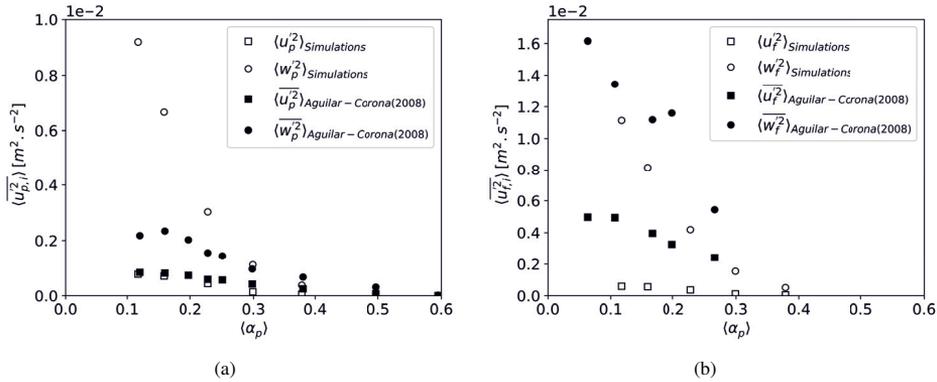


Figure 2: The horizontal and vertical contributions of the spatial mean of the velocity variance for (a) particle and (b) fluid.

Therefore, we need to improve the mathematical modeling of the phase agitation coupling by a modeling of the fluid turbulence production due to the particle motion. It is worth noticing that the Euler-Euler approach used in this paper, the fluid turbulence modeling is made in the frame of very small particles and small particle Reynolds numbers (Vermorel et al., 2003). Hence, the fluid pseudo-turbulence production in the wakes of the particles is neglected but it need to be accounted for the fluid-particle fluidized beds modelling.

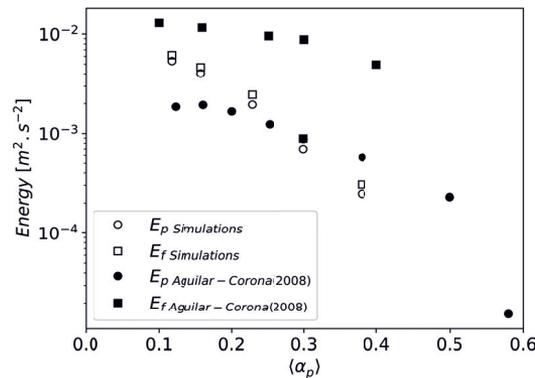


Figure 3: Comparison between particle and fluid fluctuating kinetic energy.

6. Conclusion

Numerical simulations employing the Euler-Euler approach were performed in order to investigate the bias and to understand better the model applied in a liquid-solid fluidized bed. We ran simulations in a completely 3-D fluidized bed that Aguilar-Corona (2008) had examined experimentally. Overall, the spatial mean of the velocity variance is being underestimated by the model, except for vertical particle component, which is being over-predicted for dense regime. It is interesting to note that the agitation of the two phases reduces as the solid concentration in the bed increases and the model was able to catch the flow anisotropy, reflecting the trend seen in the experiments. However, the estimation of the kinetic energy for both phase is roughly of the same order of magnitude,

which is not the case for the measured values from Aguilar-Corona (2008). As a result, we should propose a model capable to compute the pseudo-turbulence production, which ought to increase the fluid velocity fluctuation prediction without causing any influence on the particle fluctuating motion. In other words, the model may predict a higher value to fluid velocity fluctuation and consequently approaching the experimental results.

7. Acknowledgement

This work was performed using HPC resources from GENCI-CINES (Grant 2020- A0082B06938) and from CALMIP (Grant 2020-P1132). This work has been supported by the ANR–MUSCATS project, grant ANR-19-CE05-0010-02 of the French National Agency of Research (ANR).

References

- A. Aguilar-Corona, 2008. Agitation des particules dans un lit fluidisé liquide. Etude expérimentale. Ph.d thesis, Institut National Polytechnique de Toulouse, Toulouse, France.
- F. Gevrin, O. Masbernat, O. Simonin, 2008. Granular pressure and particle velocity fluctuations prediction in liquid fluidized beds. *Chemical Engineering Science* 63 (9), 2450–2464.
- A. Gobin, H. Neau, O. Simonin, J. R. Llinas, V. Reiling, J. L. Sélo, 2003. Fluid Dynamic Numerical Simulation of a Gas Phase Polymerization Reactor. *International Journal for Numerical Methods in Fluids* 43, 1199–1220.
- Z. Hamidouche, E. Masi, P. Fede, R. Ansart, H. Neau, M. Hemati, O. Simonin, 2018. Numerical Simulation of Multi-phase Reactive Flows. In: *Advances in Chemical Engineering*. Vol. 52. Academic Press, pp. 51–124.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0065237718300097>
- M. S. Khan, G. M. Evans, Z. Peng, E. Doroodchi, B. Moghtaderi, J. B. Joshi, S. Mitra, 2017. Expansion behaviour of a binary solid-liquid fluidised bed with different solid mass ratio. *Advanced Powder Technology* 28 (12), 3111–3129.
URL <https://doi.org/10.1016/j.apt.2017.09.009>
- D. Kunii, O. Levenspiel, 1994. *Fluidization Engineering*, 2nd Edition. Butterworth-Heinemann.
- H. Neau, M. Pigou, P. Fede, R. Ansart, C. Baudry, N. Méricoux, J. Laviéville, Y. Fournier, N. Renon, O. Simonin, 2020. Massively parallel numerical simulation using up to 36,000 CPU cores of an industrial-scale polydispersed reactive pressurized fluidized bed with a mesh of one billion cells. *Powder Technology* 366, 906–924.
URL <https://doi.org/10.1016/j.powtec.2020.03.010>
- G. S. Nirmala, L. Muruganandam, 2019. Hydrodynamics in a Liquid Solid Circulating Fluidized Bed—A Review. *Journal of Chemical Reviews* 1 (2), 114–129.
- P. Ren, W. Li, K. Yu, 2021. Computational fluid dynamics simulation of adsorption process in a liquid-solids fluidized bed. *Journal of Environmental Chemical Engineering* 9 (4), 105428.
URL <https://doi.org/10.1016/j.jece.2021.105428>
- J. F. Richardson, W. N. Zaki, 1954. The sedimentation of a suspension of uniform spheres under conditions of viscous flow. *Chemical Engineering Science* 3 (2), 65–73.
- S. Tenneti, R. Garg, S. Subramaniam, 2011. Drag law for monodisperse gas-solid systems using particle-resolved direct numerical simulation of flow past fixed assemblies of spheres. *International Journal of Multiphase Flow* 37 (9), 1072–1092.
URL <http://dx.doi.org/10.1016/j.ijmultiphaseflow.2011.05.010>
- O. Vermorel, B. Bedat, O. Simonin, T. Poinot, aug 2003. Numerical study and modelling of turbulence modulation in a particle laden slab flow. *Journal of Turbulence* 4.
- S. Wang, J. Sun, Q. Yang, Y. Zhao, J. Gao, Y. Liu, 2014. Numerical simulation of flow behavior of particles in an inverse liquid-solid fluidized bed. *Powder Technology* 261, 14–21.
URL <http://dx.doi.org/10.1016/j.powtec.2014.04.017>
- C. Y. Wen, Y. H. Yu, 1966. *Mechanics of fluidization*. Chemical Engineering Progress Symposium Series, 62, 100 - 110.
- J. G. Yates, P. Lettieri, 2016. *Fluidized-Bed Reactors: Processes and Operating Conditions*. Vol. 26. Springer, London.
- K. Zhang, Y. Guan, X. Yao, Y. Li, X. Fan, S. Brandani, feb 2013. Two- and three-dimensional computational studies of liquid–solid fluidization. *Powder Technology* 235, 180–191.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0032591012006651>

Non-process elements in kraft bleach plants: adsorption equilibrium aiming at reducing water consumption

Ana M. Sousa^{a*}, Carolina T. Pinheiro^a, José Granjo^b, Licínio M. Gando-Ferreira^a, Lino O. Santos^a, Margarida M. J. Quina^a

^a*University of Coimbra, CIEPQPF, Department of Chemical Engineering, Rua Silvío Lima, Pólo II – Pinhal de Marrocos, 3030-790 Coimbra, Portugal*

^b*CERENA, DEQ, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal.*

*Corresponding author: anasousa@eq.uc.pt

Abstract

The pulp industry is a high consumer of water, and efforts to reduce water consumption have led to the build-up of non-process elements (NPE), mainly in the bleaching area. These species can jeopardize the plant operability and reduce the product quality. This work aimed to develop a simulation tool to describe the concentration of NPE in pulp suspensions reliably so that it can be used to optimize bleaching sequences. The simulator is based on the Donnan model with parameters estimated from adsorption equilibrium experiments. The model predictions were compared with experimental results, and a good predictive capacity of the model was proved.

Keywords: pulp bleaching, non-process elements, adsorption, Donnan equilibrium.

1. Introduction

Over the last decades, the pulp and paper production industry has made a significant effort to reduce freshwater consumption in the process. Indeed, water consumption in pulp bleaching plants has decreased from 35-50 m³/t AD (air-dried) in the 1980s to ca. 25 m³/t AD in modern plants (Huber et al., 2014). However, further reductions have been explored to make pulp and paper production more sustainable, especially when the mills are located in water-scarce areas or if groundwater is used (Tewari et al., 2009).

Bleaching is the set of operations that follow the pulp cooking process and precede paper production. Here, the residual lignin is oxidized and removed from the pulp along with some chromophore groups responsible for the brown colour of the pulp. In most cases, the chromophore groups are functional groups of the carboxylic type (-COOH) or of the phenolic type (-PhOH) (Bygrave, 1997). Depending on the pH in each stream, these functional groups can act as active sites for the adsorption of metallic elements in the pulp (Rudie and Hart, 2006). The bleaching section comprises several bleaching and washing stages. In the bleaching stages, the pulp reacts with different oxidizing agents to degrade the chromophore groups and then proceeds to one or more washing steps, in which the dissolved compounds (organic and inorganic) are removed. For this reason, this section turns out to be the pulp mill zone that requires the greatest amount of water, generating effluent with high pollutant loads (Tewari et al., 2009).

The optimization of pulp bleaching has involved using reagents with lower environmental risks, but that guarantee high efficiency in removing chromophore groups, requiring

smaller water quantities to wash the pulp. In addition to these measures, the pulp washing is operated in countercurrent, limiting the amount of fresh water introduced into the process and consequently reducing the amount of effluent generated (Huber et al., 2014). This closure in the water circuits led to the build-up of some organic compounds (resulting from the removal of chromophore groups and lignin) and inorganic species. Most of the inorganic species accumulated in the process are denoted as non-process elements (NPE). These are metallic elements not directly involved in the pulp production process but rather introduced as wood constituents (Ulmgren, 1997). The main NPE are calcium, sodium, magnesium, and chlorine (Andrade et al., 2007). The increase in the concentration of chloride in the system, for example, can lead to corrosion problems in the equipment, limiting the possibility of burning the filtrates in the recovery boiler. On the other hand, the accumulation of the remaining metallic elements is associated with scaling problems, fouling, and loss of bleaching capacity (Stratton et al., 2004). Furthermore, metallic cations tend to adsorb on the fiber surface, binding to the functional groups in the pulp, degrading the fiber quality for downstream operations. WinGEMS 5.3 is a software developed to the pulp production processes. However, some limitations have been pointed out in terms of its ability to predict the sorption and scaling phenomena, generating results with large deviations from the industrial data. WinGEMS includes the Langmuir sorption model (Eq. 1, Langmuir, 1918), to predict the amount of sorbed components as a function of the concentration of the component in the solution phase (C , g/L), and the active sites in the pulp in an oven-dried (OD) basis (A , kg/t OD). The model also includes a parameter, B (L/g), that is specific for each NPE.

$$\text{Concentration of sorbed component} = \frac{ABC}{1 + BC} \quad (1)$$

WinGEMS 5.3 contains only the A and B parameters for the Na^+ , Ca^{2+} , and K^+ ions and does not include its variation with pH, although the amount of active sites is clearly dependent on this variable. These assumptions limit and weaken the predictive capacity of the simulator. Due to the lack of accuracy in predicting the adsorption phenomenon and the inability to include very important elements such as magnesium, it is imperative to implement alternative models that solve these limitations, such as the Donnan equilibrium model. This work aimed to develop a computer simulation tool, in MATLAB, to be further integrated with WinGEMS to improve the predictive capacity of the software, allowing to optimize the amount of water introduced in the washing steps, while restricting the formation of precipitates and the adsorption on the fiber.

2. Model development and implementation

The Donnan equilibrium model was initially developed to govern the uneven distribution of ions between two sides of a membrane (Donnan and Harris, 1911) and later applied to pulp suspensions to simulate the adsorption equilibrium (Towers and Scallan, 1996). The pulp suspension is described as two compartmentalized phases called fiber and solution. The fiber phase (F) considers the fiber (including the functional groups) and the water of the internal solution with the dissolved ions bounded to the fiber. The solution phase (S) contains the external water and the ions dissolved in it. The Donnan model considers the existence of two types of functional groups in the fiber (AH and BH , in its protonated form). Table 1 lists all the equations of the model. Depending on the pH, these groups can dissociate, generating negative charges (A^- and B^- , check Eq. 2a and 2b). The dissociation constants (K_{AH} and K_{BH}) of these functional groups are indicated according to Eq. 3a and 3b, considering the molalities (m_i^F) of the various species (i) involved in

the reactions. To calculate the Donnan equilibrium, electroneutrality is assumed in each phase of the system, that is, the net ionic charge is zero. This results in Eq. 4 and 5, where z corresponds to the valence of the ion j or the functional group k . Electroneutrality equations also take into account the ions resulting from the water dissociation in each phase, whose concentrations are calculated according to the dissociation constant ($K_W = 10^{-14}$, at 25 °C) (Eq. 6). The fiber phase is considered a homogeneous phase separated from the external solution. The presence of fixed ionic species, such as carboxylic and other functional groups in fibers induces Donnan equilibrium. Thus, the mobile metal ion concentrations inside and outside the fiber phase are not equal. The Donnan equilibrium expression is described by Eq. 7, where m_j corresponds to the molality of the ion j , z_j the valence of the ion j , and λ the Donnan constant. It is necessary to guarantee the conservation of the functional groups, both in their protonated and deprotonated state, therefore mass balances (Eq. 8a and 8b) to the total amount of groups (Q_A and Q_B) are introduced, where M^F relates to the mass of water in the internal solution, m_{AH}^F and m_{BH}^F , the molalities of the protonated groups, and $m_{A^-}^F$ and $m_{B^-}^F$, the molalities of the deprotonated groups. Finally, it is necessary to close the mass balances for each ionic species in the whole system. Thus, Eq. 9 expresses the balance for each cationic species (C_j) and Eq. 10 for each anionic species (A_j). M^S represents the mass of water in the external solution. Solving the system of Eqs. in Table 1, the λ parameter can be calculated as an implicit function of the solution pH. Commonly, as pH increases, λ also increases.

Table 1. Equations used to simulate the Donnan equilibrium model in a fiber suspension.

$AH \Leftrightarrow A^- + H^+$	(2a)	$K_W = m_{H^+}^F \cdot m_{OH^-}^F = m_{H^+}^S \cdot m_{OH^-}^S$	(6)
$BH \Leftrightarrow B^- + H^+$	(2b)	$\left(\frac{m_j^F}{m_j^S}\right)^{\frac{1}{z_j}} = \lambda$	(7)
$K_{AH} \Leftrightarrow \frac{[A^-][H^+]}{[AH]} = \frac{(m_{A^-}^F)(m_{H^+}^F)}{(m_{AH}^F)}$	(3a)	$\frac{Q_A}{M^F} = m_{AH}^F + m_{A^-}^F$	(8a)
$K_{BH} \Leftrightarrow \frac{[B^-][H^+]}{[BH]} = \frac{(m_{B^-}^F)(m_{H^+}^F)}{(m_{BH}^F)}$	(3b)	$\frac{Q_B}{M^F} = m_{BH}^F + m_{B^-}^F$	(8b)
$\sum_j z_j m_j^F + \sum_k z_k m_k^F = 0$	(4)	$C_j = M^S m_j^S + M^F m_j^F$	(9)
$\sum_j z_j m_j^S = 0$	(5)	$A_j = M^S m_j^S + M^F m_j^F$	(10)

3. Materials and methods

Pulp samples were collected at two points on the bleaching line of a Portuguese pulp and paper mill: p-O2 (after oxygen delignification) and D2 (end of a bleaching sequence). The analysis followed TAPPI procedures. The pulps were acid washed (HCl, until pulp pH of 1.5-2) before analysis. The concentration of functional groups and their dissociation constants were determined by potentiometric titration. The quantification of NPE in the pulp was determined by flame atomic absorption spectrometry (AAS), after acid digestion. The sorption tests were conducted at 60 °C, a pulp consistency of 5% wt., and varying the pH of the experiment using HCl or NaOH solutions. Single-component tests were carried out for Ca, Na, and Mg and multi-component tests were also carried out. The concentrations tested were close to 500, 1000, and 100 mg/kg OD for Ca, Na, and Mg, respectively, in single-component tests. For multi-component tests, the concentrations tested were close to 3000, 30000 and 500 mg/kg OD for Ca, Na, and Mg, respectively.

The Donnan equilibrium model was implemented in MATLAB. The adsorption for the p-O2 sample was simulated in WinGEMS in two different modes: (1) a pulp stream was simulated as the models are implemented in WinGEMS; (2) the Donnan equilibrium values were imported to WinGEMS and equations were solved to calculate the distribution of NPE between the pulp and the solution.

4. Results and discussion

The characterization of the pulps by titration allowed to determine the concentration of functional groups per unit of pulp mass (mmol/kg OD), as well as the dissociation constants of the two types of functional groups. The titrations curves (results not shown) did not reveal the presence of phenolic groups for the pulps p-O2 and D2. Thus, Eq. 2b, 3b, and 8b in Table 1 were not considered in the simulations. These results are in line with what was documented by Zhang (2006), who refers only to the existence of carboxylic groups. The results are summarized in Table 2. The concentration of functional groups significantly decreased during the bleaching process. These values were used to parameterize the Donnan equilibrium simulator in MATLAB.

Table 2. Pulp characterization in functional groups.

Pulp	k_{AH}	pK_A	Q_A (mmol/kg OD)
p-O2	2.51×10^{-4}	3.6	99
D2	1.78×10^{-4}	3.75	47

For the single-component adsorption tests, it was found that generally, the Donnan model fits well with the experimental data for Ca and Mg elements (Fig. 1 a, b, e, and f). For Na, the fit quality is lower, which can be due to technical difficulties in the AAS determination of Na, but the results are also satisfactory (Fig. 1 c and d). For the elements Ca and Na, the prediction of the Langmuir model was calculated, using the constants initially in WinGEMS and the same concentrations used in the adsorption tests and the simulation of the Donnan equilibrium. For Ca, the values of constants A and B were 1.5 kg/t OD and 4 L/g, respectively, which for the concentration of 500 mg/kg OD resulted in the adsorption of 136.4 mg/kg OD of calcium. Under typical industrial conditions, the p-O2 pulp has an alkaline pH value, so this prediction is not suitable, compared to the results presented in Fig. 1a, where all the calcium is sorbed into the pulp, for such pH values. The D2 pulp has a pH between 3.5 and 5, so the Langmuir equation is not adequate to simulate these conditions, as the experimental essays showed that sorbed calcium varies between 200 and 400 mg/kg OD for this pH range (Fig. 1b). For Na, the constants A and B took the values of 2.43 kg/t OD and 53 L/g, respectively, which for the concentration of 1000 mg/kg OD led to the adsorption of 1764.2 mg/kg OD. This value is wrong, since it considers the adsorption of a greater amount of sodium than available, which overestimates its concentration in the fiber. Analyzing Fig. c and d, it was found that, for the conditions tested, all the sodium was sorbed into the pulp. In the case of magnesium, WinGEMS does not include constants to calculate the adsorption equilibrium, but it was found that the Donnan model allowed a good approximation for this element in each of the pulps (Fig. 1e and f).

The Donnan equilibrium simulation for the multi-component system also revealed a good predictive capacity of this model for both samples (Fig. 2). This result is significant since the pulp suspension has several NPE interacting with each other in real systems, influencing the adsorption process. It was observed a slight decrease in the sorbed amount for D2, since under these conditions, the functional groups of the pulp become saturated.

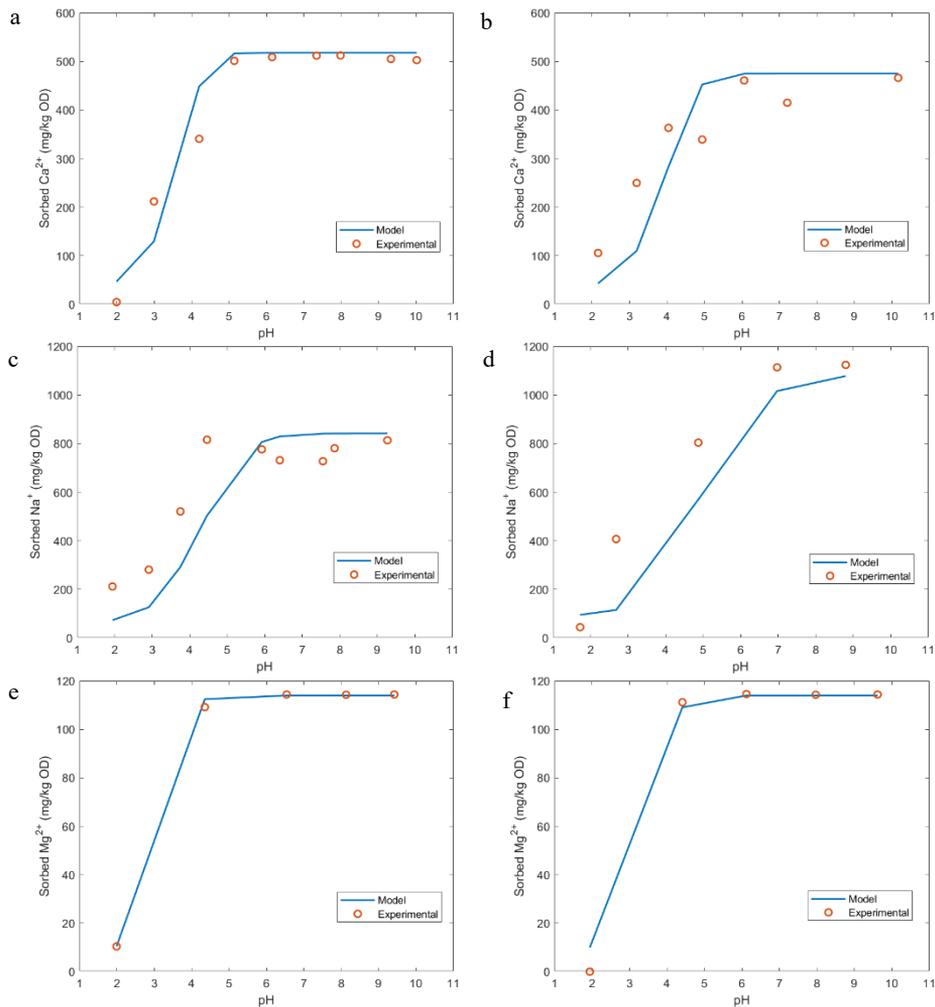


Fig. 1. Adsorption equilibrium for Ca (a and b), Na (c and d) and Mg (e and f), experimental and Donnan model for pulp p-O2 (a, c, and e) and D2 (b, d, and f).

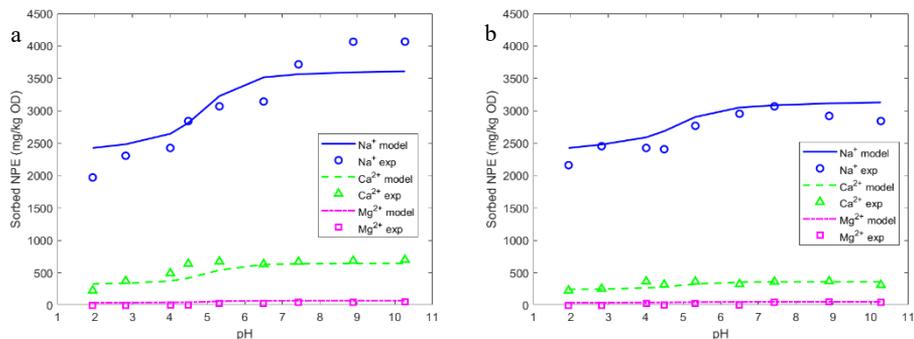


Fig. 2. Multi-component sorption essays and Donnan model: (a) – pulp p-O2 and (b) – pulp D2.

Finally, a stream containing p-O2 pulp was simulated in WinGEMS 5.3 for the experimentally tested conditions (for Ca and Na), considering a pH of 10 in two different modes. In mode (1), the simulation considered the adsorption models available in the software and the results showed that the introduced ions remained completely in the solution phase, demonstrating the limitations of the software. In mode (2), the simulation considered λ value (1.38 at pH 10) calculated in MATLAB script and Eq. 7 (Table 1) that establish the partition between the phases. In this case, the results showed that a certain fraction of these ions (Ca^{2+} and Na^+) are effectively adsorbed in the pulp p-O2. The results in Table 3 are in good agreement with the experimental data showed in Fig. 2a, at pH 10.

Table 3. WinGEMS simulations for p-O2 sample, at pH 10.

	Without Donnan model – Mode (1)		With Donnan model – Mode (2)	
	wt.%	mg/kg OD	wt.%	mg/kg OD
Ca^{2+} sorbed	0 ^a	0	0.0715 ^a	715
Na^+ sorbed	0 ^a	0	0.3692 ^a	3692
Ca^{2+} solution	0.0205 ^b	3895	0.01707 ^b	3180
Na^+ solution	0.1513 ^b	28740	0.1318 ^b	25048

a – % Suspended solids; b – % mass of liquor.

5. Conclusions

This work showed that the Langmuir model implemented in WinGEMS could not predict the adsorption phenomenon in the pulp mill, namely because the concentration of functional groups and pH varies in the different bleaching stages. The Donnan model allowed a good predictive capability for mono- and multi-component systems and it is easily implemented in WinGEMS. Implementing this model to simulate the bleaching line can be valuable for assisting in the optimization of bleaching sequences and configurations, allowing for a reduction in water consumption.

References

- A. A. Andrade, P. M. Glória, J. V. H. d'Angelo, D. O. Perissotto, R. A. Lima, 2007, *Water Sci. Technol.*, 55, 6, 199-204.
- G. Bygrave, 1997, Thermodynamic model and data on the partition of non-process elements between kraft fibres and water in a pulp suspension, MSc. from The University of British Columbia, Vancouver.
- F. M. Donnan, A. B. Harris, 1911, *J. Chem. Soc. Transactions*, 99, 1554-1577.
- P. Huber, A. Burnet, M. Petit-Conil, 2014, A review, *J. Environ. Manage.*, 141, 36-50.
- I. Langmuir, 1918, *J. Am. Chem. Soc.*, 40, 9, 1361-1403.
- A. W. Rudie, P. W. Hart, 2006, *TAPPI J.*, 5, 8, 3-15.
- S. C. Stratton, P L. Gleadow, A. P. Johnson, 2004, *Water Sci. Technol.*, 50, 3, 183-194.
- P. K. Tewari, V. S. Batra, M. Balakrishnan, 2009, *J. Environ. Manage.*, 90, 265-273.
- M. Towers, A. M. Scallan, 1996, *J. Pulp Pap. Sci.*, 22, 9, 332-337.
- P. Ulmgren, 1997, *Nord. Pulp Pap. Res. J.*, 111997, 32-41.
- D. Zhang, 2006, Characterization and enhancement of carboxyl groups in softwood kraft pulps during oxygen delignification, PhD. from Georgia Institute of Technology.

Acknowledgments:

This work was carried out under the Project inactus – innovative products and technologies from eucalyptus, Project N.º 21874 funded by Portugal 2020 through European Regional Development Fund (ERDF) in the frame of COMPETE 2020 nº246/AXIS II/2017. J. Granjo gratefully acknowledges the financial support from CERENA (strategic project FCT-UIDB/UIDB/04028/2020).

Re-design and scheduling of dairy thermal treatment processes for continuous operation

Wei-Fu Tseng^a and Sandro Macchietto^a

*^aDepartment of Chemical Engineering, Imperial College London, South Kensington Campus, London SW72AZ, UK
s.macchietto@imperial.ac.uk*

Abstract

Heat treatment processes of liquid dairy products (pasteurisation, sterilisation) are energy, water and wastes intensive. They are normally conducted in an energy integrated system of 3 plate heat exchangers (PHEs), in batch mode heating-cleaning cycles. During cleaning, the whole process is interrupted, causing loss of production. A new semi-continuous pasteurisation process is proposed here comprising a design modification (an additional plate heat exchanger and holding tube), the calculation of (time-varying) pasteurisation time and its use in a new logic for switching between heating and cleaning, and suitable scheduling of equipment in a rotation strategy. All schemes are studied using a detailed distributed dynamic model of the process. The simulation results for a typical high-temperature-short-time (HTST) case study show that the rotational scheduling of the PHEs and tubes results in a semi-continuous operation where milk production can be maintained indefinitely. Substantially higher productivity (+23% throughput), reduced energy, and total cost (-47%) are achieved relative to the traditional batch mode operation.

Keywords: interaction between design and scheduling; dynamic modelling; food processes; heat treatment; fouling and cleaning

1. Introduction and background

The heat treatment of milk, and dairy fluid in general, is typically conducted in an energy integrated system of 3 plate heat exchangers (PHEs) and connecting pipework (Fig. 1a). Cold milk is fed to a preheater (Regeneration) PHE, then a main Heating PHE which uses steam or water utilities. A Holding Tube ensures milk is kept at a high temperature for a sufficiently long time, dictated by preservation requirements. For example, in the high-temperature-short-time (HTST) treatment, milk must be held above 72-75°C for 15-20 s. From the holding tube, the hot milk returns to the preheater for energy recovery against the incoming cold milk, then on to the final Cooler PHE. Fouling requires stopping the process and carrying out a Cleaning in Place (CIP) of the entire system, according to the State Transition Network in Fig.1b. The process is conducted in batch mode, with no production during cleaning. Many studies have discussed the dynamic modelling, fouling deposition and CIP of single PHEs (e.g., Georgiadis et al 1998a, 1998b; Georgiadis & Macchietto, 2000; Manik et al, 2004; Jun & Puri, 2005; Guan & Macchietto, 2018; Sharma & Macchietto, 2021). Only some (e.g., Aguiar & Gut, 2014; Gutierrez et al, 2014; Zhu et al., 2020) integrated individual models into the dynamic simulation of a complete system. We are not aware of papers on the continuous operation of such systems in the presence of fouling. Here, a new process is proposed for continuous milk production, via a design modification (using an additional PHE and holding tube, Fig. 2), novel phase switching logic between heating and cleaning, the flexible use of PHEs and holding tubes in rotation, and suitable scheduling of all operations. For a high-temperature-short-time

(HTST) treatment, it is shown that this leads to higher production and reduced costs. A 2D-distributed dynamic model of all units in the pasteuriser unit is used, as detailed by Zhu et al, 2020. Each PHE model includes fouling kinetics, heat transfer, and deposition on each side plate of each channel and in the tube using a moving boundary for the deposit (Fig. 3). Each point along each plate and tube is characterised by (time-varying) temperature, shear rate, heat transfer rate, fouling deposition, and deposit depth. All modelling and simulations were done using gPROMS 7.0.7 (Process Systems Enterprise).

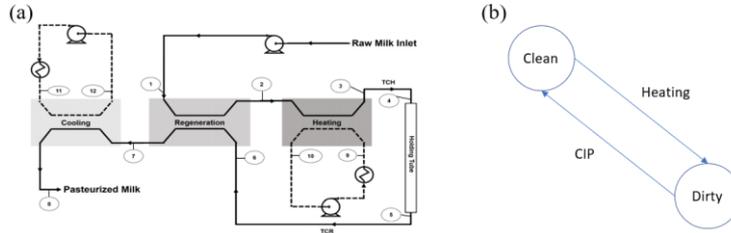


Figure 1. (a) Pasteuriser unit (Gutierrez et al, 2014; Zhu et al, 2020), TCH and TCR are connecting tubes. (b) State transition network - overall system (circles = unit states, lines = tasks).

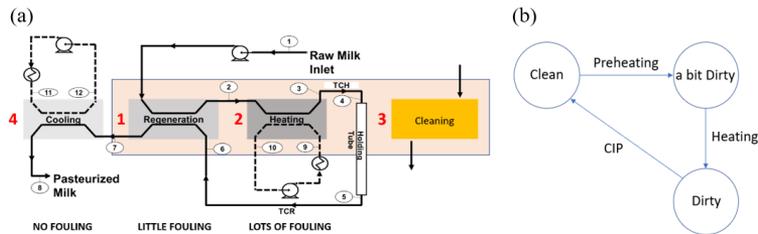


Figure 2. (a) Modified pasteuriser unit with an additional PHE (here being cleaned). Not shown are an additional holding tube and pipework for flexibly re-connecting units. (b) State transition network for individual PHEs 1-3 and holding tubes (circles = unit states, lines = tasks).

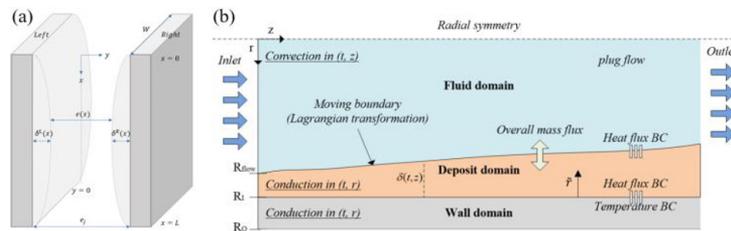


Figure 3. (a) Schematic diagram of a single PHE channel subject to fouling. (b) Cross-sectional diagram of the moving boundary model for a tube (Zhu et al, 2020).

2. New switching logic using pasteurisation time

A typical process requirement for preservation is to hold the fluid above a certain temperature for a minimum time. In the HTST treatment of milk, this is 72-75°C for 15-20 s. Such milk temperature may be reached in the end of the Regeneration PHE, Heating PHE, holding tube, and initial channels on return to the Regeneration PHE. Temperatures change due to fouling, and direct monitoring of the HTST constraint during heating is not straightforward. Following temperatures at various inlet/exit points gives a partial idea. An alternative is to estimate the amount of deposit on the plates, using simulations models. An often-used criterion (Georgiadis & Macchietto, 2000) is to switch from

heating to cleaning when a critical fouling (a deposit mass density of 16 g/m²) is reached on any plate in a PHE. These are all indirect indicators. A direct way of estimating the actual pasteurisation time is proposed as follows: First, at any given time the temperature profile of milk is established (from the simulation model) along its path from the entrance to the preheater, through the channels in the various PHEs, along the holding tube, to the exit from the cooler (points ①-⑥ in Fig. 2a). This is visualized in a novel diagram (Fig. 4a). Such a diagram can be used to also depict the local distribution at a given time of other variables, such as deposition rates (Fig 4b), heat transfer rates, shear rates, etc. The pasteurisation time (i.e., residence time of milk above the minimum temperature) in each PHE channel, tube, and overall can be calculated from the diagram graphically, and in the model as follows. From Fig. 3a, the gap, e_x , and volume, V_c , in each PHE channel are:

$$e_x(x) = e_j - \delta^L(x) - \delta^R(x) \quad (1)$$

$$V_c = \int_0^L W \cdot e(x) dx \quad (2)$$

where x is a length coordinate, e_j is the channel gap under clean condition, δ^L and δ^R are the deposit thickness on left and right plates, respectively, and W is the channel width. The corresponding equations for the tube model (Fig. 3b) are:

$$R_{flow}(z) = R_i - \delta(z) \quad (3)$$

$$V_{TB} = \int_0^L \pi \cdot R_{flow}^2 dz \quad (4)$$

where R_{flow} is the real radius of flowing fluid, R_i is the inner radius of the tube, δ is the deposit thickness, and V_{TB} is the tube volume. The volume data are used to calculate the pasteurisation time in each PHE channel, τ_c , and tube, τ_{TB} , given respectively by:

$$\tau_c = \frac{V_c}{\dot{V}} \quad (5)$$

$$\tau_{TB} = \frac{V_{TB}}{\dot{V}} \quad (6)$$

where \dot{V} is the volumetric flowrate of milk. Only those locations where the temperature in a channel or tube is higher than the minimum temperature (here, 72°C) are considered in the pasteurisation time calculations. The process fluid can be the hot fluid and/or cold fluid (in the heating PHE milk is the cold fluid, while in the regeneration PHE, milk is on both sides). The overall pasteurisation time, at any instant, is the sum of the pasteurisation times in all channels of all PHEs and in tube. In a clean system, it will start being longer than the minimum HTST requirement, but decreases over time due to fouling. A new termination condition is therefore defined, based on the time-varying pasteurisation time: TC1 (overall pasteurisation time < 15s). Here, both TC1, and the max fouling criterion, TC2 (critical deposit mass density on any PHE plate >16 g/m²) are used in a combined OR logic condition to decide when to switch from heating to cleaning.

3. High-temperature-short-time (HTST) case studies

A HTST pasteurisation treatment is considered, with milk temperature > 72°C for 15 s and initial configuration in Fig. 2a (TCH and TCR are connecting tubes). Milk is fed to PHE1 (used as regenerator), then goes to PHE2 (used as a heater), to holding Tube 1,

back to PHE1 for energy recovery and to PHE4, used as a cooler. Equipment, operation, and model parameters for this typical industrial system are in Zhu et al, 2020, based on Gutierrez et al, 2014. The new PHE and holding tube have the same geometry and configuration as the original ones. Each PHE has 12 plates with flow configuration as in Fig. 5d. Fouling is included in all models, including for TCH and TCR. Inlet flowrates and temperatures are given in Table 1 for two settings. Fig. 4 shows the temperature and deposition rate diagrams for both settings at time $t = 3.8$ h, after starting from a completely clean system. The x axis in the diagrams in Fig. 4 and Fig. 5 is the distance along the milk path in the pasteuriser unit (points ①-⑩ in Fig. 2). The ordinate shows the milk temperature (Fig. 4a) and deposition rate (Fig. 4b) at each point along this path. At the start time $t=0$ (clean conditions) the pasteurisation time for Setting 1 is 16.1 s, above the minimum HTST requirement. Due to fouling, after 3.8 h of operation, the pasteurisation time becomes 15s, thus triggering switching condition TC1. Significant fouling also occurs in the holding tube and at the higher temperature locations in PHE1 and PHE2. With Setting 2, a smaller hot utility flowrate results in a lower top milk temperature and milder fouling, however the pasteurisation time is only 8.8 s even at initial clean conditions, failing to meet the HTST requirement. A similar diagram of the heat flux from the processing fluid (negative values denoting heat transferred to the fluid) to the left-side and right-side plates in each PHE channel is shown in Fig. 5a and 5b, respectively, for Setting 1. Due to the alternance of co-current and counter-current flows in adjoining PHE channels (Fig. 5d), the two are different. A clear, strong relationship is evident between the relative flow directions on the two sides of a plate and heat flux pattern. In co-current flow plates, the heat flux shows a zigzag pattern, while in counter-current flow plates, it shows a piecewise continuous line. Adding up the heat flux on the left and right plates of each channel gives the total heat flux from the processing fluid flowing there (Fig. 5c, with adiabatic tubes). The heat transfer picture is clearly very complex. These strong interactions between equipment, operation and fouling make a detailed dynamic model essential to estimate the actual pasteurisation time and how it changes over time.

Table 1. Inlet stream flowrates (F) and temperatures (T) for two settings (based on Aguiar & Gut, 2014), and resulting pasteurisation times when all units are clean (at time $t = 0$).

Setting	F_{milk} (L/h)	F_{heat} (L/h)	F_{cool} (L/h)	Inlet Milk temperature (°C)	Inlet Hot utility temperature (°C)	Inlet cool utility temperature (°C)	Initial pasteurisation time (s)	Highest milk temperature (°C)	Severity of fouling
1	20	60	60	19.4	90	10.1	16.1	84.9	severe
2	20	20	20	19.4	90	10.1	8.8	76.1	mild

4. Semi-continuous operation with the additional equipment

In a traditional operation, upon detection of excessive fouling after a heating period, the whole unit is cleaned. Here, after the above initial Phase 1 of 3.8 h, only the main heating exchanger (PHE2) and holding tube (Tube 1), heavily fouled, are taken out for cleaning. The equipment is reconfigured in Phase 2 as follows: PHE3 (cleaned), is put in service as a Preheater; PHE1 (mildly fouled from preheater use) becomes the main Heater; The second (clean) tube (Tube 2) is used as holding tube; PHE4 continues its service as a Cooler. When PHE1 or Tube 2 violate the fouling or pasteurisation time conditions, Phase 2 ends, they are cleaned, PHE2 becomes the Preheater and PHE3 the main Heater, with Tube 1 used again. The equipment is operated in a rotation sequence of these three Phases according to the schedule in Fig. 6a. The termination criterion used for CIP is TC3 (residual deposit mass < 0.01 g) on each plate for PHEs, and overall, for tubes. In Phase

1, the CIP time for Tube 1 is 52.7 minutes, leaving negligible deposit, that for PHE2 is 24.8 minutes, leaving a small residual deposit. As the heater, previously used as a regenerator, does not start clean, Phase 2 is shortened to 3.44 h. Phase 3 lasts 3.42 h before a termination condition is triggered (in all cases, TC1). PHEs 1-3 rotate between three tasks (preheating, heating, CIP) according to the STN in Fig. 2b, tubes switch between two tasks (holding, CIP). The three-phase cycle takes 10.7 hours with the schedule in Fig. 6a. The deposit on each plate of each PHE during the full cycle is shown in Fig. 6b-d.

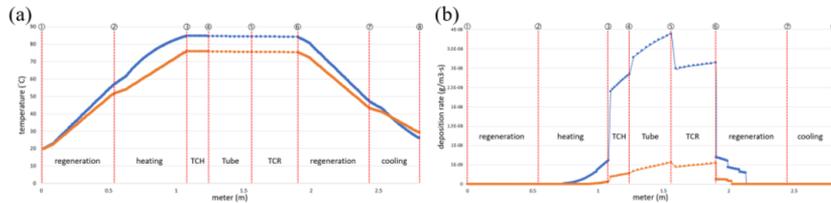


Figure 4. (a) Overall temperature diagram and (b) Total deposition rate diagram for at $t = 3.8h$. Setting 1 denoted as the blue line; Setting 2 denoted as the orange line.

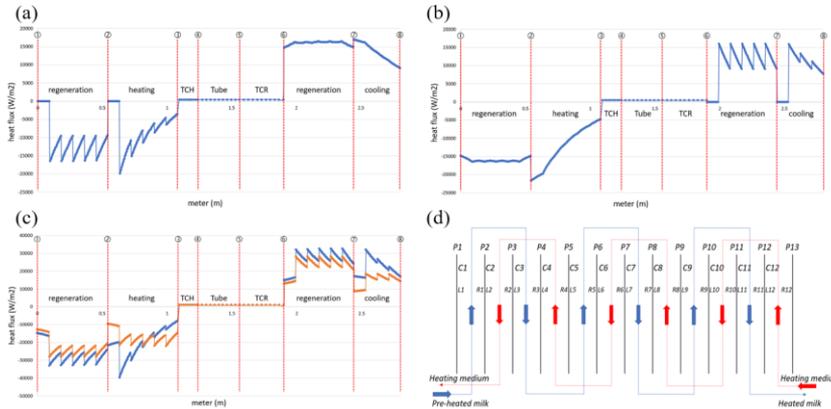


Figure 5. Heat flux from processing fluid, at $t = 3.8 h$, to (a) left plate and (b) right plate; (c) Total heat flux - blue for Setting 1, orange for Setting 2 (d) Flow patterns in the PHE configuration.

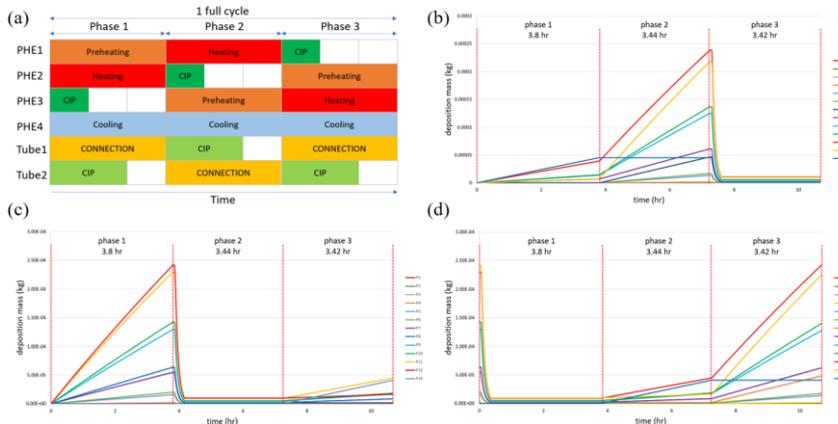


Figure 6. Setting 1 operation: (a) Schedule of 4 PHEs and 2 tubes in three phases; Deposit mass during the cycle on each plate of (b) PHE1 (c) PHE2 (d) PHE3.

After the first cycle, subsequent cycles remain stable and can be repeated indefinitely. The performance of this semi-continuous cycled operation is compared to that in batch mode, which alternates production (heating) and stoppage (cleaning) of the whole system, using the method and cost parameters in Sharma & Macchietto, 2021. For the batch operation, heating is terminated according to TC2 (deposit mass density on any PHE plate $>16 \text{ g/m}^2$) and CIP according to TC3 (residual deposit mass on each PHE plate $< 0.01 \text{ g}$). The semi-continuous mode pasteurises milk continuously with a 23% higher daily throughput, 69% lower cleaning cost, and 47% lower total operating costs.

5. Conclusions

Calculating milk pasteurisation time and its decay due to fouling, based on a detailed dynamic model of the integrated heat treatment process, allows us to monitor directly the key process constraint, and to use it in a new logic for switching between heating and cleaning. The new diagram of key properties of the process fluid along its path offers great insights into the integrated operation dynamics. Adding an exchanger and holding tube to the traditional scheme, using them flexibly in rotation, and suitable scheduling of preheating, heating, and cleaning tasks in individual equipment give additional degrees of operational freedom. Exploiting the interactions between design, dynamics, and scheduling results in a semi-continuous, cyclic operation, with uninterrupted milk production. A case study for a typical HTST heat treatment of milk shows that this semi-continuous mode delivers significantly higher throughput and lower cost than the traditional (batch) operation. As some new equipment is required, a full economic assessment depends on the trade-off between capital and operating cost for the new design. The optimisation of the semi-continuous operation, overall benefit and return on investment is beyond the scope of this paper and will be investigated in future work.

References

- Aguiar, H. F. & Gut, J. A. W. (2014) Continuous HTST pasteurization of liquid foods with plate heat exchangers: Mathematical modeling and experimental validation using a time-temperature integrator. *Journal of Food Engineering*, 123, 78-86.
- Georgiadis, M. C. & Macchietto, S. (2000) Dynamic modelling and simulation of plate heat exchangers under milk fouling. *Chemical Engineering Science*, 55(9), 1605-1619.
- Georgiadis, M.C., Rotstein, G.E. & Macchietto, S. (1998a) Optimal design and operation of heat exchangers under milk fouling. *AIChE Journal*, 44(9), 2099-2111.
- Georgiadis, M.C., Rotstein, G.E. & Macchietto, S. (1998b) Modelling and simulation of complex plate heat exchanger arrangements under milk fouling. *Computers and Chemical Engineering*, 22 (SUPPL.1)
- Guan, S. & Macchietto, S. (2018) A Novel Dynamic Model of Plate Heat Exchangers Subject to Fouling. *Computer-Aided Chemical Engineering*, 43 Part B, 1679-1684, Elsevier.
- Gutierrez, C. G. C. C., Diniz, G. N. & Gut, J. A. W. (2014) Dynamic simulation of a plate pasteurizer unit: Mathematical modeling and experimental validation. *Journal of Food Engineering*, 131, 124-134.
- Jun, S. & Puri, V. M. (2005) 3D milk fouling model of plate heat exchangers using computational fluid dynamics. *International Journal of Dairy Technology*, 58(4), 214-224.
- Manik, M. H., Bildea, C. S., Grievink, J. & Marshman, C. (2004) Modelling and optimisation of milk pasteurisation processes. *Computer Aided Chemical Engineering*, 955-960.
- Sharma, A. & Macchietto, S. (2021) Fouling and cleaning of plate heat exchangers: Dairy application. *Food and Bioproducts Processing*, 126, 32-41.
- Zhu, M., Santamaria, F. L. & Macchietto, S. (2020) A General Dynamic Model of a Complete Milk Pasteuriser Unit Subject to Fouling. *Computer Aided Chemical Engineering*, 48 Part A, 247-252, Elsevier.

Xflow modelling for investigation of fluid structure interaction of artificial reef: application to burial effect

Baptiste Oudon^a; Chanez Belaidi^b; Jessica Salaun^c; Arnaud Coutu^{a*}

a Institut Polytechnique UniLaSalle, Université d'Artois, GéoLab, 19 Rue Pierre Waguet, BP 30313, 60026 Beauvais, France

b Institut Galilée, 99 Av. Jean Baptiste Clément, 93430 Villetaneuse

c Atlantique Landes Récifs, 1 Rue du Pitey, 40130 Capbreton

**Arnaud.coutu@unilasalle.fr*

Abstract

Artificial reefs are presented as a possible response to the shrinking biological resource in coastal strip. They are intended to protect, regenerate, concentrate and enhance populations of marine organisms (Miller *et al.*, 2007). The fluid-structure interaction generates changes in the flow regime, resulting in the formation of basins around the reefs by vortices above the structure (Sumer *et al.*, 2001). The mean velocity correlated with the turbulent kinetic energy allows to study the sediment transport process in turbulent flow (Mazzuoli *et al.*, 2020). Modelling of fluid-structure seabed interaction (FSSI) was realized on the software Xflow 2020. The evolution of the burial depth leads to an increase in the turbulent kinetic energy. This evolution correlates with the mean velocity and shows an evolution of the sediment transport in time. The dynamics of a saturated sand was supposed similar to a pseudo viscous fluid (Chávez-Modena *et al.*, 2020; Ouda and Toorman, 2019; Zhou *et al.*, 2014). A two phases flow modelling with a sand defined by the Herschel-Buckley viscous model was realized. The result showed the same transport processes explained by (Mazzuoli *et al.*, 2020). The innovation of this article is to highlight turbulence as the principal reason of reef burial. A predictive method to characterize structure burial during time was showed by multiphase modelling.

Keywords: Fluid structure seabed interaction, Herschel Buckley, Lattice Boltzmann Method, Particle based tracking, Sediment transport.

1. Introduction

Marine biodiversity is weakened due to climate change leading to an increase in extreme events. Between the years 1982 and 2016, heat waves doubled in frequency and became more intense and widespread (Abram *et al.*, 2019). This type of event causes degradation and loss of habitat such as coral reefs (Ainsworth *et al.*, 2020). Coastal areas are particularly affected by this fragilization due to human overexploitation, ocean acidification, increased salinity, and rising sea levels (Abram *et al.*, 2019). Marine Protected Areas (MPAs) and artificial reefs have shown their usefulness in marine ecosystem recovery (Rees *et al.*, 2020). In the last 3 decades, a large number of artificial reefs have been immersed in coastal areas around the world, including Japan, America and Europe (Santos and Monteiro., 1997; Kim *et al.*, 2001; Jorgensen *et al.*, 2009).

On another side, marine scour is an erosion phenomenon created by the formation of turbulence during fluid-structure interaction. These turbulences cause the displacement of sediments above the structure (Sumer *et al.*, 2001). Several studies have identified different mechanisms and characteristics of scouring around wrecks, large rocks or artificial structures (Werner *et al.*, 1980; McNinch *et al.*, 2001; Quinn., 2006). Important factors influencing

scour processes include object structure and orientation, bathymetry, and seabed geomorphology (Quinn., 2006). However, these reefs get buried in the sand over time. Very few works were conducted on the characterization of artificial reef burial. However, many numerical models of pipeline burial have been studied, using different approaches such as the impermeable wall method (Fuhrman *et al.*, 2014) as well as the study of soil liquefaction (Ye *et al.*, 2013; Li *et al.*, 2018). The purpose of this study is to determine if the main cause of the burial of artificial reefs is scouring and if it is possible to estimate artificial reefs lifetime with numerical modelling. For this purpose, two types of Fluid Structure Seabed Interaction modelling (FSSI) have been performed: a modelling of FSSI of the seabed geomorphology modeled in 3D for the years 2019, 2020 and 2021 to determine the reasons of the changing geomorphology through computational fluid dynamics (CFD) and a second modelling to validate the prediction tool for FSSI modelling considering the behavior of the surface part of the saturated sand as a pseudo viscous fluid (Ouda and Toorman., 2019; Chávez-Modena *et al.*, 2020; Zhou *et al.*, 2014).

2. Environment

Atlantic Land Reef is an association based in the South-West of France. This association is installed since 1999. Three different types of artificial reefs are implanted on the Landes coast. During the years 2019, 2020 and 2021, depth data around these reefs were collected by divers. With these data, a 3D model of 20x20 meters of the geomorphology around the reef was made (**Figure 1**). The burial modelling was done on two types of these reefs with different structures, named Babel and Typi. The modified geomorphology has a bowl shape that grows from 2019 to 2021 and had not the same shape for the 2 artificial reefs studied. All six bowls were modelled in 3D: one 3D modelling was done for each year and for each reef studied. In this paper the results are focus on the Babel reef illustrated in **Figure 1**.

3. Simulation of the interaction fluid structure on the bow

The software Xflow®2020 was used for this study. It resolves the Lattice-Boltzmann Method (LBM) inspired from the Kinetic theory. The turbulence near wall was calculated by the wall function called Wall-Modeled Large Eddy Simulation (WMLES). The dynamics of sediments in turbulent environments were studied by (Mazzuoli *et al.*, 2020). The results obtained from this study could not be analytically compared with the study of Mazzuoli *et al.* (2020) because the sands studied had not exactly the same composition but the physical characteristics of these sands were similar enough to compare the transport process. The threshold of sediment transport in the ocean environment was evaluated by the shield parameter (Li *et al.*, 2020), and this parameter was compared with the turbulence intensity by Mazzuoli *et al.* (2020). The turbulence intensity at ground was displayed in percentage and was calculated using the equation (1):

$$I = \frac{\sqrt{\frac{2}{3}K}}{U} \quad (1)$$

Where U is the mean velocity and k is the turbulence kinetic energy. Numerical simulations performed on the bowls have shown strong turbulence intensity above reefs. The proportion of ground surface with a turbulence intensity of 40 % increased with time as the bowl gets larger (**Figure 1 (a),(b),(c)**). The maximum turbulence intensity also increased over time, representing 42 % in 2019 and 54 % in 2021. The turbulence intensity could be correlated with the water mean velocity at ground level. All the proportion of ground surface with a turbulence intensity higher than 40 % corresponded to a current speed of 0.03 m.s⁻¹. The equation of the turbulent intensity has shown that the maximum turbulence intensity, in correlation with the mean velocity, indicated an increase of turbulent kinetic energy as the bowl was increasing. K went from 2.1 J for the 40 % intensity zones to 11 J for the maximum turbulence intensity. Mazzuoli *et al.* (2020) has shown that a high turbulent kinetic energy correlated with a very low velocity (< 0.05 m. s⁻¹) creates a low sediment transport. Thus, sand grains roll and slide on the ground.

The 20 % turbulent intensity zones near the reef (**Figure 1**, green zone) were concerned by a 1 m.s^{-1} velocity. The turbulent kinetic energy for these areas was therefore 6 J, which represented a non negligible energy (Mazzuoli *et al.*, 2020). The bowl was deepening over the years, resulting in an increasing slope in these areas. The increasing velocity for a turbulent kinetic energy of 6 J resulted in more sediment transport for 2019 but as the bowl increased, the scour reduced for the region.

The regions away from the reef were concerned by a turbulence fluctuation considering space and time. For 2021 the turbulence intensity of regions away from the reef was constant and the geomorphology of these areas was deeper than 2020 which was deeper than 2019 as well as represented in **Figure 1 (a),(b),(c)**. The mean velocity was 0.07 m.s^{-1} for a turbulent kinetic energy of 9 J. This data was higher for regions away from the reef than near the reef which suggested a growing sediment transport in these areas. The transport process could explain a slow scouring above reefs which was occurring over several years. This modelling confirmed our hypothesis that turbulence is possibly responsible for the increase of the bowls and therefore probably plays a key role in for burial of artificial reefs.

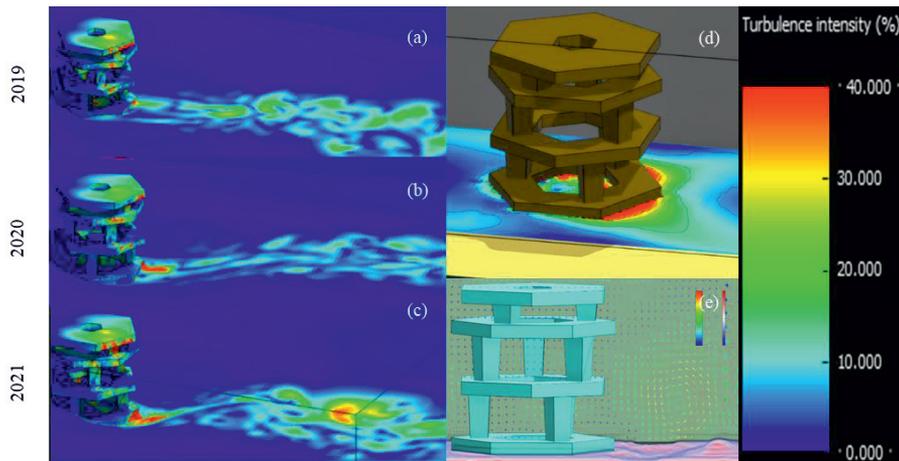


Figure 1: (a-c) Evolution of the turbulence intensity during the year 2019 to 2021, (d) Turbulence intensity at the ground, (e) Interaction turbulence with seabed

4. Multiphase modelling

A multiphase modelling was conducted to determine if the turbulence energy is responsible for burial of artificial reefs. Two types of multiphase modelling were done to validate the transport process observed in experimental measures.

4.1 Rheology of saturated sand defined as a viscous fluid

The rheology data was obtained from literature (Zhou *et al.*, 2014) and defined the behavior of a saturated sand as a viscous fluid. This fluid follows the viscosity model of a power law. The rheological data was determined following a triaxial test which possessed a soil consolidation pressure of 100 kPa. These consolidation pressures values were chosen as the weight of the Babel reef is 11094 kg. Thus, the pressure applied by the reef on the sand corresponded to a force of 109 kPa. The rheological data could be calculated with equations (2) and (3).

$$\tau = \frac{\sigma}{2} \quad \epsilon = \frac{\sigma}{E} \quad (2) \quad (3)$$

Where τ is the shear stress, σ is the stress with $\sigma = 42$ kPa. ε is the strain which can be defined as modulus of elasticity better known as Young's modulus for granular models according to the equation (4) (Massarsch and Tekn., 2015).

$$E = \frac{(1 - \nu)}{(1 - 2\nu)(1 + \nu)17.11} \quad (4)$$

Where ν is Poisson ratio which is equal to 0.49 (Massarsch and Tekn., 2015; Zhou *et al.*, 2014), These data helped to calculate the shear strain and the dynamic viscosity, defined by equations (5) and (6)

$$\gamma = (1 + \nu)\varepsilon \quad \mu = \frac{\tau}{\dot{\gamma}} \quad (5) \quad (6)$$

Where γ is the shear strain and μ is the fluid viscosity. The shear rate velocity is obtained as defined by equation (7).

$$\dot{\gamma} = \left(\frac{\mu}{k}\right)^{\frac{1}{n-1}} \quad (7)$$

$\dot{\gamma}$ is the shear rate velocity, n is the non-dimensional flow index and k is the consistency coefficient. For a consolidation pressure of 100 kPa, n is equal to -0.1 and k is equal to 10.3 according to (Zhou *et al.*, 2014). These parameters obtained by equations (2) to (7) were injected in a viscosity law. The viscosity model chosen was Herschel-Buckley, which is an extension of the power law (Kang and King., 2016), because this viscosity model defines the behavior of a non-Newtonian fluid at threshold. The model is defined by adding the threshold shear stress (τ_0) and the threshold viscosity (μ_0), as defined in the equations (8) and (9).

$$\tau_0 = \tau - k\dot{\gamma}^n \quad \mu_0 = k\dot{\gamma}^{n-1} + \tau_0\dot{\gamma}^{-1} \quad (8) \quad (9)$$

These parameters were used in the following multiphase modelling

4.2 Multiphase simulations

The modelling has shown similar results to the bowl **Figure 1 (a)**. A zone of 40 % turbulence intensity was highlighted near the reef. The water mean velocity near sand was 0.07 m.s^{-1} but the turbulence energy generated by the fluid-structure interaction leads to a velocity reduction of 0.02 m.s^{-1} . The turbulence kinetic energy is therefore 6 J, corresponding to the same transport process as observed in the bowl. The artificial reef was buried of 0.05 m during the first second of simulation and then stabilized. The saturated sand kept the same morphology indicating that no sediment transport occurred over such a short period. The burial corresponded to the stabilization of the structure in the sand. The immobile saturated sand results corresponded to the sediment transport results of the bowls modelling.

4.3 Simulation with strong turbulences

The **Figure 1 (e)** illustrates the sediment displacement and the eddy caused by turbulence energy. The mean velocity of the sand at 0.54s was of 0.11 m.s^{-1} for a turbulence intensity of 49%, which corresponds to a turbulent kinetic energy of 44J. This high turbulent energy is explained by the maximum speed of the vorticity of 2.9 m.s^{-1} . This very high velocity correlates with the turbulent kinetic energy, inducing sediments transport. A bowl is created from a depth of 0.12m at 0.54s. The transport of sediments with strong turbulence energy seemed to correspond with the conclusions of (Mazzuoli *et al.*, 2020). The reef shape induced a different eddy formation in space, creating a different scouring process according to the studied zone. The reef's behavior was the same during multiphase modelling at normal conditions as with multiphase modelling with strong turbulence energy.

7. Conclusions

The modelling of the fluid-structure interaction on the bowls allowed to highlight an evolution of the scouring according to the geomorphology. The interactions between water and reef led to an increase in the turbulent kinetic energy and a reduction of the mean velocity above the structure. These turbulence characteristics resulted in a slow sediment transport and caused bowls expansion encouraging the burial of artificial reefs. Multiphase modelling can predict the burial of artificial structure but the WMLES modelling requires too much calculation time with current technologies. In order to achieve the objective of simulating over several years a Reynolds-Averaged Navier-Stokes equations (RANS) modelling should be experienced at first. Machine learning is being developed in this area and could be implemented in this RANS simulation in order to drastically reduce the amount of computation while keeping a good accuracy on the turbulence modelling.

References

- B.M. Sumer, R.J.S. Whitehouse, A. Tørum, 2001, Scour around coastal structures: a summary of recent research, *Coastal Engineering* 44, 153–190.
- C.G. Kim, 2001, Artificial Reefs in Korea, *Fisheries* 26, 15–18.
- D. Jørgensen, 2009, An oasis in a watery desert? Discourses on an industrial ecosystem in the Gulf of Mexico Rigs-to-Reefs program, *History and Technology* 25, 343–364.
- D.R. Fuhrman, C. Baykal, B. Mutlu Sumer, N.G. Jacobsen, J. Fredsøe, 2014, Numerical simulation of wave-induced scour and backfilling processes beneath submarine pipelines, *Coastal Engineering* 94, 10–22.
- F. Werner, G. Unsöld, B. Koopmann, A. Stefanon, 1980, Field observations and flume experiments on the nature of comet marks, *Sedimentary Geology* 26, 233–262.
- H.S. Kang, Y.T. Kim, 2017, Rheological properties of loose sands subjected to upward flow, *Canadian Geotechnical Journal*, 54, 664–673.
- J. McNinch, J. Wells, T. Drake, 2001, The fate of artifacts in an energetic, shallow-water environment: Scour and burial at the wreck site of Queen Anne’s Revenge, *Southeastern Geology* 40, 19–27.
- J. Ye, D. Jeng, R. Wang, C. Zhu, 2013, A 3-D semi-coupled numerical model for fluid–structures–seabed-interaction (FSSI-CAS 3D): Model and verification, *Journal of Fluids and Structures* 40, 148–162.
- J.R. Miller, R.J. Hobbs, 2007, Habitat restoration, do we know what we’re doing?. *Restoration Ecology*, 15, 382–390.
- K.R. Massarsch, G. Tekn, 2015, Determination of Shear Modulus of Soil from Static and Seismic Penetration Testing, *Jubilee Volume, Athens 2015*, 335–352.
- M. Chávez-Modena, J. Martínez, J. Cabello, E. Ferrer, 2020, Simulations of Aerodynamic Separated Flows using the Lattice Boltzmann Solver XFlow, *Energies* 13.
- M. Mazzuoli, P. Blondeaux, G. Vittori, M. Uhlmann, J. Simeonov, J. Calantoni, 2020, Interface-resolved direct numerical simulations of sediment transport in a turbulent oscillatory boundary layer, *Journal of Fluid Mechanics*, 885, A28.
- M. Ouda, E.A. Toorman, 2019, Development of a new multiphase sediment transport model for free surface flows, *International Journal of Multiphase Flow* 117, 81–102.
- N. Abram, C. Adler, N.L. Bindoff, L. Cheng, 2019, Special Report of GIEC on the Ocean and Cryosphere in a Changing Climate.
- Q. Zhou, C. Lv, Z.H. Wang, G.X. Chen, 2014, Fluid Characteristic of Saturated Sands under Cyclic Loading, 178–1.
- R. Quinn, 2006, The role of scour in shipwreck site formation processes and the preservation of wreck-associated scour signatures in the sedimentary record – evidence from seabed and sub-surface data, *Journal of Archaeological Science*, 33, 1419–1432.
- S.E. Rees, E.V. Sheehan, B.D. Stewart, R. Clark, T. Appleby, M.J. Attrill, P.J.S. Jones, D. Johnson, N. Bradshaw, S. Pittman, J. Oates, J-L. Solandt, 2020, Emerging themes to support ambitious UK marine biodiversity conservation, *Marine Policy*, 117, 103864.
- T.D. Ainsworth, C.L. Hurd, R.D. Gates, P.W. Boyd, 2019, How do we overcome abrupt degradation of marine ecosystems and meet the challenge of heat waves and climate extremes?, *Global Change Biology*, Wiley Online Library.
- Y. Li, M.C. Ong, D.R. Fuhrman, 2020, CFD investigations of scour beneath a submarine pipeline with the effect of upward seepage, *Coastal Engineering*, 156, 103624.
- Y. Li, M.C. Ong, T. Tang, 2018, Numerical analysis of wave-induced poro-elastic seabed response around a hexagonal gravity-based offshore foundation, *Coastal Engineering*, 136, 81–95

Metamodeling of chemical engineering unit operations using Kriging and prediction error estimation

Thibault Delage^a, Sanaa Zannane^a and Thibaut Neveux^{a,*}

^a*EDF R&D Chatou, 6 quai Watier, 78400 Chatou, France*
thibaut.neveux@edf.fr

Abstract

Kriging is a metamodeling technique, consisting in interpolating between observations points, i.e. simulations with the actual model, which is able to produce analytical expressions for both the metamodel and the prediction error in the whole domain. The goal of this work is to illustrate, on two emblematic categories of unit operations (molecular separator and exothermic reactor), how to use this ability to sequentially build a metamodel, by adding new interpolating points in regions where the error between meta- and actual models is high.

Keywords: Metamodeling, Kriging, separation, reactor

1. Motivation

Metamodeling, or surrogate modeling, refers to the production of a simplified model of an actual model. Such techniques are widely used when the actual model is judged inapplicable or with difficulty within a given context, such as: informatic complexity (e.g., communication standards); issues in code coupling (e.g., different spatial or temporal scales); computing efficiency (e.g. optimization, real-time simulations); need to hide information/expertise during implementation etc. In chemical engineering, process models are in general phenomenological ones, with phenomena modelled from molecular scale to plant scale. Challenges arising are therefore to be able to mimic the 'local' behaviors such as stiff non-linearity and discontinuities due to the physicochemical nature of modelled phenomena (activation processes, reactions, thermodynamics, transfer, phase change etc.); while being appropriate in the whole physical domain.

Among metamodeling techniques, Kriging is an interpolation method between observations points, i.e., simulation results with the actual model (Sacks et al., 1989). This method, also called Gaussian process regression, is well-known in chemical engineering and applied to learn various models such as thermodynamic (Zhang et al., 2013) or unit operations ones (Al et al., 2019) in specific fields. As an interpolation method, the prediction error is zero at the observation points, i.e. data sample used for fitting the Kriging metamodel, and increases between them. Therefore, the sample distribution directly impacts the prediction quality of the metamodel, which could lead to poor prediction performances when using the metamodel. One of the main features of Kriging is the ability to provide analytical expressions for both the metamodel (mean of Gaussian processes) and the prediction error in the whole domain (using the variance of the Gaussian processes). It is therefore possible to perform adaptive sampling (see for instance Liu et al., 2017 or Boukouvala and Ierapetritou, 2012) during the construction step of the metamodel, using the 'local' estimation of the prediction error (e.g. a confidence level) in every point of the domain. Regions with behaviors observed in chemical engineering unit operations (e.g. stiff non-linearities such as reaction fronts) are thus detected with the local prediction error, allowing to efficiently explore the search space by refining optimally the design of experiments.

The goal here is to illustrate how to use the ability of Kriging to estimate the prediction error in the whole domain to sequentially build a metamodel; and to evaluate how Kriging is a suitable method to catch such behaviors using two different unit operations models.

2. Test-problems: chemical engineering models for unit operations

Two tests problems are chosen here, in order to cover two categories of unit operations encountered in industrial processes: molecular separation by membrane and exothermic catalytic reactor.

These cases exhibit various physical behaviors (see Figure 1) and numbers of input and output variables. Each case is summarized below, with the chosen parametrization of variables, and expected difficulties from a metamodeling point of view.

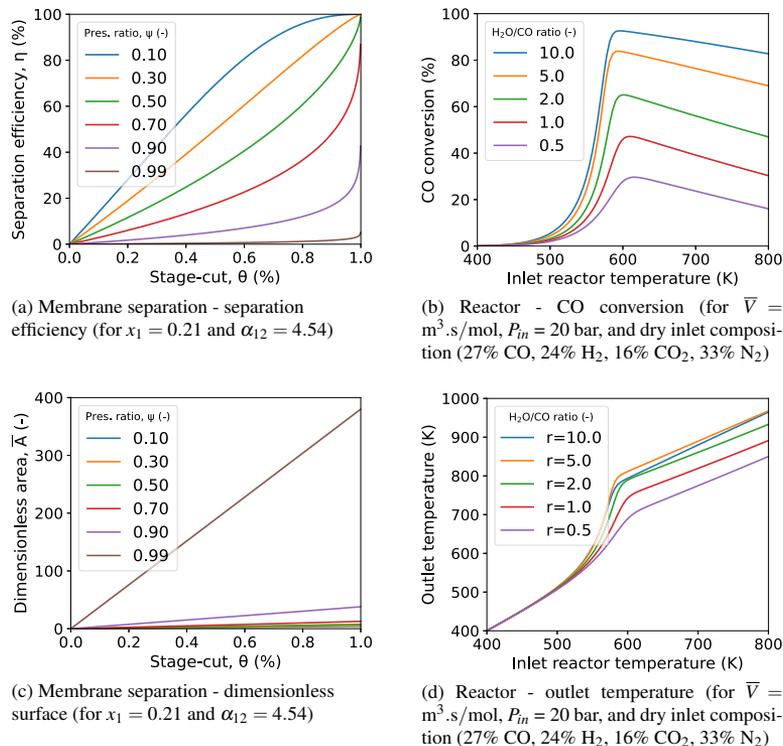


Figure 1: Example of model outputs for the two unit operations

2.1. Case 1. Molecular separation: membrane gas permeation unit (binary mixture)

In this case, a simple gas permeation model is used, assuming crossflow operations, constant permeabilities, and isothermal separation. The physical model requires 4 input variables: mole fraction of the ‘fastest’ molecule in the feed stream x_1 , membrane selectivity α_{12} (ratio of permeabilities), pressure ratio applied between permeate and feed ψ , and design stage-cut θ (desired ratio of flowrate at permeate over feed); and generates 2 output variables: separation efficiency η , and a dimensionless surface \bar{A} . For a complete discussion on gas permeation modeling and dimensionless quantities, see Bounaceur et al., 2017).

With this model, phenomena are mostly linear in the majority of the domain but nonlinear, yet monotonous, profiles occur in the domain boundaries (see for example Figure 1 (a) for high values of stage-cut θ and pressure ratio ψ).

2.2. Case 2. Exothermic catalytic reactor: water-gas shift reaction

The reactor case uses an 1D exothermic catalytic model for water-gas shift reaction ($\text{CO} + \text{H}_2\text{O} = \text{CO}_2 + \text{H}_2$), using kinetic laws from Hla et al. (2011), under plug-flow and adiabatic operation. The physical model requires 7 input variables: gas mole fractions \mathbf{y} of CO , H_2 , H_2O , CO_2 (N_2 deduced) in the inlet stream, inlet temperature T_{in} and pressure P_{in} , and reactor volume normalized by the inlet flowrate \bar{V} ; and generates 2 output variables from which all quantities can be retrieved: CO conversion rate and outlet temperature.

In this case, there is a strong coupling between all input variables as the model considers both kinetic effects (favored by temperature, stoichiometric reactant ratio and reactor volume) and thermodynamic equilibrium limitations (penalized by temperature and product/reactant ratios). It results in non-monotonous behaviors and very stiff thresholds for the CO conversion (see Figure 1 (b)) within the domain, whose location depends on all input variables.

3. Kriging for metamodeling

3.1. Gaussian processes

Let denote $z(\mathbf{x})$ the output of the numerical model we look to metamodel, with $\mathbf{x} = (x_1, x_2, \dots, x_d) \in T$ denoting the input variables, T being a subset of \mathbb{R}^d (e.g. in case 1, $z(\mathbf{x})$ would be η or \bar{A} and the input dimension is $d = 4$). Kriging framework relies on the assumption that $z(\mathbf{x})$ is the realization of a Gaussian process Z which is completely specified by its mean function $m(\mathbf{x}) = \mathbb{E}[Z(\mathbf{x})]$ and its covariance function: $\text{cov}(Z(\mathbf{x}), Z(\mathbf{x}')) = \sigma^2 r(\mathbf{x}, \mathbf{x}')$. The correlation functions r we are using in this work are the anisotropic $\frac{3}{2}$ -Matérn kernel: $r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) = \prod_{i=1}^d \left(1 + \frac{\sqrt{3}(x_i - x'_i)}{\theta_i}\right) \exp\left(\frac{-\sqrt{3}(x_i - x'_i)}{\theta_i}\right)$ and the anisotropic squared exponential kernel: $r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}) = \prod_{i=1}^d \exp\left[-\frac{1}{2} \left(\frac{(x_i - x'_i)}{\theta_i}\right)^2\right]$. The mean function $m(\mathbf{x})$ of the Gaussian process is chosen here to be a constant function $m(\mathbf{x}) = c \in \mathbb{R}$.

3.2. The Kriging metamodel

Let denote $\mathbf{D} = (x_1, x_2, \dots, x_n)$ the design of experiments (i.e. some calculations performed on the actual model) of size n used to train the metamodel, with the corresponding outputs $\mathbf{z} = (z(x_1), \dots, z(x_n))$. Then, following previous assumptions, z is the realization of a Gaussian process and the Kriging model writes:

$$z(\mathbf{x}) = c + \boldsymbol{\varepsilon}(\mathbf{x})$$

with $\boldsymbol{\varepsilon}(\mathbf{x})$ a stationary Gaussian process with zero mean and variance σ^2 and correlation function $r(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta})$. $Z(\mathbf{x})$ is a conditioned Gaussian $\mathcal{N}(\hat{z}(\mathbf{x}), s^2(\mathbf{x}))$, and its mean and variance can be computed. The mean, hyperparameters $\boldsymbol{\theta}$ and variance σ^2 are estimated from the observations \mathbf{D} and \mathbf{z} using the maximum likelihood estimator. The Kriging mean $\hat{z}(\mathbf{x})$ is the surrogate model that we use to approximate the numerical model $z(\mathbf{x})$ and the Kriging variance $s^2(\mathbf{x})$ is used as a representation of the model mean squared error. As Kriging is an interpolating metamodel, the 'local' error is consequently equal to zero for the observations points and rises between points.

3.3. Adaptive design of experiments and validation

In order to train the Kriging metamodel, we use an adaptive strategy starting from an initial Latin Hypercube Sample (LHS) optimized (through a simulated annealing routine) according to the L2-discrepancy criterion (Damblin et al., 2013). Using this initial design, a first Kriging model is constructed and, using a large validation sample $\mathbf{Y} = (y_1, \dots, y_m)$ and its response by the actual model $z(\mathbf{Y}) = (z(y_1), \dots, z(y_m))$, the quality of the metamodel is evaluated by the means of the absolute relative error $e_{rel,i} = \left| \frac{\hat{z}(y_i) - z(y_i)}{z(y_i)} \right|$ and the Q_2 predictivity factor: $Q_2 = 1 - \frac{\sum_i [\hat{z}(y_i) - z(y_i)]^2}{\sum_i [z(y_i) - \bar{z}]^2}$

with \bar{z} the statistical mean of the sample $z(Y)$. The Q_2 compares the residual variance with the total variance. The closer it is to 1, the more accurate the model is. After the initial training, and if the expected quality is not yet reached, new points are sequentially added to the design of experiments using the prediction error of the metamodel given by its variance $s^2(x)$. Various criterion can be used in order to iteratively add new points (e.g. Jones et al., 1998), here the criterion maximizes the conditional variance $s^2(x)$ i.e. $x_{n+1} = \arg \max_x s^2(x)$ as we seek to control the maximum errors as much as possible. It seeks to add the new points where the metamodel is supposed to perform poorly in order to ensure that the prediction does not fail in some specific areas of the physical space. Figure 2 illustrates this idea on a one-dimensional function ($0.5x^2 + \sin(2.5x)$).

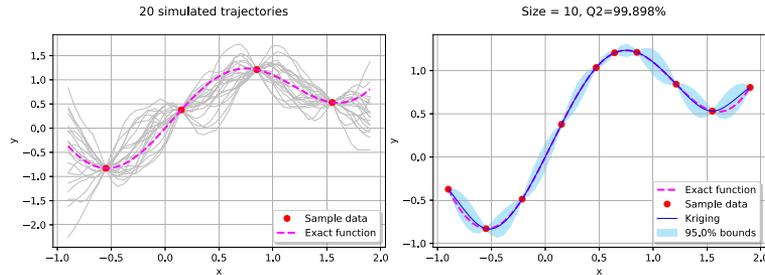


Figure 2: Example of simulated trajectories between 4 sample points (left) and Kriging results with 6 new sample points, sequentially added using prediction error (right)

3.4. Implementation

The Kriging implementation of OpenTURNS (www.openturns.org) is used here, an open-source software for data analysis, metamodeling, and treatment of uncertainties (Baudin et al., 2016).

4. Assessment using the test-problems

In this part, we will present the results obtained on the two tests problems described in 2. The Kriging metamodel is built using a squared exponential covariance for case 1 and a Matérn covariance for case 2. We use an initial LHS of size $N = 100$ and perform $M = 1000$ iterations using the criterion presented in 3.3.

4.1. Case 1: molecular separation

Figure 3 shows the initial design of experiments (blue) and the 300 first additional calculations (orange). The criterion aims at choosing new points on the edges of the physical space and especially on its corner, since the conditional variance has the tendency to strongly increase in these areas. Moreover, this quite classical behavior is reinforced by the fact that the model shows non-linearities around these boundaries, as discussed in 2.1 and shown in figure 1.

After filling these voids, the algorithm tends to add calculation in the other hard-to-predict areas of the physical space such as in the neighborhood of the stage-cut $\theta \sim 0.8$ and the membrane selectivity $\alpha \sim 5$ for instance.

In order to assess the quality of the metamodel, we use a LHS test sample of size 1000. Figure 3 shows the evolution of the Q_2 predictability factor according to the number of iteration and the increasing quality of the metamodel.

As the Q_2 only carries a global estimation of the quality of the metamodel, we also show the box plots evolution of the relative errors according to the number of iterations of the adaptive sampling algorithm for the two outputs of the model: dimensionless \bar{A} and efficiency η in Figures 4. The

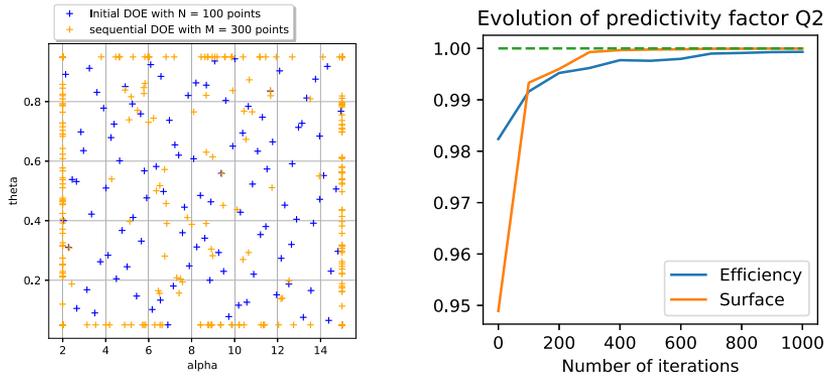


Figure 3: Incremental design of experiment (first 400 points) used to build the metamodel (left) and evolution of Q_2 for case 1 (right)

results are quite satisfactory as we reach, after 1000 iterations, a Q_2 larger than 0.999 for η and larger than 0.9999 for \bar{A} as well as a relative error well-controlled (e.g. 75% of tested points falls under a 3% and 4% relative error for \bar{A} and η respectively) even if some specific areas still suffer from a quite large error. This method has been benchmarked on a neural network constructed from a sample of size 2200 and proved to outperform it from 600 iterations.

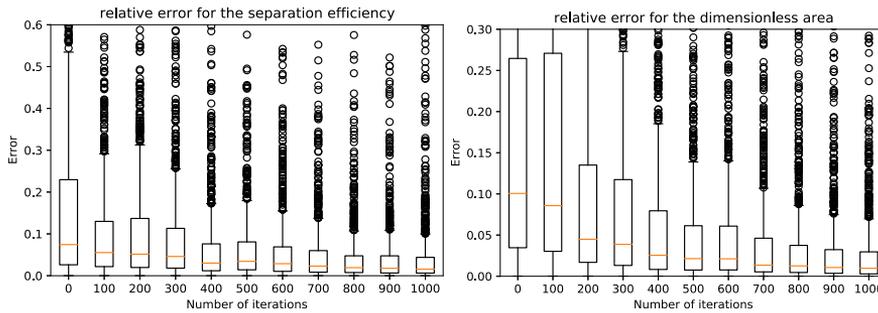
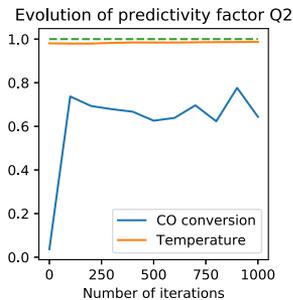


Figure 4: Evolution of the relative error for the separation efficiency η

4.2. Case 2. Exothermic catalytic reactor



Evolution of Q_2 for case 2

This second test-case proves to be more challenging, as it involves a larger number of input variables (7) and more complex physical behavior (see 2.2). If the metamodel manages to approach the outlet temperature quite well ($Q_2 > 0.98$), it reveals to be way less satisfactory to predict the CO conversion, which Q_2 does not exceed 0.8. Figure 5 shows the relative errors for both outputs and if the error on the temperature is acceptable – 75% (resp. 95%) of tested points falls under a 2% (resp. 5%) relative error – it is still far from the accuracy required to justify most metamodel usage in the case of the CO conversion.

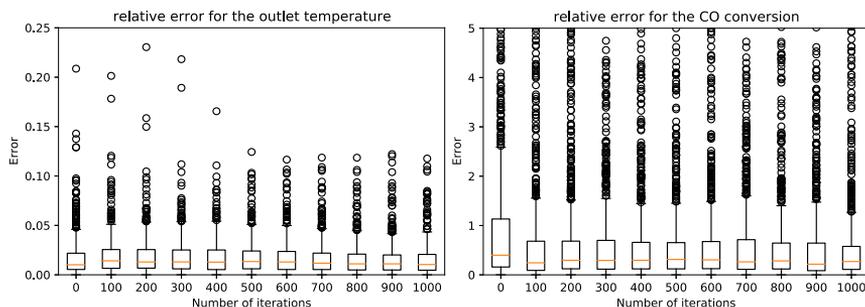


Figure 5: Evolution of the relative error for outlet temperature and CO conversion for case 2

5. Conclusion and perspectives

This work illustrates how to use the ability of Kriging to provide a prediction error in the whole domain, to sequentially add new interpolation points where the metamodel exhibits high error, aiming at reducing the size of the design of experiments (hence time/budget for the metamodel building step), and discrepancies while using the metamodel as it is more refined in regions with high non-linearities (usually difficult to catch by metamodels). This strategy performs remarkably well in the case of molecular separation by gas permeation. The second, more complex, case (exothermic reactor) highlights the limits of the methodology when facing the curse of dimensionality and the increase of the boundaries of the domain, since the algorithm privileges these before filling the areas with strong non-linearities. A coarse first solution is to increase the initial training sample, $Q_2 > 0.995$ for the outlet temperature can be reached with $N = 500$ and $M = 600$. Another possible approach would be to investigate the performance of other criteria, such as the IMSE (Integrated Mean Squared Error).

This first benchmark need to be tested with other criteria for the choice of new points, other types of unit operations, and compared with other sampling strategies (driven or not by the error) and metamodel type, to derive good metamodeling practices for chemical engineers.

References

- R. Al, C. R. Behera, A. Zubov, K. V. Gernaey, G. Sin, 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – an application to the BSM2 model. *Computers & Chemical Engineering* 127, 233–246.
- M. Baudin, A. Dutfoy, B. Iooss, A.-L. Popelin, 2016. OpenTURNS: An industrial software for uncertainty quantification in simulation. In: *Handbook of Uncertainty Quantification*. Springer International Publishing, Cham, pp. 1–38.
- F. Boukouvala, M. G. Ierapetritou, 2012. Feasibility analysis of black-box processes using an adaptive sampling kriging-based method. *Computers & Chemical Engineering* 36, 358–368.
- R. Bounaceur, E. Berger, M. Pfister, A. A. Ramirez Santos, E. Favre, 2017. Rigorous variable permeability modelling and process simulation for the design of polymeric membrane gas separation units: MEMSIC simulation tool. *Journal of Membrane Science* 523, 77–91.
- G. Damblin, M. Couplet, B. Iooss, 2013. Numerical studies of space filling designs: optimization of latin hypercube samples and subprojection properties.
- S. S. Hla, G. Duffy, L. Morpeth, A. Cousins, D. Roberts, J. Edwards, 2011. Investigation into the performance of a Co–Mo based sour shift catalyst using simulated coal-derived syngases. *International Journal of Hydrogen Energy* 36 (11), 6638–6645.
- D. R. Jones, M. Schonlau, W. J. Welch, 1998. Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 13 (4), 455–492.
- H. Liu, J. Cai, Y.-S. Ong, 2017. An adaptive sampling approach for kriging metamodeling by maximizing expected prediction error. *Computers & Chemical Engineering* 106, 171–182, ESCAPE-26.
- J. Sacks, W. J. Welch, T. J. Mitchell, H. P. Wynn, 1989. Design and Analysis of Computer Experiments. *Statistical Science* 4 (4), 409 – 423.
- C. Zhang, Z. Zhan, Z. Shao, Y. Zhao, X. Chen, X. Gu, Z. Yao, L. Feng, L. T. Biegler, 2013. Equation-oriented optimization on an industrial high-density polyethylene slurry process with target molecular weight distribution. *Industrial & Engineering Chemistry Research* 52 (22), 7240–7251.

On the integration of process engineering with metabolomics for the production of muconic acid: the case for *Saccharomyces Cerevisiae*

Stefanos Xenios^a, Daniel Weilandt^b, Hatzimanikatis Vasilis^b, Ljubisa Miskovic^b, Antonis Kokosis^a

^a*Department of Process Engineering, NTUA, Iroon Politechniou 6 Zografou, Athens, Greece*

^b*Laboratory of Computational Systems Biotechnology, EPFL, Rte Cantonale, 1015 Lausanne, Switzerland*
jirajia@hotmail.com

Abstract

Muconic acid is a high value product which has gathered interest in applications in the manufacture of new resins, bio plastics, food additives, agrochemicals and pharmaceuticals. Lots of efforts have been made for an economically viable biotechnological strategy for muconic acid production but as of yet have been fruitless. Directed evolution and DBTL cycles hold important promises for the development of future catalysts with high efficiency and productivity. However, process engineering is typically disjointed from these cycles and more often than not the mismatch of kinetics presents a major challenge and a bottleneck in the scaling up of novel bioprocesses.

The paper addresses the integration of metabolomics and experimental data using the optimization and risk analysis of complex living entities (ORACLE) platform combined with clustering and advanced analytics. The methodology consists of six steps. In the first step, the stoichiometry of the system is defined through biochemical data and experimental data are integrated into the model to further constrain it. In the second step, steady state fluxes and metabolite concentrations are calculated based on metabolomics analysis. In the third step, kinetic parameters for every reaction are sampled to fit in with the steady state fluxes based on mechanistic kinetics expressions. In the fourth step, consistency checks and pruning consider the stability of the system and the consistency with experimental data. In the fifth step, the flux control coefficients for the desired metabolite flux are calculated based on the well-established metabolic control analysis (MCA) framework.

In this project, large-scale mechanistic kinetic models for a muconic acid producing *S.Cerevisiae* strain were developed using the aforementioned ORACLE platform. A total of 23500 of potential kinetic models were generated out of which 366(1.58%) passed the pruning step and 70(0.12%) models out of the 372 passed the stability check. Enzymes such as glucose-6-phosphate isomerase (PGI), transketolase (TKT2) and enolase (ENO) had large control coefficients on muconic acid flux.

Keywords: *S.Cerevisiae*; muconic acid; metabolomics; metabolic engineering ; large-scale kinetics;

1. Introduction

Most strain design approaches make the assumption of the system being in steady state and fail to take into consideration the kinetic aspect. Although such techniques provide useful insights and are computationally cost efficient, sometimes they fail to predict potential bottlenecks or rate limiting reactions inside the metabolic network. The current thesis aims to develop large scale metabolic kinetic models for a muconic acid producing yeast using the Optimization and Risk Assessment of Complex Living Entities (ORACLE) workflow. The generated populations of kinetic models will be used as an input for the well-established MCA framework to identify enzymes closely affiliated to muconic acid flux and perform enzyme perturbations to quantify this effect. This way we will be able to develop metabolic strategies while taking into consideration stoichiometry, thermodynamics, kinetics and their interplay.

2. Problem description and workflow outline

2.1.1 Problem description

The development of large-scale metabolic models is a challenging task. The lack of experimental data, the uncertainty of some kinetics parameters available in literature and databases, the uncertainty of the types of mechanisms for every reaction, the errors in metabolomics and fluxomics data are some of the problems (Stanford et al, 2013). Moreover, for every reaction in the system a rate expression along with values of kinetic parameters are required for a kinetic model. Errors also in the thermodynamic properties hinder the ascertainment of a unique steady-state profile for metabolic fluxes and metabolite concentrations. Taking all this into account it is impossible to find a unique kinetics model which describes the physiology but it is possible to produce a population of models that agree with the physiology and statistical analysis on these models can be used to analyse and predict the metabolic responses in the system (Tokic et al, 2020).

The primary challenge will be to produce large-scale kinetic models that are physiologically relevant meaning that their dynamic behaviour is close to the experimentally observed one. The secondary challenge is to identify kinetic models that show “robust” behaviour. Finally, the population of the kinetics models will be used to target key enzymes and offer metabolic strategies for the increase of flux of the desired product.

2.1.2 Workflow outline

The proposed workflow outline is based on the Optimization and Risk Analysis of Complex Living Entities (ORACLE) methodology (Miskovic et al, 2010) (Wang et al, 2004). The ORACLE methodology consists of 7 steps:

- 1) In the first step, the stoichiometry of the system is defined either by biochemical data or genome reconstruction analysis. The genome scale model is reduced to the subsystems of interest. Then, the experimental data are integrated into the reduced model as flux bounds or metabolite concentrations bounds or reaction directionalities. Heterologous reactions are added to the reduced model and gene knockouts are expressed by setting both lower and upper bounds for the linked reaction to zero.
- 2) In the second step, the solution space from Thermodynamic Flux Balance Analysis is sampled. If the thermodynamic properties of a reaction are unknown, we employ

- a Monte Carlo method to predict the reaction ΔG value. Metabolite concentrations and reaction fluxes profiles, that agree with thermodynamics, are generated.
- 3) In the third step, the kinetic parameters for every reaction in the reduced model are sampled using a hit and run sampler. Mechanistic kinetic expressions such as reversible Michaelis Menten are used to describe enzyme kinetics in our formulation. The thermodynamic displacements are calculated during the TFA sampling step. In order to efficiently sample the K_m values, we sample the enzyme saturations. We used simple enzyme saturations expressions in our formulation; $\sigma_E = \frac{\frac{S}{K_m}}{\frac{S}{K_m} + 1}$. Lastly, the v_{max} values are calculated to fit the sampled steady state flux.
 - 4) In the fourth step, the produced kinetic models undergo pruning and stability checks. Pruning reduces the models to those that are physiologically relevant and consistent with the experimental data. The time constants for every metabolite mass balance must be within some physiological bounds to ensure that metabolism is in a quasi-steady state. Stability checks consist of sets of random perturbations to ensure that the generated models show robust model behavior and return to their reference steady state. The models are further screened out based on their stability score.
 - 5) In the fifth step, the stable models are used to calculate flux control coefficients and concentration control coefficients based on a well-established MCA framework. Flux control coefficients C_p^v and concentration control coefficients C_p^x are defined as the fractional change of metabolic fluxes and metabolite concentrations, respectively, in response to a fractional change of system parameters.
 - 6) In the sixth step, advanced statistical analysis and visualization is performed on the produced populations of control coefficients.

3. Case study: Muconic acid production from *S.Cerevisiae*

3.1 Reduced model and experimental data integration

The genome scale model for *Saccharomyces Cerevisiae* used in the current thesis was the yeast8 model⁶⁴. The reduced model contains 5 subsystems that form our metabolic network; Glycolysis, Pentose Phosphate Pathway, TCA cycle, Oxidative Phosphorylation, Shikimate pathway (aromatic amino acids).

The reduced yeast model has 226 metabolites, that makes a total of 226 mass balances. Out of those, 49 are extracellular metabolites. A total of 308 reactions describes the metabolic network out of which 121 are transport reactions, 183 are enzymatic and 1 lumped biomass that describes biomass production and subsequently yeast growth. We added the heterologous reactions for the muconic acid production pathway. According to the experimental data (Wang et al, 2020) the pathway used was the shunting of the shikimate pathway. We assumed that the reactions added took place in the cytosol.

3.2 Kinetic Models

Having generated a big number of TFA samples we screened out 47 samples that produced physiologically relevant kinetic models. We followed through generating 500 parameter sets for every TFA sample, thus resulting in 23500 kinetic models. Only 366 models were within physiological bounds (1,58%). We then proceeded with the stability checks where we performed 200 sets of random perturbations on a short range ($0.8[X_{ref}] \leq [X] \leq 1.2[X_{ref}]$) and on a wider range ($0.5[X_{ref}] \leq [X] \leq 2[X_{ref}]$).

We evaluated the models based on the times they returned to the reference steady state. The two other scenarios were either the models reaching a pathological state (escape) or another steady state.

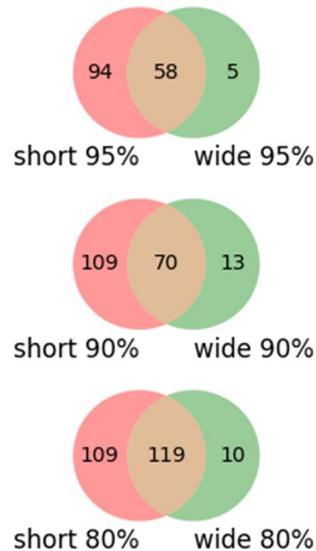


Figure 1. Venn diagrams of models for various stability scores (80%, 90%, 95%) for both ranges of perturbations.

3.3 MCA and enzyme perturbations results

We selected 70 models with stability scores 90% and calculated the muconic flux control coefficients. Based on the enzymes, hinted by the MCA results, that seem to have the most influence on muconic flux we performed some enzyme perturbations and calculated the effect on muconic flux.

We were able to identify the top 20 enzymes that affect muconic flux. Some attractive metabolic strategies that arisen were to increase flux to the Pentose Phosphate Pathway as to increase flux to e4p (PGI↓,TKT2↓,PGL↓) and to increase the dehydroshikimate flux (DHQTi↑). From the MCA results, we deduced that a bottleneck in muconic acid production is the deficiency of e4p in the cell.

For the enzymes that had large control coefficients for the muconic flux we performed some perturbations to calculate the effect on muconic flux. Although, some enzymes that are connected to the Electron Transport Chain (ETC) and are infeasible to regulate, we calculated the effect on muconic flux just for comparison. The candidate enzymes (PGI, TKT2, PGL, DHQTi) showed an increase in muconic flux with varying results. PGI downregulation by 0.5 resulted in 12.5 times bigger than reference muconic flux whereas TKT2 and PGL downregulation by 0.5 resulted in 1.4 and 1.3 times accordingly bigger flux. DHQTi upregulation by 1.5 resulted in 1.8 times bigger than reference muconic flux.

We also tested upregulating the heterologous enzymes (PaAroZ, KpAroY) of the muconic pathway but the increase in muconic flux was insignificant.

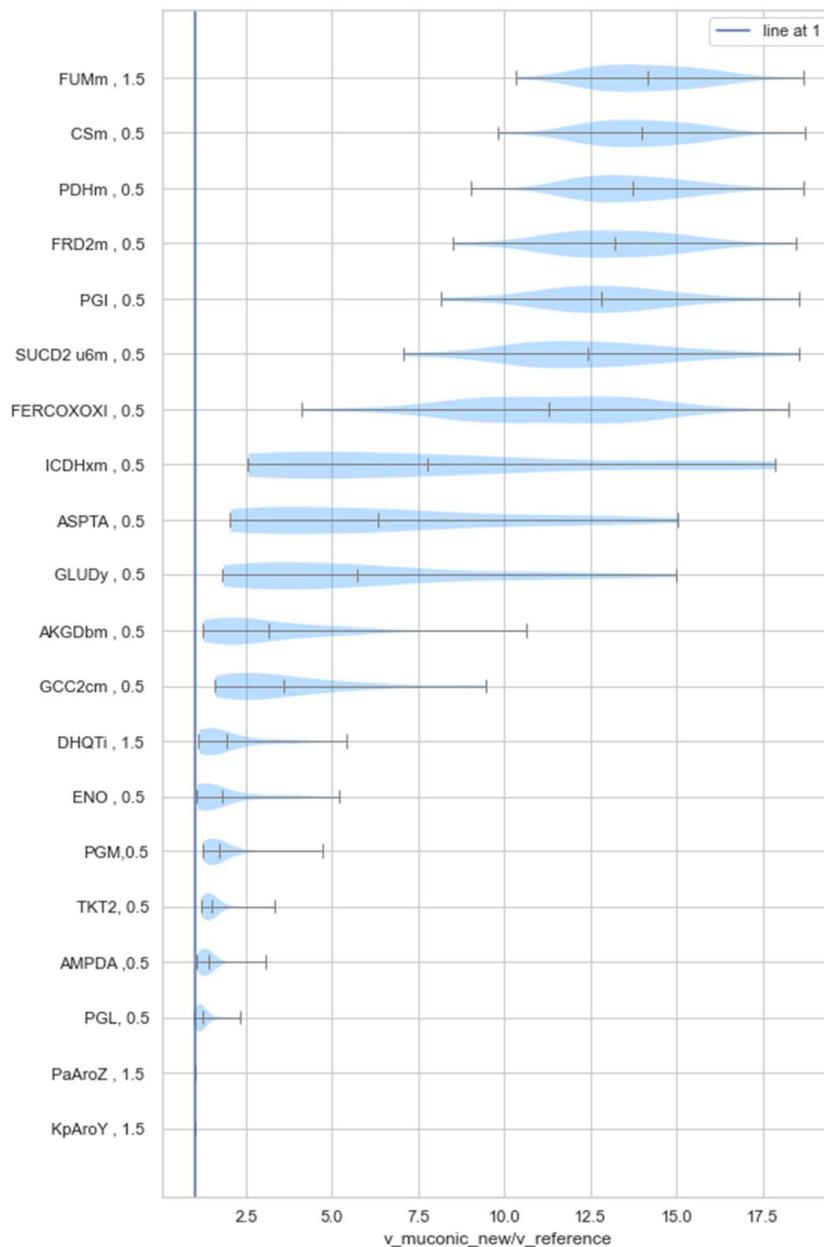


Figure 2: Violin plots of enzyme upregulations and downregulations that lead to an increase in muconic flux. Middle bar corresponds to the mean value. Both extremas also included.

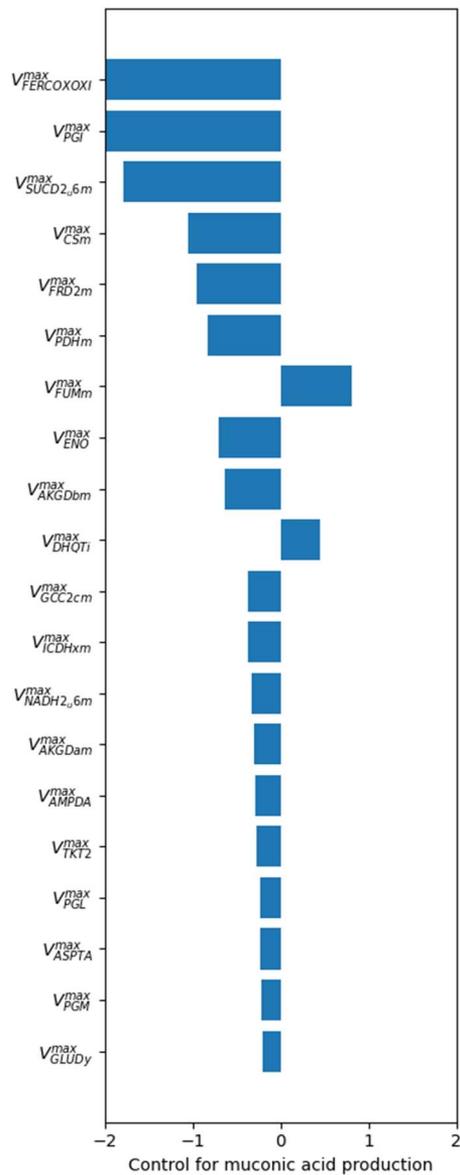


Figure 3: Bar plot of the top 20 enzymes with high muconic flux control coefficients

References

- Miskovic et al. (2010). *Production of biofuels and biochemicals: in need of an ORACLE*. Trends in Biotechnology.
- Stanford et al. (2013). *Systematic Construction of Kinetic Models from Genome-Scale Networks*. PLoS ONE.
- Tokic et al. (2020). *Large-scale kinetic metabolic models of Pseudomonas putida KT2440 for consistent design of metabolic engineering strategies*. Biotechnology of Biofuels.
- Wang et al. (2004). *Metabolic Control Analysis under Uncertainty: Framework Development and Case Studies*. Biophysical Journal.
- Wang et al. (2020). *Improvement of cis,cis-Muconic Acid Production in Saccharomyces cerevisiae through Biosensor-Aided Genome Engineering*. ACS Synthetic Biology.

Modelling and simulation of a residual lignocellulosic biomass pyrolysis pilot plant

Nezly Martelo^a, Antonio Gagliano^b, Alberto Fichera^b, Rosaria Volpe^b, Mirari Antxustegi^a, Rodrigo Llano-Ponte^a, María Gonzalez Alriols^a

^a*Biorefinery Processes Research Group (BioRP), Chemical and Environmental Engineering Department, University of the Basque Country, UPV/EHU, Plaza Europa 1, San Sebastian 20018, Spain*

^b*Department of Electrical, Electronics and Computer Engineering, University of Catania, viale Andrea Doria 6, 95125 Catania, Italy*

Abstract

The present work has been devoted to the study, modelling and simulation of a pilot scale pyrolysis plant feed with residual agricultural residues located in Catania, Sicily. A global model representing the process flowsheet was created with Aspen Plus[®] software and was fed with actual experimental data obtained in the existing pilot plant. The aim has been to develop a flexible simulation model in which different types of local agricultural residues could be fed to the reactor and establish the optimal conditions for each case. Raw materials' characterization included ultimate and proximate analysis, and biochemical composition (cellulose, hemicelluloses, lignin, and inorganics contents). All these data were implemented in the simulation for the feedstock definition. The base scenario was established by feeding 35 kg/h of biomass under a pyrolysis temperature between 500 and 900°C. The decomposition of the organic matter was defined as a two-stage process: a first decomposition of the organic fraction into a residual solid fraction (char) and a gaseous mixture, and a subsequent cooling of the mixture to separate the condensable part and to obtain the pyro-oil and the permanent gas stream. This latter stream of gas was used to generate thermal power that could be used to support the energy requirements of the pyrolysis process. To reproduce the degradation of biomass and the resulting evolution of chemical species, a mathematical model, based on the operative temperature and the reaction kinetics was developed. The simulations were designed in order to maximize the gas fraction and minimize the char one. The obtained results were discussed in terms of compound yields, and energetic content of the generated pyro-gas.

Keywords: residual lignocellulosic biomass, simulation, pyrolysis.

1. Introduction

The global energy consumption is still based on the use of fossil resources, even if the need to move towards a sustainable energy market is evident and reflected in legislation and politic agendas worldwide. The increasing energy demand is expected, thus, to be sustained in a big percentage by renewable energies soon, and, among the renewable options, biomass is one of the main pillars (Pang et al., 2021). A wide variety of biomass types can be used for bioenergy generation, but a very interesting option is to get profit of agricultural or forestry residues that are generated in different activities. Cereal straw (Zanzi et al., 2002), olive stones (Benanti et al, 2011), fruit waste (Lam et al., 2016), nut shells (Jahirul et al., 2012) or pruning (Adrados et al., 2017) are some examples of biomass wastes with potential as raw material to produce bioenergy. Thermochemical

processes, such as gasification, combustion, torrefaction, hydrothermal liquefaction, hydrothermal carbonization or pyrolysis, allow the release of the energy stored in biomass (solar energy) in the form of chemical energy (Goyal et al., 2008). The main drawback of biomass as raw material for energy purposes is its low density, which entails the handling and transporting of huge quantities of material, decreasing profitability. Pyrolysis transform biomass in a solid (bio-char), a liquid (bio-oil) and a gaseous stream, through its degradation in the absence of oxygen. Bio-oil is typically transformed into fuels for transportation or for the synthesis of value-added chemicals while bio-char and the gas fraction can be used for heat and power generation (Brown et al., 2013). Pyrolysis is considered as the fundamental thermal conversion process, as it is a well established technology and a crucial step of the combustion or gasification processes (Wang et al., 2017). Different mechanisms take place during pyrolysis, and the obtained products, specially bio-oil, present a complex composition that difficult the modelling of the process (Lam et al., 2019). Several approaches have been developed about the kinetics of the chemical reactions of biomass pyrolysis process (Ranzi et al., 2008; Ranzi et al., 2014; Di Blasi et al., 2008; Fedyukhin et al., 2017) and some researchers have implemented the kinetic parameters in models and simulations (Mabrouki et al., 2015), trying to relate initial feedstock composition, reaction mechanisms and products definition (Peters et al., 2017) in order to define a simplified but accurate way of modelling the products' yields and compositions (Agrifoglio et al., 2021). The process conditions (temperature, heating rate, residence time), strongly determine the distribution and yields of the obtained products (Wang et al., 2017), and a variety of studies cover the relationship between process conditions and products yields and compositions (Kan et al., 2016), for slow (Williams and Besler, 1996), fast (Bridgwater, 2012) and flash pyrolysis (Amutio et al., 2012). Furthermore, recent contributions indicate that the products' distribution and characteristics are heavily influenced as well by the original feedstock composition (cellulose, hemicelluloses and lignin), (Tang et al., 2020; Cheng et al., 2020; Pang et al., 2021). The present work is based on a previous one (Agrifoglio et al., 2021), in which a simulation model for biomass pyrolysis was developed based on experimental data of a pyrolysis pilot plant. The aim was to predict the products' yields and compositions and to estimate the potential of the pyro-gas stream to be used for energetic purposes. The paper proposed as further research the utilization of experimental data to relate these values with other working parameters, such as the heating rate, residence time or particle grind size, as well as to extract more information about the influence of biomass quality and types in the process. From this basis, different raw materials (other types of local biomass lignocellulosic wastes) have been tested and the generated data have been implemented in the model, in order to discuss the previously mentioned influence of the biomass type on the pyro-products distribution. The percentages and composition of the cellulose, hemicelluloses and lignin of the raw material (the predominant molecule in the hemicelluloses fraction or the functional groups in lignin structure), will define the decomposition profile of this specific biomass type with the heating rate and temperature. In particular, olive and peach stones and walnut and hazelnut shells were selected as raw materials and were defined as non-conventional products in the simulation model based on experimental values.

2. Methodology

2.1. Definition of the experimental data for the simulation model

The chemical characterization of the tested biomass types (percentages of cellulose, hemicelluloses and lignin; proximate, ultimate analysis), as well as the calorific values,

were obtained experimentally. These parameters (Table 1) were used as input in the simulation model developed with Aspen Plus[®], to define the biomass decomposition model, which included both primary (gas₁, char) and secondary reactions (gas₂, oil).

Table 1. Proximate and ultimate analyses, chemical composition (dry basis, wt. %) and higher heating values (HHV) of the analysed biomass samples.

		Olive stones	Peach stones	Walnut shell	Hazelnut shell
Proximate analysis (wt. %)	Fixed carbon	7.60	10.3	29.3	27.4
	Volatile matter	82.4	85.4	67.8	70.6
	Ash	3.30	0.30	1.20	2.50
	Moisture	7.01	4.01	9.51	9.02
Ultimate analysis (wt. %, dry basis)	Carbon	50.4	47.1	49.1	46.6
	Hydrogen	5.95	6.40	5.22	5.83
	Nitrogen	0.11	0.25	0.35	0.57
	Sulphur	0.08	0.15	0.09	0.75
	Oxygen	43.5	46.1	45.2	46.3
Chemical composition (wt. %)	Cellulose	19.1	28.2	23.8	20.4
	Hemicelluloses	36.6	23.9	29.6	26.2
	Lignin	43.0	43.3	40.3	48.7
Heating value (MJ/kg)	HHV	21.7	21.3	20.8	21.9

The mass balances were adjusted according with the results of the thermochemical decomposition analysis performed. The evolution of the species involved during the pyrolysis was defined by the derivative of the conversion rate and the kinetic equation (Arrhenius type), as described in the initial work (Agrifoglio et al, 2021). The decomposition of cellulose was defined in two steps, starting with the intermediate “active cellulose” and continuing with the subsequent degradation of it and the formation of levoglucosan (Gorensek et al., 2019). Hemicelluloses degraded into two intermediate compounds from which different decomposition routes generated several minor compounds, among them, xylosan (Ranzi et al., 2014). The lignin decomposition generated different intermediates which further degrade into coumaryl, different alcohols and phenol, among other compounds (Ranzi et al., 2008).

2.2. Design of the simulated process flowsheet

Figure 1 presents the flowsheet of the simulated process with Aspen Plus[®].

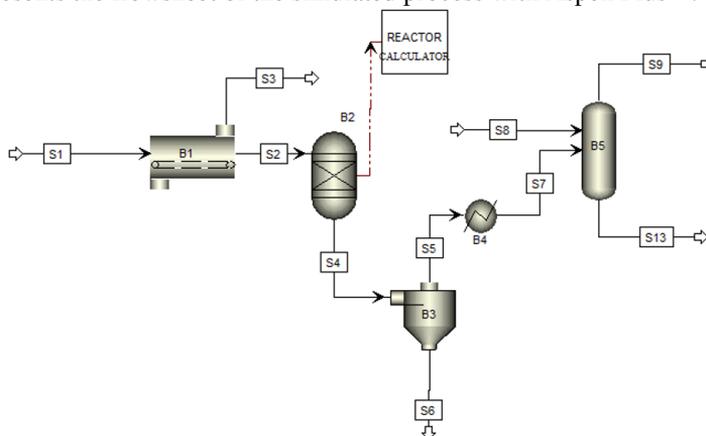


Figure 1. Layout of the simulated process flowsheet.

The base scenario was established by feeding 35 kg/h of biomass under a pyrolysis temperature between 500 and 900°C. Block B1 is a dryer that adjusts the moisture content of the raw material and preheat it before feeding it to the yield reactor (B2), which was connected to a calculator block. The excel sheet coupled to the block include the coefficients to calculate the products yield distribution. The stream exiting from the reactor, S4 is introduced into a cyclone (B3) to separate the gas (S5) and solid (S6) streams. The gas stream is conditioned by cooling (B4) and washing (B5) with fresh water (S8).

3. Results

The yields of char, liquid and gaseous fractions at different temperatures in the range 500-900 °C are reported in Table 2.

Table 2. Yields of char, oil and gas fractions at different temperatures for the studied biomass types.

Products yield wt. %	Temperature °C	Olive stones	Peach stones	Walnut shell	Hazelnut shell
Char	500	34.7	32.1	28.9	30.1
	600	23.3	22.9	26.4	28.2
	700	16.5	16.2	24.2	25.9
	800	13.4	13.3	22.9	23.8
	900	12.1	11.7	21.7	21.4
Oil	500	21.8	20.5	24.5	23.9
	600	18.8	18.0	22.7	21.0
	700	15.6	15.6	19.1	18.7
	800	12.8	12.2	16.3	16.3
	900	9.21	9.00	13.2	13.1
Gas	500	43.5	47.4	46.6	46.0
	600	57.9	59.1	50.9	50.8
	700	67.9	68.2	56.7	55.4
	800	73.8	74.5	60.6	59.9
	900	78.6	79.3	65.1	65.5

It can be seen how the generation of gas increased with temperature for all studied raw materials. The generated products percentages were quite similar for the stones, which reached a 78-79% of gas generation at 900°C. The nut shells achieved about 65-66% of gas at 900 °C. The yield of pyro-gas and char, as a function of the temperature, fitted a polynomial interpolation in all the cases, while the oil yield was described by a linear interpolation (Table 3).

Table 3. Derived analytical interpolations of the products yields.

	Char	Oil	Gas
Olive stones	$y = 0,0171x^2 - 0,1581x + 0,4854$ $R^2 = 0,9981$	$y = -0,0312x + 0,25$ $R^2 = 0,9987$	$y = -0,0166x^2 + 0,186x + 0,2686$ $R^2 = 0,9985$
Peach stones	$y = 0,0136x^2 - 0,1318x + 0,4386$ $R^2 = 0,9987$	$y = -0,0288x + 0,237$ $R^2 = 0,9982$	$y = -0,0119x^2 + 0,1503x + 0,3364$ $R^2 = 0,9998$
Walnut shell	$y = 0,0034x^2 - 0,0368x + 0,323$ $R^2 = 0,9991$	$y = -0,0288x + 0,2784$ $R^2 = 0,9947$	$y = -0,0011x^2 + 0,0531x + 0,4122$ $R^2 = 0,9974$
Hazelnut shell	$y = -0,0006x^2 - 0,0184x + 0,3202$ $R^2 = 0,9996$	$y = -0,0263x + 0,2649$ $R^2 = 0,9967$	$y = 0,0011x^2 + 0,0417x + 0,4184$ $R^2 = 0,9992$

The obtained coefficients of determination (R^2) were close to one in all cases, indicating a very good prediction capability of the model for all the studied raw materials.

As the goal has been to maximize the flow of the gas stream, the temperature of 800°C was selected to evaluate the heating value of the clean gas stream (S9) in order to be combusted and fed to a combined heat and power (CHP) system. At 800 °C, carbon dioxide CO₂ was found to be the predominant component for all studied raw materials, reaching a fraction of 23-29%, followed by hydrogen H₂ (19-26%), methane CH₄ (7-11%) and carbon monoxide CO (6-9%). The water content of this stream was about 15% in all the cases. The values of the flow and HHV of the obtained gas stream are presented in Table 4.

Table 4. Flows and energy content of the produced gas streams.

Biomass type	Pyro-gas		
	Mass flow (kg/h)	Volumetric flow (m ³ /h)	Higher Heating Value HHV (MJ/kg)
Olive stones	20.5	73.2	29.0
Peach stones	21.7	72.9	28.2
Walnut shell	16.4	60.7	16.1
Hazelnut shell	15.9	58.8	17.8

It can be seen that the stone-type raw materials generated a higher volume of gas with higher HHV so, for the purpose of generating electricity and thermal energy in a CHP system, these two raw materials would offer better performance. Nevertheless, the HHVs calculated for the shell-type raw materials were high enough to be considered as valid for the same purpose (Demirbas et al., 2006).

4. Conclusions

The simulation model for biomass pyrolysis created by (Agrifoglio et al, 2021) has been tested by using as input experimental data of four lignocellulosic wastes, two stone-type and two shell-type raw materials. The experimental data of the four tested raw materials were used as input in the simulation model and the results were accurate in all the cases with a good fit. The pyrolysis conditions were selected trying to maximize the gas stream and the obtained values indicated higher yields and HHV for the stone-type raw materials (about 20 kg/h and 28,5 MJ/kg, at 800 °C) even if the shell-type ones were found to generate a gas suitable to be used as feed stream of a CHP unit (59 16 kg/h and 17 MJ/kg, at 800 °C).

Acknowledgements

This research has been funded by the University of the Basque Country, UPV/EHU, (Project PES20/38).

References

- A. Adrados, A. Lopez-Urionabarrenechea, E. Acha, J. Solar, B.M. Caballero, I. de Marco, 2017, Hydrogen rich reducing gases generation in the production of charcoal from woody biomass carbonization. *Energy Convers. Manag.*, 148, 352–359.
- A. Agrifoglio, A. Fichera, A. Gagliano, R. Volpe, 2021, Energy analysis of a micro-cogeneration unit fed by biogas as a function of pyrolysis operating parameters, *Comptes Rendus. Chimie*, 24, S1, 39-55.

- M. Amutio, G. Lopez, R. Aguado, J. Bilbao, M. Olazar, 2012, Biomass oxidative flash pyrolysis: autothermal operation, yields and product properties. *Energy Fuels* 26, 1353–62.
- E. Benanti, C. Freda, V. Lorefice, G. Braccio, V. Sharma, 2011, Simulation of olive pits pyrolysis in a rotary kiln plant, *Therm. Sci.*, 15, 145–158.
- A.V. Bridgwater Review of fast pyrolysis of biomass and product upgrading. *Biomass Bioenergy* 2012, 38, 68–94.
- T.R. Brown, R. Thilakaratne, R.C. Brown, G. Hu, 2013, Techno-economic analysis of biomass to transportation fuels and electricity via fast pyrolysis and hydroprocessing, *Fuel*, 106, 463–9.
- F. Cheng, H. Bayat, U. Jena, C.E. Brewer, 2020, Impact of feedstock composition on pyrolysis of low-cost, protein- and lignin-rich biomass: A review. *Journal of Analytical and Applied Pyrolysis* 147, 104780.
- A. Demirbas, 2006, Effect of temperature on pyrolysis products from four nut shells, *J. Anal. Appl. Pyrolysis* 76, 285–289.
- A. Feduykhin, I. Sultanguzin, A. Gyul'maliev, V. Sergeev, 2017, Biomass Pyrolysis and Gasification Comprehensive Modeling for Effective Power Generation at Combined Cycle Power Plant, *Eurasian Chemico-Technological Journal* 19, 245–253.
- M.B. Gorenssek, R. Shukre, Ch.-Ch. Chen, 2019, Development of a Thermophysical Properties Model for Flowsheet Simulation of Biomass Pyrolysis Processes, *ACS Sustainable Chem. Eng.*, 7, 9, 9017–9027.
- H.B. Goyal, D. Seal, R.C. Saxena, 2008, Bio-fuels from thermochemical conversion of renewable resources: A review, *Renewable and Sustainable Energy Reviews* 12, 504–517.
- M.I. Jahirul, M. Rasul, A. Chowdhury, N. Ashwath, 2012, Biofuels production through biomass pyrolysis —A technological review, *Energies*, 5, 4952–5001.
- T. Kan, V. Strezov, T.J. Evans, 2016, Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters, *Renew. Sustain. Energy Rev.* 57, 1126–1140.
- S.S. Lam, R.K. Liew, X.Y. Lim, F.N. Ani, A. Jusoh, 2016, Fruit waste as feedstock for recovery by pyrolysis technique, *Int. Biodeterior. Biodegrad.* 113, 325–333.
- S.S. Lam, Y.F. Tsang, P.N.Y. Yek, R.K. Liew, M.S. Osman, W. Peng, W.H. Lee, Y.K. Park, 2019, Co-processing of oil palm waste and waste oil via microwave cotorrefaction: a waste reduction approach for producing solid fuel product with improved properties, *Process Saf. Environ. Prot.* 128, 30–35.
- J. Mabrouki, M.A. Abbassi, K. Guedri, A. Omri, M. Jeguirim, 2015, Simulation of biofuel production via fast pyrolysis of palm oil residues, *Fuel*, 159, 819–827.
- Y.X. Pang, Y. Yan, D.C.Y. Foo, N. Sharmin, H. Zhao, E. Lester, T. Wu, C.H. Pang, 2021, The influence of lignocellulose on biomass pyrolysis product distribution and economics via steady state process simulation, *Journal of Analytical and Applied Pyrolysis* 158, 104968.
- J.F. Peters, S.W. Banks, A.V. Bridgwater, J. Dufour, 2017, A kinetic reaction model for biomass pyrolysis processes in Aspen Plus, *Applied Energy*, 188, 595–603.
- E. Ranzi, A. Cuoci, T. Faravelli, A. Frassoldati, G. Migliavacca, S. Pierucci, S. Sommariva, 2008, Chemical Kinetics of Biomass Pyrolysis, *Energy Fuels*, 22, 4292–4300.
- E. Ranzi, M. Corbetta, F. Manenti, S. Pierucci, 2014, Kinetic modeling of the thermal degradation and combustion of biomass, *Chem. Eng. Sci.*, 110, 2–12.
- Q. Tang, Y. Chen, H. Yang, M. Liu, H. Xiao, Z. Wu, H. Chen, S.R. Naqvi, 2020, Prediction of Bio-oil Yield and Hydrogen Contents Based on Machine Learning Method: Effect of Biomass Compositions and Pyrolysis Conditions, *Energy & Fuels*, 34,9, 11050–11060.
- S. Wang, G. Dai, H. Yang, Z. Luo, 2017, Lignocellulosic biomass pyrolysis mechanism: A state-of-the-art review, *Prog. Energy Combust. Sci.* 62, 33–86.
- P.T. Williams, S. Besler, 1996, The influence of temperature and heating rate on the slow pyrolysis of biomass. *Renew. Energy*, 7, 233–50.
- R. Zanzi, K. Sjöström, E. Björnbo, 2002, Rapid pyrolysis of agricultural residues at high temperature, *Biomass Bioenergy*, 23, 5, 357–366.

Techno-economic-environmental analysis of biodiesel production by magnetic nanoparticles CLEAs of eversa® transform

Alves, E.S.^a, Miranda, L.P.^a, Guimarães, J.R.^a, Tardioli, P.W.^a, Giordano R.C.^a, Furlan F.F.^a

^a*Chemical Engineering Graduate Program, Federal University of São Carlos, UFSCar, São Carlos, São Paulo, Brazil*

Abstract

The urgency of reducing carbon emissions has been constantly reinforced by the scientific community. At the same time, there has been a constant growth in the global energy demand. In this sense, techno-economic-environmental studies are essential to assess the feasibility of new technologies, seeking a carbon neutral economy. The production of biodiesel through an enzymatic route has the potential to enable the manufacture of this biofuel, with a low environmental impact. The use of Crosslinked Enzymatic Aggregates (CLEAs) together with magnetic nanoparticles allows the reuse of the enzyme, possibly reducing operating costs. In this scenario, this work studied the production of ethylic biodiesel, by transesterification of soybean oil with bioethanol, catalysed by the enzyme Eversa® Transform 2.0, immobilized in the form of CLEAs. Two distinct purification processes were assessed: distillation and caustic polishing of biodiesel. Economic feasibility was assessed by the Net Present Value (NPV), with a rate of return of 11% per year, while environmental performance assessment was based on CML-IA midpoint indicators. As main results, the purification by distillation showed better techno-economic-environmental performance, since the distillation process requires fewer separation steps and material inputs, compared to the caustic polishing process. The global sensitivity analysis showed that all variables studied influence the environmental and economic indicator. The base case, with five reuses (value based on the literature), was insufficient to reach an economic performance superior to the current industrial process based on alkaline homogeneous catalysis. And the increase in the number of reuses is important to improve the environmental indicator as well.

Keywords: Biodiesel, LCA, Techno-economic analysis, eversa®transform.

Introduction

The growing global demand for energy, added to policies and legislation to regularize environmental issues, encourages the development of new bioenergy sources and processes (Tom, Sankaranarayanan, and Rodrigues 2019). Biodiesel is a biofuel that has many environmental and social advantages (Ajala et al. 2015). However, the production of biodiesel by chemical catalysis has problems related to the high cost of primary raw materials and many processes necessary for purification (Gebremariam and Marchetti,

2018). In this context, enzymes, which are biocatalysts, perform selective catalysis, which facilitates product purification and has lower energy consumption due to enzymatic reactions occurring under mild conditions of temperature and pressure. However, operational instability and difficulty in recovering free enzymes prevent the application of these biocatalysts in some industrial processes, despite all the advantages of their use (Garcia-Galan et al. 2011). One way around these barriers is to immobilize the enzyme. This technique results in greater chemical and thermochemical stability and contributes to protecting enzyme molecules from denaturation, in addition to enabling easy recovery and reuse of the biocatalyst (Santos, Puna, and Gomes 2020). Crosslinked enzyme aggregates (CLEAs) is an immobilization technique, with low execution complexity and high activity recovery (Amaral-Fonseca et al. 2018). CLEAs have low mechanical stability, which makes enzyme recovery difficult. To overcome this problem, an option is to attach the CLEAs particles to magnetic iron oxide nanoparticles. In this way, the biocatalyst separation can be performed through the use of an external magnetic field, avoiding enzyme compaction, as occurs in other separation methods such as filtration and centrifugation (Kopp et al. 2015).

Miranda et. al (2020) achieved satisfactory results in the production of biodiesel using Eversa® Transform 2.0 immobilized in CLEAs with magnetic nanoparticles in the transesterification of refined soybean oil. Using 12 Uest/g of oil enzyme load, 87.8 ± 0.7 wt % of fat acid ethyl ester (FAEE) yield was achieved after 12 h of reaction, in addition to the biocatalyst maintaining 89.6% of the biodiesel yield produced in the first batch, after five reuses (Miranda et al. 2020). The main objective of this study is to investigate the economic feasibility and environmental impact of biodiesel production catalyzed by Eversa® Transform CLEA with magnetic nanoparticles and its purification by two different routes: biodiesel distillation and caustic polishing with subsequent washing with water. The study was based on a process simulation involving all the unit operations necessary for the product to meet the international standards for biodiesel. A Sensitivity Analysis (SA) was also carried out to identify the process parameters that mainly influence the indicators of economic feasibility and environmental impact.

1. Methodology

1.1. Process Description

The production process of ethylic biodiesel was simulated based on the work of Miranda et al. 2020. Refined soybean oil is fed into the reactor, with excess of hydrous ethanol (oil/ethanol molar ratio 1:6) and the lipase immobilized in the form of CLEA in a vortex flow reactor. The reaction proceeds at 40°C for 12 h. In this step, two reactions occurs transesterification and hydrolysis. Then, the enzyme is recovered through a plate that generates a magnetic field, which collects the CLEAs that are washed with tert-butanol and reused in 5 batches. The solvent used to wash the catalysts is sent to a distillation column for purification and returns to the process. The first step in biofuel purification is decanting, which separates the raw glycerol from the light phase (mainly biodiesel). Then, both streams leaving the decanter are sent to flash for ethanol recovery. After going through the flash, the glycerol reaches a purity of 96%, of double-distilled glycerin. The ethanol goes to a distillation column for purification and returns to the reactor. At this stage, biodiesel already present low concentrations of glycerol, ethanol and water, but still contains a high concentration of Free Fatty Acid (FFA). For the separation of FFA, two process possibilities are tested. The first considers a caustic polishing as proposed in

the work of Miranda et. al (2020). The FFA is neutralized with a 4% (w/v) NaOH solution, the acid reacts with the base and forms FFA salts. A centrifugation step is performed mainly to remove the water added to the NaOH solution. The biofuel is recovered, washed with hot distilled water, and centrifuged (this protocol was repeated three times) to remove any traces of glycerol, alcohol, acid, base or soap that are still present in the biofuel. Then, the removal of water still present is carried out in a flash. The second strategy for FFA removal is a vacuum distillation column in order to remove triacylglycerides, raw materials and other components used in the separation steps (Al-Zuhair et al. 2011; Gebremariam et al. 2019; Sotoft et al. 2010).

The simulation of the process, as well as the calculation of economic feasibility and environmental assessment, were performed in the EMSO Software (Environment for Modeling, Simulation and Optimization). EMSO is an equation-oriented simulator, for academic, teaching and research that provides an environment for simulating and optimizing problems.

1.2. Kriging meta-model

To avoid convergence difficulties typical of steady-state simulations of rigorous phenomenological models of liquid-liquid equilibrium (Decantation) and Liquid-Vapor equilibrium (Distillation Column), metamodels were used to represent these stages (Carpio et al. 2018; Furlan et al. 2012). Therefore, the rigorous models of the distillation and decanter columns were simulated in Aspen Plus and were replaced by meta-models for simulation in EMSO. A universal Kring model was used, built using the "DACE" toolbox, code that adjusts and validates kriging models (Lophaven, Nuelsen, and Sondergaard 2002).

1.3. Economic and Environmental Assessment

The economic feasibility and life cycle assessment (LCA) equations were directly included in EMSO, integrated with the mass and energy balances of the process. Mass and energy balances were used to size the equipment, perform process cost calculations and calculate CO₂ emissions for the production process. This integration between simulation, economic feasibility and LCA facilitates further studies of the process, such as sensitivity analysis.

The data used in the evaluations carried out in this work was based on the Brazilian context in the year of 2020. The Plant was sized to process 17.6 ton/h of refined soybean oil. The metric for calculating economic viability was the Net Present Value (NPV), with a rate of return of 11% per year.

The LCA of this study used the CML Baseline 2000 method. Global warming potential (GWP 100) measured in kg of CO_{2eq} was the impact category chosen. Mass allocation was made between the two products of the process, biodiesel and glycerol.

1.4. Sensitivity Analysis

Sensitivity analysis can determine the parameters that have significant effects on the indicator of biodiesel production. Morris method was used with the following input variables: catalyst reuse number (5 - 30 reuse), transesterification yield (94.9 - 97.9%) and hydrolysis yield (0.87 - 1.41%). The output variables were the NPV and GWP. Yield intervals were selected according to experimental data and enzyme reuse was determined

extrapolating the range obtained experimentally. In the experimental planning, levels equal to 4 were applied and 400 calculations of elementary effects were performed.

2. Results and discussion

2.1. Economic viability

Using the results from material and energy balance together with the prices of raw materials, utilities, labor, and equipment, the techno-economic-environmental performances of the two technology options have been evaluated and presented in Table 1. The total cost of building the plant is US\$ 29.99 million for the purification of biodiesel via distillation and US\$ 30.13 million for caustic polishing. Other studies that use enzymatic catalysis point to an installation cost of 13.5 million for an annual production of 42 thousand tons (Gebremariam et al. 2019). For an annual production of 200 tons, Sotoft et al. (2010) reported an installation cost of 21 million for heterogeneous enzymatic catalysis and 30 million for homogeneous enzymatic catalysis. The values of the study presented were corrected by chemical engineering plant cost index (CEPCI) to the base year of the present work.

Table1: Result of the evaluation techno-economic-environmental of the production of biodiesel catalyzed by CLEA, for different purification methods.

Item	Purification Process	
	Distillation	Caustic Polishing
Annual Biodiesel Production (kg)	141,796,609	139,862,404
Annual Glycerol Production (kg)	13,265,238	13,280,848
Total capital investment cost (US\$)	28,503,448	33,750,667
Total Input (US\$)	914,427,719	915,003,449
Energy (US\$)	358,941	370,668
Utility cost (US\$)	11,210,988	11,758,530
Total Product Cost (US\$)	992,375,455	994,354,866
Revenue (US\$)	180,544,091	178,205,637
Net Present Value at 11% (US\$)	-5,575,617,290	-5,610,014,123
Biodiesel's GWP 100 (kg CO ₂ eq/kg)	2.68	3.73

Caustic polishing is more expensive to implement and operate, resulting in a higher total capital investment cost and total product cost, compared to Distillation. Additionally, the distillation process has a better process yield, as it requires fewer steps to reach the desired standard of purity, which causes less product loss during product purification. However, both purification alternatives have negative profit and, therefore, negative NPV which indicates that the investment is not economically feasible. The high cost of the process is mainly due to the cost of inputs and 85 % of the raw material cost is for the CLEA production. Therefore, it is necessary to increase the number of enzyme reuses, so that this process can become viable. For the process to be economically feasible, at least 475 reuses are required for the distillation and 956 for the caustic polishing process.

2.2. Environmental Assessment

The base cases including transesterification process and two different purification processes have been analyzed with the CML Baseline 2000. The GWP 100 results are shown in Table 1. The enzymatic reaction takes place under mild operating conditions and the absence of soaps and other residues, which makes it possible to apply separation processes with less impact. Vacuum distillation, despite having a higher energy consumption, does not use water and chemicals in the purification step. Therefore, it shows a lower global warming potential.

2.3. Sensitivity Analysis

The standardized elementary effects (SEE) and their mean absolute values were used to select which variables have little effect on the model's outputs. Table 2 shows the mean values (μ) of the elementary effects, the absolute value (μ^*), and the standard deviation (σ). When analyzing the values presented in Table 2, it is evident that the three selected variables have an influence on the NPV and only the hydrolysis conversion can be considered as a variable of low influence in the GWP 100 (for the caustic polishing).

Table 2: Sensitivity analysis results, based on Morris sensitivity analysis.

	NPV			GWP 100			
	μ^*	μ	σ	μ^*	μ	σ	
Distillation	Reuse Number	3.22e+8	1.71+7	8.53e+8	0.67	-0.032	1.712
	Transesterification Conv.	1.11e+9	1.30+8	1.70e+9	2.22	-0.253	3.40
	Hydrolysis Conv.	1.05e+9	-8.79e+8	1.66e+9	2.18	0.162	3.34
	NPV			GWP 100			
	μ^*	μ	σ	μ^*	μ	σ	
Caustic Polishing	Reuse Number	1.79e+9	1.79e+9	1.17e+8	3.58	-3.58	2.33
	Transesterification Conv.	2.84e+8	2.84e+8	1.70e+4	0.13	-0.13	0.05
	Hydrolysis Conv.	1.60e+5	-1.60e+5	4.95e+3	0.001	-0.001	0.001

It is observed, for the caustic polishing process, that the absolute values of μ and μ^* are equal, which indicates a monotonic behavior. Therefore, the sign of the effect of the input variables in the outputs remain the same in all space. This behavior is not observed in the distillation process, showing that the inputs have both positive and negative effects on the output variables. The values of sigma are similar in magnitude to μ and μ^* , which indicates that the SEE values varies substantially. This shows that there will be cross interactions between variables and non-linear behavior. It is important to emphasize that the interpretation of the indices is essentially qualitative and must be used for essentially for factor fixing. To determine the relative importance of the factors, a global sensitivity analysis based on variance must be used (Sin and Gernaey 2009).

3. Conclusions

The replacement of fossil fuels by ethylic biodiesel produced by enzymatic route has great potential for success. However, there are still challenges to be overcome, such as the cost

of enzymes. For the use of Eversa® Transform 2.0 lipase by crosslinked enzymatic aggregates (CLEAs) with magnetic nanoparticles to be viable, it is necessary to increase the number of reuses by 475 and 956 times, for distillation and caustic polishing respectively, or increase the specific catalytic activity of the biocatalyst. The separation of biodiesel produced by enzymatic route, due to the absence of soap and other contaminants, allows the use of different technologies, such as distillation, which proved to be more economical and environmentally friendly than the caustic polishing.

References

- Ajala, Olawale E., Folorunsho Aberuagba, Temitope E. Odetoeye, and Adejoke M. Ajala. 2015. "Biodiesel: Sustainable Energy Replacement to Petroleum-Based Diesel Fuel – A Review." *ChemBioEng Reviews* 2(3):145–56.
- Al-Zuhair, Sulaiman, Asma Almenhali, Iman Hamad, Maryam Alshehhi, Noura Alsuwaidi, and Suaad Mohamed. 2011. "Enzymatic Production of Biodiesel from Used/Waste Vegetable Oils: Design of a Pilot Plant." *Renewable Energy* 36(10):2605–14.
- Amaral-Fonseca, Murilo, Willian Kopp, Raquel de Lima Camargo Giordano, Roberto Fernández-Lafuente, and Paulo Waldir Tardioli. 2018. "Preparation of Magnetic Cross-Linked Amyloglucosidase Aggregates: Solving Some Activity Problems." *Catalysts* 8(11).
- Carpio, Roymel R., Felipe F. Furlan, Roberto C. Giordano, and Argimiro R. Secchi. 2018. "A Kriging-Based Approach for Conjugating Specific Dynamic Models into Whole Plant Stationary Simulations." *Computers and Chemical Engineering* 119:90–94.
- Furlan, Felipe Fernando, Caliane Bastos Borba Costa, Gabriel de Castro Fonseca, Rafael de Pelegrini Soares, Argimiro Resende Secchi, Antonio José Gonçalves da Cruz, and Roberto de Campos Giordano. 2012. "Assessing the Production of First and Second Generation Bioethanol from Sugarcane through the Integration of Global Optimization and Process Detailed Modeling." *Computers and Chemical Engineering* 43:1–9.
- García-Galan, Cristina, A. Berenguer-Murcia, Roberto Fernandez-Lafuente, and Rafael C. Rodrigues. 2011. "Potential of Different Enzyme Immobilization Strategies to Improve Enzyme Performance." *Advanced Synthesis and Catalysis* 353(16):2885–2904.
- Gebremariam, S. N. and J. M. Marchetti. 2018. "Economics of Biodiesel Production: Review." *Energy Conversion and Management* 168(April):74–84.
- Gebremariam, Shemelis N., Trine Hvosllef-eide, Meseret T. Terfa, and Jorge M. Marchetti. 2019. "Energies Techno-Economic Performance of Di Ff Erent Technological Based Bio-Refineries For." *Energies* 12(20):3916.
- Kopp, Willian, Felipe A. Silva, Lionete N. Lima, Sueli H. Masunaga, Paulo W. Tardioli, Roberto C. Giordano, Fernando M. Araújo-Moreira, and Raquel L. C. Giordano. 2015. "Synthesis and Characterization of Robust Magnetic Carriers for Bioprocess Applications." *Materials Science and Engineering B: Solid-State Materials for Advanced Technology* 193(C):217–28.
- Lophaven, S. N., H. B. Nuelsen, and J. Sondergaard. 2002. *A Matlab Kriging Toolbox, Version 2.0*.
- Miranda, Leticia Passos, José Renato Guimarães, Roberto Campos Giordano, Roberto Fernandez-Lafuente, and Paulo Waldir Tardioli. 2020. "Catalysts Transform and Magnetic Nanoparticles . Performance in the Ethanolysis of Soybean Oil." 10:9–11.
- Santos, Samuel, Jaime Puna, and João Gomes. 2020. "A Review on Bio-Based Catalysts (Immobilized Enzymes) Used for Biodiesel Production." *Energies* 13(11).
- Sin, Gürkan and Krist V. Gernaey. 2009. "Improving the Morris Method for Sensitivity Analysis by Scaling the Elementary Effects." *Computer Aided Chemical Engineering* 26:925–30.
- Sotof, Lene Fjerbaek, Ben Guang Rong, Knud V. Christensen, and Birgir Norddahl. 2010. "Process Simulation and Economical Evaluation of Enzymatic Biodiesel Production Plant." *Bioresource Technology* 101(14):5266–74.
- Tom, Rijo Jackson, Suresh Sankaranarayanan, and Joel J. P. C. Rodrigues. 2019. "Smart Energy Management and Demand Reduction by Consumers and Utilities in an IoT-Fog-Based Power Distribution System." *IEEE Internet of Things Journal* 6(5):7386–94.

Numerical Analysis of Impellers Hydrodynamics Performance in a Bioreactor CSTR with SPH

R. Murrieta-Dueñas^a, C.E. Alvarado-Rodríguez^b, J. Cortez-González^{a,*}, R. Gutierrez-Guerra^a

^a *Tecnológico Nacional de México / Instituto Tecnológico Superior de Irapuato (ITESI), Carretera Irapuato-Silao Km. 12.5, C.P.:36821 Irapuato, Guanajuato, México*

^b *CONACYT-Universidad de Guanajuato, Noria Alta S/N CP.36050, Guanajuato, Gto., México*

*corresponding author: jazmin.cg@irapuato.tecnm.mx

Abstract

In this paper we analyzed the hydrodynamic in a bioreactor which was simulated three-dimensionally applying the Computational Fluid Dynamics (CFD) using the Smoothed Particle Hydrodynamics (SPH) method. The developed model was validated through the comparison between the numerical and experimental data of the internal flow produced by a double Rushton stirred. The performance of the impellers analyzed in this work, named propeller, anchor and Rushton, were compared respect to the mixing time, fluid velocity profiles and mixing efficiency. In the SPH numerical simulations were considered the properties of the viscous fluid obtained in the xylitol production as density $\rho = 1240 \text{ kg/m}^3$ and viscosity $\eta = 1.587 \times 10^{-5} \text{ m/s}^2$. This simulation was performed using the DualSPHysics code based in the Smoothed Particle Hydrodynamics (SPH) method. The numerical results obtained in the validation test shown that the SPH method is capable to reproducing hydrodynamics in stirred tank bioreactors. We compared the hydrodynamic performance of three impellers: anchor, propeller and Rushton, in which it was possible to predict stagnation zones and speed profiles that guarantee a more homogeneous mixing. The optimal impeller for a stirring speed of 200 and 600 rpm is anchor and Rushton for 400 rpm.

Keywords: CFD, hydrodynamics bioreactor, SPH.

1. Introduction

The biotechnology industry has become very important in recent years, due to the high demand for products derived from biological systems. The bioreactor is the main unit process and its design must fulfill certain requirements that are needed to ensure the fermentation process efficiency, such as: the stirring speed, the aeration rate, the heating intensity or cooling rate, and the nutrients feeding rate. The homogeneity of dissolved oxygen and substrate concentration in the bioreactor is set by the impellers.

Biotechnological process depends on the biological reactor in particular on the homogenization of the substrate, oxygen and biomass inside the tank. Therefore, it is important to analyze the hydrodynamics of biological reactors, especially the stirred tank reactors used traditionally in fermentation processes.

The Computational Fluid Dynamics (CFD), is a very useful tool to analyze in detail the hydrodynamics and mixing in bioreactors of stirred tank used in microbial fermentation processes (Gelves et al., 2013). The CFD allows the numerical solution of the transport equations by computational means. It can generate a huge amount of information that in practice could not be obtained experimentally, or they are totally inaccessible by the experimental route, also is an effective strategy to improve or ensure the performance of a process. This tool allows determining axial and longitudinal flow patterns, velocity profiles, mixing times, stagnation points or vortex formation. The CFD has been used for description of mixing in bioreactors by (Tajsoleiman et al., 2019; Delafosse et al., 2014) where they propose a compartment model where the flow rates between two adjacent compartments are easily computed from the velocity fields obtained by CFD. Moreover it has been reported studies of hydrodynamic shear stress generated by different impeller combinations in stirred bioreactors and hydrodynamic performance of coaxial mixers in the mixing of yield-pseudoplastic fluids (Kazemzadeh et al., 2016). The evaluations of gas-liquid mass transfer, oxygen uptake, and dynamic oxygen distribution were reported by (Stickel et al., 2018).

The purpose of this paper is to know the distribution of nutrients within the fermentation process, as well as to determine homogeneity and velocity profiles in order to guarantee the hydrodynamics that promotes mass transfer in three different impellers traditionally biotechnology used anchor, propeller and Rushton.

2. Methodology

The analysis has focused on the hydrodynamics of the tank by means of CFD using the SPH method to describe the flow patterns and velocity profiles inside the reactor with the objective of obtain data that help to decrease the mixing time and minimizing or eliminating stagnation zones.

For this analysis we first performed the numerical validation of the SPH method by comparing the results with those obtained experimentally by De Lamotte et al., (2017). For validation, two Rushton turbines with 4 blades and a diameter of 0.1 m were used for agitation according to the dimensions reported by De Lamotte et al., (2017) in their experiment. The coupling of the momentum equation and the continuity equation was performed. Stirring speeds $\omega = 200\text{rpm}$, fluid density $\rho = 998\text{ kg/m}^3$, kinematic viscosity $\eta = 1 \times 10^{-6}\text{ m}^2/\text{s}$ and a total of 1,250,000 particles making up the fluid, tank and agitator were considered.

Subsequently, we chose 3 impellers: anchor, propeller, and Rushton with a stirring speeds (ω) 200, 400 and 600 rpm and then the hydrodynamic performance and mixing times for each impeller was analyzed. Finally, for each stirrer we determined the mixing time at the stirring speed that showed the highest homogeneity. For the simulations of the hydrodynamics of the stirred tank bioreactor, the following parameters were considered: density $\rho = 1240\text{ kg/m}^3$, kinematic viscosity $\eta = 1.587 \times 10^{-5}\text{ m}^2/\text{s}$, and a total of 500,000 particles in a time of 15 seconds. The height-diameter (h/D) scale ratio of the reactor is 1.5 times. The sizing of the bioreactor and impellers was performed in SolidWorks®, the total volume of the tank is 5 L, with a height of 0.33m, a diameter of tank 0.16m and with 4 baffles to reduce stagnation zones. The impellers have a height of 0.27m.

2.1. Impellers

Stirring and mixing is an operation used to accelerate heat and mass transfer processes. This operation introduces energy into the volume of the fluid, i.e. the kinetic energy of the agitator is transferred to the fluid causing it to move. This operation involves systems that are in a single phase or in several liquid, solid and gaseous phases. For this purpose, agitators have been designed to generate parallel flows to the impeller axis, which are called axial flow impellers; and some that generate flows in radial-tangential direction, which are called radial flow impellers.

The propeller is an axial flow agitator and is used for liquids that are not very viscous. The speed of these impeller varies depending on their size, between 400 and 800 rpm for large propellers and between 1150 and 1750 rpm for the small ones. One characteristic of the propeller is that its blades vigorously shear, or friction, the liquid (Ramirez, 2012). These agitators are used to homogenize, suspend fluids, as well as increase the heat exchange. The propeller agitator used in this study has a diameter $d = 0.06\text{m}$, and a height $h = 0.27\text{m}$ (Figure 1a).

The anchor is a radial agitator, composed of paddles rotating at low to moderate speeds in the center of the tank, impelling the liquid radially and tangentially, with no vertical movement with respect to the agitator, unless the paddles are inclined. The flows are directed towards the tank wall and then continue upward or downward. The anchor-type agitator in this study has a diameter $d = 0.06\text{ cm}$, a height $h = 0.27\text{m}$, and has a deflector plate at a 45° angle (Figure 1b).

The Rushton impeller consist of a driving component with more than four blades, mounted on the same element and fixed to a rotating shaft. They are effective for a wide range of viscosities. When used in low viscous liquids, they produce strong flows, which spread throughout the tank and destroy stagnant liquid masses. The Rushton turbine is ideal for fermentation (Verdugo, 2013). Rushton propeller blades are flat and positioned vertically along the agitation axis, producing a unidirectional radial flow. Two Rushton 4-blade turbines with a diameter of 0.1 m were used for agitation (Figure 1c).

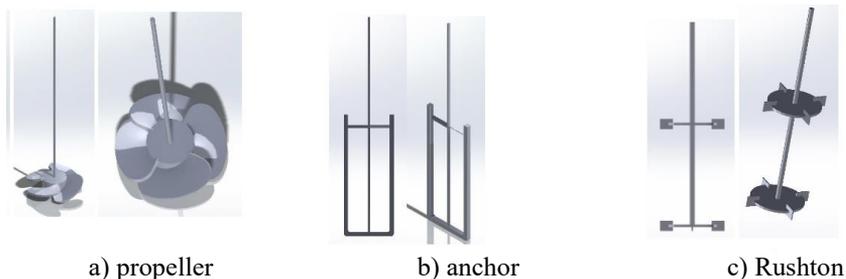


Figure 1. Design of impellers.

2.2 The SPH method

SPH is a Lagrangian, meshless method with applications in the field of Computational Fluid Dynamics. Originally invented for astrophysics in the 1970s (Monagan J, 1992) it has been applied in many different fields, including fluid dynamics (Alvarado-Rodríguez C.E. et al., 2019) and wastewater treatment (Mokos et al., 2015). The method uses points

named particles to represent the continuum and these particles move according to the governing equations in the fluid dynamic. When simulating free-surface flows, no special surface treatment is necessary due to the Lagrangian nature of SPH, making this technique ideal for studying violent free-surface motion. Moreover, the movement of the boundaries can be set easily without the necessity of update the mesh as in the Eulerian methods.

The SPH formalism used in the simulations is reported by (Dominguez et al., 2021) which is set in the DualSPHysics code, in this work only the continuity (Eq. 1), the momentum (Eq. 2) and equation of state (Eq. 3) in the SPH formalism are reported.

$$\frac{dv_a}{dt} = -\sum_b m_b \left(\frac{P_a + P_b}{\rho_a \rho_b} + \sum_b m_b \left(\frac{4v_0 r_{ab} \cdot \nabla_a W_{ab}}{(\rho_a + \rho_b)(r_{ab}^2 + \eta^2)} \right) v_{ab} + \sum_b m_b \left(\frac{\tau_{ab}^j}{\rho_b^2} + \frac{\tau_{ab}^i}{\rho_a^2} \right) \nabla_a W_{ab} \right) \nabla_a W_{ab} + g, \quad (1)$$

$$\frac{d\rho_a}{dt} = -\rho_a \sum_b \frac{m_b}{\rho_b} (v_b - v_a) \cdot \nabla_a W_{ab}, \quad (2)$$

$$P = B \left[\left(\frac{\rho}{\rho_0} \right)^\gamma - 1 \right], \quad (3)$$

where the subscripts a and b are denoted for the mean particle “ a ” and the neighbors particles “ b ”, v is the velocity, t is time, m is mass, P is pressure, ρ is density, v_0 is the kinematic viscosity, τ is the stress tensor, $B = c_0^2 \rho_0 / \gamma$, c_0 is an artificial sound speed, and $\gamma = 7$, and W is the kernel function defined in the SPH method.

3. Results

3.1 Validation of SPH method

The numerical results from the simulation were validated by comparing with the experimental results reported by De Lamotte et al., (2017). The Reynolds number, tank velocity fields and velocity variation were calculated for the case with a stirring speed $\omega = 300$ rpm. The velocity profile obtained by the SPH method versus that obtained experimentally is shown in the Figure 2. The flow distribution is similar, forming two vortices inside the tank similar to the vortices obtained experimentally, likewise the velocity field is in the same range of velocity values (0-2 m/s). With this numerical results the SPH method is validated as a tool for the simulation of the hydrodynamics of stirred tanks.

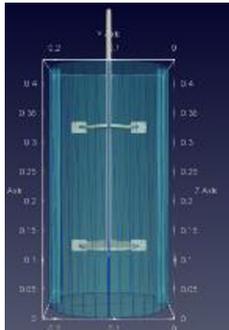


Figure 2. Sizing of the tank used for the simulations.

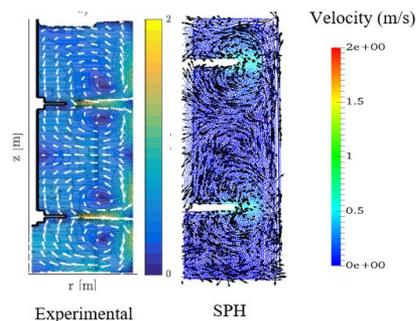


Figure 3. Comparison of velocity profile and vortex formation experimental vs SPH simulation

3.2 Impellers performance

We study the hydrodynamic performance in each speed stirring and each impellers studies for to evaluate possible stagnation zones through of flow pattern impellers. We chose three specific zones within the reactor to evaluate fluid movement within the tank: top, bottom and middle of the biological reactor. The numerical results of the SPH method of propeller, anchor and Rushton impellers for stirring speed of 200 rpm are shown in the Figure 4. This analysis was also performed for speeds of 400 and 600rpm for all stirrers.

In the propeller impeller according to the results obtained from the analysis of the flow stabilization as a function of time (selected points inside the biological reactor), the flow is not homogenized at 15 seconds using 200 rpm, while at 400 and 600 rpm it is stabilized in this time. Addition, with 200 rpm there is no movement in the upper part of the reactor, which suggests a deficient mixing in the tank (see Figure 4a). The other speeds show a better homogenization of speed inside the tank.

According results on the hydrodynamic performance of the anchor impeller, the circulation pattern manifests a radial recirculation behavior in liquid and a turbulent regime at all three speeds. No stagnation zones are shown in any of the stirring speeds studied. In the Figure 4b shown the pattern flow in anchor impeller at 200rpm. The analysis of the three points inside the reactor show flow stabilization time, obtained that for the stirring speeds at 200 and 400 rpm achieved stabilization in 9 and 11 seconds respectively, while the 600 rpm stirring speed needed more than 13 seconds.

In the Rushton impeller the results of the three points of the homogeneity of speed as a function of time shown that the circulation pattern manifests a radial recirculation behavior in liquid and a turbulent regimen (see the top in the Figure 4c). The minimum flow stabilization time in the tank is observed at 400 rpm. For Rushton impeller, there are stagnation points around the blades at the three speeds analyzed.

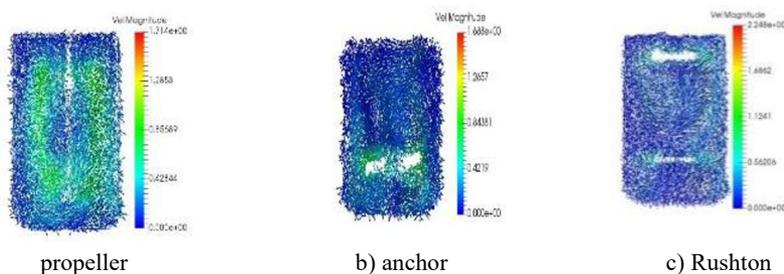


Figure 4. Flow pattern impellers, $\omega=200\text{rpm}$

4. Conclusions

This work presents the numerical validation of the SPH method and its application for the study of three different impellers used in the biotechnological industry: propeller, anchor and Rushton at three different stirred speeds, 200, 400 and 600 rpm. The validation shows good agreement between the numerical and experimental results, so it is concluded that the SPH method is a powerful and versatile tool for the analysis and design of stirred tank bioreactors. According to the results, the best impeller was the Rushton at stirring speed

of 400 rpm, because it does not present vortices and its flow stabilization time is 10 seconds. The propeller impeller shows a low flow stabilization time in the tank requiring more than 15 seconds for a speed of 200 rpm. The anchor impeller shows a similar performance to the Rushton, with flow stabilization time of 9, 11 and 13 seconds for 600, 400 and 200 rpm respectively. In addition, with the analysis of the hydrodynamics of stirred tank bioreactors in SPH, it was possible to predict stagnation zones and velocity profiles that guarantee a more homogeneous mixing.

Acknowledgements

The calculations in this work were partially supported by ABACUS under the CONACyT project EDOMEX-2011-C01-165873. The ABACUS project is thanked for the availability of computational resources.

We are grateful to the Tecnológico Nacional de México for the support for the development of the project "Hydrodynamic analysis of CSTR biological reactors using SPH".

References

- Alvarado-Rodríguez, C. E., Klapp, J., Domínguez, J. M., Uribe-Ramírez, A. R., Ramírez-Minguela, J. J., & Gómez-Gesteira, M. (2019, March). Multiphase Flows Simulation with the Smoothed Particle Hydrodynamics Method. In International Conference on Supercomputing in Mexico (pp. 282-301). Springer, Cham.
- De Lamotte, A., Delafosse, A., Calvo, S., Delvigne, F., & Toye, D. (2017). Investigating the effects of hydrodynamics and mixing on mass transfer through the free-surface in stirred tank bioreactors. *Chemical Engineering Science*, 172, 125-142.
- Domínguez JM, Fourtakas G, Altomare C, Canelas RB, Tafuni A, García-Feal O, Martínez-Estévez I, Mokos A, Vacondio R, Crespo AJC, Rogers BD, Stansby PK, Gómez-Gesteira M. (2021), DualSPHysics: from fluid dynamics to multiphysics problems. *Computational Particle Mechanics*. doi:10.1007/s40571-021-00404-2.
- Gelves, R., Benavides, A., & Quintero, J. C. (2013). Predicción del comportamiento hidrodinámico en el escalado de un reactor de tanque agitado para procesos aerobios, mediante CFD. *Ingeniería. Revista chilena de ingeniería*, 21(3), 347-361.
- Mokos, A., Benedict, D. R., Stansby, K. P., & Dominguez, J. M. (2015). Multi-phase SPH modelling of violent hydrodynamics on GPUs. *Journal Computer Physics Communications*, 196, 304-31.
- Monagan J J, "Smoothed Particles Hydrodynamics", *Annu. Rev. Astron. Astrphys.*, 30:534-74, 1992.
- Verdugo, L. R. (2013). Diseño y cálculo de un agitador de fluidos. Facultad de ingeniería. Departamento de ingeniería mecánica. , 78-91.
- Tajsoleiman, T., Spann, R., Bach, C., Gernaey, K. V., Huusom, J. K., & Krühne, U. (2019). A CFD based automatic method for compartment model development. *Computers & Chemical Engineering*, 123, 236-245.
- Delafosse, A., Collignon, M. L., Calvo, S., Delvigne, F., Crine, M., Thonart, P., & Toye, D. (2014). CFD-based compartment model for description of mixing in bioreactors. *Chemical Engineering Science*, 106, 76-85.
- Kazemzadeh, A., Ein-Mozaffari, F., Lohi, A., & Pakzad, L. (2016). Investigation of hydrodynamic performances of coaxial mixers in agitation of yield-pseudoplastic fluids: Single and double central impellers in combination with the anchor. *Chemical Engineering Journal*, 294, 417-430.
- Stickel, J. J., Sitaraman, H., Rahimi, M. J., & Humbird, D. (2018). *Computational Fluid Dynamics Study of Full-Scale Aerobic Bioreactors: Evaluation of Gas Liquid Mass Transfer, Oxygen Uptake, and Dynamic Oxygen Distribution* (No. NREL/JA-5100-68269). National Renewable Energy Lab.(NREL), Golden, CO (United States).

On the use of embedded models and advanced analytics to model complex processes in the cement industry

Alexandros Piladarinos,^a Antonis Kokossis,^{a,c} Ioannis Marinos,^b Thanassis Gentimis^c

^a*School of Chemical Engineering, National Technical University of Athens, Zografou Campus, 9, Iroon Polytechniou Str. GR-15780, Athens, Greece*

^b*TITAN SA, 22A Halkidos Str., Athens 11143, Greece*

^c*SYMBIOLABS Circular Intelligence LC, Artemidos 6 & Epidavrou str., Marousi 15125, Greece*

Abstract

The paper explains a generic and systematic approach in the development of embedded models that could be further used for model reduction. The systems approach makes a structured and systematic use of data as they are produced at three distinct stages: simulation assignments by means of spatial differential equations, optimization runs that regress parameters for each simulation, and deep learning training that converts parameters into functions of system variables. Simulation models refer to steady-state operations of closed-circuit grinding models formulated as differential equations with parameters treated as degrees of freedom. Results from this implementation present a consistent accuracy improvement over the model used as basis.

Keywords: Embedded Model, ANN, Deep Learning, Cement Grinding, Ball Mill

1. Introduction

Raw material and clinker grinding are energy intensive processes, consuming 70% of the electricity demand in a typical cement plant, which cumulatively amounts to 2% of global electricity consumption [1]. Due to energy costs, emission restrictions and tight product quality specifications the cement industry is challenged to perform at the highest possible standards; robust and reliable technology in optimization, control and sensor software are important enablers to secure high performance. Still, the industrial processes in the cement industry involve complex multi-phase flow dynamics modelled by breakage distribution function with the rate of breakage determined by probabilistic approach intended to match experimental measurements. The cement industries hold large records of data generated from long times of operation and service in their industrial processes. Deep learning models are already tested successfully to predict environmental performance, even in the absence of first principles [2]. However, one is further challenged to capitalize on the wide range of conventional, yet insightful, models as they are popular for grinding and thermal conversion processing. Model complexity is tackled by limiting assumptions that prevent such models from accuracy and restrict their use to simulation rather than to support decisions and/or assist in the re-engineering of the industrial process. A systematic and combined use of first-principles models with operational data holds a strong promise. In exploiting the underpinning engineering principles supported by the models, one is hopeful to achieve a better analysis of computations and a direct interpretation of their results.

The paper explores the systematic development of embedded models using a ball mill grinding as a pilot. The analytical model is first presented for the ball mill circuit. A conventional approach would directly regress operational data to model parameters. Instead, the paper outlines a methodology where machine learning is applied to convert parameters as functions of process variables. Using a combined stream of model-based and of operational-based datasets, a deep learning model is trained and produced by means of an embedded model. The model capabilities include options to invert the model as well as options to expand it with decision parameters for process design and reengineering.

2. Presentation of first-principle models and parameters

The set of first principles models is presented first alongside parameters to train and connect with real-life data. The feed is considered to consist of a dry mix of clinker, pozzolan, gypsum, limestone and fly-ash, ingredients of common Portland cement. The cement grinding circuit consists of a finishing ball mill that is connected to a SEPAX rotor classifier. The models for the ball mill and the classifier are presented next.

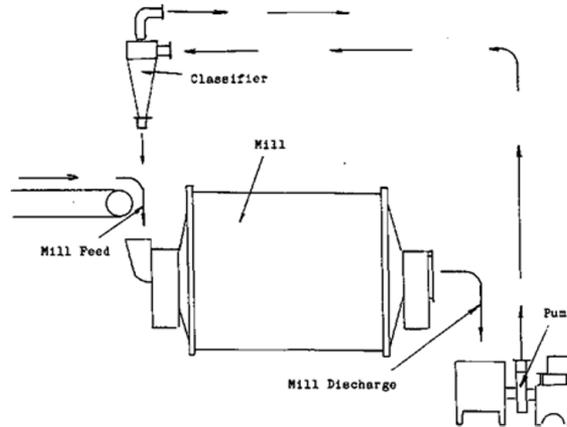


Figure 1: Diagram of a ball mill grinding circuit.

2.1 Ball Mill Model

A first principles model has been presented by Boulvin, *et al.*, 1999 [3]. The model considers cement behaviour inside the mill as solid flow of discretely sized particles. The modelling equations consist of a system of PDEs describing mass flow through convection and diffusion, and particle diameter reduction through breakage. The system of PDEs solve for linear density of cement particles, u_i for each particle size interval, axially distributed over the length of the mill:

$$-v_i \frac{\partial u_i}{\partial z} + D_i \frac{\partial^2 u_i}{\partial z^2} + \sum_{j=1}^9 b_{ij} s_j f(u_j) = \frac{\partial u_i}{\partial t} \quad i \in \mathbb{N}, i \in [1,10] \quad (1)$$

Along with the appropriate boundary conditions:

$$D_i \frac{\partial u_i}{\partial z} \Big|_{z=0} - v_i u_i(z=0) = -q_f w_{fi} - q_r w_{r,i} \quad (1a)$$

$$\frac{\partial u_i}{\partial z} \Big|_{z=L_2} = 0 \quad (1b)$$

Comminution is modelled using the Breakage and Selection function standard. Such models make use of empirical equations to determine the specific rate of breakage of each particle, s_i and the distribution of the diameters, d_i of produced particles, b_i .

$$s_i = ad_i^a \quad (2)$$

A cut-off variation of the Broadbent and Callcott equation [4] is used as breakage function to keep regression parameters to a minimum:

$$B(d_i) = 1 - (1 - d_i/d_1)^{c_3} \quad (3)$$

$$b_i(d_i) = B_i(d_i) - B_{i+1}(d_i) \quad (3a)$$

c_3 expresses the grindability of the material mix. Harder materials tend to break in larger lumps of particles instead of shattering. Linear dependency between the content of the clinker and cement grindability are assumed following.

$$c_3 = c_1 m + c_4 \quad (3b)$$

The effectiveness reduction due to the impact energy cushioning effect of fine holdup accumulation is also accounted for, in the form of the following function. [5]

$$f(u) = u(z)_i e^{-c_n * H(z)} \quad (4)$$

H accounts for the total cement particles mass holdup (of all sizes). Blaine measurements assume cement particles of spherical shape and that Blaine is proportionate to their total surface area. Likewise, product density ρ , is considered constant and independent of the cement composition.

$$Bl = \frac{6c}{\rho} \sum_{i=1}^{10} d_i w_i \quad (5)$$

q_f , $w_{f,i}$ and q_r , and $w_{r,i}$ respectively denote mass flow and weight fractions of i particle size interval in the feed and the recycle streams; they are linear expressions of $\varepsilon(x_i)$ that, in turn, is set by the classifier type. The recycle stream consists of classifier rejects following:

$$q_r w_{r,i} = \varepsilon(x_i) u_i \quad (6)$$

2.2. Classifier Model

The Tromp curve of a SEPAX rotor classifier has a distinct shape which is mostly independent of the operation variables. The classifier is commonly modelled using a Rosin-Rammler model fitted closely to experimental data in order to replicate the characteristic “fish-hook” curve of selectivity over particle diameters. A polynomial equation is fitted for each segment of the curve to match real-life performance.

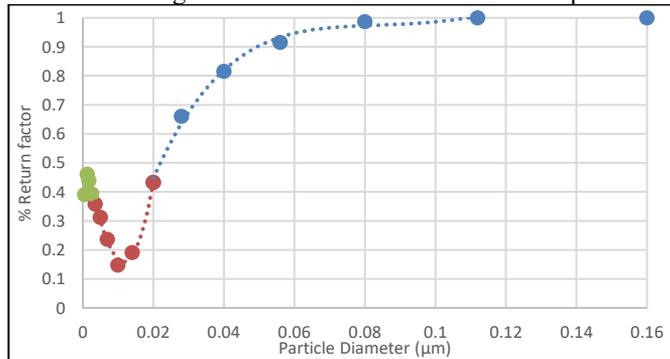


Figure 2: Classifier selectivity for fixed classifier speed and airflow.

The rotor selectivity depends mainly on the feed-rate and the rotating speed. Higher rotating speed increases the classifier’s affinity to smaller sized particles. In essence, as the rotating speed of the classifier varies, the selectivity curve shifts horizontally in a linear manner. To replicate such a behaviour, an effective diameter, z , of the product particles is developed as follows:

$$z_i = x_i + c_2(v_s - v_o) \quad (7)$$

3. Outline of methodology

Given are

- (i). First principles models ($\Sigma 1$ & $\Sigma 2$) as described by (a) the Ball mill model, $\Sigma 1$: Eq. (1-5); (b) the Classifier model, $\Sigma 2$: Eq. (6-7)
- (ii). Sets of real-life data (DS1): This is a dataset of a cement plant ball mill circuit's operation. The dataset originates from a log of hourly semi-steady state measurements for the whole of 2020. Of the log's measurements, DS1 contains: (a) Ball mill dry feed-rate, q_f ; (b) clinker content, m ; (c) Separator speed v_s ; (d) Blaine specific surface, Bl
- (iii). Sets of model parameters (PS1 & PS2). They involve regression parameters within the model $\Sigma 1$ & $\Sigma 2$: $PS1 \equiv \{a, c_1, c_2, c_3, c_4, c_n, c\}$; and a parameter to regress to single data points of DS1: $PS2 \equiv \{c_1\}$

The objective would be the development of an embedded model to convert parameter c_1 as functions of $\{q_f, m, v_s\}$.

The proposed methodology involves the following stages:

Stage 1: Model regression stage where regression parameters are fitted to in-site data.

In this stage, parameters PS1 in integrated model $\Sigma 1$ & $\Sigma 2$ are regressed to plant operation data DS1 by minimizing the objective function:

$$\Phi = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \min_{\theta} \left\{ \sum_{i=1}^{NE} \sum_{j=1}^{NV_i} \sum_{k=1}^{NM_{ij}} \left[\ln(\sigma_{ijk}^2) + \frac{(Bl_{ijk} - \bar{Bl}_{ijk})^2}{\sigma_{ijk}^2} \right] \right\} \quad (8)$$

Stage 2: Data development stage where selected regression parameters are produced to train the model. In this stage, input-output data for machine learning are produced using (a) the set of data from the plant DS1; and (b) outputs from the integration of $\Sigma 1$ and $\Sigma 2$. Dataset DS2 is produced as values of regressed parameters PS2 by minimizing objective function (8).

Stage 3: Neural network training stage using real data inputs and regressed model parameters. The selected parameters are turned to functions of input. In this stage, parameters PS2 are converted to functions of the process input through an ANN trained with (a) the physical input from DS1 and (b) the regressed parameters dataset DS2 produced in (1). $PS3 = f(DS1, DS2)$

Stage 4: Embedded model validation and testing.

4. Implementation

Two identical models of the ball mill, $\Sigma 1$ representing each compartment of the mill are integrated with classifier's model $\Sigma 2$. Simulations of the circuit model $\Sigma 1$ & $\Sigma 2$ are run for steady state operation on gPROMS Model Builder 6.0.4. Parameters PS1 of the model, are regressed to in-site dataset DS1 by minimizing the error of Blaine prediction Eq. (8). Parameters in PS2 of the regressed model (stage 1) are repetitively regressed to each of the individual data points of dataset DS1. A large number of regressions taking account of a wide range of the input values (q_f, m, v_s) take place in this step. Both en masse (stage 1) and individual point regression take place within gPROMS Model Builder Model's Validation Module, which makes use of maximum likelihood estimation algorithm (MLE) for minimizing the objective function value Eq. (8).

Input values (q_f, m, v_s) from DS1 and the complete DS2 dataset is used to train a feedforward ANN in order to determine the function by which they correlate with each other, $c_1 = c_1(q_f, m, v_s)$. The ANN DS1&DS2 are used to train, consists of two 140 neuron hidden layers and a "ReLU" activation function applied to both of them. The output layer instead has a linear activation function. Network training is coded on python using the Tensorflow module and the API Keras and done through "Adam", a stochastic gradient descent optimizer, updating the parameters as follows:

$$\theta_t = \theta_{t-1} - a \frac{\nabla(\theta_{t-1})}{\sqrt{v_t + \epsilon}}$$

While mean squared error (MSE) is used as error function. Then, weights and biases of the trained neurons are acquired and used to replicate the neural model as functions in the gPROMS Model Builder environment. Function, $c_1(q_f, m, v_s)$ replaces the respective constant parameter c_1 of the analytical model, thus forming an embedded one. The Blaine prediction accuracy of the embedded model as well as that of the analytical model with regressed parameters PS1 from stage (1) is tested. The testing of the models' performance occurs in two stages, one for interpolation and one for extrapolation. The testing dataset used for interpolation is picked to match the range and standard deviation of inputs & output in DS1. The extrapolation testing dataset consists of inputs (q_f, m, v_s) . Inputs (q_f, v_s) in specific, are outside of the training dataset's DS1 range, but within a 20% extension from it in either direction. In each testing scenario, the performance of the two models is compared to that of a pure ANN model directly trained to DS1. This ANN consists of 3 hidden layers, the first one containing 128 and the other two 256 neurons each. Both the pure and the integrated ANNs were trained with a low constant learning rate of 0.00001 and a batch size of the entire training sets as training time was not crucial. Both were trained with 2033 data points 20% of which were used for testing. The number of layers and their respective neurons, as well as the number of training cycles were tuned to the end of minimizing the training and testing sets' mean average error without them significantly diverging from one another.

5. Results

All Models' Blaine Prediction Accuracy Comparison:

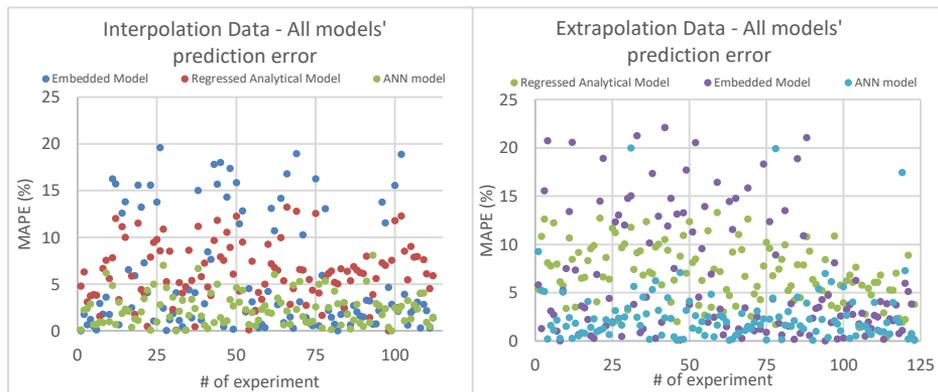


Figure 3a,3b: Comparison of absolute error of measured Blaine of all models on both interpolation and extrapolation validation datasets.

Metric	Pure ANN model/ Interpolation Data	Analytical Model/ Interpolation Data	Embedded Model/ Interpolation Data	Pure ANN model/ Extrapolation Data	Analytical Model/ Extrapolation Data	Embedded Model/ Extrapolation Data
Mean Absolute Percentage Error	2.31%	6.18%	5.57%	2.46%	6.91%	6.37%

Table 1: Comparison of error metrics of all models on both interpolation and extrapolation validation datasets.

While the embedded model outperforms the analytical model on both interpolation and extrapolation performance, it accounts for the largest deviations on both.

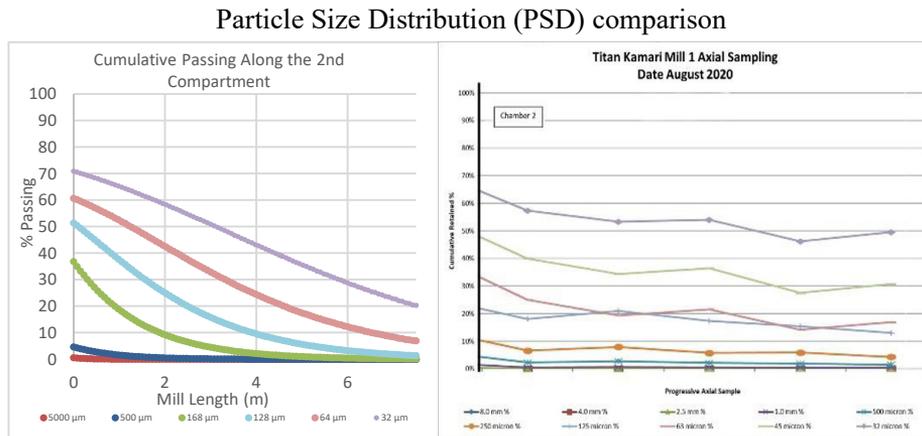


Figure 4a,4b: Comparison of simulated and experimental variability of cumulative passing along the second compartment of the mill.

As no data are available of the conditions the circuit was operating in at the time of sampling, simulations were run using ordinary routine input values. Axial variability of the particle size distribution provided by the embedded model is in agreement with of axial sampling data of the mill, as seen at figure 4a, 4b. Coarse particle passthrough in the second compartment of the mill is estimated by the model to be 4.7% while measured value is 5%. The particle diameter where 50% of passing on mill exit occurs is at 14 μ m while the simulated one is estimated at 32 μ m. This deviation may be a result of sparse discretization of particle sizes.

6. Conclusions

Embedded model is able to achieve higher prediction accuracy in both interpolation and extrapolation conditions while retaining its first principles virtues as the particle size distribution of cement both inside the mill and the product are in agreement with measured values. Further performance increase is expected with the use a larger “input – parameter” training dataset. This methodology is one that would benefit most analytical models fast enough to run enough regressions for the ANN training to be feasible.

Acknowledgement

The authors are grateful to TITAN Greece for the in-site data provided.

References

1. F. Akbar, Ardi E., V. Alireza, H. Vajihe, (2014). Simulation of clinker grinding circuits of cement plant based on process models calibrated using GA search method. *Journal of Central South University*, 21.
2. C. Chatzilenas, T. Gentimis, T. Dalamagas, A.C. Kokossis, A. Katsiaboulas, I. Marinos, 2021. Machine learning applications and process intelligence for cement industries. *Computer Aided Chemical Engineering*, Elsevier, Volume 50.
3. Boulvin M., C. Renotte, A. V. Wouwer, M. Remy, S. Tarasiewicz and P. César 1999. Modeling, simulation and evaluation of control loops for a cement grinding process. *European Journal of Control*, 5.
4. S.R. Broadbent and T.G. Callcott, 1956. *Philosophical Transactions of the Royal Society of London*, A249, 99.
5. Lepore, Renato & Vande Wouwer, Alain & Remy, Marcel & Bogaerts, Philippe. (2004). State and Parameter Estimation in Cement Grinding Circuits - Practical Aspects. *IFAC Proceedings Volumes*, 37.

Aspen Plus® modeling approach of beechwood gasification in a fluidized bed reactor using biochar as bed material

Leonela Martes Hernández, Minda Loweski Feliz, Luis Reyes Alonzo, Lokmane Abdelouahed, Bechara Taouk

Normandie Univ, INSA Rouen Normandie, UNIROUEN, LSPC-Laboratoire de Sécurité des Procédés Chimiques, 76800, Saint Etienne Du Rouvray, France

Abstract

Biomass processing is considered a strategy changer in the development of new pathways for green energy production, being the thermochemical processes as pyrolysis and gasification part of the main methods used to produce renewable fuels (Mohabeer et al., 2019). The purpose of this work is to perform a simulation of the gasification process of biomass comparing the use of sand and biochar as bed material with the software Aspen Plus®. The simulation is supported and validated by experimental results collected on a pilot fluidized bed. The model includes a biomass pretreatment process and the gasification where the gasifier (a RYield reactor) is supported by a set of correlations obtained in function of the different gasification temperatures. This process allows a good reproduction of the pilot gasifier outlet. An analysis of the effect of the use of biochar as a bed material over the tar reduction on the producer gas is made, applying different gasification agents (H₂O and CO₂) and a temperature variation (600-900°C).

Keywords: Biomass, Gasification, Aspen Plus, Biochar, Fluidized bed, Tar cracking, Simulation.

1. Introduction

The process of biomass gasification consists in the decomposition of organic material (solid or liquid) through the application of heat under the presence of a gasifying agent as CO₂, O₂, steam or a combination of them, which results in the production of biochar and syngas, a mix of gases constituted principally by CO, H₂, CH₄, other light hydrocarbons (C_{≤3}) and tars (Reyes Alonzo, 2020). Due to an incomplete process of gasification and because of the biomass composition, different undesirable products as tars are found in great proportion in the syngas. The tar is a dark, odorous and immensely viscous liquid that contains a diverse amount of aromatic species with one to five rings, oxygenated hydrocarbons and polycyclic hydrocarbons (Gao et al., 2020), they appear in different types depending on the molecular weight of the molecules (Table 1) (Molino et al., 2016). The undesirability of the presence of tar on the gasification process is due to the operational problems such as the reduction of the heating efficiency of the gas and clogging issues in the equipment, caused by the condensation of tar in the zones of low temperature, those issues raise the operational costs due to the application of cleaning devices downstream, being one of the reasons that has made difficult a wider commercialization of the gasification technology (Rakesh N & Dasappa, 2018).

In order to reduce the tar concentration on the producer gas and favorize the production of the syngas, different catalytic bed materials are used for gasification. Bed materials as quartz, feldspar, limestone, etc., have been tested with the aim to find less expensive alternatives to the widely used olivine (Mauerhofer et al., 2018). In this work, the application of biochar as a catalytic bed material is evaluated, since it is considered an economic and accessible alternative in view of the fact that it is a gasification by-product,

and also the biochar has demonstrated to have a good catalytic activity to reduce the tar yield, increasing the syngas production (Abu El-Rub, 2008).

Table 1. Classification of the different types of tar (Molino et al., 2016).

Tar type	Name	Characteristic compound	Examples
1	GC undetectable	Too heavy and cannot be detected with a GC.	Gravimetric tar
2	Heterocyclic aromatics	Water-soluble compounds and tar molecules that contain heteroatoms.	Cresol, Pyridine, Quinoline, dibenzo phenol
3	Light aromatic (1 ring)	Hydrocarbons with single ring.	Xylene, Toluene, Styrene, Fluorene, anthracene
4	Light PAH's (2-3 rings)	Hydrocarbons with two and three rings, which condense at low temperature.	Naphthalene, Indene, Fluorene, Phenanthrene
5	Heavy PAH's compound (4-7 rings)	Hydrocarbons with more than three rings, which condense at high temperature.	Pyrene, Coronene, Chrysene, Perylene

Aiming to describe the phenomena that take place during the biomass gasification, and the syngas composition after the process, different processes simulators are currently used. For this work, we use the simulator Aspen Plus® as a tool to obtain mass and energy balances of the process (Abdelouahed et al., 2012). In literature the biomass is decomposed into its constituents (C, O₂, H₂, N₂, H₂O, S) on a RYield reactor before being feed in the gasifier which could be a combination of the RStoic and RGibbs reactors or a single reactor that transforms those gases into syngas. (Dattatray & Shilapuram, 2019; Zaman & Ghosh, 2021). A significant concern of the use of those models is the fact that they do not represent the actual gasification process which comes along with tars production, leaving aside a major operational problem that could be further observed parametrically using the software, therefore the approach considered is the used by other authors that have model the gasification using empirical correlations obtained from the experimental data gathered after processing the producer gas composition. This process might be convenient, however it cannot be easily extended to other installations (Abdelouahed et al., 2012; Francois et al., 2013).

This article is focused on the modeling of a biomass gasification process through the application of the empirical correlations obtained from previous experiments (Reyes Alonzo, 2020) and the exhibition of the effects of the use of biochar as catalytic bed material over the use of inert material as sand. Also, a comparison of the performance of CO₂ over the use of steam as gasifying agents will be made. Those elements of study have their particularity.

The use of CO₂ as gasifying agent has not been intensively exploited experimentally as the steam due to its energetic downsides, an external heat source is needed in order to maintain the process temperature which is caused by the absence of the partial combustion reaction of biomass that take place when oxidizing agents are used (Sadhvani et al., 2016). However, gasification using CO₂ generates a producer gas with higher energetic values and more efficiency in the tar cracking process when compared with the use of steam (Reyes Alonzo, 2020).

2. Aspen Plus® simulation

The software Aspen Plus ® Version 10 was used in order to obtain a proper reproduction of the results of the composition of the syngas obtained during the process of gasification that was carried out using CO₂ or steam as the gasifying agent. Different units were used to represent the stages of the process (from biomass pre-treatment until solids separation). All the components involved were added on the properties section along with the non-conventional solids: biomass and biochar, which were named as WOOD and CHAR respectively. For this model, the thermodynamic method of Peng-Robinson was selected and the stream class selected was MIXNCPSD for representing the secondary streams containing conventional component, sand and non-conventional solids. The models HCOALGEN and DCOALIGT were selected for enthalpy and density of non-conventional solids were selected. The water was selected as a *moisture component* for the biomass using the *Solid Characterization* menu, where the particle size distribution (PSD) of biomass with an average size from 0.3 mm to 22 mm was considered.

Table 2. Elemental analysis for beechwood pellets and beechwood biochar.

Material	Elemental analysis (wt%)				Proximate analysis (wt%)			
	Carbon (C)	Hydrogen (H)	Nitrogen (N)	Oxygen (O)	Moisture content	Volatile material	Fixed carbon	Ash content
Beechwood pellets	49.35	6.25	<0.01	44.4	6.23	75.4	17.54	0.83
Beechwood biochar	78.24	3.13	0	18.63	0	1.59	93.83	4.58

2.1 Biomass pre-treatment

For the feeding process, the data of the proximate and ultimate analysis of the beechwood was inserted into the stream BIOMASS (Table. 2). This stream goes through the block CRUSHER where the biomass or beechwood pellets (22 mm) are turned into powder with a particle size suitable for the fluidization parameters (~0.3mm). The crushed biomass stream goes through the block DRYER which will reduce the moisture content of the biomass. The configuration ‘Shortcut’ was selected as the operative mode of this block, therefore the amount of moisture contained on the biomass after the drying process was established beforehand.

2.2 Gasification process simulation

The experimental data and the characteristics of the operating conditions of the experimental setup can be found in previous work (Reyes et al., 2020). In the simulation, the pretreated biomass is mixed with the gasification agent (CO₂ or steam) in the MIXER block before entering to the gasifier. The fluidization conditions are established with a flow rate of 1.5 L/min of gasifying agent. The gasifier is modeled using a RYield reactor customized with two calculators with their specific Fortran subroutines representing the

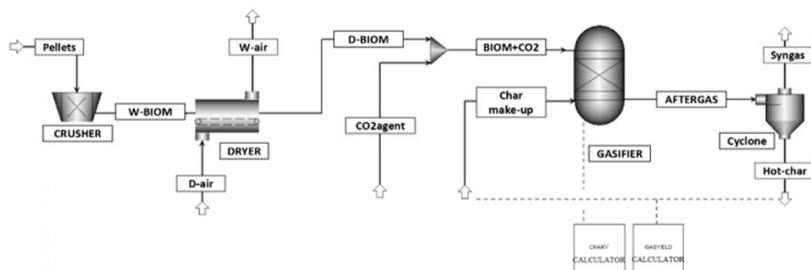


Figure 1. Process flowsheet for the biomass pre-treatment and gasification.

final composition of the gasification products and the incoming amount of biochar into the gasifier. Firstly, the yield of biochar is obtained from the GAS-Y block results and secondly, the block CHAR-Y is executed using this biochar yield as the calculation basis for the biochar entering through the CHARFEED stream as gasifier bed material (Fig. 1). The correlations used on the subroutine of the GAS-Y block represent the yields of each solid, liquid and gaseous specie produced during the gasification. Three sets of correlations were applied in order to represent the gasification with sand and CO₂ and biochar with steam and CO₂ (Table 3).

Table 3. Parameters for the correlations of biomass gasification.

Products	Sable				CO ₂				Steam				
	a	b	c	d	a	b	c	d	a	b	c	d	
Gases	H ₂	-2.24 × 10 ⁻⁸	4.83 × 10 ⁻⁵	-3.40 × 10 ⁻²	7.86	5.09 × 10 ⁻⁸	-1.02 × 10 ⁻⁵	7.10 × 10 ⁻³	-1.69	2.72 × 10 ⁻⁸	-6.06 × 10 ⁻⁵	4.71 × 10 ⁻²	-12.30
	CO	-2.87 × 10 ⁻⁷	8.21 × 10 ⁻⁴	-0.69	1.86 × 10 ²	7.49 × 10 ⁻⁷	-1.50 × 10 ⁻³	1.03	-2.40 × 10 ²	6.15 × 10 ⁻⁶	-1.38 × 10 ⁻²	10.30	-2.52 × 10 ²
	H ₄	1.15 × 10 ⁻⁶	-6.73 × 10 ⁻³	8.00 × 10 ⁻²	-25.8	1.12 × 10 ⁻⁷	-2.47 × 10 ⁻⁴	18.50	-45.90	9.07 × 10 ⁻⁷	-2.05 × 10 ⁻³	1.54	-3.82 × 10 ²
	O ₂	-1.47 × 10 ⁻⁷	-4.25 × 10 ⁻³	0.37	-1.47 × 10 ²	-1.27 × 10 ⁻⁶	2.81 × 10 ⁻³	-1.97	4.80 × 10 ²	4.69 × 10 ⁻⁸	-1.03 × 10 ⁻²	7.58	-1.83 × 10 ²
	C ₂ H ₂	-1.39 × 10 ⁻⁹	1.47 × 10 ⁻⁸	1.92 × 10 ⁻⁴	-0.340	-3.57 × 10 ⁻¹⁰	1.87 × 10 ⁻⁸	-1.64 × 10 ⁻³	4.80 × 10 ²	3.72 × 10 ⁻⁸	-8.41 × 10 ⁻⁵	6.33 × 10 ⁻²	-15.70
	C ₂ H ₄	-4.77 × 10 ⁻⁹	1.03 × 10 ⁻⁵	-7.44 × 10 ⁻³	1.86	-1.04 × 10 ⁻⁸	2.14 × 10 ⁻⁵	-1.45 × 10 ⁻²	3.24	-6.87 × 10 ⁻⁷	1.56 × 10 ⁻³	-1.15	2.79 × 10 ²
	C ₂ H ₆	-1.91 × 10 ⁻⁸	4.41 × 10 ⁻⁵	-3.37 × 10 ⁻²	8.60	-7.19 × 10 ⁻⁸	1.58 × 10 ⁻⁴	-0.113	26.90	3.53 × 10 ⁻⁸	-8.25 × 10 ⁻⁵	6.37 × 10 ⁻²	-16.10
	C ₃ H ₆	-9.48 × 10 ⁻⁸	2.27 × 10 ⁻⁵	-1.81 × 10 ⁻¹	4.78	9.02 × 10 ⁻⁹	-2.50 × 10 ⁻⁵	2.20 × 10 ⁻²	-6.08	-1.39 × 10 ⁻⁷	3.05 × 10 ⁻⁴	-0.22	52.50
	C ₃ H ₈	-9.93 × 10 ⁻⁸	2.38 × 10 ⁻⁵	-1.90 × 10 ⁻²	5.01	2.24 × 10 ⁻⁸	-5.34 × 10 ⁻⁵	4.19 × 10 ⁻²	-10.70	0	0	0	0
	Tars	CH ₃ COOH	6.30 × 10 ⁻⁷	-1.25 × 10 ⁻³	0.76	-1.31 × 10 ²	1.57 × 10 ⁻⁸	-3.30 × 10 ⁻³	2.20	-4.53 × 10 ²	-3.51 × 10 ⁻⁶	8.37 × 10 ⁻³	-6.64
C ₆ H ₆ O		-3.16 × 10 ⁻⁷	6.88 × 10 ⁻⁴	-0.49	1.18 × 10 ²	3.52 × 10 ⁻⁸	-1.24 × 10 ⁻⁴	0.12	-35.10	-2.76 × 10 ⁻⁶	6.00 × 10 ⁻³	-4.29	1.01 × 10 ³
C ₈ H ₈ O		8.63 × 10 ⁻⁸	-1.82 × 10 ⁻⁴	0.123	-25.70	1.74 × 10 ⁻⁸	-3.09 × 10 ⁻⁵	1.27 × 10 ⁻²	1.06	-2.30 × 10 ⁻⁶	1.84 × 10 ⁻⁵	7.54 × 10 ⁻³	-4.88
C ₉ H ₈ O		-2.27 × 10 ⁻⁸	5.62 × 10 ⁻⁵	-4.82 × 10 ⁻²	15.30	8.45 × 10 ⁻⁸	-1.81 × 10 ⁻⁴	0.12	-23.60	-6.86 × 10 ⁻⁷	1.50 × 10 ⁻³	-1.09	2.66 × 10 ²
C ₉ H ₈		-3.26 × 10 ⁻⁸	7.16 × 10 ⁻⁵	-5.19 × 10 ⁻²	12.40	2.25 × 10 ⁻⁸	-3.44 × 10 ⁻⁵	8.72 × 10 ⁻³	3.72	2.78 × 10 ⁻⁷	-6.22 × 10 ⁻⁴	0.45	-1.04 × 10 ²
C ₁₀ H ₈		-1.75 × 10 ⁻⁷	4.15 × 10 ⁻⁴	-0.32	84.60	-7.20 × 10 ⁻⁸	1.75 × 10 ⁻⁴	-0.14	38.30	9.08 × 10 ⁻⁸	-2.02 × 10 ⁻⁴	0.14	-33.90
C ₁₀ H ₁₀		2.00 × 10 ⁻⁷	-4.59 × 10 ⁻⁴	0.34	-85.90	-1.04 × 10 ⁻⁸	2.98 × 10 ⁻⁵	-2.81 × 10 ⁻²	8.74	-1.87 × 10 ⁻⁷	4.10 × 10 ⁻⁴	-0.29	70.70
C ₁₀ H ₈		-1.54 × 10 ⁻⁷	3.57 × 10 ⁻⁴	-0.27	69.30	-2.44 × 10 ⁻⁷	5.56 × 10 ⁻⁴	-0.42	1.06 × 10 ²	-8.40 × 10 ⁻⁷	1.82 × 10 ⁻³	-1.29	3.05 × 10 ²
C ₁₀ H ₈		-5.46 × 10 ⁻⁸	1.36 × 10 ⁻⁴	-0.10	28.60	-4.81 × 10 ⁻⁸	1.06 × 10 ⁻⁴	-7.80 × 10 ⁻²	19.50	-5.65 × 10 ⁻⁷	1.21 × 10 ⁻³	-0.85	1.99 × 10 ²
C ₁₀ H ₁₀		-1.88 × 10 ⁻⁸	4.53 × 10 ⁻⁵	-3.52 × 10 ⁻²	8.92	3.03 × 10 ⁻⁸	-6.68 × 10 ⁻⁵	4.83 × 10 ⁻²	-11.40	1.72 × 10 ⁻⁷	-4.04 × 10 ⁻⁴	0.31	-78.10
Solids	H ₂ O	3.95 × 10 ⁻⁷	-8.14 × 10 ⁻⁴	0.51	-78.10	-1.01 × 10 ⁻⁶	2.12 × 10 ⁻³	-1.48	3.57 × 10 ²	-2.14 × 10 ⁻⁶	4.49 × 10 ⁻³	-3.12	7.35 × 10 ²
	CHAR	-4.67 × 10 ⁻⁸	7.90 × 10 ⁻⁵	-6.50 × 10 ⁻²	43.10	8.00 × 10 ⁻⁸	-4.06 × 10 ⁻⁴	0.43	-1.16 × 10 ²	-8.53 × 10 ⁻⁷	1.95 × 10 ⁻³	-1.52	4.16 × 10 ²

3. Results and analysis

3.1. Effects of the biochar use with different gasifying agents

The process was modeled in Aspen Plus® adjusting temperatures from 600°C to 900°C on the RYield reactor. As expected, the correlations have an approached agreement with the experimental results, which enables different parametric evaluations for this set of temperatures and its impact over the product's yield with a sensitivity analysis. Fig. 2 shows the tar concentration for different temperatures and bed material (sand or biochar). The use of char reduces noticeably the concentration of heavy molecules; this demonstrates its catalytic activity. The good agreement of the model with the experimental process allows the obtention of a proper mass and energy balance of the system (Fig. 3).

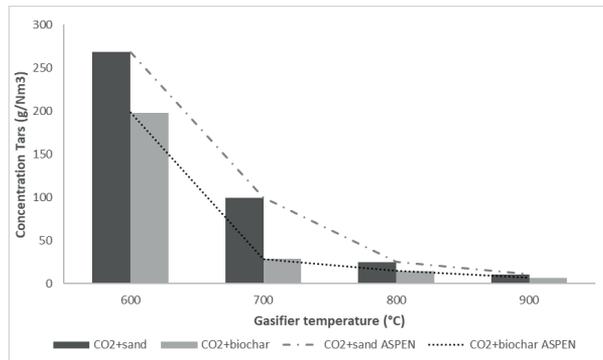


Figure 2. Tar concentration for different temperatures and bed materials.

The material and energy balance for the whole system is easily obtained through the values of the stream and block results. Fig. 3 shows the mass and energy balance for the process using sand and biochar as bed materials. This global balance was made considering external streams and the equipment that consume significant amounts of energy (dryer, crusher, and RYield reactor). There is a slight difference on the amount of heat required for the gasifier when using biochar which corresponds to the need of heat for the type of reaction (endothermic) that are taking place during the process with biochar. There are basically no changes on the yield of syngas and biochar observed in both examples, which opens the question of whether the escalation was effective or whether it is necessary to include a correction.

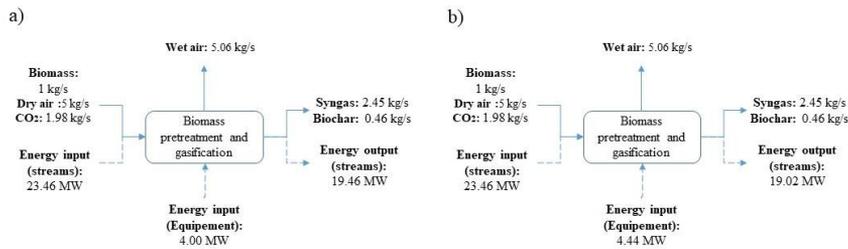


Figure 3. Mass and energy balance for gasification process using sand (a) and biochar (b) as bed materials.

3.2. Effects of the biochar as bed material varying the gasifying agents

The CO₂ and the steam are known for their use as biochar activator (Molina-Sabio et al., 1996), therefore during the experiments, both gasifying agents present a significant decrement of the tar concentration in comparison of the use of sand (Fig. 4). At lower temperatures the catalytic cracking effect over the final tar concentration is higher with the use steam. This can be explained due to the higher porosity of the steam-activated biochar, along with the possible steam reforming reaction that take place during steam gasification.

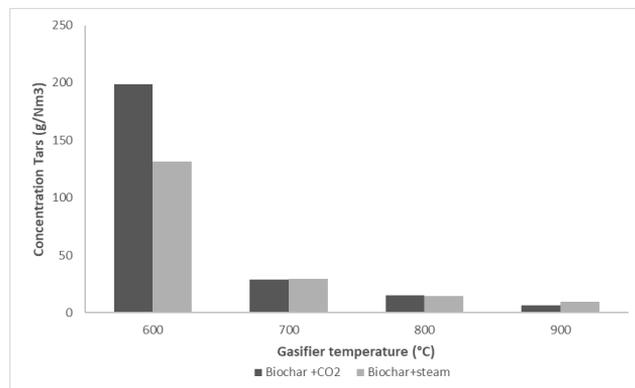


Figure 4. Effects of different gasifying agents along with biochar as bed material.

For a temperature as high as 900°C it can be observed that the CO₂ has a more suitable tar cracking behavior than steam, could be supposed that this temperature might be favorizing other reactions involving the CO₂ itself, but a right conclusion of this phenomena is out of this case of study.

Conclusion

The use of experimental correlations through Fortran subroutines in Aspen Plus® guarantees a good reproduction of experimental results. This modeling methodology allows the evaluation of different parameters such as temperature and energy spending for a specific installation, but since correlations are tailored to the process, their application in another installation would have limitations. The use of the biochar as catalytic bed material was evaluated and as expected, the catalytic activity of the biochar was higher than the activity of sand, this is due to the existence of active pores on the biochar that increase the interaction of the hydrocarbons which facilitates the conversion of tar into lighter molecules. When comparing the energy balance of both processes, a difference in the energy input for the gasifier is perceived and this can be interpreted as an effect of the befall of endothermic reactions that need more heat to be performed. The application of CO₂ and steam as gasifying agents along with the use of biochar, reduces considerably the tar concentration contained on the syngas. The steam shows a better performance in comparison to the CO₂, this is due to the fact that the biochar activated with steam has a higher porosity than the biochar activated with CO₂.

References

- Abdelouahed, L., Authier, O., Mauviel, G., Corriou, J. P., Verdier, G., & Dufour, A. (2012). <https://doi.org/10.1021/ef300411k>
- Abu El-Rub, Z. Y. K. (2008). Biomass char as an in-situ catalyst for tar removal in gasification systems. s.n.].
- Dattatray, A. D., & Shilapuram, V. (2019). <https://doi.org/10.1016/j.fuel.2018.09.023>
- Francois, J., Abdelouahed, L., Mauviel, G., Patisson, F., Mirgaux, O., Rogaume, C., Rogaume, Y., Feidt, M., & Dufour, A. (2013). <https://doi.org/10.1016/j.biombioe.2013.01.004>
- Gao, X., Wang, Z., Ashok, J., & Kawi, S. (2020). <https://doi.org/10.1016/j.cesx.2020.100065>
- Mauerhofer, A. M., Benedikt, F., Schmid, J. C., Fuchs, J., Müller, S., & Hofbauer, H. (2018). <https://doi.org/10.1016/j.energy.2018.05.158>
- Mohabeer, C., Reyes, L., Abdelouahed, L., Marcotte, S., Buvat, J.-C., Tidahy, L., Abi-Aad, E., & Taouk, B. (2019). [https://doi.org/10.1016/S1872-5813\(19\)30008-8](https://doi.org/10.1016/S1872-5813(19)30008-8)
- Molina-Sabio, M., Gonzalez, M. T., Rodriguez-Reinoso, F., & Sepúlveda-Escribano, A. (1996). [https://doi.org/10.1016/0008-6223\(96\)00006-1](https://doi.org/10.1016/0008-6223(96)00006-1)
- Molino, A., Chianese, S., & Musmarra, D. (2016). <https://doi.org/10.1016/j.jechem.2015.11.005>
- Rakesh N, & Dasappa, S. (2018). <https://doi.org/10.1016/j.rser.2018.04.017>
- Reyes Alonzo, L. C. (2020). <http://www.theses.fr/2020NORMIR23/document>
- Reyes, L., Abdelouahed, L., Buvat, J.-C., Campusano, B., Boyer, C. D., & Taouk, B. (2020). <https://doi.org/10.1016/j.cherd.2020.09.017>
- Reyes, L., Abdelouahed, L., Campusano, B., Buvat, J.-C., & Taouk, B. (2021). <https://doi.org/10.1016/j.fuproc.2020.106664>
- Sadhwani, N., Adhikari, S., & Eden, M. R. (2016). <https://doi.org/10.1021/acs.iecr.5b04000>
- Zaman, S. A., & Ghosh, S. (2021). <https://doi.org/10.1016/j.biortech.2021.125412>

Mathematical modeling of the diffusion-limited (DLA) aggregation accompanied by particles swarming in reactors

Leila Musabekova^a, Sabira Akhmetova^a, Assel Sailau^a, Arnold Brener^a

^a*Auezov University of South Kazakhstan, Tauke Khan av. 5, Kazakhstan*

Abstract

This article is devoted to the development of a new approach to modeling aggregation processes and swarming in poly-disperse systems. This approach takes into account multi-particle collisions, what is of fundamental importance at a high concentration of the dispersed phase. Due to the complex hydrodynamic picture in industrial devices the aggregation process is accompanied by multiple collisions and swarming of particles, and it occurs according to various mechanisms. The presented method is an extension of the method previously developed by the authors for the case of a three-dimensional stochastic lattice. Namely, the discrete-event-simulation paradigm (DES) extends in this work to the case of three-dimensional lattices. A three-dimensional array is used for describing aggregation in the case of kinetics limited by the particle diffusion rate (DLA) with allowance to the swarming process. The article provides a description of the algorithm and the results of a computer experiment with comments. Computer simulation of swarming and aggregation in batch and flow reactors has been carried.

Keywords: DLA aggregation, swarming, many-particle collisions, 3-D stochastic lattice.

1. Introduction

The issues of calculating the kinetics and dynamic characteristics of reactors with the formation of a polydisperse solid phase in the working volume are relevant under calculation of various types of industrial apparatuses (Carranza and Coates, 2000). Need in the reliable swarming and aggregation models covers the range from fine chemical technology and pharmaceuticals to environmental cleaning and industrial waste (Maningo et al., 2016).

However, despite the long-standing interest of researchers and the existence of many works, the theoretical analysis of many problems remains poorly developed (Rimer and Ariel, 2017). In this paper, a new approach based on the DES discrete event modeling paradigm (Zeigler et al., 2000) applying to the case of three-dimensional lattices is presented.

The novelty of this approach lies in the fact that it allows one to study the distribution of clusters of different orders over the volume of the reactor at different times, taking into account multi-particle collisions and swarming phenomenon (Carrillo et al., 2010). The contribution of the obtained results to practice lies in the fact that they expand the range of real application of the developed model for describing swarming and aggregation processes in the volume of industrial devices (Zhou et al., 2015, Schmidt et al., 2006). This makes it possible to take into account the peculiarities of the kinetics of aggregation in the apparatus volume (Mogilner and Edelstein-Keshet, 1999). Cases of batch and flow-through tubular apparatus are considered. According to DES (Zeigler et al., 2000, Brener

et al., 2017), the region in which the diffusion and aggregation processes occur is covered by a fixed spatial lattice (Brener et al., 2017). In this paper, we consider a three-dimensional array to describe aggregation and swarming in the case of kinetics limited by the rate of diffusion of particles (DLA). Thus, the characteristic time of aggregation was completely determined by the characteristic time of diffusion of particles, which was assumed to be equal to the time of drift of particles from one cell of the lattice to another. Hereinafter, the term "order of clusters" will mean the number of monomer particles in a given cluster (Wattis, 2006).

2. Algorithm description

2.1. Algorithm for batch reactor

Below, an algorithm is described that corresponds to that presented in (Brener et al., 2017). Namely, at each moment of time, a random selection of the components of the particle drift in the horizontal and vertical directions is made from a given set of characteristic displacements. This set is formed taking into account the dependence of the mobility of a cluster of particles on its order (Naldi et al., 2010).

For clusters of a higher order, the mobility decreases, which corresponds to the well-known physics of random particle drift in the volume (Satyobroto Talukder, 2011). In the model algorithm in this work, it was assumed that a random selection from the cell sequence (-3; -2; -1; 0; 1; 2; 3) for temporal drifts along each of the three spatial axes can be made for clusters of order 1 or 2, and for clusters with orders higher than two, the choice must be made from the sequence (-3; -2; -2; -1; -1; 0; 0; 1; 1; 2; 2; 3). Thus, an increase in the probability of higher-order clusters moving to nearer cells during a time step is simulated.

The behavior of clusters near the walls can also be taken into account in the general case by changing the sequence of random selection of displacements. In a numerical experiment applied to a batch reactor, it was assumed that particles captured in the boundary cells do not move beyond the boundaries of the lattice under consideration (Brener et al., 2017). A cluster of order 1 was initially placed in each cell.

Then the process of random transition of particles from one cell to another cell was simulated at each time step. The aggregation of particles occurs at the moment they enter the common cell of the lattice without any delay (DLA case).

In contrast to the planar case (Brener et al., 2017), two three-dimensional arrays were formed. The first array simulated the entire lattice with particles of different orders obtained in the process of displacement and aggregation at each time step. The second array shows the number of collisions of particles in each cell in each unit of time.

2.2. Algorithm for a flow-through tubular reactor

The algorithm is briefly described below in accordance with the work (Brener et al., 2017, Zeigler et al., 2000). During the calculation, four 3D arrays are generated.

The first array simulates the entire lattice with clusters of different orders obtained in the process of displacement and aggregation in each unit of time.

The second array simulates a similar lattice with clusters that enter the reactor with a fresh flow, which enters the reactor at a given average horizontal velocity.

The third array simulates the situation at the end of the calculated time unit and generates the initial situation for the next unit. It is formed as the sum of the two previous arrays.

The fourth array indicates the number of collisions of particles in each cell in each unit of time.

2.3. Algorithm for swarming description

To adapt the stochastic lattice method to simulating swarming, especially taking into account the aggregation process and the influence of random drift, a special lattice enlargement algorithm was developed, which is presented in this work.

Since it is this stage of modeling that is the principal contribution to the algorithm, some mathematical details are presented below.

The main steps of this algorithm in the case of DLA are as follows.

1. The entire volume of the reactor is dissected into longitudinal layers. Then the following matrices are created: the main matrix of concentrations $C(I, J)$ and a coarse matrix $CO(X, Y)$ of the total number of clusters in blocks, where M, N are the dimensions of the matrix $C(I, J)$ and R, T are the dimensions of the matrix $CO(X, Y)$. Here $R = M / a$; $T = N / b$; and a, b - the dimensions of the coarse lattice block in the each layer (that is, the number of rows and columns of the matrix $C(I, J)$ captured in the block to generate the matrix $CO(X, Y)$).

2. The block with specific coordinates (X_s, Y_s) in the coarse matrix $CO(X, Y)$ is built from the elements of the matrix $C(I, J)$, which consists of rows numbered from $(aX_{s-1} + 1)$ to aX_s and columns numbered from $(bY_{s-1} + 1)$ to bY_s .

3. The element of the matrix $CO_{X_s Y_s}$ must be calculated as the sum d_{kl} of all the elements of the matrix $C(I, J)$ that fall into the block (X_s, Y_s) . Namely:

$$CO_{X_s Y_s} = \sum_{bY_{s-1}+1}^{bY_s} \sum_{aX_{s-1}+1}^{aX_s} d_{kl} . \quad (1)$$

Here the elements d_{kl} should be calculated according to the following algorithm called as *A-algorithm*: if $c_{X_s Y_s} \neq 0$, then $d_{kl} = 1$, otherwise $d_{kl} = 0$. A coarse matrix $COI(X, Y)$ of the total number of orders of clusters in blocks should be calculated similarly to the matrix $CO(X, Y)$. The only difference is that elements $col_{X_s Y_s}$ are calculated using the formula

$$col_{X_s Y_s} = \sum_{bY_{s-1}+1}^{bY_s} \sum_{aX_{s-1}+1}^{aX_s} c_{kl} . \quad (2)$$

That is, it is not necessary to pre-calculate by algorithm (A).

It can be noted that this algorithm is inspired by the ideology of the renormalization group in the modern theory of phase transitions (Kadanoff, 2000). We believe that a universal approach that considers the processes of swarming and aggregation in disperse systems from positions similar to those used in the modern theory of phase transitions to describe ordering in many-particle systems is quite promising.

Of course, this issue is subject to further discussion, since the kinetic features of swarming and aggregation processes are not similar.

At the same time, these processes occur in a certain temporal hierarchy, since the swarming process contributes to the intensification of aggregation. This factor clearly manifested itself in computer simulations, what is shown in the next section of the article.

3. Some results of the simulation

Some results of the simulation are shown in Figures 1, 2, 3. Different color intensities of different blocks of rough matrices correspond to different concentrations of clusters, both without taking into account their orders (cold colors) and different sums of orders of cluster aggregates (warm colors). The blocks were colored using the RGB code of the additive color model (Poynton, 2012). The step of change was 20 units both in the number of clusters and in the sum of their orders. The overall range observed during the simulation was from 0 to 200 units. The dimensions of the computational matrix were 20×20 ; the dimensions of any block of the coarse matrix were 4×4 . The volume of 200×20 was divided into 40×5 horizontally and 4×5 vertically, it turned out a 5×5 coarse scheme.

The calculations were performed for three different initial carrier flow velocity profiles: uniform flow (type 1), symmetric irregular flow (type 2), and asymmetric irregular flow (type 3). The conditional flow rates W for these three cases are equal: 1- $W = 2$; 2- $W_w = 2$ at the walls, $W_c = 6$ on the central axis of the stream; 3- $W_w = 2$ at the upper wall, $W_c = 6$ at the central axis of the flow, $W_w = 4$ at the lower wall. Flow direction from left to right. The order of all particles introduced by the carrier flow into the circuit is taken to be 1. Then particles of different orders appear due to DLA aggregation (Brenner et al., 2017). The pictures presented on the figures correspond to the depth of 5 calculating blocks.

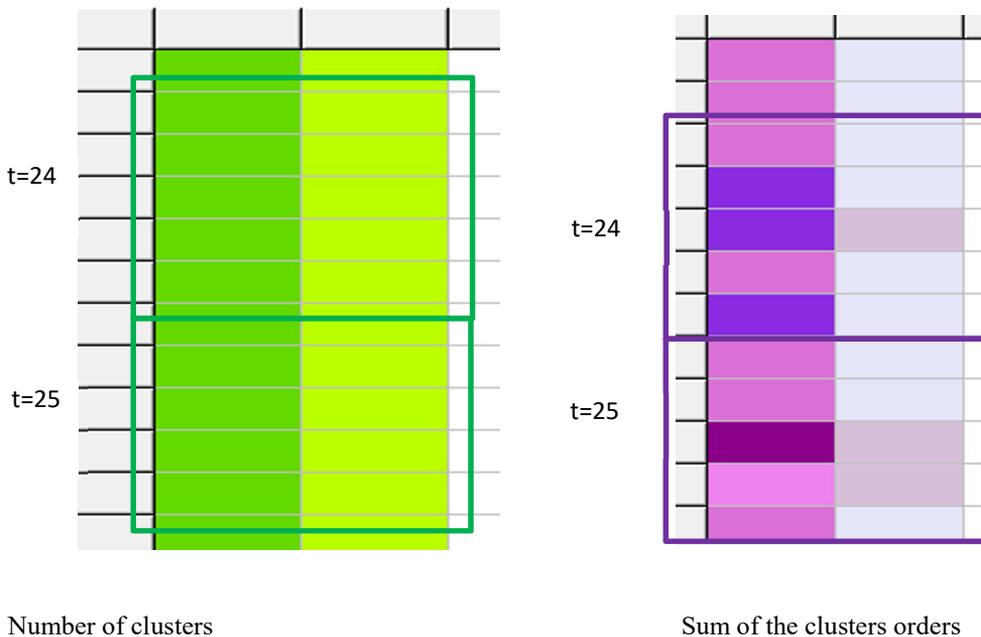


Figure 1. A typical picture with a uniform distribution for time intervals $t = 24, 25$ according to the enlarged scheme 5×5 . $W=2$.

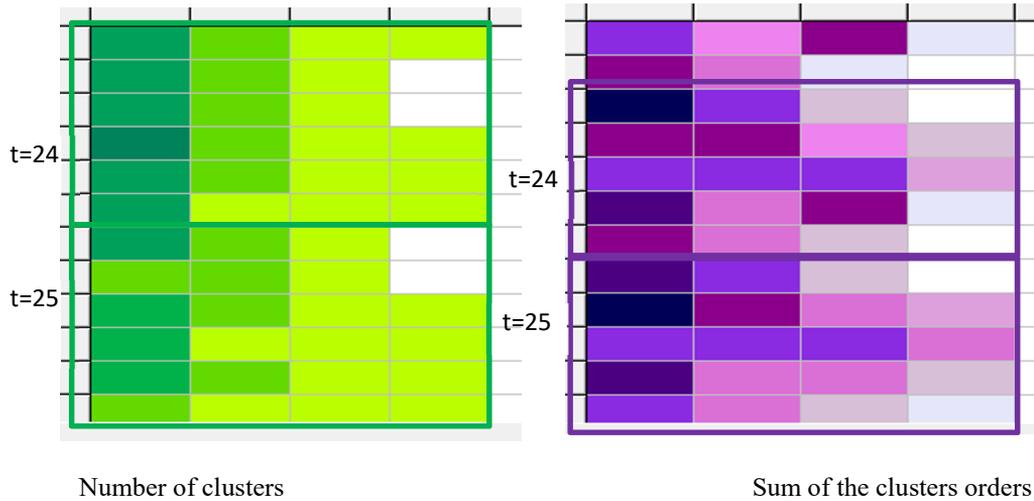


Figure 2. Under the uneven distribution of type 2 for time intervals $t = 24, 25$ according to the enlarged scheme $5X5$. $W_1=2, W_2=6$

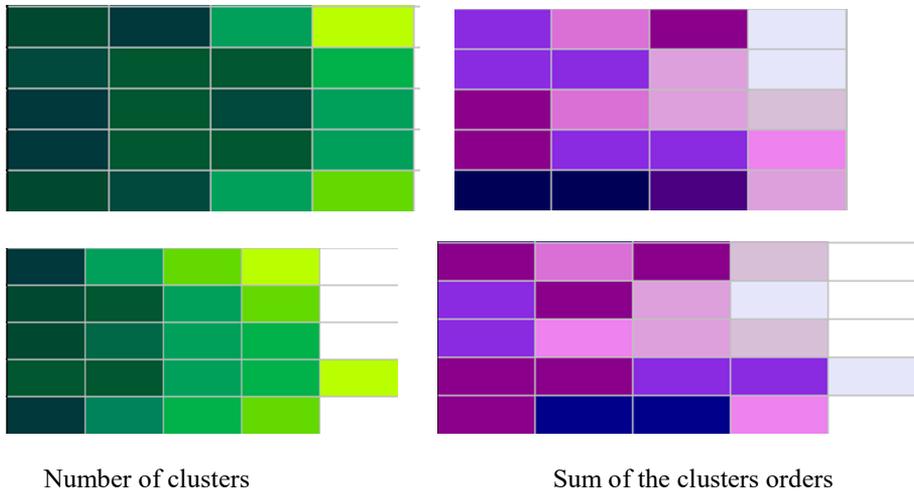


Figure 3. Under the uneven distribution of type 2 for time intervals $t = 24, 25$ according to the enlarged scheme $5X5$. $W_1=2, W_2=6, W_3=4$

As the result of visualizing the simulation results, it is possible to clearly distinguish two swarms of particles formed in a reactor with an uneven profile of the carrier flow velocity. It is clearly seen that the highest cluster orders are observed in the same areas, i.e. the aggregation process is intensified.

4. Conclusion

The results of numerical experiments showed that the change in the number of clusters in a batch reactor occurs at the initial stage of the aggregation process at a much higher rate than this phenomenon was described before by a numerical experiment on a flat lattice. The completion of the process with the formation of a big single cluster is also observed much faster. In a flow-through reactor, the cluster size distribution also changes in the initial section near the reactor entrance, but the length of the transition section is longer than the experiment on a planar grid shows.

The results of numerical experiments demonstrate a clear correlation between the localization of high concentration regions, those swarming formation and regions of a higher total order of clusters of the dispersed phase. The methodology proposed in this paper opens up opportunities for a deeper study of the detected correlation.

References

- A.M. Brenner, L.M. Musabekova, M.A. Jamankarayeva, 2017, Stochastic lattice model of aggregation in heterogeneous polydisperse media, *Chemical Engineering Transactions*, 60, 79-84.
- P.M. Carranza, P.S. Coates, P.S., 2000, Swarm modelling, *Proceedings of the 3rd Generative Art Conference*, Milan: Generative Design Lab Milan, Polytechnic University, Italy.
- J.A. Carrillo, M. Fornasier M., G. Toscani, F. Vecil, 2010, Particle, kinetic, and hydrodynamic models of swarming, *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, 297-336.
- L.P. Kadanoff, 2000, *Statistical Physics. Statics, Dynamics and Renormalization*, University of Chicago, World Scientific, Singapore, New Jersey, London, Hong Kong.
- J.M.Z. Maningo, G.E.U. Faelden, R.C.S. Nakano, A.A. Bandala, R.R.P. Vicerra, E.P. Dadios, 2016, Formation Control in Quadrotor Swarm Aggregation using Smoothed Particle Hydrodynamics, *Proceedings of IEEE Region 10 Conference (TENCON)*, 2070-2075.
- A. Mogilner, L. Edelstein-Keshet, 1999, A non-local model for a swarm, *J. Math. Biol.*, 38, 534-570.
- G. Naldi, L. Pareschi, G. Toscani (Eds), 2010, *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Birkhauser, Boston, Basel, Berlin.
- C. Poynton, 2012, *Digital Video and HD. Algorithms and Interfaces*, Morgan Kaufmann Publ. Waltham MA USA, 753.
- O. Rimer, G. Ariel, 2017, Kinetic Order-Disorder Transitions in a Pause-and-Go Swarming Model with Memory, *Journal of Theoretical Biology*, DOI: <http://dx.doi.org/10.1016/j.jtbi.2017.02.009>.
- Satyobroto Talukder, 2011, *Mathematical Modelling and Applications of Particle Swarm Optimization*, Master Thesis, School of Engineering Blekinge Institute of Technology, Karlskrona, Sweden.
- S.A. Schmidt, M. Simon, M.M. Attarakih, L.G. Lagar, H.-G. Bart, 2006, Droplet population balance modelling—hydrodynamics and mass transfer, *Chemical Engineering Science*, 61, 246–256.
- J.A. Wattis, 2006, An introduction to mathematical models of coagulation–fragmentation processes: a discrete deterministic mean-field approach, *Physica D: Nonlinear Phenomena*, 222(1), 1-20.
- B.P. Zeigler, H. Praehofer, T.G. Kim, 2000, *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems*, Academic Press, San Diego, CA.
- X.H. Zhou, B.C. Huang, T. Zhou, Y.C. Liu, H.C. Shi, 2015, Aggregation behavior of engineered nanoparticles and their impact on activated sludge in wastewater treatment, *Chemosphere*, 119, 568-576.

Verification of Neural Network Surrogates

Joshua Haddad^a, Michael Bynum^a, Michael Eydenberg^a, Logan Blakely^a,
Zachary Kilwein^b, Fani Boukouvala^b, Carl D. Laird^c and Jordan Jalving^a

^a*Sandia National Laboratories, Albuquerque, NM 87123, USA*

^b*Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA*

^c*Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
jhjalvi@sandia.gov

Abstract

Neural networks (NN)s have been increasingly proposed as surrogates for approximation of systems with computationally expensive physics for rapid online evaluation or exploration. As these surrogate models are integrated into larger optimization problems used for decision making, there is a need to verify their behavior to ensure adequate performance over the desired parameter space. We extend the ideas of optimization-based neural network verification to provide guarantees of surrogate performance over the feasible optimization space. In doing so, we present formulations to represent neural networks within decision-making problems, and we develop verification approaches that use model constraints to provide increasingly tight error estimates. We demonstrate the capabilities on a simple steady-state reactor design problem.

Keywords: Neural Networks, Verification, Optimization, Surrogate Modeling

1. Introduction

Neural networks (NNs) have seen wide success across engineering disciplines. Their excellent approximation qualities (Hornik et al. (1989)) can reduce challenging problems into tractable computational models and their training procedure can incorporate diverse data sets and expert domain knowledge (Beck et al. (2016)). Successful engineering applications have harnessed NNs in planning, design, and control (Pistikopoulos et al. (2021)) with notable process-systems applications that span: forecasting renewable energy generation (Lee et al. (2016)), predicting distillation dynamics (Sánchez-Ramírez et al. (2020)), and estimating reactor performance (Salah et al. (2016)).

Typical NN applications are concerned with making forward predictions (i.e., NNs are evaluated in a forward-mode), but they are also advantageous in the context of optimization where they can take on the form of *algebraic surrogates*. Such surrogate models can be used to reduce the complexity of underlying equations by replacing them with more tractable formulations. Neural network surrogates have seen successful application in optimizing chemical process operation (Fernandes (2006)), performing process synthesis with super-structures (Henao and Maravelias (2010)), solving stochastic optimization problems to operate distillation columns (Gutiérrez-Antonio (2016)), and representing contingency constraints in security-constrained optimal power flow (Kilwein et al. (2021)).

The *verification* of neural networks is often concerned with finding adversarial inputs (Goodfellow et al. (2015)) using techniques such as mixed-integer-linear-programming (MILP) (Tjeng et al. (2017)), Satisfiability Modulo Theories (Scheibler et al. (2015)), and Lagrangian duality

(Dvijotham et al. (2018)). This manuscript extends verification concepts and addresses the need to verify NNs that are deployed as optimization-based surrogates. In contrast to methods that verify over a specified input space, we propose formulations that verify over constraints involving both the inputs *and* outputs of the neural network. In doing so, we develop optimization problems that verify worst-case NN prediction error subject to known model physics and operational constraints, and we show that incorporating known constraints leads to tighter error estimates than simply verifying over the input space.

2. Optimization with Neural Network Surrogates

We consider the solution of an optimization problem given by, (1),

$$f^*(p) = \min_{x,y,z} f(x,y,z,p) \quad (1a)$$

$$\text{s.t. } y = h(x,p) \quad (1b)$$

$$c(x,y,z,p) = 0 \quad (1c)$$

$$g(x,y,z,p) \leq 0 \quad (1d)$$

$$x^L \leq x \leq x^U \quad (1e)$$

where p represents system parameters which are known inputs for a particular optimization instance, and x , y , and z represent optimization variables. We desire efficient solutions of this problem for different values of the parameters p (e.g., in an online context, or for multi-scenario analysis). For improved solution, we consider instead an approximate formulation where we replace a portion of the model with a neural network surrogate. We assume that the NN surrogate brings some benefit to the optimization problem by facilitating a more tractable computation. This could entail the simplification of (1b) to support more rapid or reliable optimization in an online context, or involve creating a piecewise linear approximation (e.g., using ReLU activation functions) of h to facilitate global optimization approaches. We show the new optimization formulation below, where we approximate h in (1b) with the neural network N in (2b). Here, \hat{y} represents the neural network output, and is an approximation of the original variables y .

$$f_S^*(p) := \min_{x,\hat{y},z} f(x,\hat{y},z,p) \quad (2a)$$

$$\text{s.t. } \hat{y} = N(x,p) \quad (2b)$$

$$c(x,\hat{y},z,p) = 0 \quad (2c)$$

$$g(x,\hat{y},z,p) \leq 0 \quad (2d)$$

$$x^L \leq x \leq x^U \quad (2e)$$

This hybrid modeling approach is common in engineering applications. However, before “deploying” problem (2) in a decision-making application, we wish to verify the accuracy of the neural network approximation. In the next section, we describe a verification formulation that explicitly considers the feasible region of the constraints above and the range on the input parameters p .

3. Formulation of Neural Network Verification Problem

Typical verification approaches seek to determine the maximum error between y and \hat{y} over a pre-defined input space in x and p . However, these approaches can produce errors that are larger than necessary since they allow points in x and p that may not be feasible with respect to the constraints in formulations (1) and (2). Here, we are specifically interested in the accuracy of the neural network over the feasible space of the optimization problem, and we formulate the verification problem (3)

shown below:

$$\begin{aligned}
 & \max_{x,p,y,\hat{y},z} \|y - \hat{y}\|_{\infty} && (3a) \\
 \text{s.t. } & y = h(x, p) && \text{True Model} && (3b) \\
 & \hat{y} = N(x, p) && \text{NN Surrogate} && (3c) \\
 & x^L \leq x \leq x^U && \text{Input Bounds} && (3d) \\
 & p^L \leq p \leq p^U && \text{Parameter Bounds} && (3e) \\
 & \left[\begin{array}{l} c(x, y, z, p) = 0 \\ g(x, y, z, p) \leq 0 \end{array} \right] \vee \left[\begin{array}{l} c(x, \hat{y}, z, p) = 0 \\ g(x, \hat{y}, z, p) \leq 0 \end{array} \right] && \text{Feasibility Constraints} && (3f)
 \end{aligned}$$

Here, the objective function (3a) maximizes the infinity norm over the NN prediction errors. Maximizing other error measures such as mean-squared-error, mean-absolute-error, or individual prediction errors is also possible. The constraints (3b) and (3c) are the same as (1b) and (2b), respectively, which relate the NN inputs to the true model variables y and the neural network outputs \hat{y} . As indicated above, a natural choice of verification constraints are limits on the *input space* of the neural network (i.e., bounds over x and p). While a typical verification formulation would include only (3a)-(3e), we seek to limit the verification to points in x and p that are feasible with respect to the optimization problems (1) and (2).

Our target optimization problem (2) includes feasibility constraints over \hat{y} . However, including only constraints (2c-2d) in the verification problem (rather than the disjunction in (3f)) is insufficient because of errors between y and \hat{y} . Constraints (2c-2d) alone may not sufficiently represent the feasible region in problem (1). Indeed, if the NN accuracy is low, the feasible region for x and p could be significantly underestimated or even empty. Therefore, we search for the maximum error over both the constraints (1c-1d) and (2c-2d) as represented by the disjunction in (3f). With this, problem (3) finds the maximum deviation between y and \hat{y} subject to the constraints of the optimization problem applied conservatively to either y or \hat{y} .

For the case studies in this paper, we consider three formulations to analyze the verification approach. Formulation (V1) represents the typical verification formulation used in the literature which considers only explicit constraints on the input space for x and p . In practice, while we could solve the disjunctive problem (3) directly, it is convenient to solve with each disjunct separately where the solution of (3) is given by the maximum of the solutions from (V2) and (V3) below.

$$\begin{array}{lll}
 \max_{x,p,y,\hat{y},z} \|y - \hat{y}\|_{\infty} & \max_{x,p,y,\hat{y},z} \|y - \hat{y}\|_{\infty} & \max_{x,p,y,\hat{y},z} \|y - \hat{y}\|_{\infty} \\
 \text{s.t. (3b)-(3e)} & \text{s.t. (3b)-(3e)} & \text{s.t. (3b)-(3e)} \\
 \text{(V1)} & \text{(V2)} & \text{(V3)} \\
 & c(x, y, z, p) = 0 & c(x, \hat{y}, z, p) = 0 \\
 & g(x, y, z, p) \leq 0 & g(x, \hat{y}, z, p) \leq 0
 \end{array}$$

4. Illustrative Example: Reactor Optimization with Neural Network Surrogates

We provide an illustrative reactor optimization example where we use NNs (with ReLU activations) to replace nonlinear physics with piecewise-linear approximations. We demonstrate global solution of the verification problem and compare with the input-only formulation given by (V1).

4.1. Reactor Optimization Problem

The problem of interest is a steady-state continuous-stirred-tank-reactor (CSTR) that converts feed components A and B to produce D as depicted by Figure 1 where C is an intermediate and E is a

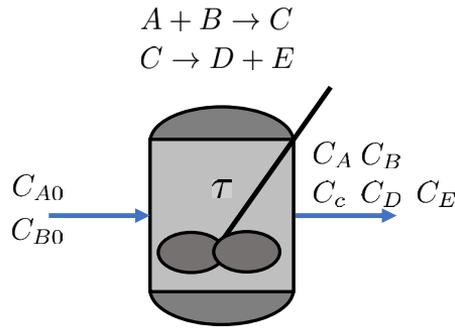


Figure 1: Simple Steady-State CSTR

side product. The overall mass balance is given by (4) (or by (5) in terms of the neural network outputs, \hat{y}).

$$C_i = C_{i,0} - \tau_1^r \quad i \in \{A, B\} \quad (4a) \quad C_i = C_{i,0} - \hat{\tau}_1^r \quad i \in \{A, B\} \quad (5a)$$

$$C_c = C_{c,0} + \tau_1^r - \tau_2^r \quad (4b) \quad C_c = C_{c,0} + \hat{\tau}_1^r - \hat{\tau}_2^r \quad (5b)$$

$$C_i = C_{i,0} + \tau_2^r \quad i \in \{D, E\} \quad (4c) \quad C_i = C_{i,0} + \hat{\tau}_2^r \quad i \in \{D, E\} \quad (5c)$$

We define variables τ_1^r and τ_2^r in (6) to represent the product of τ (the space-time) and the corresponding reaction rate (r_1 or r_2). The reaction rate constants K_i are selected as the parameters p . We also include operating constraints given by (7) which correspond to the minimum conversion and yield the reactor must achieve with $\eta_1=0.9$ and $\eta_2=0.2$.

$$\tau_1^r = \tau r_1 = \tau K_1 C_A C_B \quad (6a) \quad C_D \geq \eta_1 (C_{A,0} - C_A) \quad (7a)$$

$$\tau_2^r = \tau r_2 = \tau K_2 C_c \quad (6b) \quad C_D \geq \eta_2 (C_A + C_B + C_c) \quad (7b)$$

The true design problem as a function of p is then given by (8) where we seek to minimize the space-time τ (i.e., maximize throughput) subject to physical balances and operating requirements. Consequently, we train ReLU-based NNs to replace the associated nonlinear terms and formulate the approximate design problem (9) as an MILP (Grimstad and Andersson (2019)).

$$\min_{\tau} \quad \tau \quad (8a) \quad \min_{\tau} \quad \tau \quad (9a)$$

$$\text{s.t.} \quad \tau_1^r = \tau K_1 C_A C_B \quad (8b) \quad \text{s.t.} \quad \hat{\tau}_1^r = K_1 N_1 (C_A, C_B, \tau) \quad (9b)$$

$$\tau_2^r = \tau K_2 C_c \quad (8c) \quad \hat{\tau}_2^r = K_2 N_2 (C_c, \tau) \quad (9c)$$

$$(4), (7)$$

$$(5), (7)$$

4.2. Verification Problem

The verification problem corresponding to (3) is given by (11) with the NN input and model parameter bounds defined by (10) below.

$$\begin{aligned}
 C_A^L \leq C_A \leq C_A^U, \quad C_B^L \leq C_B \leq C_B^U, \quad C_C^L \leq C_C \leq C_C^U, \\
 \tau^L \leq \tau \leq \tau^U, \quad K_1^L \leq K_1 \leq K_1^U, \quad K_2^L \leq K_2 \leq K_2^U
 \end{aligned} \quad (10)$$

The valid range for the parameters K is given by $K_1=0.31051 \pm 10\% \frac{\ell}{\text{mol-s}}$ and $K_2=0.026650 \pm 10\% \frac{1}{s}$. For this example we seek to maximize the squared prediction error of both τ_1^r and τ_2^r (i.e., we solve verification problems for each output) but it is also possible to formulate mean-absolute

error using integer variables.

$$\max (\hat{\tau}_j^r - \tau_j^r)^2 \quad j \in \{1, 2\} \tag{11a}$$

$$\text{s.t. (8b), (8c) \hspace{15em} \text{Original Model} \tag{11b}$$

$$(9b), (9c) \hspace{15em} \text{NN} \tag{11c}$$

$$(10) \hspace{10em} \text{Input Bounds and Parameter Bounds} \tag{11d}$$

$$[(4), (7)] \vee [(5), (7)] \hspace{5em} \text{Model Constraints OR NN Constraints} \tag{11e}$$

We apply each of the presented verification formulations to our reactor example. In particular, we formulate (V1) with equations (11a-11d). We formulate (V2) and (V3) using these equations along with each of the individual disjuncts in (11e).

4.3. Results

We use TensorFlow 2.3 to train the multi-layer neural network surrogates with increasing numbers of nodes using ReLU activation functions. We use Gurobi 9.1 to solve the non-convex true design problem (8) and the verification problems. The results are presented in Figure 2 and are summarized as follows: (i) (V1) provides the expected global worst-case error over the input space, but we can obtain tighter error bounds using (V2) and (V3) which satisfy the known constraints, (ii) (V3) is tighter than (V2) with smaller networks where the accuracy is poor and the feasible region is poorly approximated, and (iii) worst-case error improves for larger neural networks, but there are likely trade-offs with performance vs accuracy. We also compare the solution

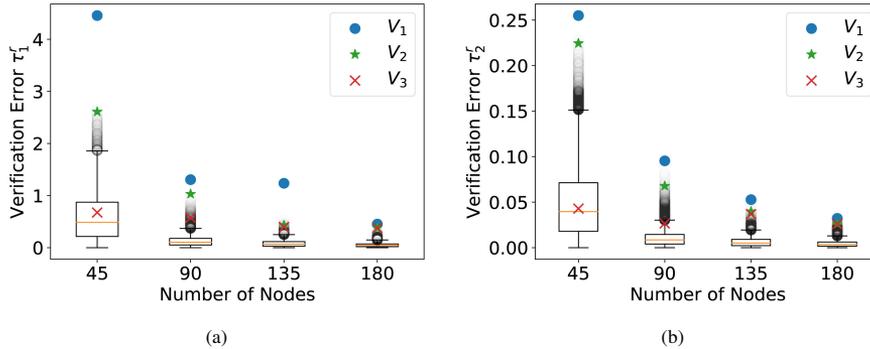


Figure 2: Results for test error (box plots) and verification (colored markers for V1, V2, and V3). Observed errors for τ_1^r (left) and τ_2^r (right) for increasingly larger neural networks.

of (9) for each NN to the true problem (8) using Gurobi 9.1. Table 1 shows each NN prediction and objective value τ which are consistent with the verification findings. Tighter worst-case error estimates correspond with closer approximations of the true problem, but even NNs with considerable error according to (V1) perform adequately as algebraic surrogates when considering the feasible region and their usage within the optimization problem.

5. Conclusions

This manuscript explores verification approaches for neural network (NN) surrogates used within optimization problems. We demonstrated how verification over known constraints produces tighter worst-case NN violations. We presented an illustrative reactor design example to elucidate verification concepts for NNs used as surrogates in an optimization setting.

Table 1: Comparison of results for reactor optimization problem solution for increasingly larger neural networks. Percentages correspond to percent difference from the true global solution.

Variable	True	Error (Relative Error %)							
		45 Nodes		90 Nodes		135 Nodes		180 Nodes	
τ_1^i	0.527	-3.9E-3	(-0.75%)	1.5E-3	(0.28%)	-8.8E-4	(-0.17%)	5.6 E-4	(0.11%)
τ_2^i	0.474	-3.5E-3	(-0.75%)	1.3E-3	(0.28%)	-7.9E-4	(-0.17%)	5.0 E-4	(0.11%)
τ	337.7	-14.5	(-4.29%)	-10.7	(-3.17%)	3.3	(0.99%)	0.17	(0.05%)

6. Acknowledgements

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This work was funded in part by the Institute for the Design of Advanced Energy Systems (IDAES) with funding from the Office of Fossil Energy, Cross-Cutting Research, U.S. Department of Energy. This work was also funded by Sandia National Laboratories Laboratory Directed Research and Development (LDRD) program.

References

- D. A. C. Beck, J. M. Carothers, V. R. Subramanian, J. Pfandner, 2016. Data Science : Accelerating Innovation and Discovery in Chemical Engineering 62 (5).
- K. Dvijotham, R. Stanforth, S. Gowal, T. A. Mann, P. Kohli, 2018. A dual approach to scalable verification of deep networks. Proceedings of the Thirty-Fourth Conference Annual Conference on Uncertainty in Artificial Intelligence.
- F. A. N. Fernandes, 2006. Optimization of fischer-tropsch synthesis using neural networks. Chemical Engineering & Technology 29 (4), 449–453.
- I. J. Goodfellow, J. Shlens, C. Szegedy, 2015. Explaining and harnessing adversarial examples. In: International Conference on Learning Representations.
- B. Grimstad, H. Andersson, 2019. ReLU networks as surrogate models in mixed-integer linear programs. Computers and Chemical Engineering 131, 106580.
- C. Gutiérrez-Antonio, 01 2016. Multiobjective stochastic optimization of dividing-wall distillation columns using a surrogate model based on neural networks. Chemical and Biochemical Engineering Quarterly 29, 491–504.
- C. A. Heno, C. T. Maravelias, 2010. Surrogate-based process synthesis. Computer-aided chemical engineering 28.
- K. Hornik, M. Stinchcombe, H. White, 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2 (5), 359–366.
- Z. Kilwein, F. Boukouvala, C. Laird, A. Castillo, L. Blakely, M. Eydenberg, J. Jalving, L. Batsch-Smith, 01 2021. AC-Optimal Power Flow Solutions with Security Constraints from Deep Neural Network Models. Vol. 50.
- S. Lee, J.-H. Ryu, B.-M. Hodge, I.-B. Lee, 2016. Development of a neural network-based renewable energy forecasting framework for process industries. In: Z. Kravanja, M. Bogataj (Eds.), 26th European Symposium on Computer Aided Process Engineering. Vol. 38 of Computer Aided Chemical Engineering. Elsevier, pp. 1527–1532.
- E. N. Pistikopoulos, A. Barbosa-povoa, J. H. Lee, R. Misener, A. Mitsos, G. V. Reklaitis, V. Venkatasubramanian, F. You, R. Gani, 2021. Process systems engineering – The generation next ? Computers and Chemical Engineering 147.
- A. Salah, L. Hanel, M. Beirov, G. Scheffknecht, 2016. Modelling ser biomass gasification using dynamic neural networks. In: Z. Kravanja, M. Bogataj (Eds.), 26th European Symposium on Computer Aided Process Engineering. Vol. 38 of Computer Aided Chemical Engineering. Elsevier, pp. 19–24.
- K. Scheibler, L. Winterer, R. Wimmer, B. Becker, 2015. Towards verification of artificial neural networks. In: MBMV.
- E. Sánchez-Ramírez, J. G. Segovia-Hernández, E. A. Hernández-Vargas, 2020. Artificial neural network to capture the dynamics of a dividing wall column. In: S. Pierucci, F. Manenti, G. L. Bozzano, D. Manca (Eds.), 30th European Symposium on Computer Aided Process Engineering. Vol. 48 of Computer Aided Chemical Engineering. Elsevier.
- V. Tjeng, K. Xiao, R. Tedrake, 2017. Evaluating robustness of neural networks with mixed integer programming. arXiv preprint arXiv:1711.07356.

Sustainable Analysis of Recent Acid Gas Treatment Schemes for LNG production

Ahmed AlNouss, Saad Al-Sobhi*

Department of Chemical Engineering, College of Engineering, Qatar University. Doha, Qatar

**Saad.al-sobhi@qu.edu.qa*

Abstract

The worldwide growing demand for natural gas consumption increases the pressure on the depletion of natural resources, consequently the significance of natural gas treatment methods. Given the huge investment associated with gas processing plants, selecting optimum and fit-for-purpose treatment methods undergoes a rigorous selection process to allocate the most cost-effective and environmentally friendly treatment scheme to remove contaminants. This study analyzes the sustainability of recent acid gas treatment schemes associated with Liquefied Natural Gas (LNG) production using Aspen simulation techniques. The acid gas treatment process is crucial in LNG production to eliminate the presence of CO₂ and H₂S from the sour feed gases for safety and environmental concerns. The recent advancement in acid gases removal units (AGRU) demonstrates the addition of an enrichment step using a low-pressure pre-flash column of Amine solvent along with a tail gas treatment (TGT) unit. The full-integrated scheme includes low BTX AGRU, Acid Gas Enrichment, Sulfur Recovery Unit, TGT unit, and common regeneration. The techno-economic evaluation demonstrates the excellence of this technology advancement in providing higher efficiency of acid gases removal with some increase in the equipment cost and huge environmental benefits.

Keywords: Acid gas treatment, LNG, Simulation, Sulfur recovery.

1. Introduction

The global energy demand continues to increase, putting more pressure on natural resources, specifically fossil fuels. The processing of these fossil fuels involves heavy treatment methods to eliminate acid gases, CO₂, H₂S, and BTX, avoiding their problems in processing lines. These gases are associated with serious negative environmental and health impacts due to the high toxicity of their combustion effluents, mainly SO₂, CO₂, and particulate matter. Environmental limits on emitting these gases into the atmosphere have been regulated by various environmental agencies and governments worldwide, emphasizing operation plants to improve the treatment methods of acid gases (Gupta et al., 2016). From another perspective, the variation of wellhead and gas processing conditions introduces a challenge for the production plants to treat the changes in acid gas content to achieve required product specifications. All of this places a huge challenge and burden on selecting the optimum scheme to meet environmental standards, respond flexibly to changes in feed gas quality, and ensure economic and technical targets of the different project circumstances. Given the huge investment associated with gas processing plants, selecting optimum and fit-for-purpose treatment methods undergoes a rigorous selection process to allocate the most cost-effective and environmentally friendly

treatment scheme to remove contaminants (Mokhatab and Poe, 2012; Korens et al., 2002). This study analyzes the sustainability of recent acid gas treatment schemes associated with Liquefied Natural Gas (LNG) production by utilizing Aspen simulation techniques. The acid gas treatment process is crucial in LNG production to eliminate the presence of CO₂ and H₂S from the sour feed gases for safety and environmental concerns. Evaluating the improvement in the acid gas treatment schemes depends mainly on the initial conditions of the sour feed gas, specifications of treated sweet gas, and environmental limitations. Performing a case-by-case optimization study is essential to achieve the optimum capital and operating costs and the largest window of operating parameters regarding sour feed gas quality. The most-effective treatment scheme must demonstrate economic effectiveness, accept different feeds, and achieve sulfur specifications in addition to the highest measures of reliability and quality performance.

The subject of selecting the optimum alternative for acid gases treatment and removal has been studied intensively in the literature. Korens et al. (2002) analyzed the performance of different processes applicable for integrated gasification combined cycle while meeting U.S. emission standards. The processes include CO₂ removal and sequestration, hot and warm gas cleanup, tail gas treating unit (TGTU), sulfur recovery unit (SRU), acid gas treatment unit (AGRU), acid gas injection, and mercury removal. They concluded that the current design of many processes would require improvement to meet the future more stringent environmental regulations. Similarly, Mokhatab and Poe (2012) studied the most commonly used processes in gas processing plants and their potential integration with expert process designs. The work attempted to establish different integration of gas processing steps to establish the optimum lineup of gas treatment while considering economics, operability, and flexibility. The processing steps included TGTU, SRU, AGRU, Mercaptans removal unit, and gas dehydration. Gupta et al. (2016) have conducted a recent comprehensive review of acid gas treatment and sulfur recovery designs. The review highlighted the advancement in Claus process design to handle increased amounts of sulfur and reduce the challenges in catalytic reactions and associated operational costs. Pellegrini et al. (2019) analyzed the energy and exergy performance of competing natural purification and liquefaction routes using Aspen HYSYS simulation based on net equivalent methane approach. Results demonstrated the excellence of the recent Dual Pressure Low-Temperature (DPLT) distillation technology over the conventional activated-MDEA technology for CO₂ removal with higher exergy efficiency and lower consumptions. Recently, Zhu et al. (2021) analyzed the absorption of H₂S and CO₂ from multiple gas feeds through sensitivity analysis and orthogonal test method. The study aimed to optimize the concentration and temperature distribution in the sweetening absorber through allocation and multi-feeding operation of gas feeds. The simulation results of the Shandong refinery revealed a 2.55% increase in H₂S purity in acid gas with a reduction in the overall exergy loss and energy consumption by 18.7% and 6.5%, respectively.

The current study objective is to compare and analyze the recent advancement in acid gas treatment with enrichment step and TGTU against the base scheme of AGRU and SRU only. This work is a continuation of the sustainability assessment in the LNG process performed for the helium extraction unit (AlNouss and Al-Sobhi, 2020), boil-off gas utilization in the liquefaction unit (Al-Sobhi et al., 2021) natural gas liquids unit (AlNouss et al., 2018; Al-Sobhi and AlNouss, 2018). Aspen HYSYS is utilized to simulate the flowsheet models of the base and advance schemes while utilizing the economic and environmental assessment features to account for and compare the capital and operating costs along with gaseous emissions.

2. Methodology and process description

The current commercial acid gas removal technologies include the physical solvent-based, the chemical solvent-based, and the mixed chemical/physical solvent-based processes. The physical solvent-based processes are associated with high expenses compared to chemical solvent-based ones, despite their capability of meeting the Claus reaction requirement of sulfur cleanup. Utilization of a mixed chemical/physical solvent stabilizes the tradeoff between the degree of physical solubility of sulfur compounds and H_2S selectivity; hence get benefits from the characteristics of both solvents. Employment of acid gas enrichment (AGE) stage enables higher efficiencies (~98%) in the Claus unit, the mainstay of SRU plant. Further, to enhance the SRU efficiency up to 99.9%, the TGTU is required to scrub out the H_2S generated from the hydrogenation step. The utilization of the AGE and TGTU steps is necessary to be able to meet stringent environmental regulations.

The base acid gas treatment scheme evaluated in this study is illustrated in Figure 1. The raw feed gas first enters the AGRU absorber and is contacted with the amine-based solvent. The solvent used in the study is Sulfinol consisting of Sulfolane (physical) and diisopropanolamine (DIPA) (chemical) mixture. This mixed solvent allows higher solubility of sulfur compounds relative to aqueous amines and better solvent loadings at higher acid gas partial pressures. The treated effluent gas from the absorber leaves from the top towards the dehydration process while the rich solvent recovered at the bottom enters a flash vessel to separate sour fuel gases. The rich solvent stream is further preheated by the lean solvent from the bottom of the regenerator column before entering the column. The acid gases are stripped from the regenerator top and sent to SRU while the lean solvent is recycled to the absorber column. Inside the SRU, the H_2S is oxidized to SO_2 before the Claus units where it is further converted to elemental sulfur and collected as a liquid through cooling. The leftover gas stream is sent to the stack.

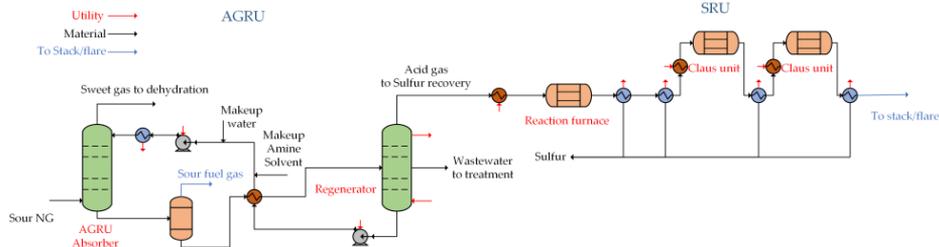


Figure 1: Base scheme of acid gas treatment facility.

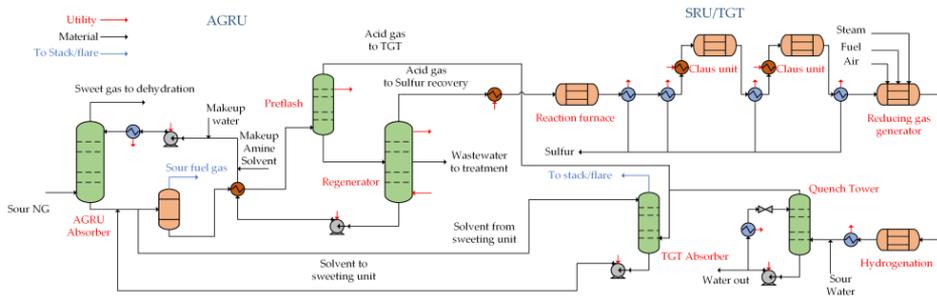


Figure 2: Advanced scheme of acid gas treatment facility.

The recent advancement in the acid gas treatment as presented in Figure 2 demonstrate the addition of enrichment step prior to the SRU and regenerator column and TGTU post

SRU to maximize sulfur recovery. The rich solvent from the flash vessel is sent to a pre-flash column to remove the acid gases partially and send them to TGTU while the remaining part is sent to the AGRU regenerator. The TGTU receives the acid gases from the pre-flash column and the effluent of the SRU unit after hydrogenation and quenching for further acid gas treatment.

The base and advanced schemes have been simulated using Aspen HYSYS software and further used to perform the economic and environmental assessment to account for and compare the capital and operating costs along with gaseous emissions. Despite the inclusion of the Sulsim tool inside Aspen HYSYS software, the SRU simulated as Sulsim sub-flowsheet can't be linked with the built-in economic and environmental analysis tools within Aspen HYSYS. Therefore, the entire system has been simulated using legacy HYSYS tools to estimate the costs and environmental emissions. The feed gas properties are presented in Table 1.

Table 1. Sour NG feed conditions and properties

Property	Value	Element	Composition (mol%)
Price [\$/T] (EIA, 2020)	211	CO ₂	4.95
Temperature [°C]	40	H ₂ S	26.73
Pressure [kPa]	4500	Methane	58.42
Molar Flow [kmole/h]	4185	Ethane	6.93
Mass Flow [kg/h]	100000	Propane	1.98
		Nitrogen	0.99

The economic and environmental emission parameters have then been used to conduct the sustainability assessment based on the sustainability weighted return on investment metric (SWROIM) formulated by El-Halwagi (2017). The conventional return on investment (ROI) is first calculated by dividing annual net economic profit (AEP) by the total capital investment (TCI). Then the equation of SWROIM is utilized to contribute to the rest of the sustainability indicators using weighing factors defined as a ratio of the relative importance of each sustainability indicator to the AEP as presented in Eq.(1). The indicator target is defined in this study to be the highest value from all alternatives. The alternatives are then benchmarked based on ROI and SWROIM.

$$SWROIM_p = \frac{AEP_p \left[1 + \sum_{i=1}^{N_{Indicators}} w_i \left(\frac{Indicator_{p,i}}{Indicator_i^{Target}} \right) \right]}{TCI_p} \quad \text{Eq. (1)}$$

3. Results and discussion

The economic and environmental assessment results revealed an enhancement in the environmental performance of the AGRU with the addition of the TGTU and AGE units despite the increase in capital and operating costs. The recovery of H₂S and CO₂ from the AGRU has increased with the pre-flash enrichment column from 68 and 61% for the base scheme to 78 and 93% for the advanced scheme, respectively. This positive recovery increase is associated with a negative increase in the capital and operating costs from 412 M\$ and 30.5 M\$/y to 433 M\$ and 56.1 M\$/y, respectively, as illustrated in Figure 3.

The addition of the AGE unit has also increased the production of elemental sulfur from 506.4 to 585.0 T/d given the higher recovery of H₂S, which has demonstrated an increase in the product sales from 118.8 to 119.1 M\$/y compared to the base scheme. The prices

of LNG and elemental sulfur used in the study are \$35/MMBTU and \$57/T (Company, 2020; FRED, 2021). Moreover, the environmental impact assessment presented in Figure 4 has demonstrated the excellence of TGTU addition in reducing the Sulfur dioxide and Carbon oxide emissions from 13.2×10^3 and 5.8×10^3 to 7.2 and 308 T/y, respectively. Whereas, the power consumption have increased from 2555.7 to 3671.4 GWh with the addition of AGE and TGTU steps.

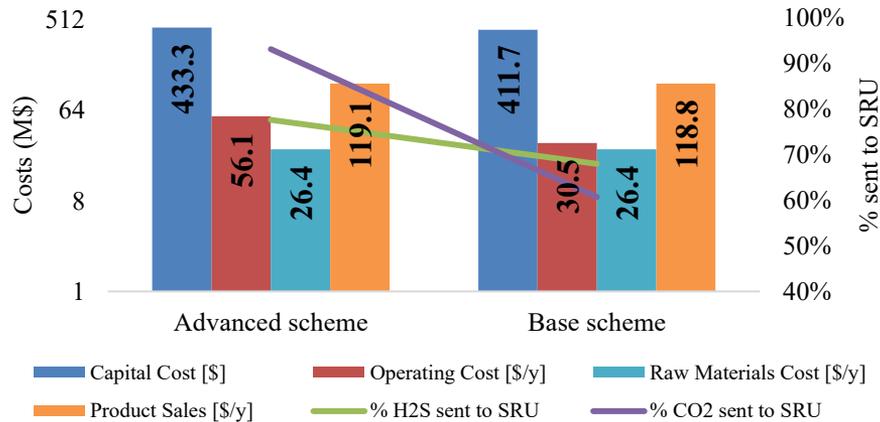


Figure 3: Economic and recovery results.

Table 2. Indicator targets

Property	H ₂ S recovery	CO ₂ recovery	SO ₂ emission [T/y]	CO emission [T/y]	Power (GWh)
Value	78%	93%	13.2×10^3	5.8×10^3	3671.4
w _i	0.25	0.25	-0.25	-0.25	-0.25

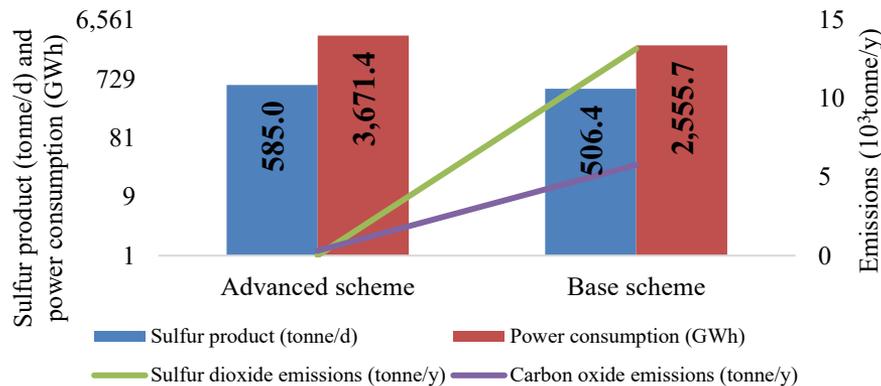


Figure 4: Environmental impacts and energy results.

These results have then been used to measure the SWROIM utilizing Eq.(1) with the indicator targets and weighting factors presented in Table 2. Implementing these weighting factors on the results obtained from the two schemes highlights the sustainability improvement in the acid gas treatment. The SWROIM results demonstrate the improvement in sustainability with the addition of the AGE and TGTU units. The conventional ROI indicated similar results from both schemes given the close capital

costs and product revenue. Whereas, taking into account the positive increase in H₂S and CO₂ recovery along with the positive decrease in SO₂ and CO emissions and the negative increase in power consumption has revealed the increase in sustainability metric for the advanced scheme from 22.1% to 27.4%, while the sustainability metric for the base scheme has decreased from 22.3% to 15.8%.

Table 1. Sustainability results

Parameter	AEP	ROI (%)	SWROIM (%)
Advanced scheme	8.7399x10 ⁷	22.1%	27.4%
Base scheme	8.7412x10 ⁷	22.3%	15.8%

4. Conclusion

Natural gas processing involves heavy treatment methods to eliminate acid gases, CO₂, H₂S, and BTX, avoiding their problems in processing lines. The current study compares the recent advancement in acid gas treatment with the addition of enrichment step and TGTU against the base scheme of AGRU and SRU only. The techno-economic evaluation demonstrated the excellence of this technology advancement in providing higher efficiency of acid gases removal with some increase in the equipment cost and huge environmental improvement.

References

- S.A. Al-Sobhi, and A. AlNouss, 2018, Applying New Sustainability Metric in Different Natural Gas Liquid (NGL) Recovery Configurations to Extend Investment Decision and Incorporate Sustainability Analysis, *Computer Aided Chemical Engineering*, 43, 145-50.
- S.A. Al-Sobhi, A. AlNouss, M. Shamlooh, K. Al-Nuaimi, A. AlMulla, and M. Khraisheh, 2021, Sustainable boil-off gas utilization in liquefied natural gas production: Economic and environmental benefits, *Journal of Cleaner Production*, 296, 126563.
- A. AlNouss, and S.A. Al-Sobhi, 2020, Circular Economy Analysis of Helium Recovery from Sales Gas Product, *Computer Aided Chemical Engineering*, 48, 913-18.
- A. AlNouss, M. Ibrahim, and S.A. Al-Sobhi, 2018, Potential energy savings and greenhouse gases (GHGs) emissions reduction strategy for natural gas liquid (NGL) recovery: Process simulation and economic evaluation, *Journal of Cleaner Production*, 194, 525-39.
- I. Company, 2020, Sulphur market prices 2020, Ishtar Company LLC, <https://ishtarcompany.com/knowledge-base/sulfur-prices/>
- EIA, 2020, Natural Gas Spot and Futures Prices (NYMEX) U.S. Energy Information Administration, https://www.eia.gov/dnav/ng/NG_PRI_FUT_S1_M.htm
- M.M. El-Halwagi, 2017, A return on investment metric for incorporating sustainability in process integration and improvement projects, *Clean Technologies and Environmental Policy*, 19, 2, 611-17.
- FRED, 2021, Global price of LNG, Federal Reserve Bank of St. Louis, <https://fred.stlouisfed.org/series/PNGASJPUSDM>
- A.K. Gupta, S. Ibrahim, and A. Al Shoaibi, 2016, Advances in sulfur chemistry for treatment of acid gases, *Progress in Energy and Combustion Science*, 54, 65-92.
- N. Korens, D. Simbeck, and D. Wilhelm, 2002, Process Screening Analysis of Alternative Gas Treating and Sulfur Removal for Gasification.
- S. Mokhatab, and W.A. Poe, 2012, Chapter 13 - Selecting the Best Gas Processing Route, *Handbook of Natural Gas Transmission and Processing (Second Edition)*, 451-72.
- L.A. Pellegrini, G. De Guido, and V. Valentina, 2019, Energy and exergy analysis of acid gas removal processes in the LNG production chain, *Journal of Natural Gas Science and Engineering*, 61, 303-19.
- W. Zhu, H. Ye, X. Zou, Y. Yang, and H. Dong, 2021, Analysis and optimization for chemical absorption of H₂S/CO₂ system: Applied in a multiple gas feeds sweetening process, *Separation and Purification Technology*, 276, 119301.

32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING

VOLUME 2

COMPUTER-AIDED CHEMICAL ENGINEERING, 51

**32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING**

VOLUME 2

Edited by

Ludovic Montastruc

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

ludovic.montastruc@ensiacet.fr

Stephane Negny

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

stephane.negny@ensiacet.fr



ELSEVIER

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

Elsevier

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

Copyright © 2022 Elsevier B.V. All rights reserved.

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

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

Notices

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

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

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

British Library Cataloguing in Publication Data

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

Library of Congress Cataloging-in-Publication Data

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

ISBN (Volume 2): 978-0-443-18632-5

ISBN (Set) : 978-0-323-95879-0

ISSN: 1570-7946

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



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Candice Janco

Acquisition Editor: Anita Koch

Editorial Project Manager: Lena Sparks

Production Project Manager: Paul Prasad Chandramohan

Designer: Mark Rogers

Typeset by STRAIVE

Contents

T2: Product/Process Synthesis and Design

- 100. Promising future for biodiesel: Superstructure optimization from feed to fuel**
Thien An Huynh, Mattia Rossi, Maryam Raeisi, Meik B. Franke, Flavio Manenti, Edwin Zondervan **595**
- 101. Model-based solvent selection for integrated synthesis, crystallisation and isolation processes**
Mohamad H. Muhieddine, Suela Jonuzaj, Shekhar K. Viswanath, Alan Armstrong, Amparo Galindo and Claire S. Adjiman **601**
- 102. Computer-aided solvent design for suppressing HCN generation in amino acid activation**
Lingfeng Gui, Alan Armstrong, Amparo Galindo, Fareed Bhasha Sayyed, Stanley P. Kolis and Claire S. Adjiman **607**
- 103. Improving sustainable CHOSYN's targets through process intensification**
Maricruz Juárez-García, Gabriel Contreras-Zarazúa, Juan Gabriel Segovia-Hernandez, Jose María Ponce-Ortega **613**
- 104. A novel process design for automated quality analysis in an integrated biopharmaceutical platform**
Simon Tallvod, Mikael Yamanee-Nolin, Joaquín Gomis-Fons, Niklas Andersson and Bernt Nilsson **619**
- 105. A multiscale approach towards the use of Urban waste within the energy transition strategy**
Sofía González-Núñez, Lidia S. Guerras, Mariano Martín **625**
- 106. Integration of thermo-chemical energy storage for flexible operation of decarbonized power plants**
Calin-Cristian Cormos, Letitia Petrescu, Ana-Maria Cormos **631**
- 107. Superstructure optimization for sustainable design of an algae biorefinery**
Maryam Raeisi, Thien An Huynh, Meik B. Franke, Edwin Zondervan **637**
- 108. Economic and Environmental Optimization and Feedstock Planning for the Renewable Jet Production Using an Intensified Process**
Raul Mauricio Rivas-Interian, Eduardo Sanchez-Ramirez, Juan Gabriel Segovia-Hernandez, Juan Jose Quiroz-Ramirez **643**

- 109. Heat exchanger networks with different shell and tube configurations**
Jose A. Caballero, Mauro A.S.S. Ravagnani, Leandro Pavao, Caliane B.B. Costa, Juan Javaloyes-Anton **649**
- 110. Considering Environmental and Social Impact in Optimization of a Supply Chain for Bio-Jet Fuel Production**
David Vallejo-Blancas, Juan Gabriel Segovia-Hernández, Eduardo Sánchez-Ramírez, Gabriel Contreras-Zarazúa, Juan José Quiroz-Ramírez **655**
- 111. Superstructure Optimization of Dimethyl Ether Process**
Emrullah Erturk, Erdal Aydin, Sahin Sarrafi, Ozgun Deliismail, Aysel Zahidova, Hasan Sildir **661**
- 112. A Comprehensive Evaluation on the COSMO-SAC-dsp model for Vapor-Liquid Equilibrium Predictions of Refrigerants**
Jiayuan Wang, Tongxin Sa, Lingyu Zhu **667**
- 113. Design and Evaluation Strategies for a Novel Hybrid Trigeration System**
Rachid Klaimi, Sabla Y. Alnouri, Mirko Stijepovic **673**
- 114. Process synthesis for desalination using superstructure optimization on membrane technologies**
Alvaro Martinez-Triana, Mohamed Kanniche, Thibaut Neveux, Christophe Castel and Eric Favre **679**
- 115. A comparison of process synthesis approaches for multistage separation processes by gas permeation**
Thibaut Neveux, Bernardetta Addis, Christophe Castel, Veronica Piccialli and Eric Favre **685**
- 116. Bifurcation analysis of combined agglomeration and layering granulation in fluidized bed spray processes**
Eric Otto, Robert Dürr and Achim Kienle **691**
- 117. Enhancement of energy saving of reverse osmosis system via incorporating a photovoltaic system**
Alanood A. Alsarayreh, Mudhar A. Al-Obaidi, Saad S. Alrwashdeh, Raj Patel and Iqbal M. Mujtaba **697**
- 118. Design and simulation of methanol synthesis using heavy residues with carbon utilization**
Siddig S. Khalafalla, Hussain A. Alibrahim, Fayez Nasir Al-Rowaili, and Umer Zahid **703**

- 119. Integrated process synthesis and design of intensified/hybrid processes**
Shuang Xu, Selen Cremaschi, Mario R. Eden and Anjan K. Tula **709**
- 120. Macroalgae-based integrated biorefinery for hydrocolloids, chemicals and advanced biofuels production**
Andrés I. Casoni, Alejandro H. Pedrozo, Fernando D. Ramos, Vanina Estrada, M. Soledad Diaz **715**
- 121. Design Space Determination of Mesenchymal Stem Cell Cultivation by Dynamic Modeling under Uncertainty**
Keita Hirono, Isuru A. Udugama, Yusuke Hayashi, Masahiro Kino-Oka, Hirokazu Sugiyama **721**
- 122. Deterministic global optimization of multistage layer melt crystallization using surrogate models and reduced space formulations**
Christian Kunde, Ronald Méndez and Achim Kienle **727**
- 123. Model-Based Design Space for Protein A Chromatography Resin Selection**
Steven Sachio, Cleo Kontoravdi and Maria M. Papathanasiou **733**
- 124. Sustainable Energy Systems Design with Heat Pump Technologies for Meeting Peak Heating Demand**
Xueyu Tian, Fengqi You **739**
- 125. Isopropanol/ n -Butanol/ Ethanol separation from diluted fermentation broth by distillation. Process optimization using MILP techniques**
Eszter Toth, Gianluca Mancino, Ludovic Raynal **745**
- 126. Automatic synthesis of hybrid processes using distillation and liquid-liquid extraction for the separation of azeotropic mixtures**
Thulasi Sasi, Mirko Skiborowski **751**
- 127. A framework for optimal and flexible schemes design under uncertainty & sustainable aspects**
Jaime D. Ponce-Rocha, Martín Picón-Núñez, and Ricardo Morales-Rodriguez **757**
- 128. Optimization-based assessment framework for CO₂ utilization to fuels strategies**
Thai Ngan Do, Chanhee You, Minseong Park, Changsu Kim, Seongju Lee, and Jiyong Kim **763**

- 129. A novel approach to continuous extraction of active ingredients from essential oils through combined chromatography**
Constantin Jurischka, Bogdan Dorneanu, Christian Stollberg, Harvey Arellano-Garcia **769**
- 130. Digital design and optimization of an integrated reaction-extraction-crystallization-filtration continuous pharmaceutical process**
Timothy J. S. Campbell, Chris D. Rielly, Brahim Benyahia **775**
- 131. Multi-objective optimization of natural gas liquefaction process simulation via kriging surrogate model**
Lucas F. Santos, Caliane B. B. Costa, Jose A. Caballero, Mauro A. S. S. Ravagnani **781**
- 132. Integrated synthesis, modeling, and assessment of waste-to-resource alternatives**
Adrián Pacheco-López, Ana Somoza-Tornos, Moisés Graells, Antonio Espuña **787**
- 133. MINLP model for work and heat exchange networks synthesis considering unclassified streams**
Lucas F. Santos, Caliane B. B. Costa, Jose A. Caballero, Mauro A. S. S. Ravagnani **793**
- 134. Planetary Boundaries Assessment of Flue Gas Valorization into Ammonia and Methane**
Sebastiano C. D'Angelo, Julian Mache, Gonzalo Guillén Gosálbez **799**
- 135. A techno-economic assessment of biochar production from date pits in the MENA region**
Aisha Al-Rumaihi, Mohammad Alherbawi, Prakash Parthasarathy, Hamish R. Mackey, Gordon McKay and Tareq Al-Ansari **805**
- 136. Design of sustainable processes for CO₂ capture**
Ana Gabriela Romero-García, Nelly Ramírez-Corona, Eduardo Sánchez-Ramírez, Heriberto Alcocer-García, Cataldo De Blasio, Juan Gabriel Segovia-Hernández **811**
- 137. Water network optimisation in chemical complexes: a refinery case study**
Francisco J.G. Patrocínio, Hugo M.D. Carabineiro, Henrique A. Matos and Nuno M.C. Oliveira **817**
- 138. Optimal design of solar-aided hydrogen production process using molten salt with CO₂ utilization for polypropylene carbonate production**
Wanrong Wang, Nan Zhang, and Jie Li **823**

- 139. Hierarchical Approach for Solvent Selection in Circular Economy**
Pranav Majgaonkar, Kai Kruber, Venkat Aryan, Ronny Hanich, Davide Pico, Mirko Skiborowski **829**
- 140. A Framework to Facilitate Decision Making for Infrastructure Options Analysis of Distribution and Utilities Systems in Chemical Production Plants**
Marcello Di Martino, Iosif Pappas, Anh Tran, R. Cory Allen, Russell R. Husfeld, Sam Eleff, Scott G. Moffatt, Styliani Avraamidou, Burcu Beykal and Efstratios N. Pistikopoulos **835**
- 141. An Optimization Model for Expansion Planning of Reliable Power Generation Systems**
Seolhee Cho and Ignacio E. Grossmann **841**
- 142. Application of an ontology-based decision support system for the design of emulsion-based cosmetic products**
Juliana Serna, Jose L. Rivera-Gil, Alex Gabriel, Javier A. Arrieta-Escobar, Vincent Boly, Véronique Falk, Paulo C. Narváez Rincón **847**
- 143. Synthesis of extractive distillation structures for the purification of ethanol**
Tomoki Nakamura, J. Rafael Alcántara Ávila, Julián Cabrera-Ruiz **853**
- 144. Optimal design for flexible operation with multiple fluctuating input parameters**
Bastian Bruns, Marcus Grünewald and Julia Riese **859**
- 145. Design and Optimization of Membrane System for Gas Separation**
Shivom Sharma and François Maréchal **865**
- 146. Assessment of Carbon Capture Technologies for Waste-to-Energy System**
Shivom Sharma, Rafael Castro-Amoedo, Jaroslav Hemrle, and François Maréchal **871**
- 147. Process Integration and Techno-Economic Assessment of a Green Biorefinery Demonstration Scale Platform for Leaf Protein Production**
Thalles A. Andrade, Morten Ambye-Jensen **877**
- 148. The effect of alternative fuels on calcium looping**
Ana Amorim, Ismail Mohamed, Luíza Marques, Rui M. Filipe, Henrique A. Matos **883**
- 149. Simultaneous Synthesis of Metabolic and Process Engineering for the Production of Muconic Acid**
Dimitriou Konstantinos, Kokossis Antonis **889**

150. Renewable hydrogen supply chain for transport application in Corsica island	
<i>T. Moustapha Mai, C. Cristofari, C. Azzaro-Pantel, E. Carrera</i>	895
 T3: Large Scale Design and Planning/Scheduling	
151. Optimal configuration of a biodiesel production network using oil from black soldier fly larvae	
<i>Dulce María Aguilar-Murguía, Sergio Iván Martínez-Guido, Juan Fernando García-Trejo, Salvador Hernández, Claudia Gutiérrez-Antonio</i>	901
152. Optimization of Sink Locations in Carbon Integration Networks	
<i>Sabla Y. Alnouri, Ilkan Sarigol, Dhabia Al-Mohannadi, Hadi Jaber</i>	907
153. Towards Efficient Bioenergy Systems: Understanding the Role of Soil Sequestration, Supply Chain Design, and Carbon Capture	
<i>Caleb H. Geissler, Eric G. O'Neill, Christos T. Maravelias</i>	913
154. A Multi-Period Planning and Scheduling Strategy for Developing Hydrogen-Based Supply Chains	
<i>R. Cory Allen, C. Doga Demirhan, Clara F. Heuberger-Austin and Efstratios N. Pistikopoulos</i>	919
155. Holistic Capacity Management and Production Planning in the Pharmaceutical Supply Chain	
<i>Simon B. Lindahl, Deenesh K. Babi, Krist V. Gernaey, Gürkan Sin</i>	925
156. Towards Resilience in Next-Generation Vaccines and Therapeutics Supply Chains	
<i>Miriam Sarkis, Kyungjae Tak, Benoit Chachuat, Nilay Shah and Maria M. Papathanasiou</i>	931
157. Network-Based Analysis of Electrified Chemical Processing with Renewable Energy Sources	
<i>Ioannis Giannikopoulos, Alkiviadis Skouteris, David T. Allen, Michael Baldea and Mark A. Stadtherr</i>	937
158. MINLP framework for systems analysis of the chemical manufacturing industry using network models	
<i>Alkiviadis Skouteris, Ioannis Giannikopoulos, David T. Allen, Michael Baldea and Mark A. Stadtherr</i>	943
159. A Novel Approach for Vehicle Fleet Sizing and Allocation under Uncertain Demand	
<i>Demian J. Presser, Vanina G. Cafaro, Diego C. Cafaro</i>	949

- 160. Fresh vs frozen: assessing the impact of cryopreservation in personalised medicine**
Niki Triantafyllou, Andrea Bernardi, Matthew Lakelin, Nilay Shah, Maria M. Papathanasiou **955**
- 161. A novel hybrid algorithm for scheduling multipurpose batch plants**
Dan Li, Dongda Zhang, Nan Zhang, Liping Zhang, Jie Li **961**
- 162. Assessment of biomass supply chain design and planning using discrete-event simulation modeling**
Helena Paulo, Miguel Vieira, Bruno S. Gonçalves, Tânia Pinto-Varela and Ana P. Barbosa-Póvoa **967**
- 163. Design of a three-echelon supply chain under uncertainty in demand and CO₂ allowance prices**
F. L. Garcia Castro, R. Ruiz-Femenia, R. Salcedo-Diaz and J. A. Caballero **973**
- 164. An industrial perspective on simulation versus optimisation decision-making capabilities**
Robert E. Franzoi, Mahmoud A. Ahmednooh, Brenno C. Menezes **979**
- 165. Decomposition of Two-stage Stochastic Scheduling Problems via Similarity Index**
Daniel Montes, José L. Pitarch and César de Prada **985**
- 166. A New MILP Formulation for Scheduling of Crude Oil Operations**
Pelin Dologlu, Sena Kurban, İrem Marttin, Nazlı Ataman, Gizem Kuşoğlu Kaya, Funda İşeri, Özgür Kabak, Y. Ilker Topcu **991**
- 167. Multiperiod optimization model for CO₂ capture, utilization and storage, Colombian case study**
Alexandra Duarte, Javier D. Angarita, Juan P. Espinosa-Cárdenas, Javier Lizcano, Rafael C. García-Saravia, Ariel Uribe-Rodríguez **997**
- 168. Towards improved scheduling: an analysis on time-steps, time-horizon, and rescheduling**
Robert E. Franzoi, Brenno C. Menezes **1003**
- 169. Pharmaceutical Industry Supply Chains: Planning Vaccines' Distribution**
Inês Duarte, Bruna Mota, Ana Paula Barbosa-Póvoa **1009**

T4: On Line Model Based Applications and Control

- 170. Plantwide control strategy for a biodiesel production process from *Jatropha curcas* oil with variable composition**
Osiris Martínez-Sánchez, Fernando Israel Gómez-Castro, Nelly Ramírez-Corona **1015**
- 171. Cascade fuzzy control of a tubular chemical reactor**
Anna Vasičkaninová, Monika Bakošová, Alajos Mészáros **1021**
- 172. Design and Implementation of an Optimal Control Framework for Post-combustion CO₂ Capture Process**
Gaurav Mirlekar, Lars O. Nord **1027**
- 173. Data Driven Process Monitoring, from Fault Detection and Diagnosis Points of View, in Industry 4.0 Context**
Hamed Ardakani, Hadiseh Hemmati, Manuel Rodriguez Hernandez **1033**
- 174. Safe Chance Constrained Reinforcement Learning for Batch Process Optimization and Control**
Max Mowbray, Panagiotis Petsagkourakis, Antonio Del Rio Chanona and Dongda Zhang **1039**
- 175. Support Vector Machine-based Design of Multi-model Inferential Sensors**
Martin Mojto, Karol Lubušký, Miroslav Fikar and Radoslav Paulen **1045**
- 176. Model predictive control for greenhouse condition adjustment and crop production prediction**
Guoqing Hu, Fengqi You **1051**
- 177. Study on the noise contents of different measurements in industrial process and their impact on process monitoring**
Tingting Tao, Cheng Ji, Chengyu Han, Jingde Wang, Wei Sun **1057**
- 178. A Comparative Study on the Influence of Different Prediction Models on the Performance of Residual-based Monitoring Methods**
Jiaying Ma, Yang Li, Fangyuan Ma, Jingde Wang, Wei Sun **1063**
- 179. Optimal operation of parallel mini-bioreactors in bioprocess development using Multi-Stage MPC**
Niels Krausch, Jong Woo Kim, Sergio Lucia, Sebastian Groß, Tilman Barz, Peter Neubauer, Mariano N. Cruz Bournazou **1069**

- 180. A platform for online system identification of dynamic ultrafiltration systems**
Oscar A. Prado-Rubio, Victor H. Grisales Diaz, Jakob Kjøbsted Huusom **1075**
- 181. Development of a Virtual Sensor for Real-Time Prediction of Granule Flow Properties**
Rexonni B. Lagare, Mariana Araujo da Conceicao, Ariana Camille Acevedo Rosario, Katherine Leigh Young, Yan-Shu Huang, M. Ziyen Sheriff, Clairmont Clementson, Paul Mort, Zoltan Nagy, Gintaras V. Reklaitis **1081**
- 182. A Hierarchical Approach to Monitoring Control Performance and Plant-Model Mismatch**
M. Ziyen Sheriff, Yan-Shu Huang, Sunidhi Bachawala, Marcial Gonzelez, Zoltan K. Nagy, Gintaras V. Reklaitis **1087**
- 183. Optimal Control Policies of a Crystallization Process Using Inverse Reinforcement Learning**
Paul Danny Anandan, Chris D. Rielly, Brahim Benyahia **1093**
- 184. Dynamic Real-Time Optimization with Closed-Loop Prediction for Nonlinear MPC-Controlled Plants**
Daniela Dering and Christopher L. E. Swartz **1099**
- 185. Dynamic Real-Time Optimization of a Solar Thermal Plant during Daytime**
Alix Untrau, Sabine Sochard, Frédéric Marias, Jean-Michel Reneaume, Galo A.C. Le Roux, Sylvain Serra **1105**
- 186. Online Bayesian Re-design of Parallel Experiments based on Asynchronous Posterior Sampling**
Martin F. Luna, M. Nicolás Cruz B., Ernesto C. Martínez **1111**
- 187. Dynamic optimisation and comparative analysis of fed-batch and perfusion bioreactor performance for monoclonal antibody (mAb) manufacturing**
Wil Jones, Dimitrios I. Gerogiorgis **1117**
- 188. Control of an industrial packed extraction column for biodiesel washing**
Manuel L. Pinho, Catarina G. Braz, Henrique A. Matos, Carla I.C. Pinheiro and José F. O. Granjo **1123**
- 189. Development of mechanistic reduced order models (ROMs) for glidant and lubricant effects in continuous manufacturing of pharmaceutical solid-dosage forms**
Sunidhi Bachawala and Marcial Gonzalez **1129**

190. Health-aware control using hybrid models applied to a gas-lifted oil well network <i>Jose Matias, Salmon Y. Ghebredngl and Johannes Jaschke</i>	1135
191. Applying Ecological Interface Design for Modular Plants: Safety-Demonstrator Case Study <i>Nazanin Hamed, Leon Urbas, Raimund Dachsel</i>	1141
192. Path-following for parametric MPCC: a flash tank case study <i>Caroline S. M. Nakama, Peter Maxwell and Johannes Jäschke</i>	1147
193. Feature engineering for neural network-based oscillation detection in process industries <i>Aswin Krishna M, Ayush Kedawat, Abhishek Bansal and Resmi Suresh</i>	1153
194. Multi-parametric Model Predictive Control Strategies for Evaporation Processes in Pharmaceutical Industries <i>Ioana Naşcu, Nikolaos A. Diangelakis, and Efstratios N. Pistikopoulos</i>	1159
195. Process as a battery: Electricity price based optimal operation of zeolite crystallization in a COBR <i>Robin Semrau, Jiadi Yang and Sebastian Engell</i>	1165
196. Leveraging Deep Learning for Efficient Explicit MPC of High-Dimensional and Non-linear Chemical Processes <i>Ahmed Shokry, Mehdi Abou El Qassime, Eric Moulines</i>	1171
197. Incorporation of error propagation into an elemental balancing based soft-sensor for improved online monitoring of microbial fed-batch processes <i>Don Fabian Müller, Konrad Lagoda, Daniel Wibbing, Christoph Herwig and Julian Kager</i>	1177
198. In-silico Formulation of Iterative Learning Control for Chromatographic Purification of Biopharmaceuticals <i>Daniel Espinoza, Niklas Andersson and Bernt Nilsson</i>	1183
199. Application of Real-Time Optimization with Modifier Adaptation to the Reactive Extrusion of Hydrophobically Modified Ethoxylated Urethanes <i>Maximilian Cegla and Sebastian Engell</i>	1189

Promising future for biodiesel: Superstructure optimization from feed to fuel

Thien An Huynh,^{a*} Mattia Rossi,^b Maryam Raeisi,^a Meik B. Franke,^a Flavio Manenti,^b Edwin Zondervan,^a

^a*Sustainable Process Technology, Faculty of Science and Technology, University of Twente, Meander, kamer 216, Postbus 217, 7500 AE Enschede, the Netherlands*

^b*Politecnico di Milano, Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Piazza Leonardo da Vinci, 32, 20133 Milano, Italy*

* t.a.huynh@utwente.nl

Abstract

A superstructure model for maximizing the profit of biodiesel production is presented in this work. The superstructure encompasses a wide range of feedstocks (e.g., waste cooking oil, tallow, rapeseed oil and algae), conventional reaction and separation equipment (e.g., continuous stirred tank reactor, decanter and vacuum distillation) and intensified operation units such as membrane reactor and reactive distillation column. The superstructure model is implemented in Advanced Interactive Multidimensional Modeling (AIMMS). The results present an optimal design of a biodiesel production process from waste cooking oil and tallow with a heterogeneous acid catalyst, a reactive distillation column and additional purification steps for producing pure glycerol which is 40% higher in price than technical glycerol. The total annual profit of the biodiesel production from waste cooking oil is 828,697 USD and from tallow is 976,450 USD. The results show that the combination of feedstock selection and implementation of advanced processing technologies to improve biodiesel production can be achieved with the superstructure optimization method.

Keywords: Superstructure, Optimization, Biodiesel, AIMMS, Design.

1. Introduction

Global climate change becomes more severe every year because of the increasing greenhouse gas (GHG) emission. The transport sector which contributes 16.2% of the global GHG emission becomes one of the targets of the European Commission (EC) renewable energies directive (2018/2001/EU) known as "RED II" (Ritchie and Roser, 2020). In RED II, the target for renewable energy in the transport sector is 14% with the share of advanced biofuel being 3.5% in 2030 (Observ'ER, 2020). However, the growth of biofuels has been slowing down recently with an increase of only 6.8% from 2018 to 2019 comparing to 12.3% between 2017 and 2018 because of high production cost and the change for non-food feedstocks (Observ'ER, 2020). Biodiesel, which shared 80.6% of total biofuel consumption in the European Union (EU) transport sector in 2019, plays an important role in the EU strategy to reduce GHG (Observ'ER, 2020). Therefore, the reduction of biodiesel production cost becomes an attractive topic for researchers.

Biodiesel is produced from renewable biobased feedstocks through chemical reactions. The selection of feedstocks is critical to the economic feasibility of biodiesel production because approximately 80% of biodiesel production cost comes from raw material costs (Zhang et al., 2003b). The biodiesel obtained from different feedstocks can be categorized

into three generations. The first generation biodiesel is derived from edible oil. The second generation biodiesel comes from non-edible oil, waste cooking oil and animal fat. The third generation is obtained from micro algae (Fazal et al., 2019). However, edible oils are generally expensive. Using edible oils to produce biodiesel leads to food-fuel debate over the concern of availability of food sources for people. Thus, the research of second- and third-generation biodiesel has become more important.

Besides the feedstocks, taking advantages of process intensification technologies in production can reduce the biodiesel cost. For example, a catalytic reactive distillation process has many benefits for biodiesel production such as lower equipment and operating cost, high productivity and reduced waste (Kiss et al., 2007). Supercritical methods present alternative biodiesel production routes without using a catalyst. Supercritical methanol as a reactant is a popular method which requires few equipment and can use both clean and used oil as feedstock (Gomez-Castro et al., 2013).

Superstructure optimization is a model-based computer-aided method of process systems engineering which can be used to optimize biofuel production process. A superstructure is a collection of numerous technical and feedstock alternatives that can form various possible process flowsheets. The superstructure can be formulated as a mathematical model by using variables, equations and constraints to describe its alternatives and flowsheets. The best flowsheet for predefined criteria and constraints can be identified by solving the model with optimization software tools (Tula et al., 2017).

Superstructure optimization has been increasingly applied in the field of biofuel to optimize the production process. AlNouss et al. (2019) developed a superstructure model of multiple biomass feedstocks and multiple biochemical products which is used to determine the optimal production process in terms of economic and environmentally friendly. Kenkel et al. (2021) presented an open-source generic superstructure optimization for modeling and optimizing of production process which is applied to design a cost optimal plant which produced methanol from captured CO₂.

However, most of published studies focused mainly on one aspect between finding inexpensive feedstocks and developing optimal processes for biodiesel production. A combined strategy of economic feedstock and innovative processes for biodiesel production has not been addressed adequately.

Therefore, the work presents a superstructure model which connects appropriate feedstock selection with process synthesis. In addition to the biodiesel production, the superstructure model comprises glycerol treatment units to increase the value of the by-product. With this superstructure model, an economic optimization is conducted for a production process from feedstock to biodiesel and glycerol. The results of this work will be compared to a conventional biodiesel production process (Zhang et al., 2003a).

2. Superstructure for biodiesel production

2.1. Problem statement

Given are units of feedstocks and reactants with initial flowrates and compositions, and reaction and separation equipment with technical specifications for producing and purifying biodiesel and glycerol. Under conditions that 1) each unit is assigned a logical decision variable and the units are connected to form possible processing routes by logical constraints, 2) the input and output flowrates of each unit are calculated by mass balance constraints, 3) The utility consumption is estimated from the heating, cooling and electricity requirement of each unit which are calculated based on flowrates and technical specifications, 4) The economics of an unit including capital investment and operating costs is calculated from flowrate and utility consumption. Deciding the feedstock and the

processing route to produce biodiesel while complying with the constraints and biodiesel purity as required in EU biodiesel standards, EN 14214. Maximizing the total profit of biodiesel production including biodiesel and glycerol sales.

2.2. Superstructure description

The superstructure model has total 67 units for biodiesel production and glycerol treatment as presented in Figure 1 A) and B). For feedstocks, the superstructure includes three biodiesel generations such as rapeseed oil and canola oil (first generation), waste cooking oil, tallow and linseed oil (second generation), and algae oil (third generation). For reaction, continuous stirred tank reactor, reactive distillation column and membrane reactor with different catalysts are included. For separation, units are vacuum distillations, decanters, acid and base neutralization reactors, water washing columns and hydrocyclones. After purification, the product stream has 96.5 wt.% FAME. Depending on its initial purity, the glycerol separated from the biodiesel production can be disposed as a waste or purified further to sell as crude glycerol (~80 - 98 wt.%), technical glycerol (~98 - 99.5 wt.%) and pure glycerol (> 99.5 wt.%) (Bart et al., 2010).

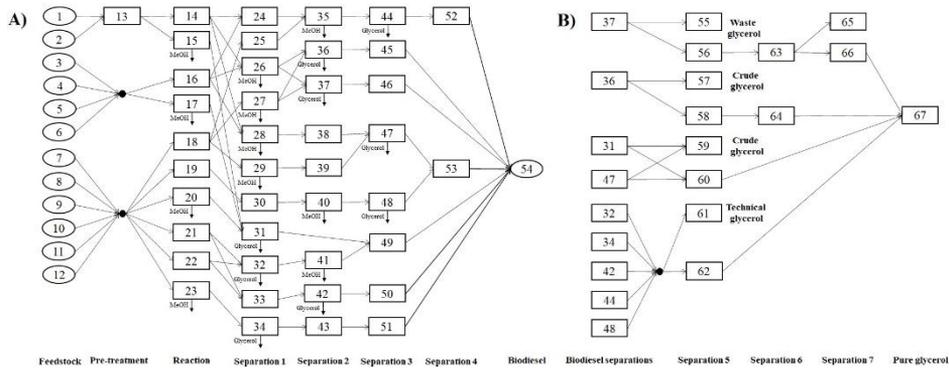


Figure 1: The superstructure for A) biodiesel production and B) glycerol treatment.

2.3. Mathematical model formulation

Each unit, j , of the superstructure has an infeed stream, F , and a reactant stream, R , which come in, and a product stream, P , and a waste stream, W , which come out, as shown in Figure 2.

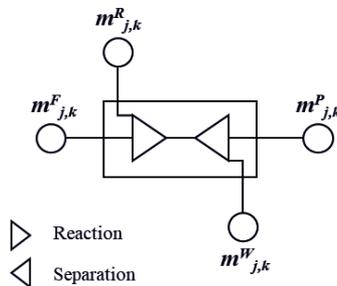


Figure 2: The illustration of streams which come in and out of a unit in the superstructure. Mass balances of a unit is described as Eq.(1) and Eq.(2).

$$m_{j,k}^P = (m_{j,k}^F + m_{j,k}^R + M_k \cdot ER \cdot \alpha_k) \cdot SF_{j,k} \cdot y_j \quad (1)$$

$$m_{j,k}^W = (m_{j,k}^F + m_{j,k}^R + M_k \cdot ER \cdot \alpha_k) \cdot (1 - SF_{j,k}) \cdot y_j \quad (2)$$

where k is a component of the streams (e.g., FAME, MeOH, glycerol, etc.), $m_{j,k}^F$, $m_{j,k}^R$, $m_{j,k}^P$ and $m_{j,k}^W$ are mass flow rates (kg/h) of component k in feed, reactant, product and waste streams, respectively. y_j is the logical decision variable which takes the value of 1 or 0 if the unit is selected or not. $SF_{j,k}$ is the split factor of component k based on how much k in the feed stream split into product stream. ER is the extend of reaction ($kmol/h$) calculated from the conversion of the main component of the feed stream into the main product of the product stream (e.g., from oil into biodiesel) as shown in Eq.(3).

$$ER = \frac{\theta_A \cdot (m_{j,A}^F + m_{j,A}^R)}{(M_A \cdot \alpha_A)} \quad (3)$$

where θ_A is the percent of component A being converted into product and α_A is the reaction stoichiometric number of A. From the extend of reaction, the conversion of another component can be calculated with the molar weight ($kg/kmol$), M_k , and the reaction stoichiometric number, α_k , of that component.

The product stream of a unit is the feed stream of the next one which directly connects to that unit on the same process route.

The energy requirement (kWh), E , of a unit is calculated as shown in Eq.(4).

$$E_j^u = (m_j^F + m_j^R) \cdot \tau_j^u \cdot H \cdot y_j \quad (4)$$

Where u is a type of utility such as heating, cooling and electricity, τ_j^u is the specific utility requirement (kW/kg) of the equipment, H is the total operating hours per year (h).

The equipment cost (USD), EC_j , of an unit is estimated based on the order of magnitude as presented in Eq.(5) (Seider et al., 2016).

$$EC_j = EC_j^{Ref,year} \cdot \left(\frac{m_j^F}{m_j^{F,Ref}} \right)^E \cdot \left(\frac{CE^{2020}}{CE^{year}} \right) \cdot y_j \quad (5)$$

where $EC_j^{Ref,year}$ is the equipment reference cost (USD), $m_j^{F,Ref}$ is the reference capacity (kg/h), CE^{year} and CE^{2020} are the Chemical Engineering Index of the reference year and 2020, respectively.

The total capital investment (USD), TCI , is calculated based on the overall factor method of Lang with the Lang factor for fluid processing plant being 5.93 and the delivery cost of equipment being 5% of total equipment cost (Seider et al., 2016). The total annualized capital investment ($TACI$) is calculated from the TCI with total project lifetime being 20 years and interest rate being 0.1.

The total annual operating cost (USD), $TAOP$, includes feedstock and reactant costs, operating and maintenance cost and total energy cost. The costs of feedstock and reactant are calculated by multiplying mass flow rate with cost per kg and operating hours per year. The operating and maintenance cost can be considered as 2% of the total annualized capital investment (Galanopoulos et al., 2019). The total energy cost is sum of energy costs of each equipment which is calculated from energy requirement, E_j^u , specific costs of electricity, heating and cooling and operating hours.

Maximizing the total annualized profit (USD), TAP , is the objective function of the superstructure optimization as presented in Eq.(6).

$$\max TAP = BDS + GLS + MES - TACI - TAOP \quad (6)$$

where the annual biodiesel sales (USD), BDS , is defined from the biodiesel price and the flow rate of product stream out of units: 45, 46, and 49-53. The annual glycerol sales (USD), GLS is defined from the glycerol grade and the mass flow rate of the glycerol

treatment. To simplify the mathematical model, the pure methanol coming out of separation unit is considered a by-product which reduces the requirement of fresh methanol. The annual methanol sales (*USD*), *MES*, is defined from the methanol stream separating from the biodiesel stream in reactive distillations and vacuum distillations. All the sales are calculated with the product market prices and the total operating hours of the refinery in a year.

The superstructure optimization problem is solved with the AIMMS Outer Approximation Algorithm (AOA) (Duran and Grossmann, 1986, and Viswanathan and Grossmann, 1990). The studies presented an algorithm to solve mixed-integer nonlinear programming (MINLP) problems which may be non-convex. The AOA is using CPLEX 20.1 as MIP solver and CONOPT 4.1 as NLP solver.

3. Results and discussion

The superstructure model is applied for two case studies: For the first case, the feedstock is only waste cooking oil to verify the superstructure model, and for the second case, the feedstocks are from different generations of biodiesel to select the most cost-effective feedstock and processing route. The biorefinery has feedstock flow rate of 1000 kg/h, 8000 operating hours per year and biodiesel output complied with EU standard, EN 14214. All the prices and costs of feedstocks, reactants and equipment are calculated in the year 2020.

In the first case study, the biodiesel is produced from the transesterification of the waste cooking oil (unit 7) by using reactive distillation with heterogeneous acid-based catalyst (unit 20) (Boon-anuwat et al., 2015). The products of the reactive distillation process are recycling methanol and biodiesel-rich stream which is going through a decanter (unit 31) to separate the glycerol. After separating the glycerol, the biodiesel product is purified to achieve purity standards by using a vacuum distillation column (unit 49). The separated glycerol is going through a vacuum distillation (unit 60) to become pure glycerol. The total annual profit of the biodiesel production in this case is 828,697 USD per year.

When comparing with the conventional biodiesel production from waste cooking oil (Zhang et al., 2003b), the superstructure optimal process has higher total annual profit. There are two reasons: First, the application of the heterogeneous acid reactive distillation process which can be used for feedstocks with high content of free fatty acids such as waste cooking oil and tallow and reduce the costs of addition separation steps for methanol and catalysts. Second, the glycerol is purified and sold as pure glycerol with higher price than in the conventional process.

In the second case study with different feedstocks, the optimal result is the biodiesel produced from tallow (unit 8) with the same processing route as the waste cooking oil. The annual profit this process is 976,450 USD per year which is higher than the first case because the price of tallow is lower than waste cooking oil.

4. Conclusion

A superstructure optimization model for biodiesel production which encompasses different generation feedstocks, conventional and process intensification technologies has been developed. The results show that the combination of waste cooking oil and tallow with reactive distillation and heterogeneous acid catalyst can open a promising future for biodiesel. The superstructure optimization is proven as a powerful tool of process systems engineering for biorefinery design by systematically and simultaneously solving multi-constraint problems. For further research, the superstructure model of biodiesel

production will be extended by a heat integration function and the ability of processing multiple feedstocks at the same time for more flexible biorefinery operation.

References

- A. AlNouss, G. McKay, T. Al-Ansari, 2019, Superstructure Optimization for the Production of Fuels, Fertilizers and Power using Biomass Gasification, *Computer Aided Chemical Engineering*, 46, 301-306, doi: 10.1016/B978-0-12-818634-3.50051-5.
- J.C.J. Bart, N. Palmeri, S. Cavallaro, 2010, chap 13 - Valorisation of the glycerol by-product from biodiesel production, *Biodiesel Science and Technology*, Biodiesel Science and Technology, Woodhead Publishing, Cambridge, UK, 571-624, doi: 0.1533/9781845697761.571.
- N. Boon-anuwat, W. Kiatkittipong, F. Aiouache, S. Assabumrungrat, 2015, Process design of continuous biodiesel production by reactive distillation: Comparison between homogeneous and heterogeneous catalysts, *Chemical Engineering and Processing: Process Intensification*, 92, 33-44, doi: 10.1016/j.cep.2015.03.025.
- M.A. Duran, I.E. Grossmann, 1986, An outer-approximation algorithm for a class of mixed-integer nonlinear programs, *Mathematical Programming*, 36, 307-339, doi: 10.1007/BF02592064.
- M.A. Fazal, S. Rubaiee, A. Al-Zahrani, 2019, Overview of the interactions between automotive materials and biodiesel obtained from different feedstocks, *Fuel Processing Technology*, 196, doi: 10.1016/j.fuproc.2019.106178.
- C. Galanopoulos, P. Kenkel, E. Zondervan, 2019, Superstructure optimization of an integrated algae biorefinery, *Computers & Chemical Engineering*, 130, doi: 10.1016/j.compchemeng.2019.106530.
- F.I. Gomez-Castro, J.G. Segovia-Hernandez, S. Hernandez-Castro, V. Rico-Ramirez, Z. Gamiño-Arroyo, I. Cano-Rodríguez, 2013, Alternatives for the Production of Biodiesel by Supercritical Technologies: A Comparative Study, *Computer Aided Chemical Engineering*, 32, 7-12, doi: 10.1016/B978-0-444-63234-0.50002-6.
- P. Kenkel, T. Wassermann, C. Rose, E. Zondervan, 2021, A generic superstructure modeling and optimization framework on the example of bi-criteria Power-to-Methanol process design, *Computers & Chemical Engineering*, 150, doi: 10.1016/j.compchemeng.2021.107327.
- A.A. Kiss, A.C. Dimian, G. Rothenberg, 2007, Biodiesel production by integrated reactive-separation design, *Computer Aided Chemical Engineering*, 24, 1283-1288, doi: 10.1016/S1570-7946(07)80238-0.
- Observ'ER, 2020, Biofuels barometer 2020, EurObserv'ER project, accessed date: 10/7/2021, URL: <https://www.eurobserv-er.org/biofuels-barometer-2020/>.
- H. Ritchie, M. Roser, 2020, Emissions by sector, CO₂ and Greenhouse Gas Emissions, Published online at OurWorldInData.org, accessed date: 13/10/2021, URL: <https://ourworldindata.org/emissions-by-sector>
- W.D. Seider, D.R. Lewin, J.D. Seader, S. Widagdo, R. Gani, K.M. Ng, 2016, *Product and Process Design Principles: Synthesis, Analysis and Evaluation - 4th Edition*, John Wiley & Sons Inc., New York, USA.
- A.K. Tula, D.K. Babi, J. Bottlaender, M.R. Eden, R. Gani, 2017, A computer-aided software-tool for sustainable process synthesis-intensification, *Computers & Chemical Engineering*, 105, 74-95, doi: 10.1016/j.compchemeng.2017.01.001.
- J. Viswanathan, I.E. Grossmann, 1990, A combined penalty function and outer-approximation method for MINLP optimization, *Computers & Chemical Engineering*, 14, 7, doi: 10.1016/0098-1354(90)87085-4.
- Y. Zhang, M.A. Dubé, D.D. McLean, M. Kates, 2003a, Biodiesel production from waste cooking oil: 1. Process design and technological assessment, *Bioresource Technology*, 89, 1, 1-16, doi: 10.1016/S0960-8524(03)00040-3.
- Y. Zhang, M.A. Dubé, D.D. McLean, M. Kates, 2003b, Biodiesel production from waste cooking oil: 2. Economic assessment and sensitivity analysis, *Bioresource Technology*, 90, 3, 229-240, doi: 10.1016/S0960-8524(03)00150-0.

Model-based solvent selection for integrated synthesis, crystallisation and isolation processes

Mohamad H. Muhieddine^a, Suela Jonuzaj^a, Shekhar K. Viswanath^b, Alan Armstrong^c, Amparo Galindo^a and Claire S. Adjiman^{a,*}

^a*Department of Chemical Engineering, The Sargent Centre for Process Systems Engineering, and Institute for Molecular Science and Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

^b*Lilly Research Laboratories, Indianapolis, IN 46082, USA*

^c*Department of Chemistry and Institute for Molecular Science and Engineering, Imperial College London, Molecular Sciences Research Hub, White City Campus, Wood Lane, London W12 0BZ, United Kingdom*

c.adjiman@imperial.ac.uk

Abstract

We present a systematic process-wide solvent selection tool based on computer-aided mixture/blend design (CAM^bD) (Gani, 2004) for the integrated synthesis, crystallisation and isolation of pharmaceutical compounds. The method proposed simultaneously identifies the solvent and/or anti-solvent mixture, mixture composition and process temperatures that optimise one or more key performance indicators. Additionally, the method entails comprehensive design specifications for the integrated process, such as the miscibility of the synthesis, crystallisation and wash solvents. The design approach is illustrated by identifying solvent mixtures for the synthesis, crystallisation and isolation of mefenamic acid. Furthermore, a multi-objective CAM^bD problem is formulated and shows that a mefenamic acid with purity of 98.8% can be achieved without significant loss of process performance in terms of the solvent E-factor.

Keywords: Synthesis, crystallisation, isolation, CAM^bD, mefenamic acid.

1. Introduction

Solvents are widely used in the chemical industry to enable a variety of processing tasks. In particular, in the manufacturing of active pharmaceutical ingredients (APIs), a large number of synthetic steps are involved, with several inter-stage isolation and purification units, which require large consumption of solvents to facilitate these process operations (Ott et al., 2014). Recently, the principles of green chemistry have been incorporated in molecular synthesis and process design in order to promote sustainable practices in the chemical industry. In particular, solvent selection has been identified as a key green chemistry research avenue given the strong influence of solvent choice on process performance and product quality (Jiménez-González et al., 2011). One barrier to choosing better solvents is that they have traditionally been selected based on heuristics or time-consuming and expensive experiments. This prevents the practitioner from exploring all possible design choices that optimise the manufacturing process, and highlights the need for more systematic solvent selection tools (Mitrofanov et al., 2012; Brown et al., 2018).

Computer-aided mixture/blend design (CAM^bD) is a promising approach for identifying solvent mixtures that best achieve a specified performance objective, given a set of atom groups

or molecules and a predefined set of target properties (Gani, 2004). CAM^bD has been successful in identifying optimal solvent mixtures for a wide range of applications, including API manufacturing. Jonuzaj et al. (2016; 2018) developed a CAM^bD formulation to select optimal solvent mixtures for separation processes, including crystallisation, in which the number, identity and composition of mixture ingredients were simultaneously identified. Watson et al. (2019; 2021) formulated a solvent mixture design problem to identify the optimal solvents, mixture composition and crystallisation temperatures for the hybrid cooling and anti-solvent crystallisation of APIs. Jonuzaj et al. (2020) proposed a comprehensive mixture design formulation for the integrated crystallisation and isolation of APIs that includes mixture property constraints for both purification stages. However, despite the ongoing efforts in developing integrated solvent selection approaches, a tool that can identify optimal solvents for end-to-end processes is still lacking.

In this work, we present a novel integrated CAM^bD formulation for the design of synthesis, crystallisation and isolation processes. This formulation allows the simultaneous identification of solvents or solvent mixtures, mixture composition and process temperatures that optimise one or more key performance indicators (KPIs). While multiple phenomena need to be taken into account to quantify the effect of solvents on key performance indicators (Folić et al., 2008), we focus on thermodynamic driving forces, i.e., species solubility, in order to obtain a shortlist of solvent candidates that can then be tested through other considerations, such as kinetics, transport properties and crystal shape. The design approach is applied to select solvents for mefenamic acid production, with the aim of maximising process yield. Additionally, a multi-objective problem is formulated to explore trade-offs between the solvent E-factor (SEF) and product purity.

2. Solvent selection approach for integrated synthesis, crystallisation and isolation processes

2.1. Problem Definition

The proposed CAM^bD problem can be used to identify optimal solvent mixtures and process conditions for the integrated synthesis, crystallisation and isolation (filtration, washing and drying) of pharmaceutical compounds. A generic end-to-end process with labelled streams is depicted in Figure 1. This process configuration can represent both batch and continuous processes; for batch processes, each stream would represent the starting or end point of the batch operation. API crystallisation is considered through either pure cooling or hybrid cooling and anti-solvent crystallisation. In order to develop the CAM^bD formulation, a number of sets are defined: set C comprises all components that may appear in any of the liquid streams, set C^K comprises all components that may crystallise throughout the process, set Q comprises all the solvents in the process, set S comprises the solvent candidates from which the process solvents are selected, and set T contains the process stream numbers shown in Figure 1.

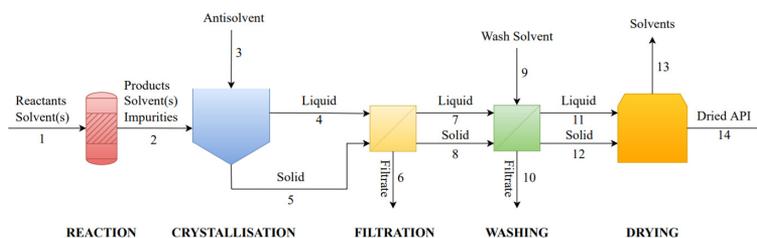


Figure 1: The conceptual flowsheet for the solvent selection problem. For crystallisation, filtration and washing, the labels Liquid and Solid refer to the liquid phase (solvents/dissolved solids) and the solid phase (crystallised solutes) of the corresponding slurry or wet filter cake, respectively.

2.2. MINLP formulation of the integrated CAM^bD problem

The proposed CAM^bD formulation is a mixed-integer nonlinear programming (MINLP) problem derived from a Generalised Disjunctive Programming (GDP) problem using the big-M approach (Jonuzaj et al., 2016). A multi-objective optimisation version is considered so that the overall mathematical structure of the problem can be represented as follows:

$$\begin{aligned}
 \min_{\mathbf{x}, \mathbf{y}} \quad & f_1(\mathbf{x}), \dots, f_P(\mathbf{x}) \\
 \text{s.t.} \quad & \mathbf{h}_0(\mathbf{x}) = \mathbf{0} \\
 & \mathbf{g}_0(\mathbf{x}) \leq \mathbf{0} \\
 & \mathbf{g}_{i,s}(\mathbf{x}) \leq \mathbf{M}_{i,s}(1 - y_{i,s}\mathbf{e}), i \in \{s1, s2\}; s \in \mathcal{S} \\
 & \mathbf{A}\mathbf{y} \leq \mathbf{a} \\
 & \mathbf{x} \in [\mathbf{x}^L, \mathbf{x}^U] \subset \mathbb{R}^n \\
 & \mathbf{y} \in \{0, 1\}^u
 \end{aligned} \tag{1}$$

where each objective function $f_p(\mathbf{x})$, $p = 1, \dots, P$ represents a KPI that needs to be optimised, \mathbf{x} is an n -dimensional vector of continuous variables such as mixture mole fractions, and \mathbf{y} is a u -dimensional vector of binary variables denoting solvent identities. The vector of constraints $\mathbf{h}_0(\mathbf{x})$ describes the models of the reaction, crystallisation and isolation units, the vector of constraints $\mathbf{g}_0(\mathbf{x})$ describes design specifications, the vector of constraints $\mathbf{g}_{i,s}(\mathbf{x})$ describes design and operating specifications that depend on the discrete choices, i.e., solvent identities, $\mathbf{M}_{i,s}$ is a matrix of positive or negative big-M values, \mathbf{e} is a unit vector, and \mathbf{A} and \mathbf{a} are a constant matrix and vector, respectively, used to express the logical relations between the binary variables.

3. Case study: solvent selection for the integrated synthesis, crystallisation and isolation of mefenamic acid

3.1. Problem Description

The synthesis of mefenamic acid (MA), a non-steroidal anti-inflammatory drug with analgesic properties, from 2,3-dimethylaniline (DMA) and 2-chlorobenzoic acid (CBA) is used as a case study. The reaction scheme is shown in Figure 2.

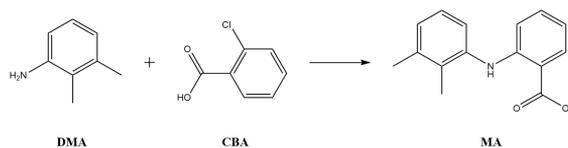


Figure 2: Reaction of 2,3-dimethylaniline (DMA) with 2-chlorobenzoic acid (CBA) to form mefenamic acid (MA)

The CAM^bD formulation identifies solvent mixtures of at most 2 components that optimise selected KPIs, while focusing on the thermodynamic aspects of the process. Two CAM^bD problems are formulated: in the first formulation the process yield is maximised while in the second formulation the SEF is minimised and product purity is maximised. Both formulations include a set of process model equations describing the material balances across the process units, as well as equilibrium constraints describing species solubility, i.e., under-saturation of CBA and DMA during synthesis, crystallisation and washing, under-saturation of MA during synthesis and that MA is at solid-liquid equilibrium during crystallisation and washing. The formulations also include a set of design specifications such as temperature constraints to ensure solvents are in the liquid phase

during operation, miscibility constraints to ensure solvents do not undergo liquid-liquid separation and constraints on the target values/ranges of the selected KPIs. In this case study, n-heptane is used as the wash solvent due to the limited solubility of MA in this solvent (Jonuzaj et al., 2020). Solvents are selected from a list of 43 compounds commonly used in the pharmaceutical industry, constituting a design space of 43 pure solvents and 903 binary solvent mixtures. The main model equations and design specifications are given in Table 1. All symbols are defined in Table 2. The liquid-phase activity coefficients are computed using UNIFAC (Fredenslund et al., 1975).

Table 1: Main CAM^bD model equations and design specifications of the MA case study

Constraint	Equation	KPI Range (where applicable)	Units (where required)
Solubility	$\ln x_{i,t} + \ln \gamma_{i,t} = \frac{\Delta H_{m,i}}{R_g} \left[\frac{1}{T_{m,i}} - \frac{1}{T_t} \right]$	–	–
Miscibility	$\frac{\partial \ln \gamma_{i,t}^{i,j}}{\partial x_{i,t}^{i,j}} + \frac{1}{x_{i,t}^{i,j}} \geq 0$	–	–
Solvent E-factor	$SEF = \frac{\sum_{i \in Q} M_{i,A} + M_{heptane,9}}{M_{MA,14}}$	3.5-20	g Solvents/g API
Crystal Yield	$Y_C = \frac{N_{MA,5}}{N_{MA,2}} \times 100\%$	90-100	%
Process Yield	$Y_P = \frac{N_{MA,14}}{N_{CBA,1}} \times 100\%$	75-95	%
Product Purity	$PP = \frac{M_{MA,14}}{\sum_{i \in CS} M_{i,14}} \times 100\%$	95-100	%
Isothermal reactor operation	$T_1 = T_2$	–	K
2:1 DMA:CBA Feed Ratio	$N_{DMA,1} = 2N_{CBA,1}$	–	mol
Reaction conversion	$x^c = 90$	–	%
Throughput	$M_{MA,5} = 80$	–	kg

Table 2: Nomenclature for the integrated CAM^bD formulation

Symbol	Description	Units (where required)
$x_{i,t}$	Mole fraction of species i in stream t	–
$\gamma_{i,t}$	Liquid-phase activity coefficient of species i in stream t	–
$x_{i,t}^{i,j}$	Mole fraction of species i in a binary mixture of i and j	–
$\gamma_{i,t}^{i,j}$	Liquid-phase activity coefficient of species i in a binary mixture of i and j	–
$\Delta H_{m,i}$	Molar enthalpy of fusion of species i	J/mol
R_g	Universal gas constant	J/mol.K
$T_{m,i}$	Melting point of species i	K
$M_{i,t}$	Mass of species i in stream t	kg
$N_{i,t}$	Moles of species i in stream t	mol
SEF	Solvent E-factor	g Solvents/g API
Y_C	Crystal Yield	%
Y_P	Process Yield	%
PP	Product Purity	%
T_t	Temperature of stream t	K
x^c	Reaction conversion	%

3.2. Results and discussion

The optimisation problems are solved using GAMS version 32.2.0 using the Simple Branch and Bound (SBB) solver (Bussieck and Drud, 2001). Integer cuts are included in the MINLP formulation to generate a ranked list of optimal solutions. For each solution, the following quantities are reported: the identities of solvent components $s1$ and $s2$, the solvent E-factor, the process yield, the product purity, the reaction temperature T_2 , the final crystallisation temperature T_4 , the composition of the solvent in the reactor, defined as the ratio of the mole fraction of solvent 1 to that of solvent 2 in the reactor outlet (stream 2), the composition of the solvent in the crystalliser, defined

as the ratio of the mole fraction of solvent 1 to that of solvent 2 in the crystalliser outlet (stream 4), and the total solvent mass used in the process, i.e., $\sum_{i \in Q} M_{i,4} + M_{heptane,9}$.

Maximising the process yield: The first CAM^bD formulation involves maximising the process yield. Three of the top ten solutions, ranked with respect to the process yield, are listed in Table 3. The first two solutions correspond to cooling crystallisation as shown by the decrease in temperature from T_2 to T_4 and the constant ratio of solvent mole fractions between synthesis and crystallisation. On the other hand, the third solution corresponds to a hybrid cooling and anti-solvent crystallisation process as shown by both the drop in temperature and decrease in the ratio of solvent mole fractions. It can be seen that in all three solutions, the SEF is greater than the lower bound of 3.5 g/g and the difference between T_2 and T_4 is large: T_2 is close to its upper bound of 403 K in the first two solutions and T_4 is at its lower bound in all solutions. This can be explained by the demanding process requirements that favour process yield maximisation. An interesting finding is that the product purity constraint is not active at these solutions, meaning that higher product purity can be achieved in other solutions, albeit at the expense of higher SEF values.

Table 3: Three top solutions of the CAM^bD problem with the objective of maximising Y_P

Solvents	Y_P (%)	SEF (g/g)	PP (%)	T_2 (K)	T_4 (K)	$\frac{x_{s1,2}}{x_{s2,2}}$	$\frac{x_{s1,4}}{x_{s2,4}}$	Total solvent use (kg)
s1: Nitrobenzene s2: n-Butyl acetate	88.10	5.58	98.56	389	290	2.08	2.08	446
s1: Anisole s2: n-Butyl acetate	87.85	7.07	98.95	389	290	0.11	0.11	566
s1: Acetic acid s2: 2,2,4-Trimethylpentane	83.94	9.24	99.30	362	290	6.50	4.13	739

Minimising the SEF and maximising product purity: To explore in more depth the trade-offs between the KPIs, the second CAM^bD formulation is a multi-objective optimisation (MOO) problem that involves minimising the SEF and maximising product purity. The problem is solved using the ϵ -constraint method (Haimes, 1971), in which the SEF is optimised while product purity is constrained by a given lower bound ϵ . The Pareto frontier representing the Pareto-optimal solutions of this MOO problem is shown in Figure 3. It can be seen that a marginal increase in product purity beyond 98.8% requires a significant increase in the SEF. Since the Pareto curve before PP = 98.8% is relatively flat, indicating a small increase in solvent consumption with increasing purity, the solution corresponding to (SEF,PP) = (5.45,98.8) would be a good compromise solution.

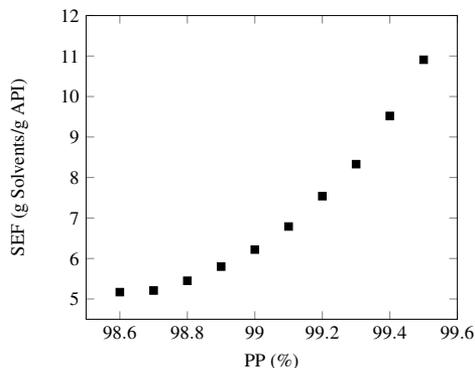


Figure 3: Pareto frontier of the MOO problem of minimising the SEF and maximising PP

4. Conclusion

In this work, a systematic solvent selection approach for the integrated synthesis, crystallisation and isolation of APIs was presented. The proposed CAM^bD formulation identifies the solvent mixtures, mixture composition and process temperatures that optimise selected KPIs, while meeting comprehensive design specifications. The approach was illustrated by identifying solvents for the synthesis, crystallisation and isolation of mefenamic acid. Furthermore, a MOO CAM^bD design problem was formulated and generated different designs with varying SEF/PP values, showing that a PP of 98.8% can be achieved without significantly deteriorating the SEF. The use of CAM^bD to guide experimental solvent screening and explore the synergies between KPIs can help accelerate pharmaceutical process development.

Acknowledgements: Funding from Eli Lilly and Company and the UK EPSRC, through the PharmaSEL-Prosperity Programme (EP/T005556/1) and from the Future Continuous Manufacturing and Advanced Crystallisation Research Hub (EP/P006965/1), is gratefully acknowledged. AG acknowledges funding from the Royal Academy of Engineering and Eli Lilly and Company for support of a Research Chair (Grant RCSRF18193).

Data statement: Data underlying this article may be accessed on Zenodo at 10.5281/zenodo.5585546 and used under the Creative Commons Attribution licence.

References

- C. J. Brown, T. McGlone, S. Yerdelen, V. Srirambhatla, F. Mabbott, R. Gurung, M. L. Briuglia, B. Ahmed, H. Polyzois, J. McGinty, et al., 2018. Enabling precision manufacturing of active pharmaceutical ingredients: workflow for seeded cooling continuous crystallisations. *Molecular Systems Design & Engineering* 3 (3), 518–549.
- M. Bussieck, A. Drud, 2001. Sbb: Simple branch and bound algorithm for mixed integer nonlinear programming. URL: <https://www.gams.com/latest/docs/solvers/sbb>.
- M. Folić, R. Gani, C. Jiménez-González, D. J. Constable, 2008. Systematic selection of green solvents for organic reacting systems. *Chinese Journal of Chemical Engineering* 16 (3), 376–383.
- A. Fredenslund, R. L. Jones, J. M. Prausnitz, 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AIChE Journal* 21 (6), 1086–1099.
- R. Gani, 2004. Chemical product design: challenges and opportunities. *Computers & Chemical Engineering* 28 (12), 2441–2457.
- Y. Haimes, 1971. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE transactions on systems, man, and cybernetics* 1 (3), 296–297.
- C. Jiménez-González, P. Poehlauer, Q. B. Broxterman, B.-S. Yang, D. Am Ende, J. Baird, C. Bertsch, R. E. Hannah, P. Dell’Orco, H. Noorman, et al., 2011. Key green engineering research areas for sustainable manufacturing: a perspective from pharmaceutical and fine chemicals manufacturers. *Organic Process Research & Development* 15 (4), 900–911.
- S. Jonuzaj, P. T. Akula, P.-M. Kleniati, C. S. Adjiman, 2016. The formulation of optimal mixtures with generalized disjunctive programming: A solvent design case study. *AIChE Journal* 62 (5), 1616–1633.
- S. Jonuzaj, A. Gupta, C. S. Adjiman, 2018. The design of optimal mixtures from atom groups using generalized disjunctive programming. *Computers & Chemical Engineering* 116, 401–421.
- S. Jonuzaj, O. L. Watson, S. Ottoboni, C. J. Price, J. Sefcik, A. Galindo, G. Jackson, C. S. Adjiman, 2020. Computer-aided solvent mixture design for the crystallisation and isolation of mefenamic acid. In: *Computer Aided Chemical Engineering*. Vol. 48. Elsevier, pp. 649–654.
- I. Mitrofanov, S. Sansonetti, J. Abildskov, G. Sin, R. Gani, 2012. The solvent selection framework: Solvents for organic synthesis, separation processes and ionic liquids solvents. In: *Computer Aided Chemical Engineering*. Vol. 30. Elsevier, pp. 762–766.
- D. Ott, D. Kralisch, I. Denčić, V. Hessel, Y. Laribi, P. D. Perrichon, C. Berguerand, L. Kiwi-Minsker, P. Loeb, 2014. Life cycle analysis within pharmaceutical process optimization and intensification: case study of active pharmaceutical ingredient production. *ChemSusChem* 7 (12), 3521–3533.
- O. L. Watson, A. Galindo, G. Jackson, C. S. Adjiman, 2019. Computer-aided design of solvent blends for the cooling and anti-solvent crystallisation of ibuprofen. In: *Computer Aided Chemical Engineering*. Vol. 46. Elsevier, pp. 949–954.
- O. L. Watson, S. Jonuzaj, J. McGinty, J. Sefcik, A. Galindo, G. Jackson, C. S. Adjiman, 2021. Computer aided design of solvent blends for hybrid cooling and antisolvent crystallization of active pharmaceutical ingredients. *Organic Process Research & Development* 25 (5), 1123–1142.

Computer-aided solvent design for suppressing HCN generation in amino acid activation

Lingfeng Gui^a, Alan Armstrong^b, Amparo Galindo^a, Fareed Bhasha Sayyed^c, Stanley P. Kolis^d and Claire S. Adjiman^{a,*}

^a*Department of Chemical Engineering, The Sargent Centre for Process Systems Engineering and Institute for Molecular Science and Engineering, Imperial College London, London SW7 2AZ, UK*

^b*Department of Chemistry and Institute for Molecular Science and Engineering, Imperial College London, Molecular Sciences Research Hub, White City Campus, London W12 0BZ, UK*

^c*Synthetic Molecule Design and Development, Eli Lilly Services India Pvt Ltd, Devarabeesanahalli, Bengaluru-560103, India*

^d*Synthetic Molecule Design and Development, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana 46285, United States*
c.adjiman@imperial.ac.uk

Abstract

A highly toxic compound, hydrogen cyanide (HCN), was discovered to result from the reaction between Ethyl cyano(hydroxyimino)acetate (Oxyma) and diisopropylcarbodiimide (DIC), a popular reagent combination for amino acid activation. The reaction solvent has been found to influence the amount of HCN produced so that judicious solvent choice offers a route to suppressing HCN formation. Given the safety implications and the time-demanding nature of experimental solvent selection, we employ a methodology of quantum mechanical computer-aided molecular design (QM-CAMD) to design a new reaction solvent in order to minimize the amount of HCN formed. In this work, we improve on the original QM-CAMD approach with an enhanced surrogate model to predict the reaction rate constant from several solvent properties. A set of solvents is selected for model regression using model-based design of experiments (MBDoE), where the determinant of the information matrix of the design, known as D-criterion, is maximized. The use of a model-based approach is especially beneficial here as it links the large discrete space of solvent molecules to the reduced space of solvent properties. The resulting surrogate model exhibits an improved adjusted coefficient of determination and leads to more accurate predicted rate constants than the model generated without using MBDoE. The proposed DoE-QM-CAMD algorithm reaches convergence in one iteration. In the future, the main reaction of amino acid activation will be considered to design a solvent that maintains the rate of the main reaction while minimizing HCN generation.

Keywords: computer-aided molecular design, design of experiments, solvent effects

1. Introduction

In the pharmaceutical industry, most reactions take place in the liquid phase, with a solvent used to facilitate the reaction. It is desired not only to achieve a high reaction rate but also to achieve high selectivity and avoid the generation of side products that become impurities, but it can be difficult to achieve these objectives simultaneously. The idea of increasing selectivity by rationally designing a solvent that decelerates the side reaction(s) offers a potentially powerful approach to

address this challenge. However, the process of solvent selection often requires time and resource consuming experiments. Aiming to accelerate the solvent screening process as well as to save resources, quantum-mechanical computer-aided molecular design (QM-CAMD) has emerged as a promising computational tool in recent years (Struebing et al., 2013; Zhou et al., 2015; Grant et al., 2018; Gertig et al., 2019). It relies on the use of a surrogate model for quantum-mechanical (QM) modelling to predict the rate constant of a reaction as a function of solvent properties. This surrogate model is incorporated into an optimization problem to identify a list of promising solvents for further experimental tests. Only a small number of QM rate constant data are used in the training set and the solvents used to generate these data are typically chosen based on qualitative arguments such as the diversity of functional groups. This often results in a surrogate model with limited accuracy. Consequently, the QM-CAMD algorithm often requires several iterations to converge, which increases the computational expense and may lead to a suboptimal solution due to model uncertainty. In a case study of solvent design for a Menschutkin reaction, the surrogate model predicted a rate constant that is 15 orders of magnitude larger than the QM rate constant at the first iteration of the algorithm; it took five iterations for the surrogate model to become accurate enough for the QM-CAMD algorithm to converge (Struebing et al., 2013). To reduce the computational expense, an efficient method that enables the rapid construction of more representative surrogate models is required.

Model-based design of experiments (MBDoe) is a statistical approach used to maximise the information content from a set of experiments by performing them at optimal conditions according to a quantitative statistical criterion (Franceschini and Macchietto, 2008). In this work, MBDoe is used to design an optimal set of computational experiments in which the reaction rates are obtained from transition state theory (TST) by calculating activation free energies (Ho and Ertem, 2016) on a QM-derived free energy surface for the reaction system in different solvents. The only experimental condition (factor) that needs to be determined is the solvent environment that is represented by a continuum solvation model (Marenich et al., 2009) used during the course of the QM calculations. The rate constants are used to regress a multi-variable linear regression (MLR) model as a surrogate for the QM method. Here the model-based approach links the large discrete space of solvent molecules to the reduced space of solvent properties. We apply this MBDoe-enhanced QM-CAMD algorithm in an industrial case study where amino acid activation is accompanied by a side reaction in which hydrogen cyanide (HCN) is formed, with the aim to abate safety concerns (McFarland et al., 2019) by taking the advantage of the effect of solvent choice on the amount of HCN produced in this reaction (Erny et al., 2020).

2. Methodology

2.1. QM-CAMD

The workflows of the original and the enhanced versions of the QM-CAMD approach are shown in Figure 1, which highlights the differences between the two versions. The enhanced QM-CAMD (DoE-QM-CAMD) algorithm uses MBDoe for the selection of the initial solvents, rather than a set of solvents based on intuition. Furthermore, they differ from each other in that the original QM-CAMD includes a conditional operation that determines whether more iterations are required for the algorithm to converge, whereas the enhanced version terminates after one iteration. In the first step of DoE-QM-CAMD, the design objective and constraints of the solvent design problem are defined; a set of initial solvents for surrogate model regression is selected using MBDoe. In the second step, rate constants k_s^L are calculated in the selected solvents $s = 1, \dots, N_s$ using transition state theory (TST) and a model combining quantum mechanics (here M062X/6-31+g(d)) and a continuum solvation model to obtain $\Delta^\ddagger G_s^{o,L}$, the liquid-phase activation energy for the reaction in a solvent s . $\Delta^\ddagger G_s^{o,L}$ is calculated using the thermodynamic cycle (TC) approach (Ho and Ertem, 2016). The rate constants calculated in the initial solvents are then used to regress the

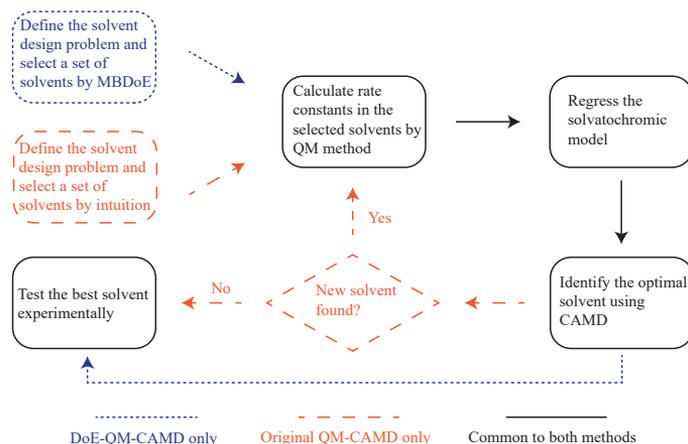


Figure 1: Workflow of QM-CAMD algorithm incorporating MBDoe

solvatochromic model Eq (1) that relates rate constants with the properties of reaction solvents using the following MLR relationship, where the subscript s denotes the solvent:

$$\ln k_s^L = c_0 + c_A A_s + c_B B_s + c_S S_s + c_\delta \delta_s + c_H \delta_{H,s}^2 \quad (1)$$

where $\ln k_s^L$ is the logarithm of the rate constant; A_s is hydrogen bond acidity; B_s is hydrogen bond basicity; S_s is polarity; δ_s is a correction parameter denoting whether the molecule is halogenated ($\delta_s = 0.5$), aromatic ($\delta_s=1$), neither ($\delta_s=0$) or both ($\delta_s=1$); $\delta_{H,s}^2$ is the squared Hildebrand solubility parameter; $c_0, c_A, c_B, c_S, c_\delta, c_H$ are the corresponding reaction-specific coefficients that need to be estimated. In the fourth step, the solvatochromic model is incorporated into a mixed-integer linear programming (MILP) problem, and an optimal reaction solvent is obtained by solving the MILP problem. Because MBDoe is used, the generated solvatochromic model from the MBDoe-selected solvents is expected to be adequately accurate so that the identified optimal solvent can be directly tested in experiments. In the original QM-CAMD algorithm, several iterations through steps 2-4 are often needed: newly found optimal solvents are added to the solvent set for model regression so that the accuracy of the solvatochromic model can be improved.

2.2. Selection of solvents using the D-optimality criterion

MBDoE using the D-optimality criterion is applied to maximize the information content that can be obtained from the QM calculations in the initial set of solvents. For a generic MLR model:

$$Y_i = \beta_0 + \sum_{j=1}^n \beta_j x_{i,j} + \varepsilon_i \quad i \in \{1, 2, 3, \dots, m\} \quad (2)$$

where $x_{i,j}$ are the $j = 1, \dots, n$ explanatory variables corresponding to the i th experiment; β_0 is the intercept; β_j are the coefficients corresponding to each of the explanatory variables; Y_i and ε_i are the measured response variables and random errors at the i th experiment respectively; m is the total number of experiments. In this work, the experiments are computational, so the noise is numerical and depends on the convergence tolerances used. The so-called Fisher information matrix I can be then constructed (Atkinson et al., 2007). The D-optimality criterion consists of maximizing the determinant of the information matrix, $\det I$, which is equivalent to minimizing the volume of the ellipsoid representing the joint confidence region of the model coefficients β_i , thus reducing a measure of the error in the parameters.

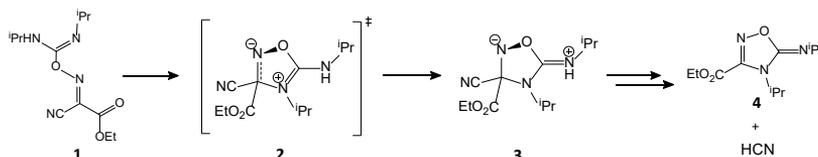


Figure 2: The reaction of Oxyma/DIC adduct **1** into a five-membered ring product **4** and hydrogen cyanide, HCN. **2** is the transition state and **3** an intermediate

Table 1: Solvents generated by MBDDoE and in Struebing et al.'s work ("Div") and the corresponding rate constants for the reaction in Figure 2, as computed by the QM method

MBDDoE			Div		
No	Solvent	$\ln k^L$	No	Solvent	$\ln k^L$
1	2,2,4-Trimethylpentane	-11.7	1	Toluene	-10.4
2	3-Fluorophenol	-3.8	2	Chlorobenzene	-6.4
3	1-Phenyl-1-propanol	-3.9	3	Ethyl acetate	-5.7
4	2-Methoxyethanol	-1.6	4	Tetrahydrofuran	-5.0
5	Adiponitrile	-2.3	5	Acetone	-2.9
6	Nitrobenzene	-2.8	6	Acetonitrile	-2.3
7	N-Methylformamide	0.0	7	Nitromethane	-2.3

3. Results and discussion

3.1. A case study: the HCN generation in amino acid activation using Oxyma and DIC

The DoE-QM-CAMD approach is applied to an industrial case study in which the choice of solvent greatly impacts on the safety of the reaction. It has been observed (McFarland et al., 2019) that during amino acid activation using Ethyl cyano(hydroxyimino)acetate (Oxyma) and diisopropylcarbodiimide (DIC), a popular reagent combination, the generation of HCN occurs, raising significant safety concerns when the reaction is scaled up to a manufacturing level. HCN formation has been found to proceed according to the Scheme in Figure 2, resulting from the decomposition of **1**, an adduct of Oxyma and DIC (McFarland et al., 2019). In this reaction, the imine-type nitrogen first attacks the oxime carbon nucleophilically and forms a zwitterionic intermediate **3** via the transition state **2**, the formation of which is also identified as the rate-determining step (RDS) of the whole reaction sequence. Species **3** further decomposes into **4** and HCN via the cleavage of cyanide and a proton transfer. Given the undesirability of this side reaction, our design objective is to minimize the rate constant of the RDS so that HCN generation can be effectively suppressed.

3.2. Solvatochromic model

A list of seven solvents, shown in Table 1, is generated by MBDDoE using the D-optimality criterion. The rate constants of the RDS are calculated using the QM method in the MBDDoE solvents as well as in the solvents used in Struebing et al.'s work for comparison. This latter set of solvents was chosen on the qualitative basis of molecular diversity, and is referred to as Div. By using the QM rate constants as the training data, a MBDDoE solvatochromic model Eq (3) and a Div solvatochromic model Eq (4) are obtained with adjusted R^2 values of 0.83 and 0.26, respectively.

$$\log k_s^L = -13.99 - 2.96A_s + 6.37B_s + 0.12S_s + 2.07\delta_s + 5.47\delta_{H,s}^2 \quad (3)$$

$$\log k_s^L = -16.59 - 5.03A_s + 7.62B_s + 16.26S_s - 0.55\delta_s - 1.81\delta_{H,s}^2 \quad (4)$$

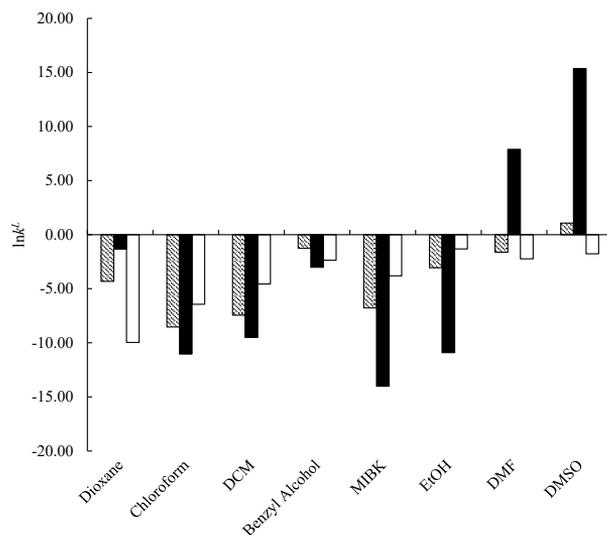


Figure 3: The logarithm of rate constants of the RDS in 8 solvents calculated using the MBDoE model Eq (3), hatched bars, the Div model Eq (4), solid bars and the QM model, empty bars

The larger adjusted R^2 of the MBDoE solvatochromic model demonstrates its better correlation between the fitted model and the input rate constants. Further, a validation set of 8 solvents is used to test the accuracy of the MBDoE and Div solvatochromic models against the QM model. As seen in Figure 3, for all the tested solvents except benzyl alcohol, the MBDoE model predicts closer results to the QM model with a mean average deviation (MAD) of 2.48 log units while the Div model yields a MAD of 8.23. In the case of benzyl alcohol, the predicted $\ln k^L$ by both solvatochromic models (MBDoE: $\ln k^L = -1.25$, Div: $\ln k^L = -3.01$) are in relatively close proximity to the QM model ($\ln k^L = -2.36$).

3.3. CAMD

The MBDoE and Div models are further incorporated into two CAMD problems denoted as MBDoE CAMD problem and Div CAMD problem, respectively. The CAMD problem formulation of Grant et al. (2018), in which additional property constraints are imposed on the boiling points, melting points, toxicities, octanol-water partition coefficients and flash points of designed solvents assembled from 45 atomic groups. The resulting MILP problems are solved in the GAMS software (<https://www.gams.com/>). The results are summarized in Table 2. The Div CAMD model identifies 2,3-dimethylpentane as the optimal solvent. However, the predicted rate constant deviates from that of the QM model by three orders of magnitudes. In contrast, the DoE CAMD model identifies 2,3,4-trimethyl-2-pentene as the optimal solvent and yields a rate constant which is much closer to the value given by the QM model with a deviation of only one order of magnitude. The observed contrast between the two CAMD models is consistent with the model validation results in Section 3.2. In addition, according to the QM model, 2,3,4-trimethyl-2-pentene yields a smaller rate constant than 2,3-dimethylpentane for the RDS of the side reaction that needs to be suppressed. These findings reinforce the hypothesis that the DoE model can provide much more reliable design results.

Table 2: Design results of the MBDoE CAMD problem and Div CAMD problem

	DoE CAMD	Div CAMD
Optimal solvent name	2,3,4-trimethyl-2-pentene	2,3-dimethylpentane
Optimal solvent structure	CH ₃ x 5 CH x 1 C=C x 1	CH ₃ x 4 CH ₂ x 1 CH x 2
k^L Solvatochromic model	$3.37 \times 10^{-6} \text{ s}^{-1}$	$1.13 \times 10^{-8} \text{ s}^{-1}$
k^L QM model	$1.34 \times 10^{-5} \text{ s}^{-1}$	$2.26 \times 10^{-5} \text{ s}^{-1}$

4. Conclusion

In conclusion, an enhanced QM-CAMD framework has been proposed by incorporating MB-DoE at the stage of selecting initial solvents for surrogate model regression. The MBDoE solvatochromic model exhibits better correlation and accuracy when compared to the Div model when using a validation solvent set. The MBDoE CAMD model also provides a better optimal solvent and a more reliable prediction of the QM rate constant for the optimal solvent identified. In the future, the case study will be further extended to the amino acid activation and amidation reaction so that an optimal solvent can be designed and used in peptide synthesis. The proposed approach has the potential to help choose solvents for many organic reactions, and to help improve reaction selectivity with minimal experimental effort.

5. Acknowledgements

Funding from Eli Lilly and Company and the UK EPSRC, through the PharmaSEL-Prosperity Programme (EP/T005556/1), is gratefully acknowledged.

References

- A. C. Atkinson, A. N. Donev, R. D. Tobias, 2007. Optimum experimental designs, with SAS. Oxford University Press, Oxford.
- M. Erny, M. Lundqvist, J. H. Rasmussen, O. Ludemann-Hombourger, F. Bihel, J. Pawlas, 2020. Minimizing HCN in DIC/Oxyma-Mediated Amide Bond-Forming Reactions. *Organic Process Research & Development* 24 (7), 1341–1349.
- G. Franceschini, S. Macchietto, 2008. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science* 63 (19), 4846–4872.
- C. Gertig, L. Kröger, L. Fleitmann, J. Scheffczyk, A. Bardow, K. Leonhard, 2019. Rx-cosmo-camd: Computer-aided molecular design of reaction solvents based on predictive kinetics from quantum chemistry. *Industrial & Engineering Chemistry Research* 58 (51), 22835–22846.
- E. Grant, Y. Pan, J. Richardson, J. R. Martinelli, A. Armstrong, A. Galindo, C. S. Adjiman, 2018. Multi-objective computer-aided solvent design for selectivity and rate in reactions. In: M. R. Eden, M. G. Ierapetritou, G. P. Towler (Eds.), 13th International Symposium on Process Systems Engineering (PSE 2018). Vol. 44 of *Computer Aided Chemical Engineering*. Elsevier, pp. 2437–2442.
- J. Ho, M. Z. Ertem, 2016. Calculating free energy changes in continuum solvation models. *The Journal of Physical Chemistry B* 120 (7), 1319–1329.
- A. V. Marenich, C. J. Cramer, D. G. Truhlar, 2009. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *The Journal of Physical Chemistry B* 113 (18), 6378–6396.
- A. D. McFarland, J. Y. Buser, M. C. Embry, C. B. Held, S. P. Kolis, 2019. Generation of Hydrogen Cyanide from the Reaction of Oxyma (Ethyl Cyano(hydroxyimino)acetate) and DIC (Diisopropylcarbodiimide). *Organic Process Research & Development* 23 (9), 2099–2105.
- H. Struebing, Z. Ganase, P. G. Karamertzanis, E. Sioungkrou, P. Haycock, P. M. Piccione, A. Armstrong, A. Galindo, C. S. Adjiman, 2013. Computer-aided Molecular Design of Solvents for Accelerated Reaction Kinetics. *Nature Chemistry* 5, 952–957.
- T. Zhou, K. McBride, X. Zhang, Z. Qi, K. Sundmacher, 2015. Integrated solvent and process design exemplified for a diels–alder reaction. *AIChE Journal* 61 (1), 147–158.

Improving sustainable CHOSYN's targets through process intensification

Maricruz Juárez-García^a, Gabriel Contreras-Zarazúa^b, Juan Gabriel Segovia-Hernández^b, Jose Maria Ponce-Ortega^a

^a*Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán 58060, Mexico*

^b*Department of Chemical Engineering, Universidad de Guanajuato, Campus Guanajuato, Guanajuato, 36050, Mexico*
jose.ponce@umich.mx, gsegovia@ugto.mx

Abstract

The Carbon-Hydrogen-Oxygen symbiosis networks (CHOSYN) are a set of hydrocarbon processing plants integrated, which pursue a proper use of mass and energy resources, this concept moves towards more sustainable designs for facing the current problem of environmental and climate change by meeting performance targets as environmental concerns and profitability enhancement. This work addresses CHOSYN synthesis using process intensification, which has been widely used for enhancing the performance of industrial processes in terms of economy, environment, and safety. The purpose is to analyze the impact that can be produced through the implementation of process intensification in CHOSYN's performance targets focusing on the intensification of distillation sequences due to the intensive energy use of this separation process. A case study is proposed, the methodology is presented in three stages: intensification of selected flowsheets of the case study, integration of the plants into the CHOSYN, and analysis of results. Thermally coupled equivalent configurations for selected distillation sequences are determined and optimized through the dynamic method known as the moving section method, once obtained these sequences, the different configurations for the CHOSYN are built through a sequential integration method. The performance analysis of the conventional CHOSYN configuration and one intensified option are addressed by the metric known as Sustainability Weighted Return on Investment Metric (SWROIM), which integrates the conventional ROI and sustainability indicators. In this case, the performance targets used for the term of sustainability in the metric are thermal energy savings and CO₂ emissions. In general, the resulting gap between SWROIM values for the conventional network and the intensified alternative measures the total positive impact by intensification over the cost and sustainability indicators of the CHOSYN.

Keywords: process intensification, process integration, resources management

1. Introduction

Carbon-Hydrogen-Oxygen symbiosis networks (CHOSYN) have been recently introduced as a macro system with multi-plant integration among a set of processing plants that deals mainly with hydrocarbon compounds (Noureldin & El-Halwagi, 2015). The integration is carried out through an interception network, where the exchangeable

streams are heated, cooled, pressurized, mixed, or divided as required to make feasible the exchange of streams, chemical conversions are used to obtain needed compounds which are more usable in the network. The main purpose of the integration in the CHOSYN is meeting a proper use of the different mass and energy resources available in the system, through the recycling of the exchangeable streams some sustainability performance targets of the network are improved, for example the reduction of fresh material flowrates, the amount of waste generation, water footprint and heat requirements, or the increase the production of value-added products and the profitability of the network. Nowadays, there is a special attention in the enhancement of all these targets in any process due to the ongoing problems related to the environment, and international agreements against climate change place special emphasis on the reduction of CO₂ emissions and energy efficiency, which force many countries to take action on this matter. In addition to the benefits of the integration to improve the performance targets of the CHOSYN, which have been widely studied and proved, this work proposes to use process intensification (PI) to further improve these sustainability targets. According to the definition of PI, which implies any chemical engineering development that leads to a substantially smaller, cleaner, safer, and more energy-efficient technology (Costello, 2004), it is logical to think that using intensification methodologies can impact the design of the CHOSYN and help to improve these objectives, which is implemented in this work. This work is focused on the intensification of distillation sequences due to the intensive energy use of this separation process, and since thermodynamic efficiency ranges for conventional distillation is 5-20%, using thermally coupled configuration this work seeks important improvements in energy efficiency.

The sustainability enhancement is measured by the SWROIM (Sustainability Weighted Return on Investment Metric), which involves simultaneously the conventional ROI and sustainability indicators (El-Halwagi, 2016), the value of this parameter is relative to the profit of the project and it is used to rank a set of projects according to the economic and environmental performance.

2. Methodology

This work aims to show the impact produced on the CHOSYN performance by applying PI in the distillation sequences of the different processes included in the network, the proposed methodology is divided into three stages: first the intensification of the flowsheets of the selected processes, then the coupled of the intensified processes to the network and finally the evaluation of the targets and the analysis of the results.

2.1. Case study

The case study proposed in this work has been taken from Juárez-García et al. (2021). It is the solution with the lowest cost for a CHOSYN configuration formed by five existing plants, four new plants, and the interception network (see Figure 1). Table 1 shows a summary of the plants and their process capacities.

Table 1. Summary of the Case study

Existing Plant	Processing capacity (ton/day)	Suitable for intensification
ATR (auto thermal reforming of methane)	13400 (syngas)	
Ethylene (steam ethane cracking)	110 (ethylene)	*
PDH (propane dehydrogenation to propylene)	2200 (propylene)	
MTP (methanol to propylene)	300 (propylene)	*
VAM (monomer of vinyl acetate from HOAc)	150 (HOAc)	

New plants	Processing capacity (ton/day)	
Methanol from syngas	882 (MeOH)	
Methanol from CO ₂	2152 (MeOH)	
Carbonylation of MeOH	202 (HOAc)	
Steam methane reforming	253 (syngas)	

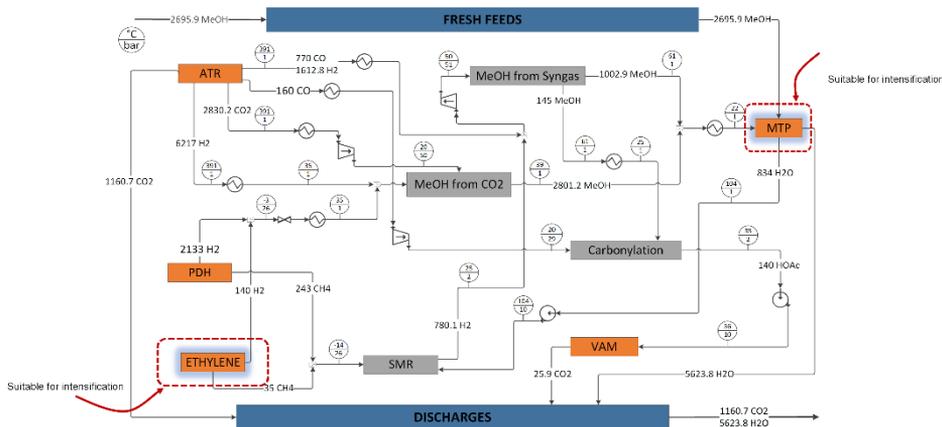


Figure 1. Case study

2.2. Intensification

The case study was simulated using Aspen plus, from the nine processes, two distillation sequences were suitable for intensification, first in the Ethylene production process and second in the methanol to propylene process (MTP). In Figure 2a, the conventional sequence for ethylene purification consists of three distillation columns, in the first column IC-201 the methane and lighter traces are separated by the top stage, the liquid bottom stream goes to the second column IC-202, where butenes and heavier hydrocarbons are separated by the bottom stage, on the top stage the light stream contains ethylene and ethane that are separated in the third column IC-203, the light stream has a composition of 99.9% ethylene. The conventional distillation sequence for purification of propylene is shown in Figure 2b, in the first column CA-401, butenes are eliminated in the bottom stream, and the top stream passes to the second column CA-402, where the propylene reaches a composition of 99.9%.

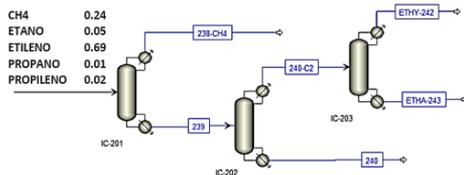


Figure 2a. Conventional sequence distillation of ethylene process

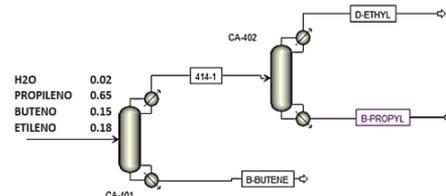


Figure 2b. Conventional sequence distillation of MTP process.

In both cases it was used the method proposed by Hernandez & Jimenez (1996) to obtain the intensified arrangements, according with the efficiency reported in the literature for similar sequences, for the ethylene purification sequence it was selected a arrangement of three columns with two recycles with an energy saving of 20% in regard of the

conventional sequence. For the propylene separation in MTP process the literature indicates that Petluyk column has the best efficiency for similar sequences, despite this, in this case the side stripper arrangement presents a greater energy savings with 20%. The intensified sequences, and the optimization of the inter-connection flows are shown in Figures 3.

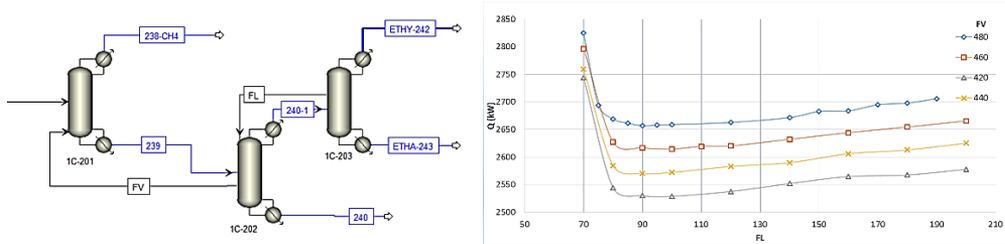


Figure 3. Intensification and optimization of the distillation sequence of ethylene process.

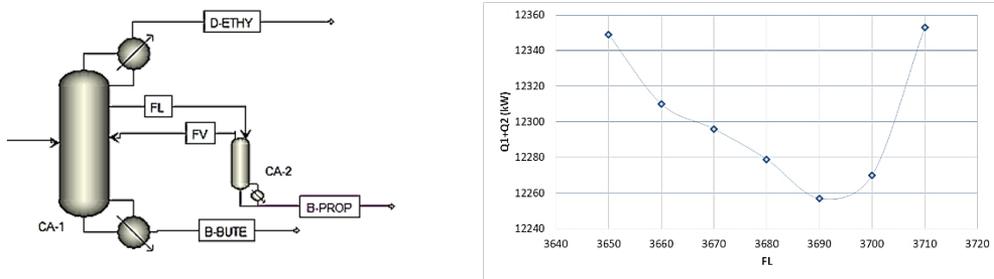


Figure 4. Intensification and optimization of the distillation sequence of MTP process.

2.3. SWROI evaluation

The Sustainability Weighted Return on Investment Metric (SWROIM) allows to measure the economic viability of a project through the conventional ROI (return on investment) and also to include sustainability aspects through sustainability indicators. The conventional ROI can be determined by E.1, which involves the annual economic profit (AEP) and the total capital investment (TCI):

$$ROI = \frac{AEP}{TCI} \quad E.1$$

The SWROI besides includes the term of sustainability as follows:

$$SWROIM = \frac{AEP \left(1 + \sum_{i=1}^{N_{indicators}} w_i \left(\frac{Indicator_i}{Indicator_i^{Target}} \right) \right)}{TCI} \quad E.2$$

Where $Indicator_i$ is the value of the i^{th} sustainability indicator and the $Indicator_i^{Target}$ is the target value for the indicator, it could be the maximum value determined by optimization or the largest value from all project options, w_i is the weighting factor which represents the relative importance for the indicator given by the company and it is a ratio to the economic profit. In particular, this work uses the reduction of CO₂ emissions and the heating savings as sustainability indicators, the targets are set in 20% of heating savings of the conventional configuration since the intensification allows to reach this percentage, and for the reduction of CO₂ emissions 22% of the total emissions of the conventional

configuration taking into account the compromise that Mexico acquired under Paris agreement. The selection of weighting factor is a non-trivial task because it can skew the results, we propose first equal weighting factors for both indicators of 0.1 in scenario A, then give more weight to the emissions indicator with a factor of 0.25 in scenario B, and vice versa in scenario C, this will allow observing how much the decision of this value influences on the final value of the SWROIM. The summary of the sustainability targets and their relative importance is shown in Table 2.

3. Results

Once obtained the intensified options for the distillation sequences a new configuration of the CHOSYN is built using Aspen plus (see Figure 5). For the SWROIM evaluation, it is needed the value of the total capital investment and the annual economic profit. As the case study was described, the cost and the profit are only related to the new plants which represent the new project investment, these values are shown in Table 3.

In this section are presented the resulting values for the SWROIM for the configuration of the CHOSYN when conventional arrangements of distillation columns are used and for the same configuration but using intensified distillation sequences. The value of the conventional ROI is shown in Table 3, the gap between both values is barely appreciated, which means that the investment cost does not greatly differ from one to another option, and the decision-maker could choose indifferently from one to another. When the sustainability targets of reduction of CO₂ emissions and heating savings are included with equal relative importance of 0.1 in scenario A, the SWROIM of the conventional CHOSYN has a value of 67.14 % and this is kept for scenarios B and C because it has the largest values for both indicators, in this same scenario A the SWROIM for the intensified configuration grows to 68.4%. For scenario B where the reduction of emissions has the biggest weight of 0.25 in regards the profit, the SWROIM is 68.81%, with a relative difference of 1.2%. Meanwhile, the SWROIM of scenario C, where the heating savings has 0.25 relative importance, is 69.34% with a relative difference of 2.2%.

Table 2. Targets of the sustainability indicators and their relative importance

Sustainability Indicator _i	Target (Indicator _i ^{Target})	Units	Weighting factor (A)	Weighting factor (B)	Weighting factor (C)
Heating savings	216918	kW	0.1	0.1	0.25
CO ₂ emissions reduction	35021	Ton of CO ₂ /year	0.1	0.25	0.1

Table 3. Results summary for the CHOSYN configurations.

	TCI (10 ⁶ USD)	AEP (10 ⁶ USD/year)	ROI	SWROIM (A)	SWROIM (B)	SWROIM (C)
Conventional configuration	181.219	121.358	67.08%	67.14%	67.14%	67.14%
Intensified configuration	181.055	122.358	67.58%	68.4%	68.81%	69.34%

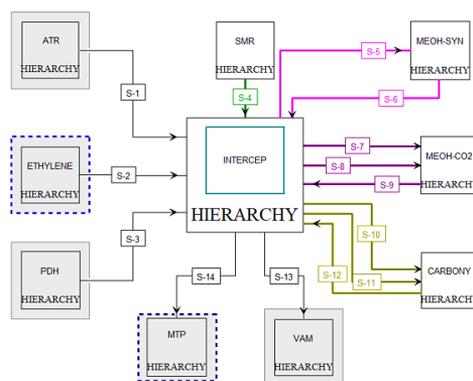


Figure 5. Aspen simulation of the CHOSYN configuration.

Table 4. SWROIM values for the different scenarios and configurations.

Despite the difference between both configurations in the different scenarios seems to be not too important, it is worth mentioning that a difference of 1.2% in the SWROIM only in economic terms represents 2.172 MMUSD and 2.2% 3.98 MMUSD what clearly shows what option is better. Whatever, the relative importance of the sustainability indicators can be changed according to the decision-maker and the company's goals in environmental concerns and provide a range of options suitable to this goals.

4. Conclusions

Process intensification has been implemented for the successful improvement of the sustainability performance targets of the CHOSYN. A case study was proposed to measure this improvement through the SWROIM which involves simultaneously economic conventional metric ROI and a sustainability term, for this work the reduction of the CO₂ emissions and energy savings are the targets to be improved. Notwithstanding in the individual processes, the intensification of the distillation sequences means energy savings up to 20%, this impact is greatly reduced in the macroscopic point of view of the CHSOYN with relative differences of the SWROIM up to 2.2% wich still represents high economic savings mainly related with the operating cost, and an enhancement over the environmental performance targets due the emissions reduction.

References

- M. M. B. Noureldin, M. M. El-Halwagi, 2015, Synthesis of C-H-O symbiosis networks. *AIChE Journal*, 61(4), 1242–1262.
- R. C. Costello, 2004, Process intensification: think small. *Chem. Eng.* 111 (4):27–31.
- M.M. El-Halwagi, 2016, A return on investment metric for incorporating sustainability in process integration and improvement projects, *Clean Technologies and Environmental Policy*, 19, 611–617.
- M. Juárez-García, J. M. Ponce-Ortega, M. M. El-Halwagi, 2021, Systematic Approach for Synthesizing Carbon–Hydrogen–Oxygen Networks Involving Detailed Process Simulations. *Industrial & Engineering Chemistry Research* 2021 60 (45), 16378–16393.
- S. Hernández, A. Jiménez, 1996, Design of optimal thermally-coupled distillation systems using a dynamic model. *Chemical engineering research & design*, 74(3), 357–362.

A novel process design for automated quality analysis in an integrated biopharmaceutical platform

Simon Tallvod, Mikael Yamanee-Nolin, Joaquín Gomis-Fons, Niklas Andersson and Bernt Nilsson

*Dept. of Chemical Engineering, Lund University, Naturvetarvägen 14, Lund 221 00, Sweden
bernt.nilsson@chemeng.lth.se*

Abstract

This contribution describes a novel process design consisting of the implementation of a high performance liquid chromatography (HPLC) system into an existing integrated biopharmaceutical purification platform in order to extend the possibilities of analysis. This setup is controlled by a custom-made set of software tools, namely the combination of Orbit along with a C# based program called Satellite, which is used to link the HPLC system to Orbit. By using this process design, automated analysis of the continuous process flow is made possible, generating full wavelength spectra, and automated pool quality analysis. These spectra contain at least an order of magnitude more data compared to a standard chromatogram, which can be utilized for further analysis, while the pool quality analyses can be used for making decisions regarding how to proceed with a particular pool. The results of the study show the successful implementation of automated analytic tools to the purification process, which also indicates that the process carries promise as a generalizable platform for Orbit-controlled designs for different analysis purposes.

Keywords: novel process design, chromatographic purification, automated analysis, integrated downstream processing

1. Introduction

We are currently seeing a great need for global accessibility and lower prices of a wide range of vital biopharmaceuticals, e.g. monoclonal antibodies for cancer treatment. With *upstream* processing having seen recent productivity improvements, the on-going paradigm shift towards continuous and integrated *downstream* processes is therefore critical in order to make new, life-saving pharmaceuticals more readily available at a global scale (Ötes et al., 2017). Concrete examples of this paradigm shift is the process design for automated lab-scale production of a recombinant protein presented by Gomis-Fons et al. (2019), as well as the process data utilization to improve data-driven modelling methods by Brestrich et al. (2015). However, while continuous processes may offer many advantages, the optimal performance thereof demands high-speed decision-making based on precise measurements and constantly updating data for real-time monitoring and control (Lee, 2017). Therefore, there is a critical need to implement automated analysis capabilities in order to achieve autonomous production of biopharmaceuticals.

The purpose of the current work is to expand on the process presented by Gomis-Fons et al. (2019), and satisfy the need for advanced analysis by presenting a novel process design, where an HPLC with a diode-array detector was integrated with the aforementioned process and its control software. This is achieved by means of physical integration of the flowpaths of the two machines, as well as integrating the HPLC system control software into Orbit (Andersson et al., 2017) via a

custom-made C# based program. The proposed process design provides two tools with different capabilities: (i) one to analyze the individual eluting peaks from the chromatographic purification, and (ii) one to analyze the continuous process flow over all available wavelengths utilizing the diode-array detector. Thereby, the setup allows for automated quality analysis of specific pools or entire system flows, generating large amounts of data to be used for e.g. data-driven modelling and autonomous control. In order to highlight the advantages of the presented process design and the capabilities of the two added analytical tools, two case studies similar to the one presented in Gomis-Fons et al. (2019) were performed and the results thereof are presented and discussed.

2. Process design

This section describes the hardware configuration of the proposed design and the software setup used to control the process. The work was carried out by connecting the flow-paths of an ÄKTA pure 25 and an Agilent 1260 HPLC by utilizing the loop and injection valves, as well as a flow splitter. In order to integrate the HPLC system and the ÄKTA system, ancillary software was developed and used to link the HPLC hardware with the previously developed control software, Orbit.

2.1. Hardware configuration

In the process described by Gomis-Fons et al. (2019), a recombinant protein is purified in four steps using three bind-and-elute chromatographic columns and one flow-through chromatographic membrane. The purification steps consist of a multimodal chromatography column (Column 1 in Figure 1), a cation exchange column (Column 2), a salt-tolerant interaction chromatography anion exchange membrane (Column 3; operated in flow-through mode) and an anion exchange column (Column 4). The purpose of the loop (Column 5) is to collect fully representative samples of the eluting pools; in order to keep the samples homogeneous, the loop also contains a magnetic stir bar. It should be noted that the ultrafiltration-diafiltration step used by Gomis-Fons et al. (2019) was omitted in the current work in order to reduce complexity.

The complete system of the current work consists of two separate liquid chromatography systems, as presented in Figure 1; an ÄKTA pure 25 running an integrated protein purification process (Gomis-Fons et al., 2019) and an Agilent 1260 HPLC for running online analytics using an ion exchange column and a linear salt gradient. The Agilent HPLC (left dashed box) was connected to the ÄKTA (right dashed box) in two ways. Firstly, the loop valve of the ÄKTA ('LV') and the injection valve of the HPLC ('HPLC InjV') are connected in series, making it possible to fill the HPLC injection loop (Blue 'Inj Loop') from the ÄKTA. The HPLC injection valve was implemented using a two-position/six-way valve. Secondly, a splitter valve (Agilent G1968D Active Splitter) was installed in the flow path of the ÄKTA directly after the UV, conductivity and pH sensors. This splitter valve separates the flow in a ratio of approximately 1:1000, and small off-stream of the splitter is sent to the HPLC sensors through its column valve. Furthermore, a setup of two computers connected over LAN, each running an instance of Orbit, is used to control the physical system and automate the process.

This setup allows for automated HPLC analysis of the peaks during the run, which provides information regarding e.g. in which steps of the purification impurities are removed, and to which extent. Such information of the sample quality can then be used for purification process control. Furthermore, the HPLC system was equipped with a diode-array detector. The purpose of adding a diode-array detector to the process is to greatly increase the wavelength range for absorption measurements. Compared to the standard single or triple wavelengths in an ÄKTA system, the diode array detector in the Agilent 1260 system is capable of measuring every wavelength between 190 nm and 950 nm. This makes it possible to differentiate between species, provided they have different absorption spectra. If the spectra of all pure eluting components or known mixtures thereof

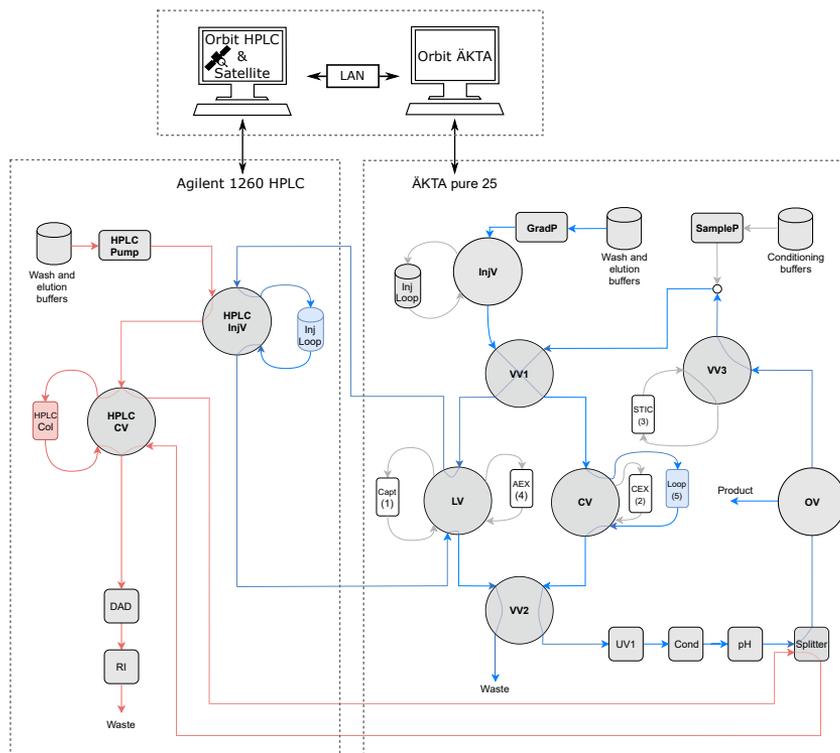


Figure 1: A schematic of the physical system layout and the control software setup.

are known beforehand, it is possible to determine the individual concentrations in real time, subsequently enabling dynamic decision-making regarding pooling (Brestrich et al., 2015), e.g. via partial least square regression (PLS). It should be noted that due to the two analytical tools being connected in series, continuous spectra cannot be taken while a peak analysis is being performed.

Through these two connections, the proposed process design thus provides two important analytical tools: online analytic chromatography on the HPLC, and the use of the HPLC sensors as additional real-time sensors in the process. This hardware configuration thereby allows for automated online analysis of specific peaks as well as continuous multi-wavelength measurements to be taken during a preparative chromatographic run. However, in order to control and/or automate the process, particular software was needed.

2.2. Control software

The control software setup is used to automate peak analysis, in order to find information of the species present in each pool, and diode-array detector measurements of the continuous process flow, which can be used to reveal much more information regarding the contents thereof. The main control software used to automate the process is a generic Python-based Process Analytical Technology (PAT) software interface (Holmqvist and Sellberg, 2016) called Orbit (Andersson et al., 2017), which has been successfully applied in research and development of biopharmaceutical downstream processing (e.g. Gomis-Fons et al. (2019, 2021)). Orbit provides an open-structure framework based on a real-time core, providing opportunities to implement add-ons to be used for process automation and to control a range of instruments. In order to integrate the HPLC-system

into the Orbit framework, a program called Satellite was developed using the Agilent-developed API known as Instrument Control Framework (ICF). Satellite was written in C# and based on a software development kit provided by Agilent. The Satellite program functions as a link between the HPLC hardware and the Orbit software, allowing Orbit to read sensor data and control the functionality of the HPLC.

In order to assert full control of the setup, two separate instances of Orbit were used to run the process. One instance was used to control the ÄKTA system, and the other to control the HPLC. Each Orbit controller instance was defined by a script dictating the specific system behavior. However, the two systems were not operating completely independently. For example, the ÄKTA system needed status information from the HPLC, i.e. whether the HPLC was ready to receive a sample in its sample loop, and the HPLC subsequently needed to know whether to begin an analysis run. This synchronization of the two systems was carried out via a global flag variable, i.e. by writing to and reading from the same variable to indicate whether a particular system is ready to perform an action.

3. Results and discussion

The major results from the two cases are presented as the preparative chromatogram of the first case followed by the analytic chromatograms of the peak quality, and the spectra generated by the HPLC diode-array detector analyzing the continuous process flow during the second case. This spectrum highlights the expected purification, and contains at least an order of magnitude more data than a standard chromatogram, which can be utilized for further analysis, data-driven modeling, and subsequent autonomous control.

In Figure 2, the preparative chromatogram of the first case study is presented, with the three eluting peaks highlighted by shaded areas under the curve. The first peak (close to the 150 minute mark) is from eluting the capture column (Column 1); the second peak is eluted from the second column, flowing through the third column; and the third peak is from the fourth column. It should be noted that the STIC (column 3) is run in flow-through mode, with the flow not passing by the UV detector, which explains why there is no visible peak in the preparative chromatogram corresponding to this column as with the others. However, due to the process design, it is still possible to perform an analytic HPLC run of the sample leaving this column.

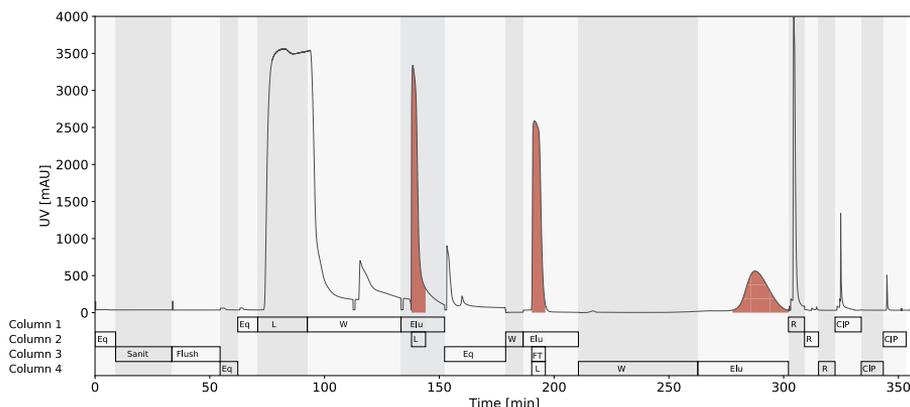


Figure 2: The chromatogram generated by the UV detector of the ÄKTA during the first case, with eluting peaks close to 150 min, 200 min, and 300 min.

The analytic chromatograms of the three peaks are presented alongside the analyses of the feed and the pure product in Figure 3. As can be seen, each step performs a vital function in isolating the product. Firstly, the capture column drastically reduces the impurities present in the feed, and the subsequent columns can be seen to each perform polishing and finishing by further eliminating the majority of the remaining impurities, leaving a single major peak after the AEX column (column 4) at the same eluting time as the pure product. The reason for the reduced peak height compared to the previous is due to the wider final pool during the chromatographic purification, meaning that the concentration in the analysed sample will be lower compared previous, narrower peaks. These results show that the proposed process design is capable of providing the automated analysis capabilities necessary to support continuous manufacturing.

Finally, the continuous spectra from the second case is presented in Figure 4, with the corresponding 280 nm chromatogram highlighted with a bold line. This wavelength is the gold standard for absorbance measurements, but proteins contain different peptide groups highly active in absorbance ranges of 180-230 nm (Carta and Jungbauer, 2010), which can be seen in Figure 4. By measuring at not only a single wavelength, but a range of wavelengths between 210-310 nm, more information can be retrieved regarding which species are present at different times, meaning that the spectra can be used to continuously analyse e.g. mixture compositions. While the nuances are difficult to see by the naked eye, the information existing in the continuous spectra remains for mathematical analysis, e.g. as presented by Brestrich et al. (2014). Utilizing this technology in the proposed process design thus enables automated monitoring of the product quality at a highly detailed level, providing further opportunities to automate decisions in continuous manufacturing.

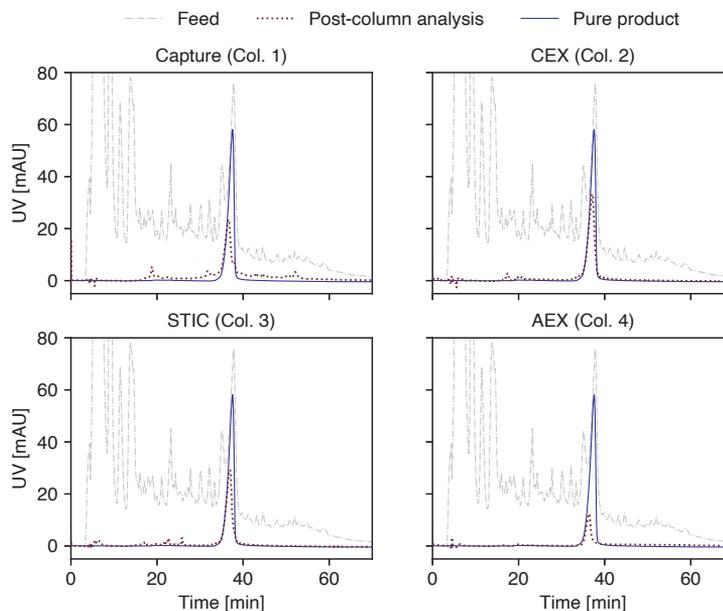


Figure 3: The results from the automated HPLC analysis of the colored peaks in the chromatogram along with the feed and the pure product, showing the performance of each step in the chromatographic purification performed on the ÄKTA pure 25.

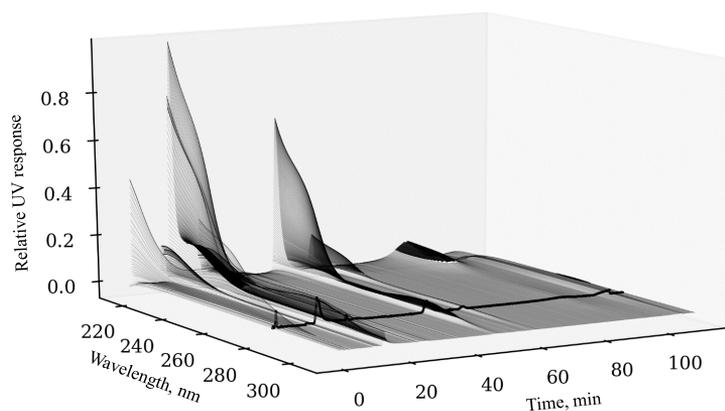


Figure 4: Continuous spectra providing information regarding the composition of the process flow at many more wavelengths than just at 280 nm, marked in bold to highlight the corresponding chromatogram.

4. Conclusions

In conclusion, the study shows that the proposed process design is capable of providing quality analysis information and large amounts of data in an automated fashion, which is a necessary step towards autonomous biopharmaceutical production. Based on the results from the peak analysis, an interesting next step would be to develop a rule-based decision-making algorithm to automate also the decisions on how to proceed with a particular pool. Furthermore, the use of the diode-array detector was shown to have potential in generating data that can also be used for e.g. identifying different species, and data-driven modelling. However, challenges pertaining to having both analysis tools active simultaneously, as well as determining the next step towards autonomous control, remain for the future. Finally, based on the presented results, the concept carries promise as a generalizable platform where customized setups for different purposes can be fully automated and controlled by Orbit.

References

- N. Andersson, A. Lofgren, M. Olofsson, A. Sellberg, B. Nilsson, P. Tiainen, 2017. Design and control of integrated chromatography column sequences. *Biotechnol Progress* 33, 923–930.
- N. Brestrich, T. Briskot, A. Osberghaus, J. Hubbuch, 2014. A tool for selective inline quantification of co-eluting proteins in chromatography using spectral analysis and partial least squares regression. *Biotechnology and bioengineering* 111 (7), 1365–1373.
- N. Brestrich, A. Sanden, A. Kraft, K. McCann, J. Bertolini, J. Hubbuch, 2015. Advances in inline quantification of co-eluting proteins in chromatography: Process-data-based model calibration and application towards real-life separation issues. *Biotechnology and bioengineering* 112 (7), 1406–1416.
- G. Carta, A. Jungbauer, 2010. *Protein Chromatography: Process Development and Scale-Up*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- J. Gomis-Fons, A. Lofgren, N. Andersson, B. Nilsson, L. Berghard, S. Wood, 2019. Integration of a complete downstream process for the automated lab-scale production of a recombinant protein. *Journal of Biotechnology* 301, 45–51.
- J. Gomis-Fons, M. Yamane-Nolin, N. Andersson, B. Nilsson, 2021. Optimal loading flow rate trajectory in monoclonal antibody capture chromatography. *Journal of Chromatography A* 1635, 461760.
- A. Holmqvist, A. Sellberg, 2016. A generic pat software interface for on-line monitoring and control of chromatographic separation systems. In: *Computer Aided Chemical Engineering*. Vol. 38. Elsevier, pp. 811–816.
- S. Lee, 2017. Present and future for continuous manufacturing: FDA perspective. <https://pqri.org/wp-content/uploads/2017/02/1-Lee-PQRI-for-CM-2017.pdf>, date accessed: 2021-08-22.
- O. Ötes, H. Flato, J. Winderl, J. Hubbuch, F. Capito, 2017. Feasibility of using continuous chromatography in downstream processing: Comparison of costs and product quality for a hybrid process vs. a conventional batch process. *Journal of Biotechnology* 259, 213–220.

A multiscale approach towards the use of Urban waste within the energy transition strategy

Sofia González-Núñez, Lidia S. Guerras, Mariano Martín

Department of Chemical Engineering, University of Salamanca, Plz Caidos 1-5, 37008, Salamanca, Spain

Abstract

The use of municipal solid waste and sludge within the waste to power initiative is evaluated following a multiscale approach. At process level, a techno-economic analysis based on systematic process design is developed to determine the optimal operating conditions and to estimate the yield and investment cost of the facilities as a function of the processing capacity. The facility consists of the boiler, the flue gas treatment chain, the steam turbine, and the cooling tower that are modelled using an equation-based approach. Experimental data and first principles are used to model the boiler and the pollutant abatement techniques, while detailed thermodynamics and transport phenomena are used to model the turbine and the cooling tower. Next, at country level, an extended facility location problem is formulated to select the location of facilities for a given budget, aiming at an electricity production cost and water consumption, considering the economies of scale and social indexes to promote the development jobs in regions suffering from depopulation. To produce power at a competitive price, below 0.06 €/kWh, a facility has to process the waste generated at cities above 250 k habitants. For this critical size, the investment cost is beyond 25 M€. At country level, for the major cities in Spain (over 65 across the mainland) it would be possible to produce up to 235 MW from the waste generated with a production cost of 0.05 €/kWh.

Keywords: Waste to energy, pollutants abatement, process design, multiscale

1. Introduction

Waste management is a major issue of our society due to the amount generated annually as well as the challenge its management represents. It is estimated that by 2021, 6.1 Mt/d of municipal solid waste (MSW) will be generated in urban areas (World bank, 2012). In addition to MSW, sludge that reaches a production of 13 Mt in 2020 (EU, 2008), can also be used as a resource. Both show a high potential to contribute to the energy transition. Incineration of MSW presents high potential for electricity generation using a Rankine cycle since the yield of thermal power plants can reach 40%. However, while most of the process studies focus on the thermodynamic yield, a facility that processes MSW requires a flue gas treatment to avoid emitting NO_x, SO₂ but above all heavy metals, dioxins and furans (Hubner et al 2000). Dried sludge can also be used as a resource. The literature is detailed evaluating the power island to compute the energy produced from incineration. However, the flue gas treatment was included in the process in a simplified manner where the models for pollutants removal were based on simple removal rates (Fodor and Klemes, 2012), and no process optimization of such facilities has been addressed. Guerras and Martín developed detailed surrogate models to evaluate the performance of the flue gas treatment technologies for NO_x, SO₂ removal (Guerras and Martín, 2019). This work

focuses on the design of incineration facilities for MSW and sludge evaluating the effect of the scale. A mathematical optimization approach is used for the optimal operation of integrated facilities including waste furnace, gas treatment, the power island, and the cooling system. After a techno-economic analysis, the effect of the scale on the cost is evaluated, to formulate a facility location problem aiming at deciding where to install this kind of facilities at different cities across a country using the available sludge and MSW within range.

2. Methodology

The methodology comprises the analysis at two scale levels. At process level the techno economic evaluation of the transformation of waste, MSW and sludge, into power is analyzed. A scale up evaluation study is required to formulate a facility location problem where the results are required for the exploitation of the waste at country level.

2.1. Process synthesis

The process consists of four sections: the boiler, the flue gas treatment chain, the turbine, and the cooling. The boiler processes the mix of wastes, municipal solid wastes (MSW) with 50% moisture and sludge, 83% moisture, whose composition is found in Lin and Ma (2012), in a blend that can be managed by the boiler, below 10% sludge to avoid issues due to the concentration in water. The model of the boiler provides the energy generated as well as the composition of the flue gas, validated using literature data. The flue gas is processed to remove NO_x, using selective catalytic reduction (R02), next, an electrostatic precipitator (PE01) removes particles followed by Lime Dry Spray process (LSD) for the removal of SO₂. A filter (FM01) followed by an adsorbent bed to remove heavy metals complete the gas treatment chain. The steam generated in the boiler feeds a Rankine cycle with reheating and regeneration. A cooling tower cools down the water used in the condensation of the steam. Figure 1 shows the process flowsheet.

The investment cost is computed using the factorial method based on the equipment cost (Sinnot and Towler, 2019) The estimation of the facility is performed by considering that piping, isolation, instrumentation, and utilities represent 20%, 15%, 20% and 10% of the units. Land and buildings cost is estimated to be 8 M€ (Martín and Martín, 2013). The fees represent 3% of the fixed cost, other administrative expenses and overheads and the plant layout represent 10% of the direct costs (fees plus fixed capital) and 5% of the fixed cost respectively. The plant start-up cost represents 15% of the investment.

2.2. Scale up method

Because of the properties of the factorial method to estimate the costs, the scale up or down as a function of the processing capacity, the cost of the units must be correlated as a function of a characteristic variable of its size or processing capacity (i.e. area, heat load, mass flow rate processed...) If the processing capacity is exceeded, additional units in parallel must be purchased (Sánchez and Martín, 2018)

2.3. Macro-scale analysis

A facility location problem is formulated to select the size and location of the incinerator plants towards the waste treatment and the social impact for a certain budget availability

and targeting a maximum electricity cost. The water consumption of the cooling tower is estimated from the correlations developed in previous work as a function the facility location (Guerras and Martín, 2020).

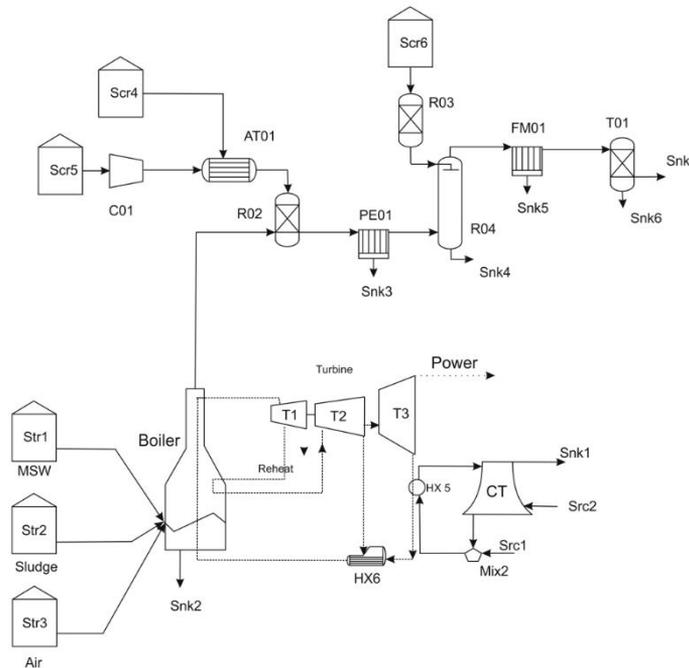


Figure 1.-Process flowsheet

3. Problem definition

3.1. Process modelling and optimization

The process is modeled unit by unit using a combination of mass and energy balances, thermodynamics, rules of thumb and experimental data. The boiler uses experimental data to determine the flue gas composition. To predict the effect of the processing variables on the removal yield of the flue gas treatment processes, such as NO_x catalytic reduction, dioxins and furans removal, the electrostatic precipitator, the SO₂ dry removal (Guerras and Martín, 2021), the filter, and the heavy metal packed tower, surrogate models are developed. In addition, the Rankine cycle is modeled units by unit the three turbines, high, medium and low pressure, where detailed thermodynamics is implemented using surrogate models for the enthalpy and entropy as a function of the pressure and temperature (Martín and Martín, 2013). The cooling tower model is based on evaluating the gas-liquid contact (Guerras and Martín, 2021).

3.2. Macro analysis

The selection of the location of the incinerator plants is a facility location problem (Grossmann, 2021). The aim is to produce power at a competitive price, using the opportunity to create wealth and jobs in regions that are suffering underpopulation issues or/and underdeveloped with respect to the neighbors. The social indexes relative population density, DH, relative unemployment, Ru, and development ratio, DR, of the

location are defined in Heras and Martín (2020) are used to build a composed objective function, eq. (1) where I is the investment.

$$OBJ = \text{Max} \sum_i \left(\text{Jobs}_i \cdot \frac{1}{2} \left(DH_i + \frac{Ru_i}{Ru_{i,\text{max}}} \right) + DR_i \cdot I_i \right) \quad (1)$$

4. Results

This section summarizes the results related to process design and the allocation problem

4.1. Process level

4.1.1. Process yield

The optimization of the process yields the ratios presented in Table 1 in terms of power ratio, turbine operation and cooling requirements. Note that water consumption is location dependent and is not as much as process driven beyond the total cooling load.

Table 1.- Summary of performance values for the operation of the facilities.

Variable	Value	Variable	Value
Boiler		Cooling cycle	
Flue-gas (kg/kg waste)	23.2	Cooling (kJ/kg)	1579
Ashes (kg/kg waste)	0.10		
Power island			
W1/W2/W3 (kJ/kg)	620/178/538	Turbine Extraction	0.16
P1/P2/P3/P4 (bar)	154/11/5.8/0.08	P(kJ/kg)	1335.3

4.1.2. Economics

The economics of the facilities are a function of the scale. Therefore, several cities across the territory are used examples such as Madrid (M), Barcelona (B), Santiago (Sa), Valencia (V), Seville (S), Gijon (Gi), Granada, (G), Pamplona (P), and Salamanca (Sal). Table 2 summarizes the results. We consider only the city and not the rest of the province, since waste collection since it is assumed that the recollection infrastructure is already in place and these facilities only change the end of use, from landfill to incineration.

Table 2.-Summary of economic parameters of facilities in different places

City	M	B	V	S	Gi	G	P	Sal	Sa
Habs (k)	3266	1637	794	689	270	232	201	144	96
P(MW)	39.2	18.8	12.6	11.8	3.8	4.0	1.9	2.0	1.3
I (M€)	67.3	46.6	39.0	37.9	27.1	26.9	23.7	24.0	22.8
E (€/kWh)	0.013	0.018	0.023	0.025	0.049	0.05	0.086	0.08	0.112
Water (L/kWh)	1.55	1.59	1.67	1.72	1.50	1.56	1.47	1.45	1.51
Jobs	325	227	189	184	130	131	115	117	111

The investment ranges from 23 M€ for a 100k habitants place to the 67 M€ of Madrid. Smaller facilities can create up to 111 jobs, including indirect ones, while larger ones can generate over 325 jobs. The production costs range within an order of magnitude from 0.013 €/kWh to 0.11 €/kWh. The contribution of the cooling section increases as the facility is smaller, while that of the gas treatment or the boiler decreases. The contribution of the turbine is always within 15%

4.2. Distributed production

Based on the previous study, it is possible to develop surrogates to estimate the production and investment costs of the facility as a function of the waste processed as well as the jobs generated and the water consumption as a function of the climate of the location. The facility location problem is formulated as in eq. (2). The results can be seen in Figure 2.

Max Obj

$$\begin{aligned}
 I_i &= f_1(W_i(\text{kg} / \text{s})) = -0.0172(W_i(\text{kg} / \text{s}))^2 + 2.0797(W_i(\text{kg} / \text{s})) + 20.829 \cdot y_i \\
 0.25 \cdot W_{\text{Avail},i} &\leq W_i \leq y_i \cdot W_{\text{Avail},i} \\
 P_i(\text{MW}) &= f_2(W_i(\text{kg} / \text{s})) = 1.336(W_i(\text{kg} / \text{s})) \\
 \text{Jobs}_i &= 0.57 \cdot (1 + 7.5) \cdot I_i \\
 C_{p,i}(\text{M€} / \text{yr}) &= f_3(W_i(\text{kg} / \text{s})) = 1.2533(W_i(\text{kg} / \text{s}))^{0.32} \\
 C_{\text{electricity}} \sum_i P_i(\text{MW}) \cdot 1000 \cdot 8760 &= \sum_i C_{p,i} \\
 C_{\text{electricity}} &\leq C_{\text{market}} \\
 \sum_i I_i &\leq \text{Budget} \\
 \text{Water}_{\text{consumption}}(\text{L} / \text{kWh}) &= -2.297 \cdot 10^{-4} \cdot T^2 + 0.798 \cdot H^2 + 7.090 \cdot p^2 + 2.200 \cdot 10^{-2} \cdot T \cdot H \\
 &+ 2.993 \cdot 10^{-2} \cdot T \cdot p - 0.515 \cdot H \cdot p - 1.533 \cdot 10^{-2} \cdot T - 1.417 \cdot H - 12.574 \cdot p + 7.6256
 \end{aligned} \tag{2}$$

Figure 2a shows that for a budget of 250 M€ and an electricity cost of 0.06 €/kWh, most of the locations make full use of the waste, but the largest city selected, Seville, only processes 83% of the waste for a power production of 9.8 MW. There are places with just 25% usage whose selection is more towards the social benefit of creating jobs, i.e. Toledo and Almería. The water consumption of the system is 1.673 L/kWh and the production capacity reaches 27 MW. For a budget of 500 M€, a larger number of locations was selected to the north, due to the depopulation suffered over the last years and the combination with lower GDP. Most of the same ones already chosen for the lower budget remain but the capacity of processing increases such as in the case of Seville, while Murcia and Cartagena were selected within the region of Murcia. The total power production reaches 55 MW while the water consumption decreases to 1.649 L/kWh

5. Conclusions

A multistage approach is presented for the valorization of urban waste. At process stage, a techno economic evaluation of a facility consisting of waste incineration, flue gas treatment to remove pollutants such as NO_x, SO₂, heavy metals and PDDFs, a steam turbine and a cooling tower to evaluate the water energy nexus. The units are modelled in detailed based on thermodynamics, first principles, industrial data and rules of thumb. Next, the investment and production costs are estimated.

A 1.336 MW/kg of mixed waste can be produced. For the facility to produce power at a competitive price, it has to process the waste generated at cities above 200k habitants. At country level, a facility location problem is formulated using a combined social index to favor economic growth. For the major cities in Spain (over 65 across the mainland) it would be possible to produce 235 MW from the waste generated with a production cost

of 0.05€/kWh and a total investment of 1700 M€. While the average costs are competitive, at regional level, some regions would produce at a high cost and only the national strategy would justify the production of power from waste beyond a waste management technology.

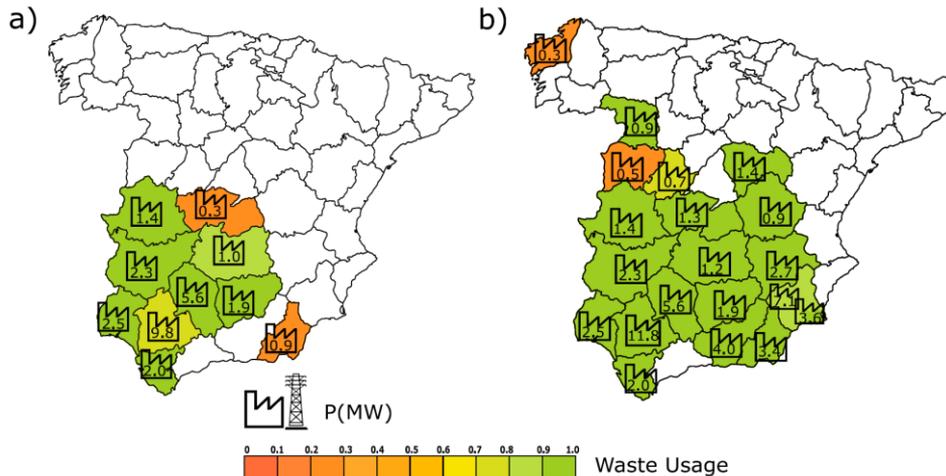


Figure 2- Facility location as a function of target cost of electricity and budget. a) I=250 M€; 0.06 €/kWh; b) 500 M€; 0.06 €/kWh

Acknowledgments

The authors would like to acknowledge MINECO, PID2019-105434RB-C31 as well as S.N. acknowledges the Undergraduate research fellowship

References

- Z. Fodor, J.J. Klemes. (2012) Waste as alternative fuel – Minimising emissions and effluents by advanced design. *Proc. Saf Environ Prot.* 90: 263-284
- L.S. Guerras, M. Martín (2020) On the water footprint in power production: Sustainable design of wet cooling towers. *Applied energy.* 263, 114620
- J. Heras, M. Martín (2020) Social issues in the energy transition: Effect on the design of the new power system. *Applied Energy* 278, 115654
- C. Hübner, R. Boos, J. Bohlmann, K. Burtscher, H. Wiesenberger (2000). *State-of-the-Art Measures for Dioxin Reduction in Austria.* Viena, Austria.
- L. Martín, M. Martín (2013) Optimal year-round operation of a Concentrated Solar Energy Plant in the South of Europe App.. *Thermal Eng.* 59., 627-633.
- A. Sánchez, M. Martín (2018) Scale up and Scale down issues of renewable Ammonia plants: Towards modular design. *Sust. Prod. Consump.* 16, 176-192
- R. Sinnott, G. Towler (2019). *Chemical Engineering Design. Principles, practice and economics of plant and process design.* 2nd ed. Elsevier Butterworth-Heinemann, Oxford.
- World Bank (2012) *What a waste: a global review of solid waste management.*
- European Commission. (2008) *Environmental, Economic and Social Impacts of the Use of Sewage Sludge on Land, Part I: Overview Report;* European Commission: Brussels, Belgium.

Integration of thermo-chemical energy storage for flexible operation of decarbonized power plants

Calin-Cristian Cormos, Letitia Petrescu, Ana-Maria Cormos

Babes-Bolyai University, Faculty of Chemistry and Chemical Engineering, Chemical Engineering Department, Arany Janos 11, Cluj-Napoca, RO-400028, Romania

Abstract

This work evaluates the integration of an innovative thermo-chemical energy storage system based on calcium-derived sorbents for flexible operation of decarbonized coal-fueled super-critical power plant. The reactive gas-solid cycle is used for both power plant decarbonization as well as a time-flexible thermo-chemical energy storage system in conjunction to a 500 MW net power output plant with 90% carbon capture rate. Overall techno-economic and environmental implications of flexible decarbonized power plant with calcium-based thermo-chemical energy storage system were evaluated using process flow modeling and thermal integration. For comparison reason, similar decarbonized power plant operated in base-load conditions as well as a non-capture power plant were also considered. As the results show, the utilization of calcium looping cycle for time-flexible thermo-chemical energy storage system in conjunction to a decarbonized fossil-based power plant bring significant benefits in term of reducing the specific capital cost (down to about 8%), the electricity production cost (down to about 3.6%), the CO₂ capture costs (down to about 3.5%), all compared to the base-load operation of similar decarbonized power plant. Although very promising in delivering better techno-economic performance indicators, the calcium looping technology still requires significant scale-up efforts from the current development level (up to 10 MW) to full industrial sizes.

Keywords: Thermo-chemical energy storage, Calcium looping cycle, Decarbonized power plants, Flexible power plant operation, Techno-economic evaluation.

1. Introduction

The energy sector is facing fundamental challenges in the transition to a low carbon economy as well as in relation to the higher penetration of renewable energy sources. The need to reduce the carbon emissions can be done by various methods e.g., switch to renewables, improving energy efficiency along the whole chain from production to utilization and deployment of Carbon Capture, utilization and Storage (CCUS) technologies. Since fossil fuels are still the dominant energy source for key utility and industrial sectors (e.g., heat and power, metallurgy, construction materials etc.), the CCUS options have a great potential to significantly contribute to the reduction of greenhouse gas emissions to meet the global environmental targets.

The renewable energy sources show a accelerate increase of their share but solar and wind applications are highly time-intermittent which put an additional operational burden to the back-up capacities most of them based on conventional fossil fuels. The energy storage technologies represent a very promising solution to accommodate the time-variability of the renewable energy sources. Various energy storage methods are available, some of them being already commercial to large scale (e.g., pumped hydro power, compressed air) others being in research stage (e.g., thermo-chemical storage).

Thermo-chemical energy storage has a great potential for development due to significantly higher energy density and improved energy efficiency (Yan et al., 2020). This paper is assessing the techno-economic implications of the Calcium Looping (CaL) cycle as both an energy-efficient CO₂ capture technology and a thermo-chemical energy storage option. As evaluated case study, a super-critical coal-based power plant operated in both base-load and load-following (flexible) conditions was considered. The key innovative elements of this paper in respect to the current state of the art are: evaluation of CaL technology for both decarbonization and thermo-chemical energy storage purposes and detailed techno-economic evaluation of 500 MW decarbonized power plant with thermo-chemical energy storage facility to improve its load following capabilities.

2. Integration of calcium looping system for power plant decarbonization

The calcium looping cycle for the post-combustion CO₂ capture is based on the following reversible chemical reaction:



The conceptual layout of CaL cycle has two reactors: one carbonation reactor in which the flue gases coming from the power plant are decarbonized according to above reaction and one calcination reactor in which the calcium carbonate is decomposed back to calcium oxide (sorbent regeneration) releasing the captured CO₂. Since the calcination reaction is endothermic, additional fuel has to be oxy-combusted to cover the reaction heat duty. Both reactors are operated in a Circulated Fluidized Bed (CFB) mode. The thermo-chemical energy storage based on CaL system uses the sorbent storage facilities in both regenerated and carbonated forms (Astolfi et al., 2019). The environmental impact of CaL cycle is minimal considering the usage of natural limestone as sorbent.

The CaL cycle was integrated into a coal-based super-critical power plant. The non-captured power plant concept (noted as Case 1) is based on current development stage. The design of decarbonized power plant integrating calcium looping cycle (noted as Case 2) is presented in Figure 1. The decarbonized power plant was assessed in two distinct situations: Case 2.a - base-load condition (according to the design presented in Figure 1) and Case 2.b - flexible (load-following) condition in which two high-temperature storage facilities (located between the carbonator and calcination reactors) were introduced for regenerated (calcinated) and carbonated sorbent material (Cormos, 2020).

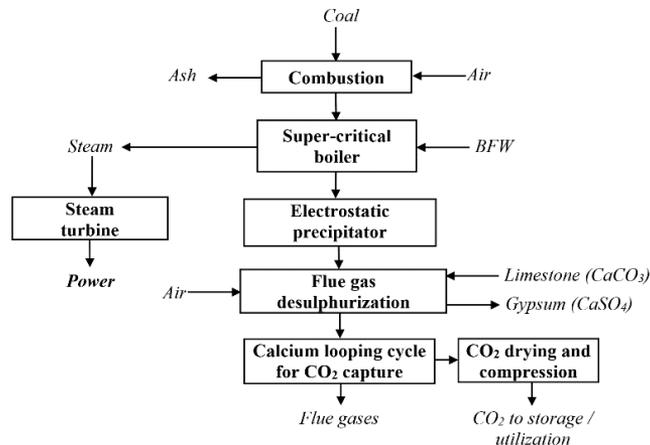


Figure 1. Decarbonized super-critical power plant integrating the CaL cycle

The design assumptions are presented in Table 1 (Cormos, 2020). All evaluated concepts were simulated using process flow modelling software ChemCAD (version 7.1.8). The CaL model was validated by comparison against experimental data from Astolfi et al., 2019 (see Figure 2 - left). For energy optimization, a detailed pinch thermal analysis was performed. Figure 2 - right presents the hot and cold composite curves for the CaL unit.

Table 1. Design assumptions of evaluated gasification-based poly-generation systems

Unit	Design assumptions
Coal composition (dry) and its calorific value	72.30% C, 7.45% O, 1.69% N, 4.11% H, 0.56% S, 13.89% ash; Moisture: 8%; Calorific value: 25.35 MJ/kg
Super-critical steam cycle	Steam temperature: 582°C / 580°C / 580°C Steam pressure: 290 bar / 75 bar / 20 bar Steam turbine efficiency: 88 - 90% Final steam expansion pressure: 45 mbar
Flue gas desulphurization	Wet (limestone slurry) desulphurization unit Desulfurization efficiency: >98%
Calcium looping cycle for post-combustion CO ₂ capture	Carbonation reactor temperature: 500 - 600°C Calcination reactor temperature: 800 - 950°C Carbon capture rate: 90%
CO ₂ processing unit (drying and compression)	Moisture removal unit: Tri-ethylene-glycol (TEG) Final compressing pressure at plant gate: 120 bar CO ₂ purity (vol.): min. 95%, max. 2000 ppm CO, max. 250 ppm water, max. 4% non-condensable gases
Auxiliary units	Energy consumption for coal handling: 0.5% of thermal input Heat exchanger minimum temperature difference: 10 °C Heat exchanger pressure drops: 1 – 3% inlet pressure

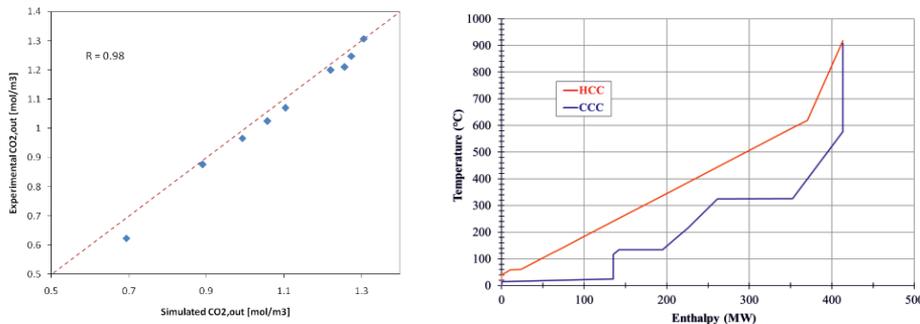


Figure 2. Validation of CaL unit (left); Thermal integration of CaL unit (right)

3. Techno-economic assessment of decarbonized power plant concepts

The first investigated power plant operation scenario was based on base-load conditions. In this case, no sorbent storage is needed. The evaluation of techno-economic performances for both decarbonized and non-capture power plant concepts was based on integrated assessment methodology as presented in a separate paper (Cormos and Dinca, 2021). Table 2 presents the main techno-economic performance indicators for base-load operation condition. It can be observed that introduction of decarbonization feature using CaL cycle implies an energy penalty for CO₂ capture of about 6.1 net efficiency percentage points. This energy penalty for post-combustion CO₂ capture is significantly reduced compared to other carbon capture technologies. For instance, the reactive gas-liquid absorption using alkanolamines (benchmark method considered in the literature) has an energy penalty of about 10 net percentage points (Koronaki et al., 2015).

Table 2. Base-load operation: Key performance indicators

Performance indicator	UM	Case 1	Case 2.a
Coal flow rate	t/h	165.00	200.01
Coal lower heating value (LHV)	MJ/kg	25.35	
Coal thermal energy	MW _{th}	1,161.87	1,408.46
Gross power output	MW _e	529.10	596.50
Fuel processing consumption	MW _e	5.80	7.04
CO ₂ capture and conditioning unit	MW _e	-	47.21
Power block consumption	MW _e	23.30	22.45
Ancillary power consumption	MW _e	29.10	76.70
Net power output	MW _e	500.00	519.80
Net power efficiency	%	43.03	36.90
Carbon capture rate	%	0.00	90.00
Specific CO ₂ emissions	kg/MWh	800.40	70.15
Specific primary energy consumption for CO ₂ avoided	MJ/kg CO ₂	-	1.95
Capital cost	M€	747.25	1,237.26
Specific capital investment cost	€/kW	1,490.00	2,474.00
Operational & maintenance cost	€/MWh	28.70	35.12
Levelized cost of electricity (LCOE)	€/MWh	70.99	80.52
CO ₂ removal cost	€/t	-	39.75
CO ₂ avoided cost	€/t	-	46.98

The economic penalty for decarbonization is also significant, all economic indicators showing increases e.g., the specific investment cost by about 66%, the operational & maintenance cost by about 22%, the levelized cost of electricity by about 14%. The CO₂ avoided cost is about 47 €/t which underline once more the attractiveness of CaL technology considering the current CO₂ emission tax which is currently in the range of 50 - 60 €/t (Ember, 2021). Following the evaluation of power plant base-load operation, a flexible load-following pattern was considered. The load-following is very important considering the integration of time-intermittent renewables. Table 3 shows the considered flexible operational cycle over one week. This operational cycle was chosen based on daily load-following pattern which is common in Central Europe (Astolfi et al., 2019).

Table 3. Weekly time operation cycle of decarbonized power plant

Monday to Friday		Saturday to Sunday
6 AM to 1 PM	10 PM to 5 AM	0 AM to 12 PM (all day)
7 PM to 9 PM	2 PM to 6 PM	
100%	50%	50%

To comply with this flexible operation scenario, the calcium looping cycle was modified by introduction of sorbent storage facilities in both regenerated and carbonated forms. Figure 3 presents the conceptual design of a flexible CaL cycle with high-temperature solid sorbent storage. Each sorbent storage has a storage temperature close to the reactor to which the solid is fed. Considering the weekly time operation cycle of decarbonized power plant as presented in Table 3, the flue gases flow to be treated for CO₂ capture is about 73% for the Monday to Friday period, 50% for the weekend and about 66.5% on average for the whole week. For base-load operation, the flue gases flow generated by the power plant to be treated for decarbonization is about 2,000 t/h with 12.4 % vol. CO₂.

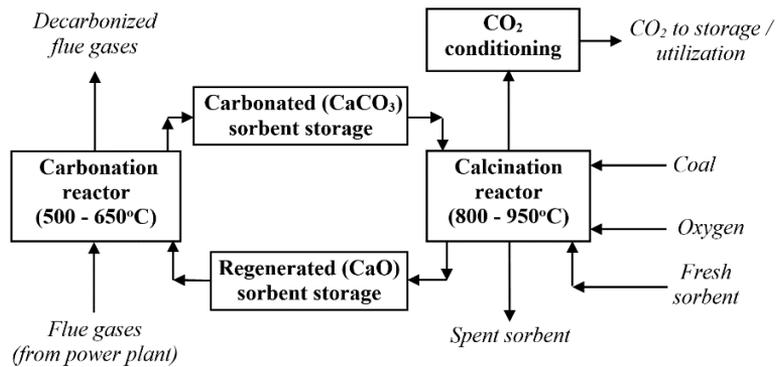


Figure 3. Flexible operation of calcium looping cycle

When the grid power demand is reduced, the power plant output as well as the flue gas to be treated for decarbonization are also reduced. For time-flexible operation, the same carbon capture rate (90%) was considered as in the case of base-load operation. To cope with power load-related variations, the calcium looping unit is to be operated in a flexible scenario considering the sorbent storage. The overall size of CaL unit as well as solid storage facilities were calculated based on the weekly operation cycle. The size of CaL unit for flexible operation was reduced to about 74% of the nominal base-load case leading to a significant capital cost reduction. A part of this capital cost reduction was used for the sorbent storage but considering the complexity of CaL design in comparison to the storage unit, there is an overall investment cost reduction for the flexible design. The time variations of power plant load and carbonated and regenerated solid storages are presented in Figure 4. As operation strategy, when the power plant operates at nominal load (100 %), the carbonated storage facility is gradually filling up with the correspondent discharge of the regenerated storage facility. At reduced power plant output (e.g., during the night or during the weekend), the process is reversed with accumulation of regenerated sorbent material for the next peak-time operation of the plant.

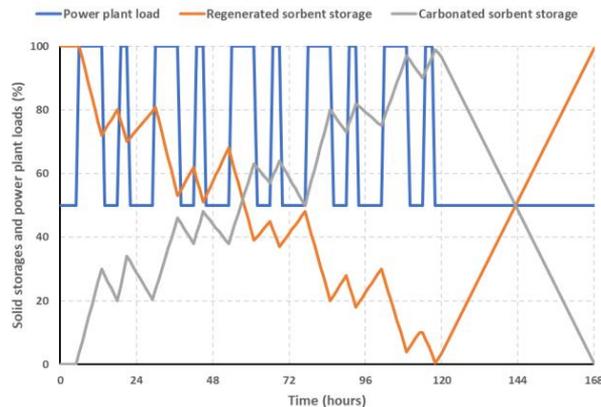


Figure 4. Weekly variations of plant load and regenerated / carbonated solid storages

As explained above, the reduction of CaL size leads to an overall capital cost reduction (although some of this reduction is used for the sorbent storage facility) with positive economic results. The comparison of key techno-economic indicators for the base-load (Case 2.a) and flexible-load (Case 2.b) decarbonized power plant is presented in Table 4

Table 4. Comparison of nominal and part-load operation of decarbonized power plant

Performance indicator	UM	Case 2.a	Case 2.b
Calcium looping cycle size factor	%	100.00	74.00
Specific power plant investment cost	€/kW	2,474.00	2,278.00
Operational & maintenance cost	€/MWh	35.12	34.85
Levelized cost of electricity (LCOE)	€/MWh	80.52	77.60
CO ₂ avoided cost	€/t	46.98	45.30

As shown, the flexible operation of decarbonized power plant fitted with sorbent storage option has better performance than base-load operation e.g., reduced investment cost by about 8%, reduced electricity production cost by about 3.6% and reduced CO₂ avoided cost by about 3.5%. However, it must be stated that further technological developments are needed for both calcium looping system as well as for high-temperature storage.

4. Conclusions

The present work assesses the potential techno-economic benefits of flexible thermo-chemical energy storage system based on the calcium looping cycle applied for a decarbonized super-critical power plant. The solid sorbent storage feature was used to increase the load-following capability of the power plant. As the results show, the flexible power plant operation brings significant advantages in comparison to the base-load case (reduction of all assessed economic indicators).

Acknowledgements

This work was supported by two grants of the Romanian Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project numbers: PN-III-P4-ID-PCE-2020-0032 and PN-III-P2-2.1-PED-2019-0181, within PNCDI III.

References

- M. Astolfi, E. De Lena, M.C. Romano, 2019, Improved flexibility and economics of Calcium Looping power plants by thermochemical energy storage, *International Journal of Greenhouse Gas Control*, 83, 140-155
- C.C. Cormos, 2020, Techno-economic implications of flexible operation for super-critical power plants equipped with calcium looping cycle as a thermo-chemical energy storage system, *Fuel*, 280, 118293
- C.C. Cormos, C. Dinca, 2021, Techno-economic and environmental implications of decarbonization process applied for Romanian fossil-based power generation sector, *Energy*, 220, 119734
- EMBER - Coal to clean energy policy, Daily carbon prices, 2021, <https://ember-climate.org/data/carbon-price-viewer/>
- I.P. Koronaki, L. Prentz, V. Papaefthimiou, 2015, Modeling of CO₂ capture via chemical absorption processes - An extensive literature review, *Renewable and Sustainable Energy Reviews* 50, 547-566
- Y. Yan, K. Wang, P.T. Clough, E.J. Anthony, 2020, Developments in calcium/chemical looping and metal oxide redox cycles for high-temperature thermochemical energy storage: A review, *Fuel Processing Technology*, 199, 106280

Superstructure optimization for sustainable design of an algae biorefinery

Maryam Raeisi, ^{a*} Thien An Huynh,^a Meik B. Franke, ^a Edwin Zondervan ^a

^a*Sustainable Process Technology, Faculty of Science and Technology, University of Twente, Meander, kamer 216, Postbus 217, 7500 AE Enschede, the Netherlands*

**m.raeisi@utwente.nl*

Abstract

In this study a superstructure of an algae biorefinery to produce added value products (pigments, omega-3, glycerol, biodiesel, biogas, and fertilizer) from microalgae is developed. From the superstructure optimization follows a cost optimal production pathway that consists of an open pond, sedimentation and flotation, flocculation/centrifugation without a dryer, hydrothermal liquefaction, organic solvent pigment extraction, N-butanol lipid extraction, lipid production, and anaerobic digestion. The profits of the algae biorefinery depend on the types of wastewaters. 107 million Euros income can be earned annually using 0.2 million tons of influent wastewater. The total profit of an algae biorefinery that uses influent wastewater as feedstock is approximately two times higher than the wastewater of wheat straw biorefinery. 63 million Euros income can be earned annually using 0.6 million tons of influent wastewater. Furthermore, the total profits of algae biorefinery in each season are calculated by using real data of sunlight periods of the Netherlands. Summer is the best season with more than 36 and 22 million Euros profits for algae biorefineries when using influent wastewater and wastewater of wheat straw biorefinery, respectively. Calculating total profits of algae biorefinery by considering fix value for durations of sunlight has more than 30% error.

Keywords: Algae biorefinery, added-value products, superstructure, MINLP.

1. Introduction

Among various types of biomass, microalgae are considered valuable due to their high growth rate and photosynthesis efficiency (Fernández-Linares et al., 2017). Microalgae are oceanic microorganisms that grow in wastewater using carbon dioxide and sun or artificial light. The capability to capture carbon dioxide and treat wastewater makes this microorganism one alternative to solve environmental issues (Assis et al., 2019). In addition, algae have a high potential to serve as source for various added-value products. Despite microalgae's vast environmental and economic benefits, high-cost extraction process from them must be overcome to extend its industrial application. Two approaches (heuristics and superstructure optimization) are typically employed to design an added value production process. The heuristic method is founded on rules acquired from experience and comprehension of unit operations. Although, this method is very fast, it does not propose a certain route for evaluating different cases (Mencarelli et al., 2020). The superstructure method relies on optimization algorithms and mathematical models to determine a cost-effective production pathway of all alternatives (Tula et al., 2017). Rizwan et al. (2013) proposed an optimal production pathway of producing biodiesel from microalgae by optimizing superstructure. The goal of this study was to maximize

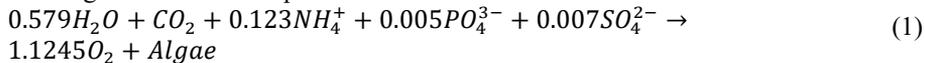
biodiesel yield/ operating margins with(out) minimizing waste. The feedstock was defined as one block of the superstructure (Rizwan et al., 2013). Galanopoulos et al. (2019) formulated a superstructure for an integrated algae biorefinery. The wastewater stream of a wheat straw biorefinery was linked to the superstructure to produce glycerol, biodiesel, bioethanol and levulinic acid (Galanopoulos et al., 2019). The duration of sunlight was assumed 12(h) per day.

The role of wastewater as an important feedstock for the algae biorefinery has not been considered sufficiently. Choosing an appropriate wastewater source will improve the chance of applying the algae biorefinery concept at an industrial scale. In this study, a superstructure is developed to optimize the pathway for producing added-value products from microalgae using two different types of wastewaters. The profits of the algae biorefinery for each wastewater composition are calculated. Total profits of algae biorefinery in each season are calculated based on actual duration sunlight data.

2. Superstructure optimization

2.1. Process description

First, carbon dioxide and wastewater with nutrients are fed to the cultivation interval to grow microalgae. Microalgae (*Haematococcus Pluvialis* ($CO_{0.378}H_{1.65}N_{0.123}P_{0.005}S_{0.007}$)) (Ba et al., 2016) consume wastewater nutrients for regenerating and growing with the following molar reaction equation.



Four options are available in this interval (open pond, flat plate photobioreactor, bubble column photobioreactor, turbo column photobioreactor). To separate water from microalgae, it should pass a harvesting (gravity sedimentation and filtration or flotation), a dewatering (filter press, centrifugation, flocculation), and a drying interval. Before going through pigment extraction (chemical solvent or supercritical fluid extraction), the microalgae cells are disrupted (by for example: hydrothermal liquefaction, high-pressure homogenization, bead beating, microwaving, sonication). Subsequently, the lipids of the microalgae are extracted by Hexane, n-butanol, or supercritical carbon dioxide extraction. The products of lipid extraction are then transferred to a lipid production interval to produce omega-3, biodiesel, and glycerol. At the remnant intervals biogas and fertilizers are produced. To produce various bioproducts from microalgae, different technologies and processes can be selected. A mapping of all possible combinations of these technologies lead to a superstructure, as shown in figure 1. The superstructure can be converted to a mathematical model by defining an objective function and various variables and constraints.

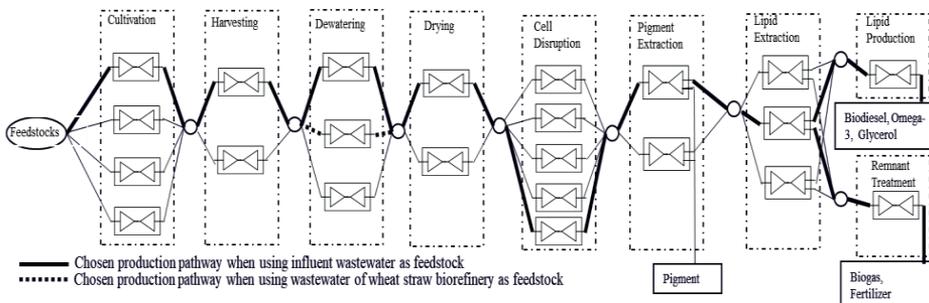


Figure 1: developed superstructure with chosen production pathway

2.2. Problem statement

Given are the composition of wastewater and products and the topology of the superstructure with all pathways. Furthermore, each interval's economic specifications, technical properties (split factors and yields), and utility requirements are given. Under conditions of logical constraints (only one technology option at each interval is selected) and where steady state mass and energy balances hold. The operating and investments costs are based on the flow rates. The decisions to be made are which technology options should be selected for each interval that maximize the revenues (profit minus costs).

2.3. Mathematical model

The developed superstructure consists of 9 intervals and 23 options (as described in process description section). A mixed-integer nonlinear programming (MINLP) was developed to optimize it. Steady state mass and energy balance and also economic scale and logical constraints are applied for each option of the superstructure.

2.3.1. Mass balance

Flows with a component (k) can enter or leave of each interval (j) as shown in figure 2. The $m_{k,j}^U$, is the mass flow (t/h) that comes from previous interval or feedstock (for the first interval of biorefinery). This flow combines with reactant flow ($m_{k,j}^R$)(t/h), if solvents are needed, to produce input flow($m_{k,j}^{IN}$) (t/h) (Eq.(2)). The reactant flow can be calculated with concentration factor $x_{k,j}$ as demonstrated in Eq.(3). This fraction factor is defined by amount of basic component in the upstream flow.

$$m_{k,j}^{IN} = m_{k,j}^U + m_{k,j}^R \quad (2)$$

$$m_{k,j}^R = x_{k,j} \cdot m_{k,j}^U \quad (3)$$

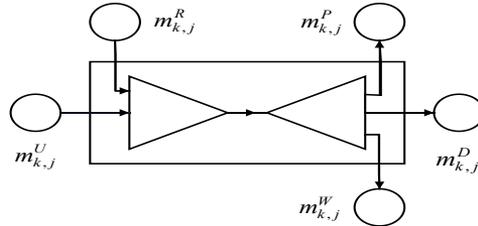


Figure 2. mass flows enter/leave the intervals (j)

The flow after reacting and distribution ($m_{k,j}^{OUT}$) (t/h) can be calculated by Eq.(4). In this equation, the $CF_{k,j}$, $S_{k,j}$ and $D_{k,j}$ are conversion factor (yield), mass stoichiometric coefficient and distribution coefficients, respectively. If there is no reaction or distribution in the block, the sum of these two parameters becomes zero. Thus, the input mass flow equal to output mass flow.

$$m_{k,j}^{OUT} = m_{k,j}^{IN} + m_{k,j}^{IN} \cdot CF_{k,j} \cdot (S_{k,j} + D_{k,j}) \quad (4)$$

After reacting or distributing, the output flow is divided to downstream flow ($m_{k,j}^D$) (t/h), waste flow ($m_{k,j}^W$) (t/h), and product flow ($m_{k,j}^P$) (t/h) by defining split factor(SF)(as shown in Eq.(5)).

$$m_{k,j}^{OUT} = SF_{k,j}^D \cdot m_{k,j}^{OUT} + SF_{k,j}^W \cdot m_{k,j}^{OUT} + SF_{k,j}^P \cdot m_{k,j}^{OUT} \quad (5)$$

2.3.2. Energy balance

Electricity U_j^E (kWh), heating U_j^H (kWh), and cooling U_j^C (kWh) are considered as utilities in this model. The required energy is calculated in a similar manner as with the mass balance. Eq.(6) shows energy balance, where SUC_j is an utility factor.

$$U_j = \sum_k m_{k,j}^{IN} \cdot SUC_j \quad (6)$$

2.3.3. Logical constraints

The logical constraints are setup to ensure the selection of only one technology at each interval. A variable (y_j) is defined and should be multiplied by each downstream flow. The summation of this variable for each stage of biorefinery should be an equal one.

2.3.4. Economy of scale

The annualized investment cost (AIC) (*Euros*) is calculated by Eq. (7)-(9), where $TIPC$, LT , IR , K^{ENG} , EC_j^{ref} , m_j^{ref} , IDX_j^{2020} , IDX_j^{ref} are total interval installation cost (*Euros*), the lifetime, the interest rate, engineering coefficient, reference cost (*Euros*), reference mass flow, cost index in 2020, reference cost index, respectively. Furthermore, the P^{Land} (*Euros*) is land price.

$$AIC = TIPC \cdot \frac{IR \cdot (IR+1)^{LT}}{(IR+1)^{LT}-1} \quad (7)$$

$$TIPC = K^{ENG} \cdot \sum_j EC_j^{ref} \cdot \left(\frac{\sum_k m_{j,k}^{IN}}{m_j^{ref}} \right)^{f_j} \cdot \left(\frac{IDX_j^{2020}}{IDX_j^{ref}} \right) + LC_j \quad (8)$$

$$LC_j = P^{Land} \cdot \frac{m_{Algae,1-4}^{OUT}}{Productivity_{Algae,1-4}} \quad (9)$$

The operating cost is a summation of raw material, utility, operating and maintenance, and waste treatment costs. The selling value is calculated by multiplying the total product mass flow, the product prices, and the operating hours per year. The selling values are decreased from the summation of annualized investment and operating costs to obtain total profits. The objective of this study is to maximize the total profits.

3. Results

Durations of sunlight determine the growth of the microalgae and subsequently the quantities of extracted products. These values are not the same in different seasons. The average sunlight periods in the Netherlands are 12.5, 15, 11.5, 8.5 (h) for spring, summer, autumn, and winter, respectively (Slegers, 2014). These sunlight periods are used to calculate the total profits of algae biorefinery. Furthermore, two types of wastewaters are considered in this study. The average compositions of influent wastewater of Netherlands in 2018 and wastewater of a wheat straw biorefinery (Galanopoulos et al., 2019) can be found in Table 1. 10% of mass flow of each type of wastewater are entered into the superstructure to consider the role of feedstocks on total profits and cost-effective production pathway. Approximately, 0.2 million tone influent wastewater and 0.6 wastewater of wheat straw biorefinery's are considered as the feedstocks of these biorefineries, annually.

Table1: composition of wastewater (Galanopoulos et al., 2019, stateline, 2021)

Component	Influent wastewater (t/h)	Wastewater of wheat straw biorefinery(t/h)
H_2O	224.59	824
NH_4^+	76.37	0.949
PO_4^{2-}	11.88	0.988
SO_4^{2-}	-	1.166

A mixed-integer non-linear programming (MINLP) model is implemented in the Advanced Interactive Multidimensional Modelling (AIMMS) software version 4.82.3.29 64-bit. The Outer Approximation Algorithm (AOA) combines the CONOPT 4.1 solver for the non-linear part and the CPLEX 20.1 solver for the integer part.

The cost-effective production pathways of producing added value products from microalgae using two types of wastewaters are shown in figure 1. The open pond, sedimentation and flotation, centrifugation without a dryer, hydrothermal liquefaction, organic solvent pigment extraction, N-butanol lipid extraction, lipid production, and anaerobic digestion are chosen when using influent wastewater. The mass percentage of water in wastewater of wheat straw biorefinery (824 t/h) is more than mass percentage of water in wastewater of influent wastewater biorefinery (224.59 (t/h)). It is economically beneficial to choose flocculation for dewatering intervals when using wastewater of wheat straw. Thus, the cost-effective production pathway for producing added value products when using wheat straw biorefinery wastewater is the open pond, sedimentation and flotation, flocculation, hydrothermal liquefaction, and organic solvent pigment extraction, N-butanol lipid extraction, lipid production, and anaerobic digestion.

The total profits of the algae biorefinery for each season when using each type of wastewaters are calculated (as shown in figure 3). Summer is the best season with the highest profit for both algae biorefineries due to long sunlight periods and a high amount of microalgae and extracted products. Using influent wastewater of Netherlands makes more profits for the algae biorefinery compared to another type of wastewater. The total profit of an algae biorefinery when using influent wastewater of Netherlands as feedstocks can be reached to 107 million Euros annually. The profit of this biorefinery is approximately two times higher than the algae biorefinery when using wastewater of wheat straw biorefinery as feedstock. Microalgae (*Haematococcus Pluvialis*) can be grown in typical Dutch influent wastewater more effectively than another type of wastewater due to the high ratio of nutrients. Furthermore, the low amount of water of influent wastewater is helped to have low investment and operating costs.

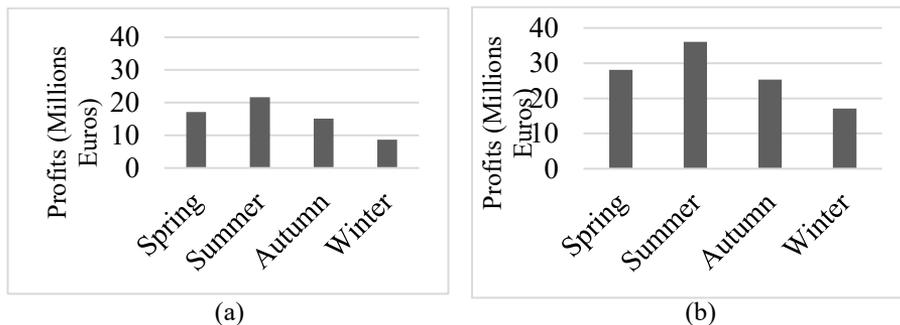


Figure3: profits of algae biorefinery in each season when using a) wastewater of wheat straw biorefinery b) influent wastewater as a feedstock

It is more common to assume fix value for sunlight periods (12 (h)) when calculating the total profits of algae biorefinery (Galanopoulos et al., 2019). With this assumption, the total profits of algae biorefineries when using wastewater of wheat straw biorefinery and influent wastewater are 93 and 163 million Euros, respectively. By considering real data of sunlight periods, the total profits of algae biorefineries are estimated 63 and 107 million Euros when using wastewater of wheat straw biorefinery and influent wastewater, respectively. Approximately 30% error are made by considering fix value for sunlight duration. Thus, this parameter has an important role in calculating the total profit of algae biorefinery.

4. Conclusion

The pathway of open pond, sedimentation and flotation, flocculation/centrifugation without a dryer, hydrothermal liquefaction, organic solvent pigment extraction, N-butanol lipid extraction, lipid production, and anaerobic digestion is chosen as the cost-effective production process for added-value products from microalgae. Types of wastewaters determine the appropriate technology of dewatering interval (flocculation/centrifugation). The incomes of 107 million Euros and 63 million Euros can be obtained annually using 0.2 million tons and 0.6 million tons of influent wastewater and wheat straw biorefinery's wastewater, respectively. Total profits of algae bioenergy with real data of sunlight duration of different seasons and an assumption fixed value (12 h) of sunlight periods are compared. This common assumption increases the estimated value of the total profits of algae biorefinery with different feedstocks.

References

- F. Ba, A.V. Ursu, C. Laroche, G. Djelveh, 2016. Haematococcus pluvialis soluble proteins: Extraction, characterization, concentration/fractionation and emulsifying properties. *Bioresour. Technol.* 200, 147–152.
- T.C. de. Assis, T.C. de, Calijuri, M.L., Assemany, P.P., Pereira, A.S.A. de P., Martins, M.A., 2019. Using atmospheric emissions as CO₂ source in the cultivation of microalgae: Productivity and economic viability. *J. Clean. Prod.* 215, 1160–1169.
- Fernández-Linares, L.C., Guerrero Barajas, C., Durán Páramo, E., Badillo Corona, J.A., 2017. Assessment of *Chlorella vulgaris* and indigenous microalgae biomass with treated wastewater as growth culture medium. *Bioresour. Technol.* 244, 400–406.
- Galanopoulos, C., Kenkel, P., Zondervan, E., 2019. Superstructure optimization of an integrated algae biorefinery. *Comput. Chem. Eng.* 130, 106530.
- Mencarelli, L., Chen, Q., Pagot, A., Grossmann, I.E., 2020. A review on superstructure optimization approaches in process system engineering. *Comput. Chem. Eng.* 136, 106808.
- Rizwan, M., Lee, J.H., Gani, R., 2013. Optimal processing pathway for the production of biodiesel from microalgal biomass: A superstructure based approach. *Comput. Chem. Eng.* 58, 305–314.
- Slegers, P.M., 2014. Scenario studies for algae production. Wageningen University. stateline, 2021. Urban waste water treatment per province and river basin district.
- Tula, A.K., Babi, D.K., Bottlaender, J., Eden, M.R., Gani, R., 2017. A computer-aided software-tool for sustainable process synthesis-intensification. *Comput. Chem. Eng.* 105, 74–95.

Economic and Environmental Optimization and Feedstock Planning for the Renewable Jet Production Using an Intensified Process

Raul Mauricio Rivas-Interian,^a Eduardo Sanchez-Ramirez,^a Juan Gabriel Segovia-Hernandez,^a Juan Jose Quiroz-Ramírez^b

^a*Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, Mexico*

^b*CIATEC, Omega No.201 Colonia Industrial Delta, Leon 37545, Mexico*

Abstract

Biojet fuel is able to be produced from lignocellulosic biomass having ethanol as intermediate. However, the whole process faces challenges related to high production costs and environmental impact as well as the feedstock annual availability. In this paper, an intensified process for biojet fuel production is presented, which includes sugarcane bagasse and corn stover planning within a superstructure scheme with steam explosion and diluted acid pretreatments. By employing stochastic optimization, the total annual cost (TAC) and the ecoindicator-99 (EI99) were minimized to meet sustainability goals. The purification zone of ethanol was intensified and schemes of vapor side stream column (VSSC) and dividing wall column (DWC) were evaluated and compared against the conventional extractive distillation. TAC savings of 5.56% and 5.02% for VSSC and DWC schemes, respectively, were achieved. Also, EI99 reductions of 1.72% and 2.92% were recorded for VSSC and DWC schemes, respectively. As results of the planning, 95.3% of annual feedstock is sugarcane bagasse, from which 54% is sent to acid pretreatment. Regarding to the corn stover, 72% is sent to acid pretreatment.

Keywords: stochastic optimization, process intensification, biojet fuel, lignocellulosic biomass, ethanol.

1. Introduction

In recent years, the production and consumption of fossil jet fuel have increased as a consequence of a rise in the number of passengers and goods transported by air. Despite the low demand caused by COVID-19 pandemic, a rise in the services offered by the sector is expected again. In an economic context still dependent on scarce oil, this represents a problem as well as the inherent environmental impact throughout the life cycle of this fuel. Given this, a promising solution is the use of biojet fuel as renewable aviation fuel. Previous research has widely explored the route of oils to obtain biojet (Rivas-Interian, 2021). However, the feedstock for the oil path, their water and land requirements, and maturation times have proven not to be entirely sustainable. In contrast, the route of alcohols offers an alternative since they can be produced from sugars present in the widely available lignocellulosic biomass. Thus, in a framework of circular economy, the use of lignocellulosic residues allows the production of alcohols necessary to obtain biojet fuel. Besides, the tools provided by process intensification make it possible to design a sustainable process with low environmental impact and capable to achieve energy savings. However, the use of biomass to produce biofuels is limited by the availability of the feedstock, which affects the supply chain. For this reason, an

adequate feedstock planning is necessary to take into account the time and frequency of the harvest. The goal of this work was to carry out the feedstock planning and the optimization of the entire biorefinery in terms of economic and environmental objectives.

2. General description of the process

The overall process of biojet production from lignocellulosic biomass requires the production of intermediate alcohols. The general scheme of alcohols production from lignocellulosic agro-waste involves steps of i) biomass pretreatment, ii) enzymatic hydrolysis, iii) fermentation of sugars, and iv) purification of the alcohols produced.

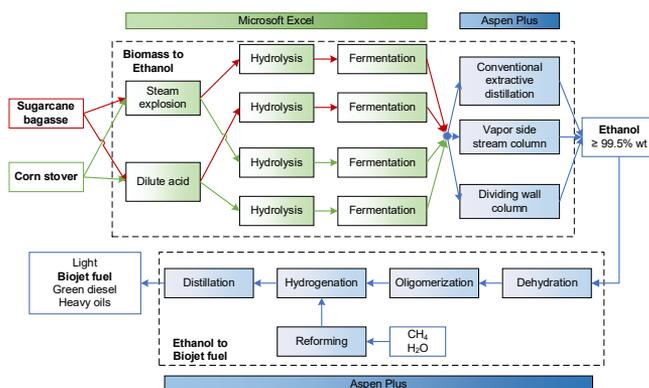


Figure 1. Complete superstructure for biojet fuel production process (Rivas-Interian, 2021)

Once the alcohols have been obtained, they are sent to the alcohol-to-jet (ATJ) process to be upgraded to biojet fuel through steps of v) dehydration to ethylene, vi) oligomerization to olefins, vii) hydrogenation to paraffins, and viii) distillation of hydrocarbon fractions where biojet is contained. Ethanol has recently been approved as raw material allowing the biojet fuel produced from it to be mixed with conventional jet fuel up to 50%.

An area of opportunity concerns to the purification of ethanol. Extractive distillation has relatively high costs and, despite this, remains as the main choice for large-scale ethanol production. Its high costs and energy consumption have been widely demonstrated. For this reason, a first approach on the path to sustainability is its intensification. By intensifying this zone, economic and environmental improvements can be achieved and an energy-efficient, cheaper, and greener process can be obtained as reported (Kiss & Suszwalak, 2012).

3. Case of study

To create a feasible market of biojet fuel in Mexico, it is necessary to meet a demand of at least 5.5% conventional jet fuel with biojet fuel (SENER, 2017), which equals 258 million liters. The production of biojet fuel from ethanol involves a series of steps that implicates the production of ethanol from lignocellulosic biomass. The availability of biomass limits the whole process, making it necessary to consider the temporality of the crops and, therefore, of their waste generated within the planning of the feedstock. To satisfy the demand of biojet fuel during a period, more than one feedstock has to be considered. Also, to leverage the number of available pretreatment technologies and their ability to break biomass, more than one pretreatment has to be evaluated regarding its

efficiency and profitability in order to find the most optimal route. The planning of the feedstock and the diverse combination of feedstocks and pretreatments make it possible to organize the process within a superstructure scheme.

3.1. Feedstock and pretreatment selection

Based on the information provided by the Mexican Agricultural and Fisheries Information Service, sugarcane bagasse and corn stover were selected as feedstocks, as they were the most abundant biomasses in Mexico during 2018. Additionally, they can be potential substrates for ethanol production since they have high sugar content and are renewable, cheap, and readily available feedstocks. On the other hand, steam explosion and dilute sulfuric acid were selected as pretreatments based on their low energy consumption per tonne of dry biomass and per gallon of bioethanol reported by Conde-Mejía, et al. (2012).

3.2. Process modeling

According to a modular manufacturing scheme, the whole process was modeled separately in two parts: lignocellulosic biomass to ethanol, and ethanol to biojet fuel. This makes possible to locate the plant of ethanol near to harvest sites and the plant of biojet fuel near to the airports thus reducing supply chain costs, increasing the flexibility of the whole process, and allowing the production network to react to dynamic supply and demand developments (Finkbeiner, et al., 2020).

The biomass-to-ethanol process was designed according to the superstructure scheme by combining the two feedstocks and pretreatments. To model the pretreatment, enzymatic hydrolysis, and fermentation of this module, kinetic data were obtained from literature and fitted to regression equations. This reaction train was modeled in MS Excel. For the fermentation, the *S. cerevisiae* strain was considered. The purification zone of ethanol as well as the module ethanol-to-biojet fuel were modeled in Aspen Plus V8.8, by employing the thermodynamic models of NRTL and ENRTL, respectively. Since the cost of separation represents a well-known area of opportunity, it was intensified by proposing a column sequence with vapor side stream and a column sequence of dividing wall (DWC) as intensified schemes.

As could be seen, the previous superstructure, along with the ATJ process, was modeled with highly non-linear and potentially non-convex equations. Besides, the existence of degrees of freedom allows solving the design problem as an optimization problem. Finally, the superstructure to be optimized is shown in Figure 1.

4. Process optimization

The search of the design and operation parameters that maximize savings and minimize environmental impact implies an optimization work. Such objectives as well as the optimization technique are described below.

4.1. Objective functions

To assess the sustainability of the process, the total annual cost (TAC) and eco-indicator-99 (EI99) were selected as objective functions. TAC allows quantifying the economic performance of a process based on its characteristics, while EI99 measures its

environmental impact in terms of human health, ecosystem quality and resource depletion. They were calculated with the Equations 1 and 2, respectively:

$$TAC = \frac{\frac{\text{Capital cost}}{\text{Recovery time}} + \text{Operation cost}}{F_k} \quad (1)$$

$$EI99 = \frac{\sum_b \sum_d \sum_{k \in K} \delta_d \omega_d \beta_b \alpha_{b,k}}{F_k} \quad (2)$$

Where F_k is the reference mass flow, β_b is the total amount of chemical b released per unit of reference flow due to direct emissions, $\alpha_{b,k}$ is the damage caused in category k per unit of chemical b released to the environment, ω_d is a weighting factor for damage in categories d , and δ_d is the normalization factor for damage of category d . The unit of measurement employed for EI99 is the ecopoint, which represents one-thousandth of the annual environmental load of an average European inhabitant. The capital cost was calculated by employing the Guthrie's method. Likewise, for the production process of ethanol, the annual production of alcohol was included to satisfy a demand of 5.5% of conventional jet fuel in Mexico with biojet fuel (SENER, 2017). The objective function is stated as follow:

$$F_{obj}(\vec{X}) = \begin{cases} \text{Min}(TAC, EI99) \\ \text{Max}(EtOH) \end{cases} \quad (3)$$

Pressures, temperatures, acid/enzyme concentrations, residence times, reflux ratios, and solvent/feed ratios were considered as continuous decision variables. Also, total stages, feed/solvent/side stream stages were considered as discrete decision variables. In total, 132 continuous and 21 discrete variables were counted for biomass-to-ethanol process. For ethanol-to-biojet process, 25 continuous and 5 discrete variables were also counted.

4.2. Stochastic optimization

To optimize the case of study, the Differential Evolution with Tabu List (DETL) (Srinivas & Rangaiah, 2007) stochastic method was employed. This is acknowledged for its ability to locate the global optimum regardless of the parameters of the problem, its small number of evaluations of the target function, and its efficiency in terms of computation times. This method was responsible for choosing the most suitable design parameters, operating conditions, and monthly feedstock to achieve an optimal design. It was implemented using a hybrid platform that incorporated MS Excel and Aspen Plus. There, a vector of decision variables is sent to MS Excel by using a dynamic data exchange with COM technology. Also knowing that a distributed configuration as lower cost than a centralized configuration (see Section 3.2), the process was optimized separately in two blocks: biomass-ethanol and ethanol-biojet. The following parameters were used: 120 individuals, 1000 generations, tabu list of 50% of the total individuals, tabu radius of 0.0001, and crossover and mutation probabilities of 0.9 and 0.3, respectively.

5. Results

5.1. Feedstock planning

As a result of the optimization, to produce enough ethanol to reach the biojet demand, 8 357 524 tons/year of sugarcane bagasse should be required of which 54% is sent to acid pretreatment; and 408 970 ton/year of corn stover, of which 72% is subjected to acid pretreatment. Only 15% and 0.85% of the bagasse and corn stover available in Mexico were employed. The most of the feedstock used annually is acid-pretreated sugarcane

bagasse (Figure 2) since the bagasse has a higher content of hexoses and it is the cheapest feedstock, so the optimization method tends to choose it as the most suitable. It is important to highlight that in some months one pretreatment is preferred over the other one. This is partially due to the randomness of the stochastic method to choose the best alternative. However, an additional objective would be necessary, either for security, occupational health or controllability for the results to show homogeneity.

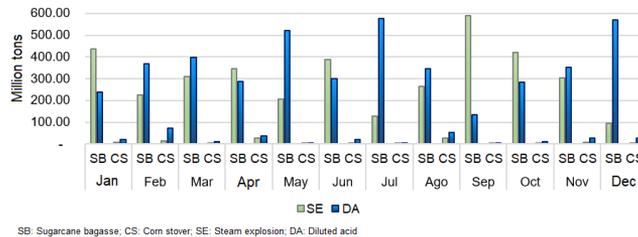


Figure 2. Annual feedstock planning

5.2. Optimization of ethanol process

The maximum ethanol production was achieved at 79 894 kg/h (679 100 ton/year) with a purity greater than 99.5% by weight, at a minimum TAC of 0.656 USD/kg ethanol and a minimum EI99 of 0.414 ecopoints/kg ethanol (Figure 3). In this case, more ethanol production is expected to demand larger equipment and more energy and feedstock, which becomes an increase in the total annual cost and in the ecoindicator.

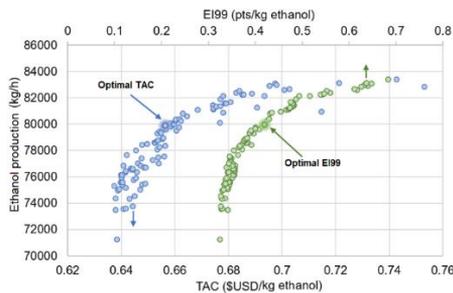


Figure 3. Pareto front for biomass - ethanol process (no separation sequence).

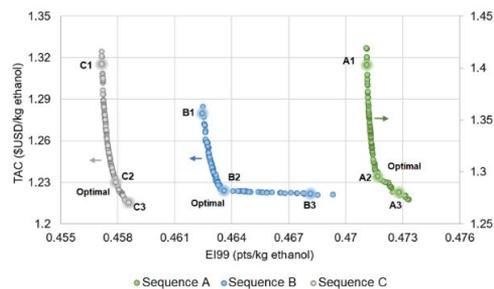


Figure 4. Pareto front for biomass - ethanol process (with separation sequence).

Figure 4 shows the Pareto front for optimal ethanol production with the separation sequence, according to Figure 3. For the three sequences, the high EI99 values result from the increasing steam required in the reboilers, while the high TAC values are related to large equipments and amounts of solvent. The point where these objective functions reach their minimum is discussed. It can be also observed that for the conventional sequence (A) a minimum TAC and EI99 of 1.295 USD/kg ethanol and 0.4716 ecopoints/kg ethanol were achieved, respectively. Taking the conventional extractive distillation as reference, the column sequence with vapor side stream (B) achieved reductions of 5.56% in the TAC and 1.72% in the ecoindicator while the dividing wall column sequence (C) achieved savings of 5.02% in the TAC and a decrease of 2.92% in the ecoindicator. It was also observed that the sequence B used less solvent than the sequence C. The amount of solvent contributed with about 75% to the TAC of the separation sequences.

5.3. Optimization of biojet process

Figure 5 shows the Pareto front ethanol-to-biojet process optimization. The high EI99 values are attributed to the greater environmental impact inherent of hydrocarbons in the process. Finally, the optimal design is capable of producing 224 206 ton/year (266 912 m³/year) of biojet fuel, which meets a 5.72% demand for conventional jet fuel in Mexico.

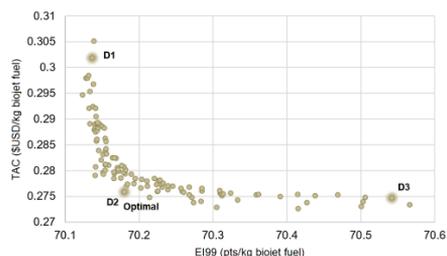


Figure 5. Pareto front for ethanol – biojet fuel process.

To have an idea of how profitable the design is, the minimum selling price of biojet was calculated. Thus, at the optimal point it was equal to 1.653 USD/liter, while the sale price of conventional jet fuel in Mexico in 2020 was 0.414 USD/liter, which indicates that the entire process is still not profitable.

6. Conclusions

The sugarcane bagasse was the most promising feedstock due to its low price and high sugar content. Also, to separate the produced ethanol the intensified scheme of vapor side stream provided the major savings and ecoindicator reductions respect to the conventional sequence, which was attributed to the minor use of solvent. However, despite the efforts for reducing total annual cost, the biojet fuel is not capable of competing with conventional jet fuel in Mexico. To achieve this, the intensification of reaction zone in biomass-to-ethanol process, the cofermentation of glucose and xylose, and the intensification of ethanol-biojet process is suggested. Besides, with the intensification of the second process, a reduction in ecoindicator values would be achieved.

References

- Conde-Mejía et al., 2012. A comparison of pretreatment methods for bioethanol production from lignocellulosic materials. *Process Saf. Environ. Prot.*, 90, 189-202.
- Finkbeiner, M., 2020. Modular Production with Bio-Based Resources in a Decentral Production Network. *Chem. Ing. Tech.*, 92, 2041-2045.
- Kiss, A. & Suszwalak, D., 2012. Enhanced bioethanol dehydration by extractive and azeotropic distillation in dividing-wall columns. *Sep. Purif. Technol.*, 86, 70-78.
- SENER, 2017. *Mapa de Ruta Tecnológica Bioturbosina*. [Online] Available at: https://www.gob.mx/cms/uploads/attachment/file/324219/MRT_Bioturbosina_Final.pdf
- Rivas-Interian, R.M., 2021. *Optimización de Sistemas Intensificados para la Producción de Bioturbosina por la Ruta de Alcoholes de Forma Sustentable*. [Online] Available at: <http://repositorio.ugto.mx/handle/20.500.12059/4806>
- Srinivas, M. & Rangaiiah, G., 2007. Differential Evolution with Tabu List for Global Optimization and Its Application to Phase Equilibrium and Parameter Estimation Problems. *Ind. Eng. Chem. Res.*, 46, 3410-3421.

Heat exchanger networks with different shell and tube configurations

José A. Caballero,^a Mauro A.S.S. Ravagnani,^b Leandro Pavao,^b Caliane B.B. Costa,^b Juan Javaloyes-Anton^a

^a*Institute of Chemical Process Engineering, University of Alicante, Carretera de S. Vicente s.n. 03690, Spain*

^b*Department of Chemical Engineering, Universidade Estadual de Maringá. Av Colombo, 5790 Maringá. Brazil*

Abstract

In a growing energy consumption world, energy efficiency has become mandatory. In this context, the design of heat exchanger networks (HEN) is of crucial importance, but most of the approaches to HEN design consider only shell and tube (S&T) heat exchangers with a perfect countercurrent heat transfer. However, the 1-2 heat exchanger, (one shell pass, and two tubes pass), is likely the most common in the chemical industry. In this work, we present a two steps sequential algorithm that allows the design of HEN to capture the main details of the heat exchangers (number of tubes pass, number of shells, logarithmic mean correction factors) that influence the cost estimation. The first stage is based on an extended transportation model. It uses the concept of temperature intervals and considers the possibility of heat transfer between the hot and cold streams inside those intervals. Then, it is possible the a priori calculation of the logarithmic mean temperature difference between matches, the efficiency factor, and the number of shells in series, maintaining the area linear in the model. The second step uses a superstructure with all the possible alternatives in which the heat exchangers predicted by the first stage model can exchange heat to design the final heat exchanger network.

Keywords: HEN, Energy Integration, Shell and Tube heat exchangers.

1. Introduction

Due to the importance of the efficient use of energy and its economic and environmental implications heat integration has consolidated as a topic of great interest in process engineering. However, the optimal design of a HEN is a very difficult problem for at least two reasons: the models are highly non-linear and non-convex, and the number of possible matches grows exponentially with the number of process streams. Obtaining a globally optimal solution is very difficult with the actual computational capacity. Furman and Sahinidis, (2001) showed that this is an NP-hard problem in the worst sense. Of remarkable importance was the appearance of pinch analysis (Flower and Linnhoff, 1980). Pinch analysis uses the first and second principles of thermodynamics to determine the minimum energy consumption without the necessity of specifying any HEN. Energy-related targets were extended with area heat transfer estimations that could lead to an a priori annualized cost estimation of the network with a 5-10% accuracy (Linnhoff and Ahmad, 1990). Finally, heuristics rules were developed to synthesize a feasible network with the minimum utility consumption and try to approach the minimum number of heat exchangers.

The increase in computer processing capacity motivated the development of automated strategies for HEN design. Particularly, mathematical programming models were a success in getting better solutions than heuristic approaches, at least for small to medium size problems. Initially, optimization methods followed the steps of pinch analysis: Minimum energy consumption, determination of the minimum number of matches, and final HEN design (Biegler et al., 1997)

A drawback of the sequential approaches is that decisions in a given stage cannot be modified in the following ones. To overcome that drawback, simultaneous approaches were proposed, in which the energy and the structure of the HEN are done in a single model. The more successful simultaneous model was developed by (Yee and Grossmann, 1990). The model known as SYNHEAT is a stage-wise approach that sacrifices some alternatives to maintain the model robust. Since its appearance a good number of improvements have been proposed, for example, to include isothermal streams, multiple utilities (Ponce-Ortega et al., 2009), non-isothermal mixing (Huang et al., 2012), algorithms for reaching the global optimum (Chang et al., 2020) or a large number of works using metaheuristic (Pavão et al., 2018). Despite its success, the major limitation of SYNHEAT is that the superstructure does not contain some heat exchanger arrangements that could eventually form part of the optimal solution.

Implicitly in most of the works on HEN, it is the assumption that all heat exchangers are shell and tube (S&T) heat exchangers with perfect countercurrent flow. While S&T are ubiquitous in the chemical process industry, most of them include multiple tube passes and eventually more than a single shell placed in series. The inclusion of detailed heat transfer parameters, pressure drops, and mechanical aspects like shell and tube bundle diameters, internal and external tube diameters, number of tubes, number of baffles, number of shells in series, tube arrangement, fluid locations, etc. are commonly not included, but they can have an important impact of the HEN performance and total cost. Some models that simultaneously design the HEN and the equipment details are, for example, (Ravagnani and Caballero, 2007) or the excellent work by Kazi et al., (2021) In this work, we present a two-stage sequential model inspired by the works by Nemet et al., (2018) and Caballero et al., (2021) that tries to capture the most important parameters of each heat exchanger into a HEN without performing a rigorous design of all the mechanical details to get a cost-optimal HEN. The first stage relies on the concept of temperature intervals and takes the form of an extended transportation model. With this approach is possible to calculate a priori the logarithmic mean temperature difference, its correction factor (F_t), and the number of shells in series for any combination of heat exchange between a set of intervals for a hot stream and another set of intervals for a cold stream. So, the area estimation is maintained linear in the model. In the second stage, we solve an NLP in which we postulate all the possible ways in which the matches predicted by the first model can exchange heat.

2. Methodology

The problem we are dealing with can be stated as follows: given is a set of hot and cold process streams for which the inlet and outlet temperatures, heat flows, and individual heat transfer coefficients are known and assumed constant. Given is also a set of hot and cold utilities for which the inlet and outlet temperatures; individual heat transfer coefficients; and unitary costs are also known. The objective is to design a practical heat exchanger network formed by 1-2 S&T heat exchangers with minimum total cost.

The model consists of two sequential stages. In the first one, we determine the streams that are exchanging heat, including utilities, as well as a good estimation of inlet and

outlet temperatures, heat exchanged, number of shells in series, logarithmic mean temperature, and its correction factor for each heat exchanger. In the second stage, using as starting values the results from the first one, we determine the structure of the HEN.

2.1. Stage 1: Extended transportation model overview.

A comprehensive description of the model cannot be included here due to space restrictions. Instead, the main characteristics of the model are presented in Figure 1.

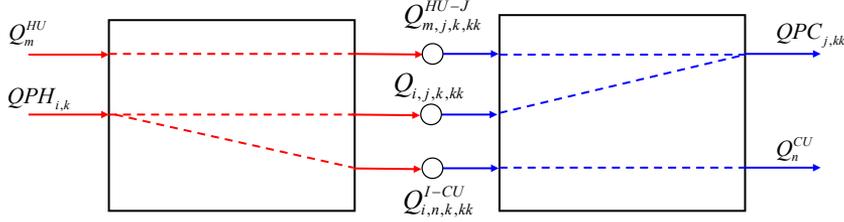


Figure 1. Representation of the transportation model. Different possibilities of heat transport between a hot stream i and a hot utility m with a cold stream j and a cold utility n between intervals k and kk .

Once the temperature intervals have been established, it is possible to calculate the amount of heat provided to each interval by all the hot process streams or demanded by the cold process streams. A hot stream can exchange ('transport') heat to any cold stream in intervals at lower temperatures ($k^{\text{hot}} \geq k^{\text{cold}}$). The transportation model is equivalent to performing energy balances in points 1, 2, 3, and 4 of Figure 1:

$$QPH_{i,k} = \sum_{\substack{kk \in INT \\ kk \geq k}} \sum_{j \in CK_{j,kk}} Q_{i,j,k,kk} + \sum_{\substack{kk \in INT \\ kk \geq k}} \sum_{n \in CUK_{n,kk}} Q_{i,n,k,kk}^{I-CU} \quad \forall (i,k) \in HK_{i,k} \quad (1)$$

$$QPC_{j,kk} = \sum_{\substack{k \in INT \\ kk \geq k}} \sum_{i \in HK_{i,k}} Q_{i,j,k,kk} + \sum_{\substack{k \in INT \\ kk \geq k}} \sum_{m \in HUK_{m,k}} Q_{m,j,k,kk}^{HU-J} \quad \forall (j,kk) \in CK_{j,kk} \quad (2)$$

$$Q_m^{HU} = \sum_{k \in HUK_{m,k}} \sum_{(j,kk) \in CK_{j,kk}} Q_{m,j,k,kk}^{HU-J} \quad \forall m \in HU \quad (3)$$

$$Q_n^{CU} = \sum_{kk \in CUK_{n,kk}} \sum_{\substack{(i,k) \in HK_{i,k} \\ kk \geq k}} Q_{i,n,k,kk}^{I-CU} \quad \forall n \in CU \quad (4)$$

The minimum energy cost can be calculated by minimizing the cost of utilities in previous equations. If we assume pure countercurrent heat exchangers, the calculation of the area is straightforward, because it can be calculated as the summation of the areas of heat exchanged by two streams in the different temperature intervals (Caballero et al., 2021; Nemet et al., 2018). However, the efficiency factor F_t and the Number of shells in series for a 1-2 S&T heat exchanger, depend on the inlet and outlet temperatures of the hot and cold streams. To deal with this problem we introduce a variable W_{i,j,k_1,k_2,k_3,k_4} that takes the value of 1 if the hot stream i exchanges heat starting in interval k_1 and finishing in interval k_2 with cold stream j in the intervals k_3 to k_4 . And zero otherwise. Dealing with a six-index binary variable makes the problems intractable except for small instances. Fortunately, it is not necessary to define the variable W as binary because it can be forced to take integer values using the binary variables $YH_{i,j,k}$ and $YC_{i,j,k}$. that take value 1 if hot

stream i /(cold stream j) is exchanging heat with cold stream j /(hot stream i) in interval k , using a set of logical relationships.

As we know the inlet and outlet temperatures of each one of the temperature intervals we can calculate a priori the logarithmic mean temperature ($LMTD_{k1,k2,k3,k4}^P$), its efficiency factor ($Ft_{k1,k2,k3,k4}^P$), and the number of shells in series ($NShells_{k1,k2,k3,k4}$) for each combination of feasible inlet and outlet temperature intervals, and each hot and cold stream. Now, the determination of those values for each heat exchange can be written as:

$$LMTD_{i,j} = \sum_{k1} \sum_{k2} \sum_{k3} \sum_{k4} LMTD_{k1,k2,k3,k4}^P W_{i,j,k1,k2,k3,k4} \tag{5}$$

$$Ft_{i,j} = \sum_{k1} \sum_{k2} \sum_{k3} \sum_{k4} Ft_{k1,k2,k3,k4}^P W_{i,j,k1,k2,k3,k4} \tag{6}$$

$$NShells_{i,j} = \sum_{k1} \sum_{k2} \sum_{k3} \sum_{k4} NShells_{k1,k2,k3,k4}^P W_{i,j,k1,k2,k3,k4} \tag{7}$$

Then the area of a given heat exchanger can be calculated as:

$$Area_{i,j} = \frac{1}{U_{i,j} LMTD_{i,j} Ft_{i,j}} \sum_{k \in INT} \sum_{kk \in INT} Q_{i,j,k,kk} \tag{8}$$

The cost of a heat exchanger is calculated using the following expression.

$$Cost_{i,j}^{HE} = NShells_{i,j} Cf_{i,j} + NShells_{i,j} Cv_{i,j} \left(\frac{Area_{i,j}}{NShells_{i,j}} \right)^\beta \tag{9}$$

To maintain the model linear the cost equation is substituted by a linear piecewise approximation.

2.2. Stage 2. Superstructure optimization.

The first stage calculates the matches between streams, but it does not provide information on the arrangement of streams (i.e. series, parallel, or any combination of them). See Figure 2. The resulting MINLP model is highly non-linear and non-convex, however, the initial values obtained from the first stage help to get an initial feasible solution from which to start the optimization. We cannot guarantee a globally optimal solution, but numerical tests have shown that usually high-quality solutions are obtained.

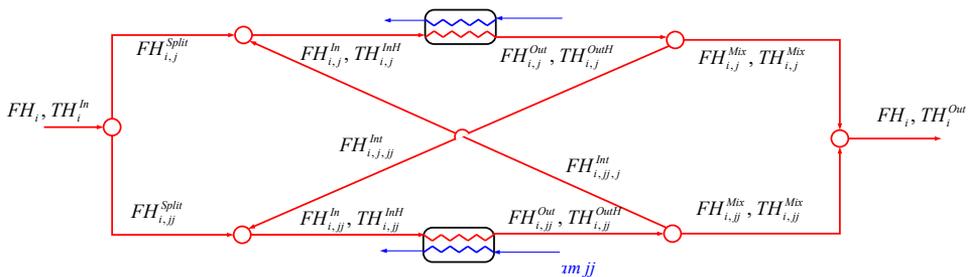


Figure 2. Scheme of the superstructure in which a hot stream i can exchange heat with two cold streams j and jj . The superstructure can be extended to cold streams and any number of heat exchanges.

3. Example

To illustrate the model, we present the results of a well-known benchmark problem modified to deal with 1-2 S&T heat exchangers. The problem was proposed by Linnhoff and Ahmad, (1990) and is one of the most used benchmark problems in the HEN synthesis literature. Being a medium-scale problem, it has five cold and four hot streams, a hot oil as hot utility and water as a cold utility. All relevant data are presented in Table 1.

Table 1. Data for the example.

Stream id.	Inlet T (°C)	Outlet T (°C)	CP (kW/°C)	h (kW/m ² K)	Cost (\$/kW _Y)
H1	327	40	100	0.5	
H2	220	160	160	0.4	
H3	220	60	60	0.14	
H4	160	45	400	0.3	
C1	100	300	100	0.35	
C2	35	164	70	0.7	
C3	85	138	350	0.5	
C4	60	170	60	0.14	
C5	140	300	200	0.6	
HU1	330	250	-	0.5	60
CU1	15	30	-	0.5	6

HE cost = NShells [Cf + Cv(Area/NShells) ^β]	Cf	Cv	β
	2000	70	1.0

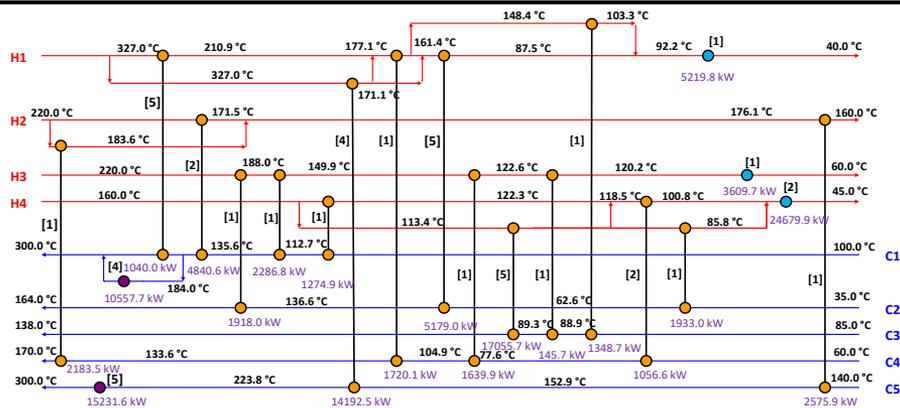


Figure 3. Heat exchanger network obtained. Square brackets indicate the number of Shells.

The models were implemented in GAMS (GAMS Development Corporation, 2021) using CPLEX as MILP solver and BARON as MINLP solver on a PC machine working under windows (i7 2.90 GHz, 32.0 GB). The MILP problem was solved to global optimality within a 5% optimality gap. In the case of MINLP problems BARON is not able to close the optimality gap, however, it can find very good solutions in the first minutes of execution. We stopped the BARON execution after 3600 seconds of CPU time. The

optimal solution has a total annualized cost of 3085 k\$/year. 1547.4 k\$/year in hot utilities; 201.1 k\$/year in cold utilities and 1336 k\$/year in an annualized cost of capital expenditures. Figure 3 shows the final configuration of the HEN and the main results obtained after the optimization.

Conclusions

In this work, we have presented a model that explicitly includes the most relevant parameters to accurately estimate the cost of a HEN that includes 1-2 S&T heat exchangers. We use a two-stage algorithm, in the first stage, using the concept of temperature intervals, it is possible the a priori calculation of the logarithmic mean temperature difference between matches, the efficiency factor, and the number of shells in series. The second step uses a superstructure with all the possible alternatives in which the heat exchangers predicted by the first stage model can exchange heat to design the final heat exchanger network.

Acknowledgments

The authors acknowledge support from the «Generalitat Valenciana» project PROMETEO 2020/064 and from the “Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq-Brazil, projects 428650/2018-0 and 311807/2018-6.”

References

- Biegler, T.L., Grossmann, I.E., Westerberg, A.W., 1997. *Systematic Methods of Chemical Process Design*. Prentice Hall PTR., Upper Saddle River, NJ. USA.
- Caballero, J.A., Pavão, L. V., Costa, C.B.B., Ravagnani, M.A.S.S., 2021. A Novel Sequential Approach for the Design of Heat Exchanger Networks. *Front. Chem. Eng.* 3, 1–13.
- Chang, C., Peccini, A., Wang, Y., Costa, A.L.H., Bagajewicz, M.J., 2020. Globally optimal synthesis of heat exchanger networks. Part I: Minimal networks. *AIChE J.* 66, 1–15.
- Flower, J.R., Linnhoff, B., 1980. A thermodynamic-combinatorial approach to the design of optimum heat exchanger networks. *AIChE J.* 26, 1–9. <https://doi.org/10.1002/aic.690260102>
- Furman, K.C., Sahinidis, N. V., 2001. Computational complexity of heat exchanger network synthesis. *Comput. Chem. Eng.* 25, 1371–1390.
- GAMS Development Corporation, 2021. *General Algebraic Modeling System (GAMS) Release 36.1.0*. Fairfax, VA. USA.
- Huang, K.F., Al-mutairi, E.M., Karimi, I.A., 2012. Heat exchanger network synthesis using a stagewise superstructure with non-isothermal mixing. *Chem. Eng. Sci.* 73, 30–43.
- Kazi, S.R., Short, M., Isafiade, A.J., Biegler, L.T., 2021. Heat exchanger network synthesis with detailed exchanger designs—2. Hybrid optimization strategy for synthesis of heat exchanger networks. *AIChE J.* 67, 1–16.
- Linnhoff, B., Ahmad, S., 1990. Cost optimum heat exchanger networks—1. Minimum energy and capital using simple models for capital cost. *Comput. Chem. Eng.* 14, 729–750.
- Nemet, A., Isafiade, A.J., Klemeš, J.J., Kravanja, Z., 2018. Two-step MILP/MINLP approach for the synthesis of large-scale HENs. *Chem. Eng. Sci.*
- Pavão, L. V., Costa, C.B.B., Ravagnani, M.A.S.S., 2018. A new stage-wise superstructure for heat exchanger network synthesis considering substages, sub-splits and cross flows. *Appl. Therm. Eng.* 143, 719–735.
- Ponce-Ortega, J.M., Serna-González, M., Jiménez-Gutiérrez, A., 2009. Synthesis of heat exchanger networks with optimal placement of utilities. *AIChE Annu. Meet. Conf. Proc.* 2849–2856.
- Ravagnani, M.A.S.S., Caballero, J.A., 2007. Optimal heat exchanger network synthesis with the detailed heat transfer equipment design. *Comput. Chem. Eng.* 31, 1432–1448.
- Yee, T.F., Grossmann, I.E., 1990. Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis. *Comput. Chem. Eng.* 14, 1165–1184.

Considering Environmental and Social Impact in Optimization of a Supply Chain for Bio-Jet Fuel Production

David Vallejo-Blancas,^a Juan Gabriel Segovia-Hernández,^a Eduardo Sánchez-Ramírez,^a Gabriel Contreras-Zarazúa,^{a,b}was Juan José Quiroz-Ramírez^b

^a*Department of Chemical Engineering University of Guanajuato, Noria Alta S/N, Noria Alta, 36000, Guanajuato, Gto., Mexico.*

^b*CONACyT-CIATEC A.C Center of Applied Innovation in Competitive Technologies, Omega 201, Industrial Delta, 37545 León, Gto., Mexico.*

Abstract

The growing contribution of aviation industry to climate change has led to the search of new energy sources in the short to medium term. Biofuels appear to be a potential alternative to the fossil fuels, also they are a promising solution to mitigate climate change. Nowadays, the agricultural residues (AR) are one of the most promising feedstocks to produce biofuels due to their high availability. However, their use at industrial scale requires overcoming some challenges such as their seasonal availability and their dispersed geographic distribution. In order to face these challenges, in this work has been proposed a multi-period supply chain (SC) to produce bio-jet fuel in Mexico through the alcohol to jet (ATJ) process using corn stover and sugarcane bagasse as raw materials. The supply chain considered implementing a centralized or decentralized configuration as a result of an optimization based on sustainability criteria, such as economic, environmental and social aspects, which are net profit, eco-indicator 99 (EI99) and jobs generation, respectively. The results show that it is possible to cover 5.5% of jet fuel consumption in Mexico, also an economic incentive of \$115 million is required for the process to operate at break-even with an environmental impact of 7.48 billion ecopoints per year and 5,594 jobs generated.

Key words: Supply chain, Bio-jet fuel, Agricultural residues, Biorefinery.

1. Introduction

Air transport is a cornerstone in current globalization, it supports about 4 billion passengers yearly, around 61 Mt of freight, which represents US\$3.5 trillion in world economic activity, before the SARS-CoV-2. In the specific case of Mexico, the aviation industry represents 3.5% of country's GDP, also in 2018 the air transport mobilized around 98 million passengers generating 1.4 million jobs. The International Air Transport Association (IATA) estimates that the air transport will double in the next 20 years, representing a considerable environmental impact, as global jet fuel consumption generates around 2-3% of carbon dioxide (CO₂) emissions due to human activity. As a result, IATA has proposed to reduce CO₂ emissions from air transport up to 50% by 2050, using 2005 as a baseline (Domínguez-García et al., 2017). In this sense, Bio-jet fuel represents a promising alternative to conventional jet fuel since its properties are very similar to those of the latter, which would allow taking advantage of the existing storage

and distribution infrastructure, also they are even currently used in aircraft up to 50% mixed with conventional jet fuel according to the norm ASTM D7566-11.

The ATJ process appears to be a promising route to produce bio-jet fuel as the technologies involved are widely known in the fuel and chemical industry, this process transforms alcohol into jet fuel through several steps such as dehydration, oligomerization, and hydrogenation (Wang and Tao et al., 2016). Moreover, bioethanol is currently the most produced biofuel worldwide and cellulosic biomass, such as AR have gained great attention as a promising raw material in biofuel production due to its wide availability, do not compete with food needs and are low cost. Typically, one part of AR are used as fodder for livestock or to protect the soil from erosion while the other is burned, generating even more CO₂ emissions than air transport. However, the high cost of transporting and storing biomass due to its heterogeneous geographic and temporal distribution, as well as its low energy density and high moisture content have greatly limited the incursion of biomass in large-scale biofuel production.

Therefore, to overcome the inherent problems of using biomass, in this work was proposed a SC to cover between 5.5-50% of jet fuel consumption in Mexico using the most common lignocellulosic residues of the country. The SC considers the production of bio-jet fuel from the ATJ process considering the production of ethanol from corn stover and sugarcane bagasse available in the Mexican field. As well as sustainability criteria such as the economic, environmental, and social aspects via the maximization of net profit, the minimization of EI99 and the evaluation of job generated by the SC, respectively.

2. Methodology

Considering that in Mexico around 586 Mt of AR are generated per year, this work evaluated the potential of two of the most abundant residues to be used as raw materials to produce jet-fuel. These residues are corn and sugarcane bagasse. A centralized and a decentralized configuration were proposed for the design of the SC (see Figure 1). Firstly, the biomass is sent directly from the harvest sites to the biorefineries for processing, while in the second biomass is densified in pretreatment depots to reduce the transportation costs prior to final processing in a biorefinery.

To consider the seasonal availability of AR, a multi-period inventory planning with an one year of time horizon was considered, which was divided into twelve time periods ($t \in T$) one for each month of the year. On the other hand, the geographic distribution of AR was considered through the production of corn and sugarcane in the 32 states of the Mexican Republic, taking a yield for residues of 110% and 34% per ton for grain corn and sugarcane, respectively (Muro Reyes et al., 2013). The 32 states of Mexico were considered as regions, in which the AR are available, also it is considered that in the centroid of each region is located all the biomass of the region, at the same time in these regions lies the potential location for pretreatment depots and biorefineries at 10 and 20 km from the centroid, respectively.

The Rhumb line method was employed to compute the distances between centroids, pretreatment depots, biorefineries and markets, while the latitude and longitude required by the method were obtained from Google Maps.

The supply chain design is considering the bio-jet fuel production by the ATJ process, in which biomass is first converted into ethanol and then into jet fuel. This process can be divided into four main stages: the pretreatment stage using mineral acids stage, which release the sugar contained in the biomass, the second stage includes the sugars fermentation and ethanol purification, the third stage includes the ethanol dehydration to

ethylene and finally the last stage consists of the oligomerization and hydrogenation of ethylene into bio-jet fuel, gasoline and diesel. The date for ATJ process was taken from (Rivas) This supply chain considers that this four stages can exist in a single facility called biorefinery, but at the same time it considers that the pretreatment stages can be located in a different place called preterm facility. The aim of pretreatment facilities is to convert biomass from nearby harvesting centers to sugars which are more energy dense products to reduce the transportation costs. Therefore, a decentralized scheme could be a more efficient scheme than traditional biorefinery scheme. The bio-jet fuel, gasoline, diesel as well as the sales of ethanol as an oxygenate for gasoline haven been considered the main products. The current markets of these products are also considered the potential markets of this supply chain, owing to they have existing infrastructure for oil derived products. Figure 1 shows the superstructure considered for this supply chain.

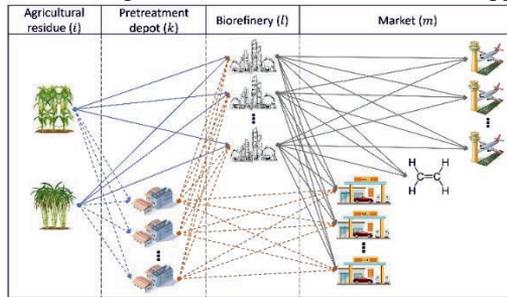


Figure 1. Supply chain superstructure

In this work the maximization of the net profit and minimization of eco Indicator 99 (EI99) have been considered as objective functions to evaluate the performance of the supply chain solutions. The net profit was chosen to determine the economic feasibility of the supply chain, whereas the EI99 evaluates the environmental impact of supply chain. Although there are other newer methodologies for assessing environmental impact, the EI99 is easy to implement and the data for this indicator are easy to obtain. A cradle-to-grave assessment with hierarchical perspective for EI99 was considered, this perspective offers an offset between the effects at short and long term effects. Additionally, other different parameter are presents and evaluated in order to have a broader and more sustainable perspective for the different supply chain solutions, this additional metrics are the CO₂ emissions and the jobs generated by the supply chain. The Jobs and Economic Development Impact (JEDI) method was used to calculate the jobs generated.

3. Mathematical model

The mathematical model for the supply chain consists of a multiperiod Mixed Integer Lineal Programming (MILP) model, which was implemented in the software Gams in order to solve it.

The model considers the biomass storage at the harvest sites, which was modeled by a mass balance as shown in the following equation:

$$A_{j,i,t}^{RM} = A_{j,i,t-1}^{RM} (1 - \gamma_{i,t}^{RM}) + C_{j,i,t}^{RM} - \sum_l BER_{j,i,l,t}^{RM} - \sum_k BEDP_{j,i,k,t}^{RM} \quad i \in I^{RM}, j, t \quad (1)$$

Where $A_{j,i,t}^{RM}$ and $A_{j,i,t-1}^{RM}$ represent the inventory level of biomass at a harvest site for a specific time and a previous period, respectively. $\gamma_{i,t}^{RM}$ represents the loss factor coefficient by AR degradation, while $BER_{j,i,l,t}^{RM}$ and $BEDP_{j,i,k,t}^{RM}$ are the shipment of AR i from the

harvest site j to the biorefinery l and to pretreatment depot k , respectively. Finally, $C_{j,i,t}^{RM}$ is the AR consumption at the harvest site, which is restricted as follows:

$$C_{j,i,t}^{RM} \leq \varphi_{j,i,t}^{RM} \quad i \in I^{RM}, j, t \quad (2)$$

Where $\varphi_{j,i,t}^{RM}$ is the availability of AR i in harvest site j in a period period t . The shipments of biomass to biorefineries are constrained as follows:

$$BER_{j,i,l,t}^{RM} \leq DBCC_{j,i,t}^{RM} \cdot y1_{j,i,l}^{RM} \quad i \in I^{RM}, j, l, t \quad (3)$$

Here $DBCC_{j,i,t}^{RM}$ is the sum of the AR stored level and the available AR at harvest site j at time t and $y1_{j,i,l}^{RM}$ is the binary variable to enable the shipment of AR across the arc $j \rightarrow l$. However, the product $DBCC_{j,i,t}^{RM} \cdot y1_{j,i,l}^{RM}$ represents a nonlinearity that was reformulated according to the following mathematical arrangement proposed by Floudas (1995):

$$BER_{j,i,l,t}^{RM} \leq h1_{j,i,l,t}^{RM} \quad i \in I^{RM}, j, l, t \quad (4)$$

$$DBCC_{j,i,t}^{RM} - \sum_t \varphi_{j,i,t}^{RM} (1 - y1_{j,i,l}^{RM}) \leq h1_{j,i,l,t}^{RM} \quad i \in I^{RM}, j, l, t \quad (5)$$

$$h1_{j,i,l,t}^{RM} \leq DBCC_{j,i,l,t}^{RM} \quad i \in I^{RM}, j, l, t \quad (6)$$

$$h1_{j,i,l,t}^{RM} \leq \sum_t \varphi_{j,i,t}^{RM} \cdot y1_{j,i,l}^{RM} \quad i \in I^{RM}, j, l, t \quad (7)$$

Where $h1_{j,i,l,t}^{RM}$ is the bilinear product to maintain the linearity of the model. The mass balance at biorefineries and pretreatment depots was carried out in a similar way. In order to guarantee a constant supply a cyclic inventory planning was considered, which consists of that the storage level in the last period ($t=T$) is equal at the beginning period ($t=0$), according to the following equation for AR:

$$A_{j,i,t=0}^{RM} = A_{j,i,t=T}^{RM} \quad i \in I^{RM}, j \quad (8)$$

Finally, the objective functions were formulated as follows

$$\max Z = \left(\overbrace{CAR + TC + PC}^{Netprofit} + \overbrace{EIAR + EIT + EIP}^{EJ99} \right)$$

Where CAR , TC , PC are the costs by biomass use, transportation cost and processing cost, respectively. On the order hand, $EIARE$, EIT and EIP represents the environmental impact by biomass uses, transportation and processing, respectively. This multi-objective optimization problem was solved using the ε -constrained method to find a tradeoff between the objectives.

4. Results

The model proposed consists of 27,264 binary variables, 672,013 continuous variables and 1,313,649 constraints. It was solved in an average time of 51,173s using a computer with a 3.00GHz i5-8500 processor with 4GB of RAM using CPLEX method with a relative gap of 1%. Figure 2 shows the Pareto front of SC optimization. In all cases, a 5.5% demand for aviation fuel is covered, as this represents the lowest environmental impact and the lowest level of subsidy required or the highest netprofit. Therefore, the difference between the upper and lower bound in the Pareto front corresponds to the AR transporting activities in each SC design. The upper region the SC design considered two biorefineries while in the lower region only one biorefinery is located, as shown in Figure 3.

Since Jalisco and Veracruz are the main producers of sugarcane, which is the most used biomass in the ATJ process, the biorefineries are installed near the harvest sites to reduce the transportation cost and the environmental impact. because the supply chain is not cost-effective a decentralized scheme is not selected for the model in order to reduce the transportation costs.

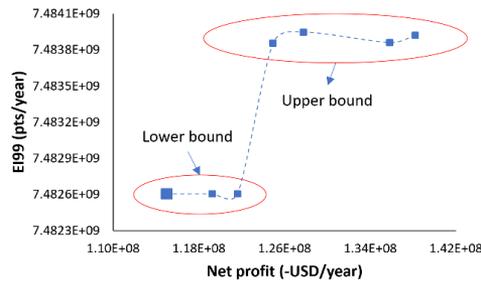


Figure 2. Pareto front for bio-jet fuel SC

It is important to mention, that for economically feasible supply chain, this decentralized scheme has proved reduce the cost and environmental impact of the process. Owing to the transportation activities increase in the upper region the environmental impact increase too, thus the benefit of the economy of scale is reduced because these two have a lower processing capacity than the biorefinery in the lower bound, which explains the behavior of the Pareto front. Moreover, the centralized configuration was preferred in the resulting SC designs because AR pretreatment and ethanol production, operations considered in the pretreatment depots, are the main drawbacks in the ATJ process, these steps represent 59% of the entire SC cost (see Figure 4a).



Figure 3. (Left) SC design lower bound (Right) SC design in upper bound.

The big dot in the lower region corresponds to the best SC solution. This solution considers a centralized configuration where only one biorefinery is installed. This solution has 5,594 jobs generated and the EI99 is around 7.48 billion points per year. Additionally, it requires a subsidy of US\$115 million to be profitable. On the other hand, the CO₂ emissions per kilogram of product commercialized as well as the energy return on investment (EROI) are 2.38 kgCO₂/kg and 1.14, respectively.

Figure 4b shows the cyclic inventory of AR in the harvest sites, which shows that the AR available during the first half of the year is stored at the harvest sites for later processing during the months when it is not available, and storage finished products in the biorefinery. Finally, Figure 6 shows the flow of bio-jet fuel supplied in each month of the year continuously and stably, unlike the availability of feedstock that shows constant fluctuations, which do not have an impact on the production of the SC.

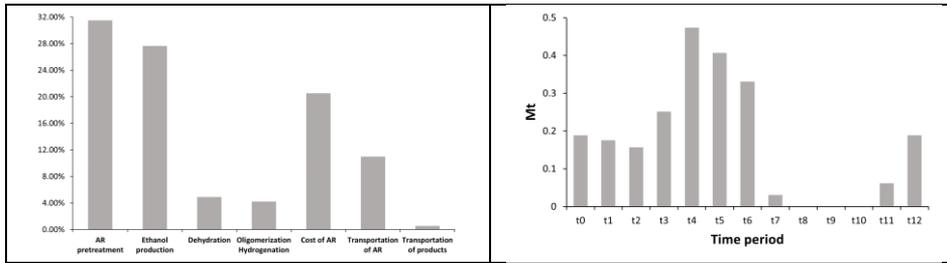


Figure 4. a) SC cost distribution, b) Cyclic inventory of AR

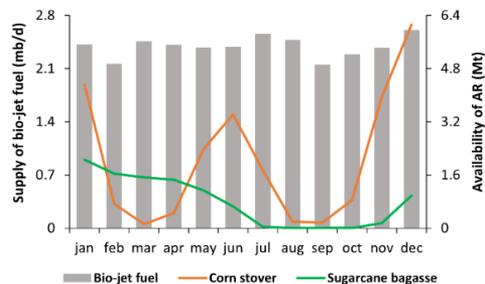


Figure 5. Monthly supply of bio-jet fuel and feedstock availability

5. Conclusions

Agricultural residues are a promising source of raw materials by biofuel production and through the design and optimizing of the SC associated with biofuel production is possible to overcome its geographical and stational dispersion. The proposed model had the purpose of identifying the areas of opportunity in the production of bio-jet fuel from the ATJ process as well as the level of coverage that is possible to achieve with the materials available in Mexican territory today. Based on the results obtained, it has been determined that the main process disadvantages are the first two stages of the process (pretreatment and ethanol production), although despite these disadvantages, the process has an EROI positive, so the improvement of the process will not only bring economic advantages, but also energy and environmental advantages. On the other hand, in this case a centralized configuration was preferred due to the processes carried out in the pretreatment depots are the costliest, and its separation from the other process stages result in a reduction of the effect of the economy of scale, which the transportation cost reduction could not compensate.

References

- Domínguez-García, S., Gutiérrez-Antonio, C., De Lira-Flores, J. A., & Ponce-Ortega, J. M. (2017). Optimal planning for the supply chain of biofuels for aviation in Mexico. *Clean Technologies and Environmental Policy*, 19(5), 1387-1402.
- Floudas, C. A. (1995). *Nonlinear and mixed-integer optimization: fundamentals and applications*. Oxford University Press.
- Muro Reyes, L., Camacho Villa, T. C., & Guevara-Hernández, F. (2013). Rastrojos: manejo, uso y mercado en el centro y sur de México.
- Wang, W. C., & Tao, L. (2016). Bio-jet fuel conversion technologies. *Renewable and Sustainable Energy Reviews*, 53, 801-822.

Superstructure Optimization of Dimethyl Ether Process

Emrullah Erturk^a, Erdal Aydın^{b,c}, Sahin Sarrafi^d, Ozgun Deliismail^d, Aysel Zahidova^d, Hasan Sildir^{a,*}

^a*Department of Chemical Engineering, Gebze Technical University, Kocaeli 41400, Turkey*

^b*Department of Chemical and Biological Engineering, Koç University, Istanbul 34457, Turkey*

^c*Koç University TUPRAS Energy Center (KUTEM), Koç University, Istanbul 34450, Turkey*

^d*Socar Turkey, R&D and Innovation Co., Aliaga, Izmir, 35800, Turkey*
hasansildir@gtu.edu.tr

Abstract

Integration of process flowsheet simulators and optimization algorithms is a prominent approach to address simultaneous design and optimization of processes, which is represented by a mixed integer nonlinear programming (MINLP) formulation. In this study, DWSIM, a free and rarely used simulator, is used as a black-box function for the evaluation in genetic algorithm in MATLAB. Proposed approach is implemented to a dimethyl ether process, calculating optimum processing conditions in addition to structural decision variables including the feedstock type, reactor, and separation unit selections. Results show that syngas has the major impact on the process economics and is significantly more economical feedstock although high number of additional processing units are required.

Keywords: superstructure optimization; mixed integer nonlinear programming; process synthesis; dimethyl ether process; black-box optimization; DWSIM.

1. Introduction

Chemical processes include complex and integrated pathways combined by several unit operations in which conversion occurs from feedstocks to valuable products. Such a conversion, in general, can be obtained through high number of alternative paths, which hinders the process architecture and operating condition selection under economic considerations and tightly ensured process constraints. Thus, a superstructure optimization problem formulation addressing simultaneously the aforementioned issues has become an important research area over the past decades.

A mixed-integer nonlinear problem (MINLP) is formulated, in general, to obtain a smaller architecture from a larger superstructure that is predefined and includes all alternative units, flows and many other considerations. The resulting MINLP formulation includes integer variables to account for sequences of events, alternative candidates, and the existence of units, whereas the continuous variables represent states. The MINLP problem is flexible and can include various user-defined considerations from economic and environmental aspects (Edgar *et al.*, 2001).

Significant developments in theory and computation have provided substantial capabilities to solve various complex optimization problems over the past decades (Chen and Grossmann, 2017). With many alternatives and modifications, MINLP solvers can be classified roughly into two major groups as rigorous, and heuristic based on solution approach. Rigorous solvers utilize all mathematical tools to exploit optimality through the explicit formulation of all superstructure elements and might ensure global optimality for many cases. As an alternative, black-box or evolutionary algorithms benefit from the patterns from successive function evaluations, rather than focusing on derivative-based iterations, to provide optimality. Evolutionary algorithms, although suffering from significant computational load and local optimality issues, might deliver practical and satisfactory solutions to complex MINLPs once the associated plant models are challenging to derive and requires a priori knowledge due to numerous unknown driving forces and their nonlinearly interacting nature (Mencarelli *et al.*, 2020). In contrast, despite significant computational developments and theoretical advancements, rigorous solvers have scalability problems for large superstructures. In addition, integration of those to process flowsheet simulators which have widespread industrial use is a challenging task, hindering rigorous solver implementations. A significant fraction of superstructure optimization studies focus on commercial software (Ploskas and Sahinidis, 2021). On the other hand, some free alternatives, including DWSIM, might provide similar performance to commercial ones (Tangsrivong *et al.*, 2020). Teerapat and Amata used a DWSIM simulation environment to validate an optimized DME process (Laiwatthanaphaisarn and Anantpinijwatna, 2018).

This study focuses on the simultaneous design and optimization of a comprehensive DME process under several scenarios including different feedstock and unit operation prices using a free process flowsheet simulator, DWSIM, which is automated through Python. The superstructure involves different feedstocks which are processed through different reactors and separation units to be determined by the optimization problem in addition to some operating conditions. A genetic algorithm is used for the evolutionary solution of MINLP problem, which accounts for the existence of the units and plant operating conditions. The case studies show different process configurations are required for profitability based on feedstock and product prices.

2. Superstructure of DME Production Process

Dimethyl ether is one of the prominent candidates for relatively clean and sustainable alternatives to fossil fuels. The increased demand for DME can be met from synthesis gas or methanol from biological waste (Saebea *et al.*, 2019).

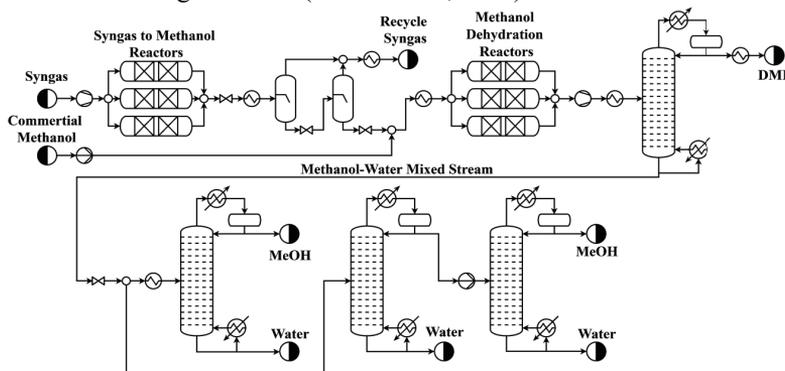


Figure 1. Superstructure of DME production process.

The proposed superstructure in this study is presented in Fig. 1. DME production might be obtained with dehydration of methanol and from various feedstocks, primarily from syngas, as an emerging trend to favor carbon management and biomass utilization. The conditions specified in the kinetic study of Nestler *et al.* (2020) were used for modelling methanol synthesis at three different catalyst density yields high to low conversion (Nestler *et al.*, 2020). Sequential flash units separate syngas from methanol with minimal loss at different pressures. Next, DME is produced with preheated methanol in catalytic dehydration reactor with three conversion preferences. Pressurized and cooled reactor effluent consists of mainly water, methanol and DME, which is obtained using 30 bar distillation column at high purity. Next, unreacted methanol and side product water recovered with two different distillation column arrangements. Side products are marketable/recyclable as well.

3. MINLP Formulation and Solution Method

Superstructure of DME process optimized with MINLP problem formulated as:

$$\begin{aligned}
 & \min_{m,y,T,P} && IC + (OC - S \cdot (1 - f)) \cdot t_f \\
 \text{s. t.} &&& y_1 + y_2 = 1 \\
 &&& y_3 + y_4 + y_5 = 1 \\
 &&& y_6 + y_7 + y_8 = 1 \\
 &&& y_9 + y_{10} = 1 \\
 &&& m_6 = y_6 \cdot (m_1 y_1 + m_{F2} y_2) \\
 &&& m_7 = y_7 \cdot (m_1 y_1 + m_{F2} y_2) \\
 &&& m_8 = y_8 \cdot (m_1 y_1 + m_{F2} y_2) \\
 &&& m_9 = y_9 \cdot m_{T101,B} \\
 &&& m_{10} = y_{10} \cdot m_{T101,B} \\
 &&& y_2 \cdot y_3 \cdot 6 \leq m_3 \leq y_2 \cdot y_3 \cdot 10.5 \\
 &&& y_2 \cdot y_4 \cdot 6 \leq m_4 \leq y_2 \cdot y_4 \cdot 12.5 \\
 &&& y_2 \cdot y_5 \cdot 6 \leq m_5 \leq y_2 \cdot y_5 \cdot 15.5 \\
 &&& y_1 \cdot m_{1,min} \leq m_1 \leq y_1 \cdot m_{1,max} \\
 &&& y_1 \cdot m_{2,min} \leq m_2 \leq y_2 \cdot m_{2,max} \\
 &&& 301 \leq T_1 \leq 308 \\
 &&& 360 \leq T_2 \leq 380 \\
 &&& 46 \leq P_1 \leq 50 \\
 &&& y_j \in \{0, 1\} \quad j = 1 \dots 10
 \end{aligned} \tag{1}$$

where t_f is the project lifetime which is 20 years for this case; f is the tax rate; y_1 and y_2 are binary variables for commercial methanol and syngas feedstock selection, respectively; y_3 , y_4 , and y_5 represent the selection of syngas to methanol reactor with corresponding mass flow rates m_3 , m_4 and m_5 ; y_6 , y_7 , and y_8 methanol dehydration reactor selection variables; m_{F2} is the flash unit effluent; $m_{T101,B}$ is the bottom product of DME distillation column; y_9 and y_{10} are one column and sequential two-column distillation route selection binaries, respectively; S is the income from the products. Equality constraints which include binary variables solely ensure selection of a single path among many. Commercial methanol and syngas stream mass flow rates are represented with m_1 and m_2 . High to low DME reactor inlet stream mass flow rates are m_6 , m_7 and m_8 , respectively. m_9 is 30-stage column and m_{10} is sequential two 15-stage columns mass flow

rates. $m_{F102,B}$ is second flash unit bottom product stream that contains mainly produced methanol, mass flow rate. All produced or purchased methanol fed to DME reactor and this feed stream, $m_{R2,in}$, expressed with fifth equality constraint. $m_{T101,B}$ is bottom product mass flow rate of DME recovery column which fed to selected separation route. T_1 , P_1 are temperature and pressure of flash unit, and T_2 DME recovery column feed stream temperature, respectively. Inequality constraints limit syngas feed mass flow for prevent exceeding equipment sizes which assumed fixed for practical calculation of capital cost investments. The capital cost investment, IC , the operating cost, OC , is calculated from:

$$IC = F_{LANG} \cdot \left(\sum_i C_{BM,i} \cdot y_i + \sum_j C_{BM,fixed,j} \right) \quad (2)$$

$$OC = m_1 \cdot y_1 \cdot C_{MeOH} + m_2 \cdot y_2 \cdot C_{Syngas} + \sum_j Q_j C_{utility,j} \quad (3)$$

where F_{LANG} is the correction for the liquid phase operations; $C_{BM,i}$ is the bare module cost of i^{th} equipment which is related to the y_i ; $C_{BM,j}$ is the bare module cost of j^{th} equipment which exist in the structure independent of the optimization problem; C_{MeOH} and C_{Syngas} are methanol and syngas prices, respectively; Q_j and $C_{utility,j}$ is the quantity and the unit cost of j^{th} utility. Equipment purchases and operational cost calculations performed using (Bailie et al., 2018) and updated to current prices using CEPCI. The distillation column fixed cost provided from ChemSep. Process equipment sizing considers the dimensions that could meet the highest production capacity. For this reason, even if there is a high investment cost in low-capacity production preferences, meeting the upper limit production capacity is ensured.

The MINLP problem in Eq. 1 is solved using Genetic Algorithm (GA) which treats the process flow sheet simulation as a black-box function which is evaluated in DWSIM. Python has a key role in the implementation of the MINLP problem as it provides the communication between DWSIM and MATLAB's GA. A simplified information flow diagram is shown in Fig. 2.

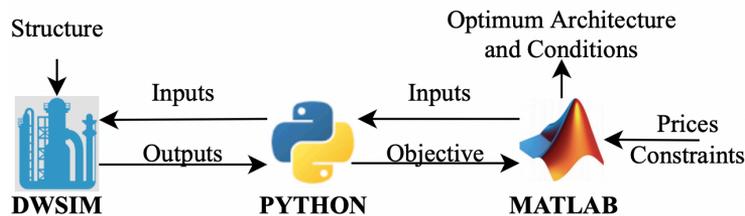


Figure 2. Superstructure optimization information flow diagram

4. Results

MINLP problem in Eq. 1 has been solved using MATLAB R2021a and DWSIM v6 on i5 9400 CPU 8 GB RAM Windows 10 x64 PC. In our case, the commercial methanol feedstock route results in an unprofitable process due to high purchase costs. The non-zero results of the solution are presented in Table 1.

In our case, syngas is selected as the primary raw material thanks to its low cost despite subsequent unit operations and associated operating costs. The resulting architecture proposes the highest conversion reactor for methanol production to favor the DME production and methanol which can be sold as a side product in our MINLP formulation.

Table 1. Non-zero process variables and economic measures of optimal results of Eq. 1

Plant variables									Economic Measures [M\$]		
m_2 [kg/s]	m_4 [kg/s]	m_{10} [kg/s]	$m_{r2,in}$ [kg/s]	$m_{F102,B}$ [kg/s]	$m_{T101,B}$ [kg/s]	T_1 [K]	T_2 [K]	P_1 [bar]	IC	OC	S
10.5	6.4	3.2	6.4	6.4	3.2	301	380	50	92	16	86

Temperature and pressure of flash unit maximize methanol yield while removing unreacted syngas. Produced methanol fed to medium conversion dehydration reactor to balance the DME conversion and related *IC* and *OC*. Sequential two 15-stage columns ensure the removal of high latent heat components while benefiting from the installation of large and single distillation column operations with high costs. The temperature of the feed stream as well as the temperature of the reboiler and condenser affect the efficiency of the separation process. However, feed stream temperature manipulation allows us to tune DME recovery column performance and operational costs. Overall process operation is performed at a maximum raw material rate within the defined operating window.

5. Conclusion

Optimum process design and synthesis under economic and environmental consideration is one of the contemporary challenges of the process systems engineering field. The idea of combining the advantages of process flowsheet simulators to obtain predictions without undergoing a significant modelling approach is useful once it is used in the black-box optimization algorithms which do not need first principal expressions explicitly. Integration of open source DWSIM simulation environment with black-box optimization solver enables us to simultaneously synthesis of process and optimization of conditions, unlike most studies which focus on commercially well-known and widely used software packages on a popular process for the architecture and processing condition calculation which accounts for feedstock selection, unit operation selection, catalysis selection, and related operating regimes. Such a sophisticated integration of high number of process equipment alternatives and feedstocks results in a nonconvex and significantly nonlinear mixed-integer nonlinear optimization problem. As a result, the global optimality of the proposed structure is beyond the scope.

A major decision variable on the plant profitability is determined by the feedstock, which is economically syngas over a wide price spectrum under current economic specifications. Furthermore, the proposed process flow diagram is comprehensive as it enables the selection of different-size unit operations to address the capacity, installation, and operating costs to deliver a realistic evaluation to some extent. Note that, the actual process is more complex than Fig. 1 and requires a more advanced MINLP formulation which also takes uncertainties of plant variables and time-varying prices into consideration in addition to other economic considerations such as inflation and other realistic details including depreciation, land prices, and many other issues. A more detailed process synthesis and the including the impact of the price uncertainties which are characterized by logistic issues and market demands is beyond the scope. Thus, the

simulations are performed for a particular and acceptable price value under marketable feed and product compositions. The architecture is flexible to address those considerations and the impact of those variables on the ultimate plant design and operations conditions can also be calculated using the proposed scheme.

References

- Bailie, R.C., Whiting, W.B., Shaeiwitz, J.A., Turton, R., Bhattacharyya, D., 2018. Analysis, Synthesis, and Design of Chemical Processes, 5th editio. ed, Prentice Hall, Boston.
- Chen, Q., Grossmann, I.E., 2017. Recent Developments and Challenges in Optimization-Based Process Synthesis. *Annu. Rev. Chem. Biomol. Eng.* 8, 249–283.
- Edgar, T.F., Himmelblau, D.M., Lasdon, L.S., 2001. Optimization of Chemical Processes, 2nd ed. ed, McGraw-Hill, New York.
- Laiwatthanaphaisarn, T., Anantpinijwatna, A., 2018. Optimization of dimethyl ether production process synthesis using superstructure analysis. *MATEC Web Conf.* 192, 03018.
- Mencarelli, L., Chen, Q., Pagot, A., Grossmann, I.E., 2020. A review on superstructure optimization approaches in process system engineering. *Comput. Chem. Eng.* 136, 106808.
- Nestler, F., Schütze, A.R., Ouda, M., Hadrich, M.J., Schaadt, A., Bajohr, S., Kolb, T., 2020. Kinetic modelling of methanol synthesis over commercial catalysts: A critical assessment. *Chem. Eng. J.* 394, 124881.
- Ploskas, N., Sahinidis, N. V., 2021. Review and comparison of algorithms and software for mixed-integer derivative-free optimization. *J. Glob. Optim.*
- Saebea, D., Authayanun, S., Arpornwichanop, A., 2019. Process simulation of bio-dimethyl ether synthesis from tri-reforming of biogas: CO₂ utilization. *Energy* 175, 36–45.
- Tangsrivong, K., Lapchit, P., Kittijungjit, T., Klamrassamee, T., Sukjai, Y., Laoonual, Y., 2020. Modeling of chemical processes using commercial and open-source software: A comparison between Aspen Plus and DWSIM. *IOP Conf. Ser. Earth Environ. Sci.* 463, 012057.

A Comprehensive Evaluation on the COSMO-SAC-dsp Model for Vapor-Liquid Equilibrium Predictions of Refrigerants

Jiayuan Wang, Tongxin Sa, Lingyu Zhu*

College of Chemical Engineering, Zhejiang University of Technology, Hangzhou, 310014, China
zhuly@zjut.edu.cn

Abstract

The performance of the COSMO-SAC-dsp model combining with the Peng-Robinson (PR) equation-of-state (EoS) via several mixing rules is assessed for vapor-liquid equilibrium (VLE) predictions of refrigerants. In particular, the original Huron-Vidal (HVO), the modified Huron-Vidal (MHV1) and the Wong-Sandler (WS) mixing rules are applied. Binary mixture combinations of 51 refrigerants with diverse molecule types are investigated. 447 valid experimental VLE datasets for 233 refrigerant pairs are collected from NIST and by an exhaustive literature search. The results show that the combination of the COSMO-SAC-dsp with the PR yields satisfied and comparable prediction accuracy on VLE in most cases comparing to that from two well established EOS- G^{ex} methods VTPR and PSRK. However, large errors could appear between chlorinated hydrocarbons + hydrocarbons and fluorinated hydrocarbons + chlorinated hydrocarbons. Moreover, errors become more significant and the VLE calculation cannot converge at all for NH_3 and CO_2 including mixtures respectively. The performance of the mixing rules for the studied refrigerant VLE database follows the decreasing order $\text{MHV1} > \text{HVO} > \text{WS}$.

Keywords: refrigerant, mixing rule, COSMO-SAC, Peng-Robinson

1. Introduction

Phase behavior, especially vapor-liquid equilibria, is important in the refrigerant mixture design. Various mixing rules have been developed combining EoS with excess Gibbs free-energy models (i.e. EoS- G^{ex} methods), which enable improved VLE predictions on nonideal mixtures¹. The EoS- G^{ex} methods offer greater extrapolation capability, flexibility and reliability of predictions than the conventional Van Der Waals one-fluid mixing rule and the direct use of G^{ex} models. Because the former typically requires refrigerant pair dependent correction binary parameters and the latter usually fails to describe the phase behavior when approaching the critical region.

G^{ex} models based on group-contribution methods are the most widely used and have been successfully combined with EoS for VLE predictions. Two well-known examples on such group contribution EoS are the VTPR² and PSRK³ methods from Gmehling and his coworkers. However, the G^{ex} models based on group contribution may have limited applicability on refrigerants as 1) many group parameters are missing or only available from commercial sources; 2) group contribution inherently does not work well for small molecules. A more predictive alternative to the group contribution method, is the G^{ex} model based on quantum mechanical conductor-like screening model (COSMO)

calculations. The COSMO models only contain a few universal parameters and do not require a prior knowledge on molecule group assignment. A comprehensive literature review on the development of the COSMO models and implementation details could be found by Bell et al.⁴

An evaluation on phase equilibria of refrigerant systems using COSMO models has recently been reported by Mambo-Lomba and Paricaud⁵. However, the VLE calculation is evaluated based on only 20 binary pairs and one mixing rule. A more systematic and dedicated evaluation using COSMO models on refrigerant VLE is carried out in this work to reach a better understanding on the reliability of predictions. In particular, the COSMO-SAC-dsp model⁶, a state-of-art version of the COSMO models, is combined with the PR EoS as the EoS-G^{ex} method for VLE predictions. As significant extensions to the previous work, 1) a much larger experimental VLE database is constructed and assessed; 2) different mixing rules are applied and discussed; 3) a comparison to the well-established group contribution EoS is added.

2. Methods

2.1 VLE data

51 molecules covering typical types of commonly used refrigerants are selected, which are categorized in to 6 classes: “CH” for hydrocarbons; “Cl”, “F” and “Cl-F” for molecules containing purely Cl, F and both respectively; “Et” for ethers; “In” for inorganics (Table 1). Such classification is mainly based on the consideration that the parameterization in the COSMO-SAC-dsp model is atom type-based.

Table 1. A classification of 51 refrigerant molecules selected in this work.

Type	Chemical Name	Ashrae Number	Type	Chemical Name	Ashrae Number
CH	methane	R50	F	2,3,3,3-tetrafluoro-1-propene	R1234yf
	ethane	R170		1,1,1,3,3-pentafluoropropane	R245fa
	propene (propylene)	R1270		Hexafluoropropylene	R1216
	propane	R290		cis 1,3,3,3-tetrafluoropropene	R1234ze(Z)
	cyclopropane	RC270		trans 1,3,3,3-tetrafluoropropene	R1234ze(E)
	Butane	R600		3,3,3-trifluoroprop-1-ene	R1243zf
	Isobutane	R600a	1,1,1,2,3,3-hexafluoropropane	R236ea	
Cl	dichloromethane	R30	1,1,1,3,3-pentafluorobutane	R365mfc	
	chloromethane	R40	octafluorocyclobutane	RC318	
	1,2-dichloroethane	R150	trichlorofluoromethane	R11	
	chloroethene	R1140	dichlorodifluoromethane	R12	
F	tetrafluoromethane	R14	Cl-F	chlorotrifluoromethane	R13
	trifluoromethane	R23		dichlorofluoromethane	R21
	difluoromethane	R32		chlorodifluoromethane	R22
	fluoromethane	R41		1,1,2-trichloro-1,2,2-trifluoroethane	R113
	hexafluoroethane	R116		1,2-dichloro-1,1,2,2-tetrafluoroethane	R114
	pentafluoroethane	R125		chloropentafluoroethane	R115
	1,1,1,2-tetrafluoroethane	R134a		2,2-dichloro-1,1,1-trifluoroethane	R123
	1,1,1-trifluoroethane	R143a		2-chloro-1,1,1,2-tetrafluoroethane	R124
	1,1-difluoroethane	R152a		1,1-dichloro-1-fluoroethane	R141b
	fluoroethane	R161		1-chloro-1,1-difluoroethane	R142b
	1,1,2,2-Tetrafluoroethane	R134	Et	Dimethyl Ether	RE170
	1,1,2,2-Tetrafluoroethene	R1114		Ethyl ether	-
	octafluoropropane	R218	In	ammonia	R717
	1,1,1,2,3,3,3-heptafluoropropane	R227ea		carbon dioxide	R744
	1,1,1,3,3,3-hexafluoropropane	R236fa			

Valid experimental VLE data for 233 refrigerant binary pairs are collected from NIST and by an exhaustive literature search. Two P-x-y datasets at distinguished temperatures

are taken for each pair if available, which results in 447 datasets with 4387 data points in total. Experimental uncertainties are not considered for model evaluation.

2.2 EoS-G^{ex} method

An overview of the VLE prediction method is illustrated by Figure 1. An improved PR EoS with the Twu α function is adopted⁷. The volume translation is not considered, as it does not affect the VLE calculation. The COSMO-SAC-dsp model⁶ is used as the G^{ex} model, and it is combined with the PR EoS via mixing rules HVO, MHV1 and WS¹. The three resulted EoS-G^{ex} methods are named as PR-HVO-COSMO, PR-MHV1-COSMO and PR-WS-COSMO respectively. VTPR² and PSRK³ methods are implemented as well for comparison.

All group parameters are retrieved from the website “<http://www.ddbst.com/>” (accessed on 08/01). The molecule group assignment is performed using the online tool from the same website as well. Sigma profiles, area and volume parameters for 51 refrigerants are generated using Material Studio conjugated with Python following the steps provided by Bell et.al⁴. The VLE is calculated iteratively in MATLAB with fixed liquid composition and temperature (i.e. bubble-point pressure flash calculation). In case of convergence difficulties, GAMS is applied to solve the VLE in an equation-oriented manner.

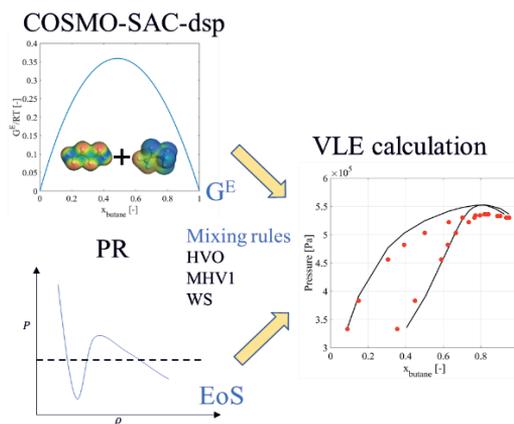


Figure 1. A schematic illustration of the VLE prediction method.

3. Results

Figure 2 provides an overview on the prediction errors from different methods. The number in each block gives the average absolute deviation (AAD) on the pressure and the vapor composition for binary mixtures of corresponding classes, which is defined by Eq 1.

$$AAD = \frac{1}{N} \sum_{i=1}^N \frac{1}{2K_i} \sum_{k=1}^{K_i} \left| \frac{P_{k,i}^{cal} - P_{k,i}^{exp}}{P_{k,i}^{exp}} \right| + \left| \frac{y_{k,i}^{cal} - y_{k,i}^{exp}}{y_{k,i}^{exp}} \right| \times 100\% \quad (1)$$

i is the index for the binary pair from 1 to N . The data point index of the binary pair i is given by k ranging from 1 to K_i . In addition, colors are associated to different ranges for better illustration. The number in the bracket is the number of refrigerant pairs used for error calculation.

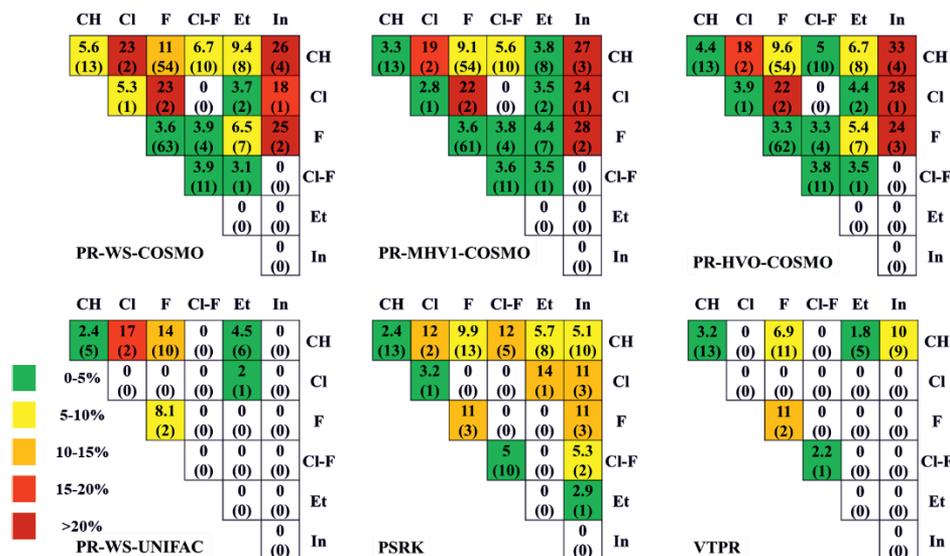


Figure 2. An overview of the prediction errors applying different EoS-G^{ex} methods.

3.1 A comparison between group contribution EoS

PSRK, VTPR and PR-WS-UNIFAC are implemented for the VLE prediction with relative satisfied accuracy (Figure 2). The PR-WS-UNIFAC method combines the PR with the UNIFAC(Do) via the WS mixing rule. Only 26 out of 233 pairs could be calculated due the missing group parameters. Large deviations appear between classes “Cl”+“CH”, “F”+“CH” and “F”+“F”, with three typical examples illustrated in Figure 3. It is clear that azeotropes cannot be well predicted. The group parameters in the UNIFAC(Do) are mainly regressed from the activity information. Dedicated sets of group parameters are proposed later to better fit the EoS-G^{ex} methods in the PSRK and VTPR. However, no significant improvement is achieved, at least on the tested refrigerant VLE database (Figure 2 and 3). The applicability of the group contribution EoS is still very limited due to undefined groups and missing parameters, in particular, on F containing groups. Moreover, proposing groups assignment methods for small molecules like refrigerants to distinguish isomerism is still a challenging task.

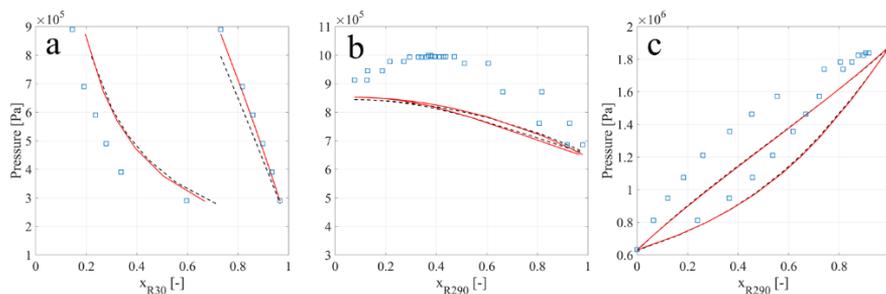


Figure 3. P-x-y plots of a) R30+R1270 at 333 K; b) R290+R143a at 283 K; c) R116+R143a at 273 K. The black dashed lines denote PR-WS-UNIFAC predictions. The red solid lines denote PSRK predictions. The square symbols denote experimental values.

3.2 A comparison between mixing rules

The values in Figure 2 indicate a decreasing order of the mixing rules on the prediction performance: MHV1>HVO>WS. Qualitative differences between three mixing rules cannot be clearly found from P-x-y plots, as exemplified by Figure 4b. The tested refrigerant pairs involving small molecules without association or highly asymmetry, which may not strongly distinguish the weaknesses and strengths of each mixing rule. However, significant improvement on the VLE predictions can still be identified by viewing the AAD values, in particular, at diluted conditions (Figure 4a).

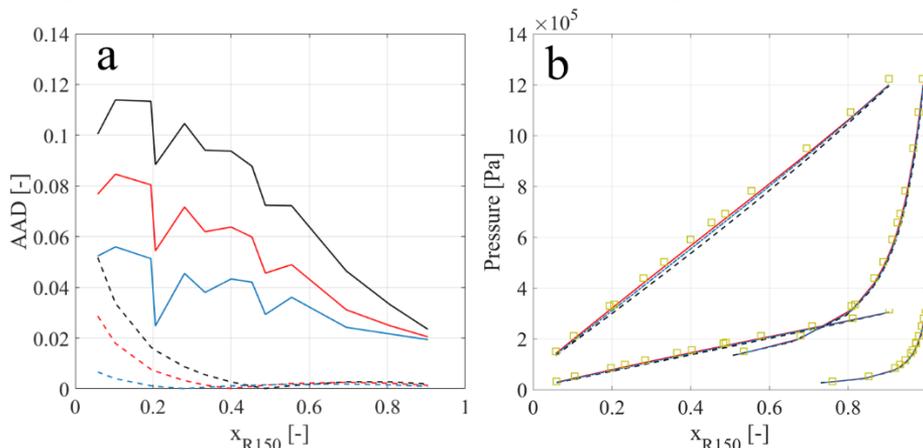


Figure 4. VLE prediction results on the R150+R1140. a) AAD on the pressure (solid lines) and vapor composition (dashed lines) at 364 K (black: WS, red: HVO, blue: MHV1); b) P-x-y plots at 293 K and 364 K (symbols: experimental values, black dashed lines: WS, red solid lines: HVO, blue dotted lines: MHV1).

3.3 A comparison between group contribution and COSMO methods

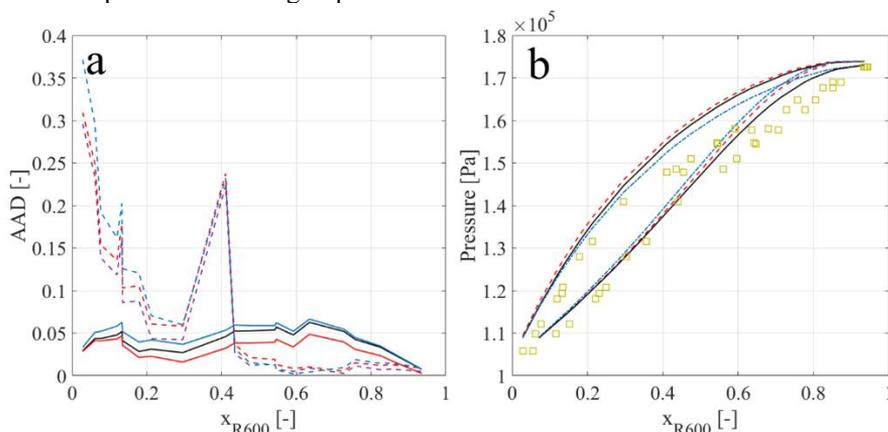


Figure 5. VLE prediction results on the R600+R1140. a) AAD on the pressure (solid lines) and vapor composition (dashed lines) at 273 K (black: WS, red: HVO, blue: MHV1); b) P-x-y plots at 273 K (symbols: experimental values, black solid lines: WS, red dashed lines: HVO, blue dotted lines: MHV1).

The EoS-G^{ex} methods based on the COSMO-SAC-dsp model outperform the group contribution EoS in terms of the extrapolation capability. The VLE of all pairs can be calculated using the COSMO-SAC-dsp, but the majority are not applicable for the group contribution EoS. This is attributed to the atom-based parameterization scheme in the COSMO models, which helps to avoid ambiguities in group assignment and to reduce the adjustable parameters. The COSMO-SAC-dsp model enables comparable prediction accuracy for most combinations of 6 molecule classes, or even better for some cases. However, large errors could appear for “Cl”+“CH”, “Cl”+“F”. Although the P-x-y shapes can be well approximated, significant errors can be found at diluted conditions (Figure 5). The large deviation around $x = 0.4$ may due the experimental data error in Figure 5. Prediction errors above 20% are typically identified for pairs including NH₃. Moreover, infeasible solutions always appear when CO₂ is involved. Convergence on VLE calculations could be achieved only when the dispersion term of the CO₂ is neglected.

4. Conclusions and outlooks

The COSMO-SAC-dsp model is combined with the PR EoS via several mixing rules for VLE predictions on 233 refrigerant pairs, the results show that:

1. The performance of the mixing rules generally follows the order of MHV1>HVO>WS, but without a significant difference.
2. The predictive performance using the COSMO-SAC-dsp model is comparable to that of group contribution EoS in most cases, but not for NH₃ and CO₂ involving mixtures.

A revision on the dispersion parameters for atoms Cl, F, N and O would possibly improve the COSMO-SAC-dsp model performance on the VLE calculation, which is still in progress and would be published separately.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 22008217).

References

1. H. Orbey, S.I. Sandler, 1998, Modeling vapor-liquid equilibria. cubic equation of state and their mixing rules. Cambridge University Press.
2. B. Schmid, A. Schedemann, J. Gmehling, 2014, Extension of the VTPR Group Contribution Equation of State: Group Interaction Parameters for Additional 192 Group Combinations and Typical Results. *Ind. Eng. Chem. Res.* 53 (8), 3393-3405.
3. S. Horstmann, A. Jabłoniec, J. Krafczyk, K. Fischer, J. Gmehling, 2005, PSRK group contribution equation of state: comprehensive revision and extension IV, including critical constants and α -function parameters for 1000 components. *Fluid Phase Equilib.* 227(2), 157–164.
4. I.H. Bell, E. Mickoleit, C.M. Hsieh, S.T. Lin, J. Vrabec, C. Breitkopf, A. Jäger, 2020, A Benchmark Open-Source Implementation of COSMO-SAC. *J. Chem. Theory Comput.* 16 (4), 2635-2646.
5. D. Mambo-Lomba and P. Paricaud, 2020, Predictions of thermodynamic properties and phase equilibria of refrigerant systems with cosmo approaches. *Int. J. Refrig.* 124,52-63.
6. C.M. Hsieh, S.T. Lin, J. Vrabec, 2014, Considering the dispersive interactions in the cosmo-sac model for more accurate predictions of fluid phase behavior. *Fluid Phase Equilib.* 367, 109-116.
7. I.H. Bell, M. Satyro. E.W. Lemmon, 2018, Consistent Two Parameters for More than 2500 Pure Fluids from Critically Evaluated Experimental Data. *J. Chem. Eng. Data*, 63, 7, 2402–2409.

Design and Evaluation Strategies for a Novel Hybrid Trigeneration System

Rachid Klaimi,^{a,b} Sabla Y. Alnouri,^{c*} Mirko Stijepovic^d

^a*Mechanical Engineering Department, American University of Beirut, P.O. Box 11-0236, Beirut, 1107-2020, Lebanon*

^b*Baha and Walid Bassatne Department of Chemical and Petroleum Engineering, American University of Beirut, P.O. Box 11-0236, Riyad El-Solh, Beirut, Lebanon*

^c*Gas Processing Center, College of Engineering, Qatar University, Doha, Qatar*

^d*Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, 11000 Belgrade, Serbia
sabla@qu.edu.qa*

Abstract

Renewable energy sources that involve the use of Concentrated Solar Power systems (CSP) are receiving increased attention due to many of the prominent features that it can provide. Most importantly, the simultaneous production of thermal and electrical energy that can be achieved via CSP systems is of great value and constitute many cogeneration processes. Moreover, CSPs have been proven to be an effective source of energy for industrial applications that require high amounts of clean energy, such as desalination. Hence, combining water production activities together with thermal and electrical energy production options via CSP technologies results in novel hybrid tri-generation systems. The possibility of locating such tri-generation systems on the eastern shore of the Mediterranean Sea is investigated in this paper. As such, several different design aspects such as the effect of direct normal irradiance (DNI) on the production freshwater via desalination has been studied. The daily and seasonal intermittency of sunlight, which consists a major challenge for CSP, necessitates the implementation of hybrid systems to ensure a continuous supply of energy for desalination purposes. This paper studies the design of a hybrid trigeneration system under different weather conditions, including four different seasons. The trigeneration system consists of a CSP plant as the main energy source. The system is usually coupled with a conventional fuel source, mainly natural gas, which serves as a backup energy source in the absence of solar energy. A desalination plant comprising of both thermal and membrane technology options for freshwater generation is also included. The design problem is translated into a Mixed Integer Non-Linear Program (MINLP) that can be solved for the optimal network structure based on appropriately selected technologies for steam, power and freshwater production. The environmental aspect of the system has also been studied by quantifying the total amount of carbon dioxide emissions. As such, the integration of a carbon capture technology to regulate those emissions has been also studied, subject to an appropriate value for the overall net carbon reduction target of the tri-generation system.

Keywords: Concentrated Solar Power, desalination, design, hybrid, trigeneration system

1. Introduction

The global energy sector is currently in a transitional phase towards achieving a higher energy share from renewable sources. Concentrated Solar Power (CSP) is one of the

renewable energy technologies with a high potential for deployment in many countries, especially those located in the MENA region, due to the high availability of direct normal irradiance (DNI). Therefore, studying the factors affecting its performance has been extensively addressed in literature (Klaimi et al., 2021). However, the incorporation of CSP in developing countries apparently faces many challenges, such as high capital investment costs, and the lack of supporting regulations. Moreover, any financial incentives are against the low cost of fossil fuels (Elmorsy et al., 2020). Therefore, the hybridization of CSP plants using conventional energy sources provides more flexibility and reliability by gearing it away from solar intermittency problems. Many have studied the integration of CSP into different industrial applications. For instance, Elmorsy et al. (2020) have worked on a novel natural gas-fired integrated solar-combined cycle power plant in Egypt, where the levelized cost of electricity (LCOE) was found to be 40 \$/MWh. Rashid et al. (2020) have presented a dynamic simulation of a CSP plant incorporating both thermal energy storage and natural gas. The hybrid plant with storage resulted in a higher LCOE value (86.32 \$/MWh) when compared to a single natural gas plant (74.92 \$/MWh). Abdelhay et al. (2020) have proposed a solar driven poly-generation system of CSP coupled with a backup natural gas heater for power, desalination and cooling. Their proposed system resulted in a relatively low unit water price of 1.247 \$/m³. Hybrid energy systems are also associated with greenhouse gas (GHG) emissions which must be taken into consideration while studying the environmental impact of cogeneration and tri-generation processes. Moreover, there exist very few studies that focus on hybrid tri-generation systems. Therefore, this paper studies the effect of solar availability on the performance of a novel CSP-trigeneration system coupled with a natural gas boiler for the production of three different pressure levels of steam, power and freshwater. The novelty of this paper lies in the multiperiod approach presented to study the performance of the CSP-desalination system across different seasons, in addition to the incorporation of carbon capture option and imposing taxation on uncaptured CO₂ which allows the estimation of associated environmental costs.

2. Methodology

The proposed CSP-natural gas hybrid tri-generation system is shown in Figure 1. First, a boiling feed water (BFW) stream is split into two streams: BFW-b that enters a utility boiler and BFW-sg entering a solar steam generator. A natural gas stream is fed into the utility boiler and undergoes combustion, which in turn results in heat transfer to the BFW-b stream. As a result, high pressure (HP) steam at 41 bars (S-b) is generated, in addition to a blowdown stream. On the other hand, the HP steam produced by the steam generator (S-sg) is due to the heat exchange between a molten salt stream circulating in the CSP plant and boiling feed water. A portion of the HP steam produced (S-ms) is utilized by the molten salt system to maintain a temperature above molten salt freezing point. The two HP streams are then mixed into one stream (S-d) that enters the shaft and electric power generation block which consists of several energy generation options (turbogenerators, drivers and throttle valves). This results in the production of low pressure (LP) steam at 2 bars and electric power to be utilized for desalination, in addition to very low pressure (VLP) steam at 0.12 bar to be condensed, deaerated and pumped to HP steam pressure in order to obtain boiling feed water. Any excess in electricity can be exported to the grid. Two different desalination technologies are available in the system for the generation of freshwater, which is the main source of revenue. As for GHG emissions from natural gas combustion, a carbon capture unit with a specific efficiency is incorporated in the system, and the uncaptured CO₂ will be released to the atmosphere.

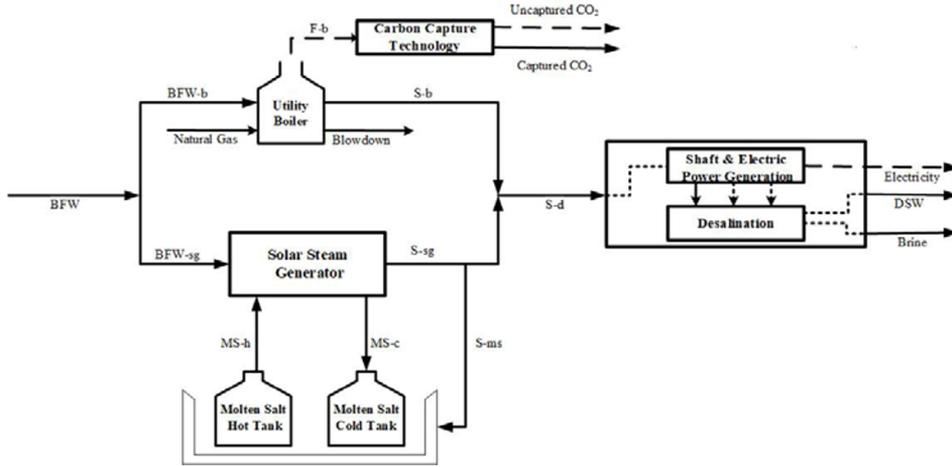


Figure 1: Hybrid Tri-generation System Structure

3. Mathematical Formulation

The proposed hybrid tri-generation system has been converted into a Mixed Integer Non-linear Problem (MINLP) that aims at minimizing the total cost of the system. The objective function is shown in Eq. (1) below, where $Cost^{CSP,energy}$ and $Cost^{NG,energy}$ are the energy production costs from CSP and natural gas respectively, $Cost^{CO_2}$ is the carbon related cost which includes carbon capture cost and taxes on the uncaptured CO_2 , $Cost^{Desal}$ is the desalination cost and Rev^{Water} is the total revenue earned from selling the freshwater produced.

$$\text{minimize } Cost^{CSP,energy} + Cost^{NG,energy} + Cost^{CO_2} + Cost^{Desal} - Rev^{Water} \quad (1)$$

Equality constraints have been added to the model to ensure the satisfaction of mass and energy balances around all the units in the system. Eq (2-5) represent the mass balances on the utility boiler and solar steam generator, where the variable G is the mass flowrate, and the subscripts S , b , sg , d and ms refer to steam, boiler, steam generator, desalination and molten salt respectively. The mass flowrate of steam required for the molten salt system G_{S-ms} and the amount of CO_2 emitted from the boiler G_{F-b} can be calculated using Eq. (6) and (7), respectively, where ω_{ST} is the fraction of HP steam produced by the generator and allocated to the storage tanks heating system, and ε_{NG} is a parameter corresponding to the amount of CO_2 emitted per unit mass of HP steam produced using natural gas. The energy balances around the boiler and steam generator are described in Eq. (8) and (9), respectively, where the variables H and Q are the specific enthalpy and thermal energy, respectively, while HV , η_b and η_{sg} are the parameters corresponding to the calorific value of natural gas, boiler and steam generator efficiencies, respectively.

$$G_{BFW} = G_{BFW-b} + G_{BFW-sg} \quad (2)$$

$$G_{BFW-b} = G_{S-b} + G_{blowdown} \quad (3)$$

$$G_{BFW-sg} = G_{S-sg} \quad (4)$$

$$G_{S-d} = G_{S-b} + G_{S-sg} - G_{S-ms} \quad (5)$$

$$G_{S-ms} = \omega_{ST} G_{S-sg} \quad (6)$$

$$G_{F-b} = \varepsilon_{NG} G_{S-b} \quad (7)$$

$$(G_{BFW-b} H_{BFW-b} + G_{NG} H_{NG}) \eta_b = G_{S-b} H_{S-b} + G_{blowdown} H_{blowdown} \quad (8)$$

$$Q_{absorbed} = \eta_{sg} Q_{released} \quad (9)$$

The inequalities of the model, described in Eq. (10-13) below address the limitations on certain aspects of the system, such as the capacities of the embedded technologies, land and solar availability, in addition to the minimum production of HP steam. In these equations, b represents a binary variable, $A_{SF,t}$ and, $A_{SF,max}$ are the required and available solar field areas in season t , DNI_t is the daily average direct normal irradiance in season t , and G_{S-dmin} is the minimum required amount of HP steam to be produced depending on the amount of energy required to generate a specific flowrate of freshwater.

$$b_i G_{min} \leq G_{BFW-i} \leq b_i G_{max} \quad \forall i \in b, sg \quad (10)$$

$$A_{SF,t} \leq A_{SF,max} \quad (11)$$

$$G_{S-sg} (H_{S-sg} - H_{BFW-sg}) \frac{24}{3600} \leq DNI_t A_{SF,t} \quad (12)$$

$$G_{S-dmin} \leq G_{S-d} \quad (13)$$

It should be noted that the nonlinearities of the model lie in the costing of some of the equipment incorporated within the system, the constraint preventing temperature cross-over in solar steam generator and the utilization of energy generation ratios of CSP and natural gas in some of the equations. Moreover, although the equations of the proposed multiperiod problem seems simple, the actual model is large and complex, consisting of 88 decision variables (80 continuous, 8 binaries) and more than 80 constraints.

4. Case Study

The proposed model was implemented on a tri-generation system located near the Mediterranean Sea in a region whose direct normal irradiance is similar to that of Beirut city in Lebanon. Since the aim of the model is to minimize the total cost of HP steam production that will be used for electricity and freshwater generation, the same selection of optimal technologies, in the shaft and electric power generation and desalination blocks, reported in the optimal case scenario in Klaimi et al. (2021) will be considered. This scenario corresponds to a large-scale desalination plant having a capacity of 100,000 m³/d with an inlet seawater salinity of 35 g/L and HP steam requirement of 7,748 t/d. Thus, the decision variables are mainly the mass flowrates of inlet boiling feed water streams, natural gas, molten salt and the land area of CSP facility. Consequently, the costs associated with the power block and desalination in the objective function were kept constant. Moreover, the daily averaged DNI values for the four seasons, which in turn were divided into four different time periods, are 146.9, 101.2, 170.2 and 229.6 kWh/m²d for fall, winter, spring and summer, respectively (Solar Atlas), whereas the maximum solar field area available for the CSP facility has been set at 20,000 m². The MINLP optimization problem has been implemented using “Whats’Best 17.0” LINDO Global Solver for MS-Excel 2016 via a laptop with Intel Core i5 Duo processor, 8 GB RAM and a 64 bit operating system. Figure 2 below shows the flowrates of HP steam produced from each source of energy, in addition to the amount of CO₂ emissions resulting from the utilization of natural gas, by the optimal solutions attained.

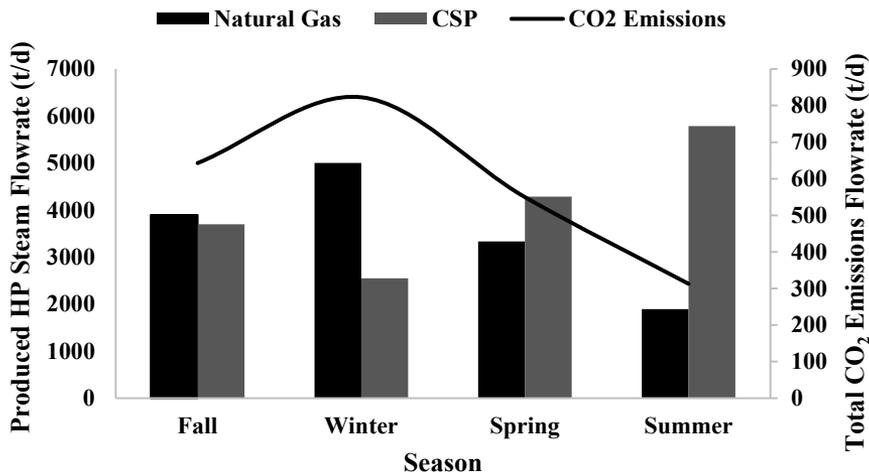


Figure 2: HP steam and CO₂ flowrates in different seasons

The obtained results confirm that the amount of HP steam generated from each energy source is highly dependent on solar availability. The highest DNI in summer resulted in the highest deployment of CSP which generates 5,785 t/d of HP steam in this period with a generation of 1,895 t/d of steam from natural gas. As DNI decreases to 146.9 kWh/m²d in fall, a lower amount of energy could be generated per unit surface area, which explains the reduction in the amount of steam generated from CSP to 4,288 t/d and the increase in the amount of steam generated from natural gas to 3,332 t/d. Winter season witnesses the lowest DNI value among all seasons of the year, which necessitates a high requirement for natural gas, in order to compensate for the reduction in the capacity of the CSP facility. This resulted in the production of around 66.23% of the total required amount of steam from natural gas in winter compared to 24.67% in summer. However, the DNI value increased again in spring, which also increased the share of renewable energy in steam production. It was also very important to quantify the amount of carbon dioxide emissions that resulted from the utilization of natural gas. Since the highest and lowest consumption of natural gas occur in winter and summer, respectively, Figure 2 shows that carbon emissions are indeed the highest in winter (with a flowrate of 823 t/d). This amount decreased when a reduction in natural gas consumption was realized in subsequent seasons, achieving 313 t/d in summer. Although CSP is known for its high capital cost, it has been selected in all seasons, even in winter which has the lowest solar availability. Despite this, a continuous selection of CSP technology was attained throughout the year. In order to further explain this aspect, Figure 3 shows a breakdown of water production cost across the four different time periods.

According to Figure 3, total water production cost from the hybrid system ranges from 1.25 USD/m³ in summer to 1.47 USD/m³ in winter. Moreover, the fraction of steam generation cost from CSP is greater than that of natural gas in all seasons, except in winter when the contribution of natural gas for energy production is much higher than that of CSP. In fact, this was expected due to the high capital cost of CSP compared to the low cost of natural gas in developing countries. However, when serious regulations regarding the environmental impacts of energy production are implemented, such as the use of carbon capture and carbon taxation on CO₂ emissions, the scenario will definitely change.

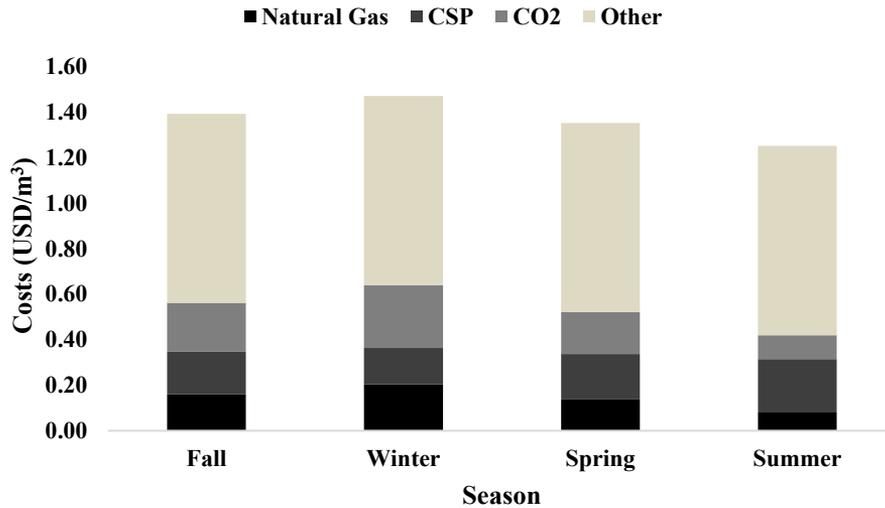


Figure 3: Water production cost breakdown

When comparing the costs related to carbon capture and taxation in winter which has the highest CO₂ emissions (0.28 USD/m³) to the cost of steam generation from CSP in summer (0.24 USD/m³), it was noted that the cheapest energy source is not always the best candidate, as the costs associated with its impacts might be higher than the cost of other options. In addition, the results showed that the available solar field area has been totally invested in all four cases, which means that when a higher area is available, the contribution of CSP could be higher, resulting in a lower total water production cost.

5. Conclusion

In this study, a mathematical model that minimizes water production cost of a hybrid tri-generation system using different seasons has been presented. A minimum water cost of 1.25 USD/m³ was attained in summer. Moreover, significant financial incentives for investing in CSP were realized.

References

- A. Abdelhay, H. S. Fath, S. A. Nada, 2020, Solar driven polygeneration system for power, desalination and cooling, *Energy*, 198, 117341.
- K. Rashid, K. Mohammadi., K. Powell, 2020, Dynamic simulation and techno-economic analysis of a concentrated solar power (CSP) plant hybridized with both thermal energy storage and natural gas, *Journal of Cleaner Production*, 248, 119193.
- L. Elmorsy, T. Morosuk, G. Tsatsaronis, 2020, Exergy-Based Analysis and Optimization of an Integrated Solar Combined-Cycle Power Plant, *Entropy*, 22(6), 655.
- R. Klaimi, S. Y. Alnouri, M. Stijepovic, 2021, Assessment strategies of different molten salts for heat transfer in tri-generation systems, *Computer Aided Chemical Engineering*, 50, Elsevier.
- R. Klaimi, S. Y. Alnouri, M. Stijepović, 2021, Design and thermo-economic evaluation of an integrated concentrated solar power–Desalination tri-generation system. *Energy Conversion and Management*, 249, 114865.
- Solar Atlas, 2021, Direct normal irradiance in Port of Beirut, retrieved from: <https://globalsolaratlas.info/detail?c=34.025348,35.46936,10&s=33.918292,35.505066&m=site>

Process synthesis for desalination using superstructure optimization of membrane technologies

Alvaro Martinez-Triana^{a,b}, Mohamed Kanniche^a, Thibaut Neveux^a, Christophe Castel^b and Eric Favre^b

^a*EDF R&D Chatou, 6 quai Watier, 78400 Chatou, France*

^b*Université de Lorraine, CNRS, LRGF, F-54000 Nancy, France*
alvaro.martinez-triana@univ-lorraine.fr

Abstract

Reverse Osmosis is considered today as the best available technology at industrial scale for desalination as an alternative process to conventional thermal technologies. Nevertheless, there is still opportunity to evaluate and optimize other membrane technologies like membrane distillation. Vacuum membrane distillation has been studied at pilot scale and it is a promising technology to treat seawater in areas with un-expensive heat sources. This study presents a modified method of superstructure optimization to evaluate key performance indicators of reverse osmosis and vacuum membrane distillation. To achieve this objective, Computer aided process simulation is deployed; mathematical model was coded in FORTRAN and added to a superstructure defined in a process simulation software (ProSimPlus). Geometry and material parameters are set according to the commercial module Dow SW30HR-380 data. Finally, superstructure's operational conditions (pressure and in-process temperature) and specific number of modules in series and parallel are optimized within an ant colony algorithm (MIDACO).

Keywords: Process synthesis, optimization, desalination

1. Introduction

Desalination processes are important for different industries and for fresh water supply. Multiple stage flash (MSF) is the most deployed thermal technology at industrial scale according to Subramani and Jacangelo (2015). However, Reverse Osmosis (RO) has been implemented as an alternative process to improve the energy performance (EP). Nevertheless, there is still opportunity to evaluate and optimize other membrane technologies regarded as an alternative to MSF and RO. Membrane distillation (MD) has not been tested at industrial scale yet. Pagliero et al. (2021) identifies a growing interest on these technologies due to the capacity to treat high concentration brines, but there are still some technical challenges related to the EP. Vacuum membrane distillation (VMD) has been tested at pilot scale for desalination. VMD research has been focused on new membrane materials and geometries for contactors. Nevertheless, simulation tools to gather the material characteristics, module geometries and operating conditions at industrial scale remains a challenge.

This work presents a modified method of Zhao et al. (2018), to identify a membrane process (VMD and RO) adapted to desalination based on process simulation, taking into account the mass and thermal transfer modelled in previous researches. Besides, this method can be deployed as a tool to evaluate membrane technologies giving an estimation of the energy and/or heat resources needed to accomplish a production objective at industrial scale. A plot for key process indicators of two membrane technologies is presented as a main result.

2. Methodology

Computer-aided process simulation is the main tool used in the present work, based on the optimization of a superstructure following the methodology presented by Zhao et al. (2018). The resultant method is composed of three blocks: a FORTRAN code to simulate each module, the industrial scale process simulation and an optimization block.

2.1. RO and VMD Modelling

Modelling of mass transfer through the membrane is coded in FORTRAN according to previous published works e.g Li et al. (2021), resolution of the mass balance for a module is based in the calculation of differential equations over the membrane area. The FORTRAN subroutine solve the differential equations system giving as a result the retentate and permeate composition, flowrate, temperature and pressure.

Mass transfer through a membrane is modelled as a gradient in the chemical potential between two homogenous phases at each side of the membrane (retentate and permeate). For RO, this driving force is due to a pressure difference. In addition, pressure drop is included into the mathematical model following the work of Lian et al. (2018).

For VMD, chemical potential difference is due to the vapour pressure at each side of the membrane. Mathematical model is based on the work done by Mendez et al. (2019); it calculates a water vapour flux through the membrane (hydrophobic membrane is assumed, thus there is not liquid phase flux). Besides, temperature polarisation effects are evaluated following the thermal balance proposed by Ibrahim and Alsahy (2013) calculating the heat transfer through the membrane.

For both technologies, concentration polarisation effect is calculated based on the fluid hydrodynamic. The commercial module simulated is spiral wound type used for brackish seawater desalination. Boudinar et al. (1992) presents a mathematical model for a commercial spiral module (Filmotec) validated in the range of flowrate, pressures and temperatures used in this work. These effects have an effect on the mass transfer coefficient modelled according to Kim and Hoek (2005), this model also allows to calculate the NaCl concentration at the wall side, then evaluate risk of fouling. Physical and chemical properties are calculated using the commercial software Simulis Thermodynamics. Vapour- liquid equilibrium (VLE) is modelled using an equation of state approach corrected by NRTL model.

2.2. Process modelling

To achieve a comparison between membrane technologies at industrial scale, it has been set a fresh water flowrate to be produced (constraint). Then, a superstructure is defined to simulate different scenarios governed by design and operational variables. These variables affect directly in the process synthesis, therefore, in the key process indicators. Table 1 present a summary of inputs, variables and key indicators used for the optimization problem.

Superstructure conception and evaluation were done with the commercial process simulation software (ProSimPlus). Heat exchangers, pumps, flash and mixers were taken directly from the library (ProSimPlus). Logic switch units, to control the number of stages, were set according to Zhao et al. (2018). VMD and RO FORTRAN subroutines were added as an external module (external DLL). Product salinity is not set as a constraint, the importance of this parameter depends on the industry; nevertheless, it is evaluated as a process indicator for RO, and it is reported as the ratio between NaCl concentration at the product and the feed.

RO superstructure is an adaptation of the process presented by Alsarayreh et al. (2021). It simulates a process structure containing up to four stages for RO and six stages for VMD. A process stage includes a pump (for RO) or heat exchanger (for VMD) to set fluid's conditions and a configuration of modules in parallel. Then, the number of modules in parallel is evaluated in a wide

Table 1: Summary inputs, variables and key indicators for optimisation problem. RO material information from Alsarayreh et al. (2021), VMD material information from Mendez et al. (2018). AA – Active area per module, WP – Water permeability, WS- salt permeability, PVDF- Polyvinylidene fluoride, NMP – modules in parallel, NMS – stages

		RO	VMD
Inputs	Water production requirement	4500 m ³ /d	
	Feed water salinity	35 g/L NaCl	
	Membrane module characteristics	Polyamide Thin-film composite	PVDF thin-film Supported
		AA: 35 m ²	
WP: 2.69X10 ⁻¹² m ³ /m ² .Pa.s		WP: 1.27X10 ⁻⁷ kg/m ² .Pa.s	
	SP: 3.2X10 ⁻⁸ m/s		
Key variables	Feed and in-process temperature	50°C – 90°C	
	Number of modules in parallel and stages	NMP: 60 – 2000	NMP: 40 – 1000
		NMS: 1- 4	NMS: 1- 6
	Vacuum pressure	0.1 bar	
Feed pressure	20 – 80 bar		
Key Process indicators (KPI)	Electricity	X	X
	Heat		X
	Membrane Area	X	X
	Product salinity	X	

range for each step independently. Fig. 1 shows the superstructure to simulate RO and VMD processes.

Superstructures contains three user scripts corresponding to the following calculation:

“Power”: Calculate total electricity in MW and heat in MW(thermic) resulting from the pumps and heat exchangers requirements.

“Area”: Calculate the total membrane area (*m*²) according to each simulated scenario.

“Polarisation”: Verify risk of precipitation inside the modules due to the increase of the concentration above the solubility of NaCl in water. In case of exceeding the limit concentration, this script acts as a constraint for the process, then the solution is excluded.

This superstructure allows the evaluation of different process configuration while changing the key variables into the ranges defined in the table 1. The KPIs are calculated as a result of each simulated scenario.

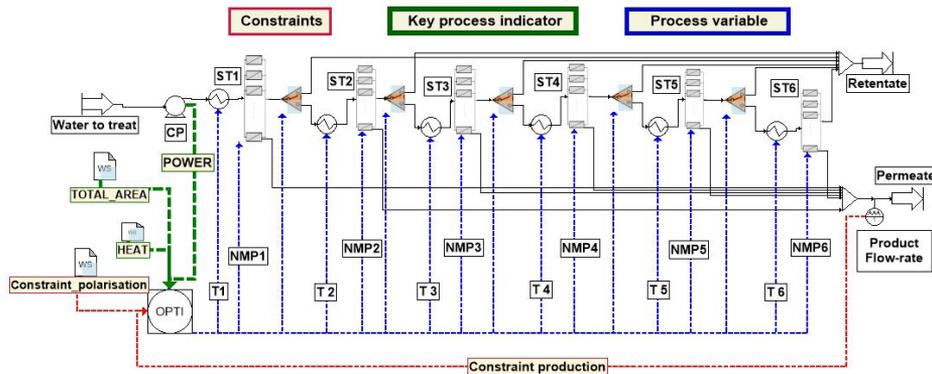


Figure 1: Superstructure for RO and VMD processes. ST - Stage, CP - Centrifugal pump, T - Temperature.

2.3. Process optimization

Optimization of the superstructure is deployed in two steps:

2.3.1. Sensitivity analysis

This analysis is evaluated using the “spec” module of ProSimPlus, which it generates an action over the inlet pressure (RO) or the inlet temperature (VMD) to accomplish the production constraint. Simulated cases are defined by uniform experiment plan to evaluate different module configurations.

2.3.2. Optimization

Parameters of superstructures are optimized using an ant colony algorithm (MIDACO). For this optimization, a multi-objective problem is defined according to the results of the sensitivity analysis. Then, a front-Pareto is obtained in function of the KPI.

3. Results and discussion

Results from the sensitivity analysis are sorted by the value of the performance indicators for each technology. Results for EP are plotted in function of the stages (x-axis) and the average modules in parallel (y-axis) for RO and VMD. As seen in the figure 2, there are configurations for RO being able to accomplish the required production in one or two stages; however, these configurations require more electrical power due to high operational pressure requirements.

For VMD, result of the sensitivity analysis does not show an optimal structure to increase EP when sorting the results by heat as presented in figure 2, however heat required increases at low number of modules in parallel. Besides, to get the production constrain is needed at least a two stages process.

This partial result validates ranges and limits defined in the optimization problem. In addition, it shows a difficulty to identify a group of desired structures while having two objectives. Then the figure 3 shows the resultant front-Pareto solutions for the multi-objective optimization.

For RO, it is obtained a front-Pareto covering a range of values for key process indicators that can

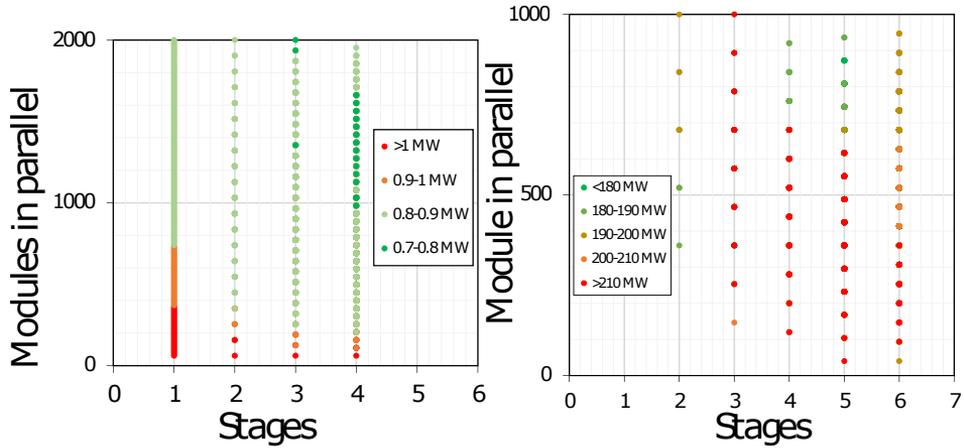


Figure 2: Effect of process structure over total electricity required for RO (left) and over total heat required for VMD (right)

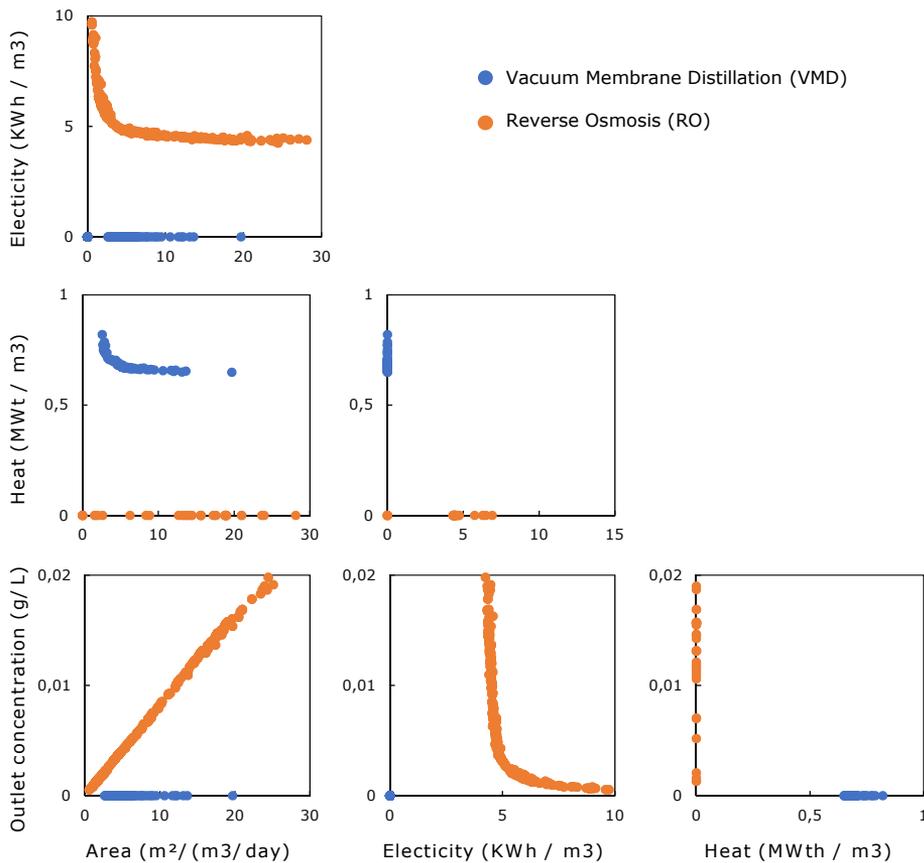


Figure 3: Pareto-front obtained from the superstructure optimisation for RO and VMD

be post-treated to balance the impact of each indicator and identify an optimal structure. Besides, the output concentration is proportional to the area. For VMD, the resultant front-Pareto shows the effect of the area over heat as EP, electricity consumption remains constant for all the cases with lower consumption than RO.

This result can be an input for an economical model allowing the identification of process and operating parameters according to the available sources (heat and electricity). VMD presents a potential over RO due to its low electricity requirement, VMD potential to treat seawater increases in case of product salinity requirements and it can be capitalised if there is a source of heat available at low cost.

Figure 3 provides a tool aimed to guide process synthesis by presenting the expected key parameter for each technology according to the salinity constraints. Besides, this methodology can be deployed with other operating conditions limited to available resources. Associated results will support the selection of one technology in particular.

As a perspective, this evaluation can provide a baseline for the technologies comparison at industrial scale. The suggested post-treatment is the addition of prices for electricity, heat and membrane area according to the industry; in this way, this work can be used as an aid to decision tool providing a more detailed OPEX and CAPEX calculation.

4. Conclusion

The modified method for superstructure optimizations applied to membrane technologies links phenomena models with process synthesis. This tool allows the user to have a deep understanding of the KPI at industrial scale but taking into account the local effects like concentration and temperature polarization and material characteristics.

This work can be used as a tool for technical and economical evaluation adapted to real prices and financial aspects to guide the process synthesis based in membrane technologies. Then, results can support the potential of membrane distillation in certain industrial cases.

References

- A. Alsarayreh, M. Al-Obaidi, S. Farag, R. Patel, I. Mujtaba, 2021. Performance evaluation of a medium-scale industrial reverse osmosis brackish water desalination plant with different brands of membranes. A simulation study. *Desalination* 503, 114927.
- M. Boudinar, W. Hanbury, S. Avlonitis, 1992. Numerical simulation and optimisation of spiral-wound modules. *Desalination* 86, 273–290.
- S. Ibrahim, Q. Alsalhy, 2013. Modeling and simulation for direct contact membrane distillation in hollow fiber modules. *AIChE J.* 59, 589–603.
- S. Kim, E. Hoek, 2005. Modeling concentration polarization in reverse osmosis processes. *Desalination* 186, 111–128.
- F. Li, C. Fang, W. Liu, L. Yang, B. Guo, S. Zhang, 2021. High recovery, energy efficient wastewater desalination. *J. Membr. Sci.* 631, 119317.
- B. Lian, G. Blandin, G. Leslie, P. Le-Clech, 2018. Impact of module design in forward osmosis and pressure assisted osmosis: An experimental and numerical study. *Desalination* 426, 108–117.
- D. Mendez, C. Castel, C. Lemaitre, E. Favre, 2019. Improved performances of vacuum membrane distillation for desalination applications: Materials vs process engineering potentialities. *Desalination* 452, 208–218.
- D. M. Mendez, C. Castel, C. Lemaitre, E. Favre, 2018. Membrane distillation (MD) processes for water desalination applications. Can dense self-standing membranes compete with microporous hydrophobic materials? *Chem. Eng. Sci.* 188, 84–96.
- M. Pagliero, M. Khayet, C. García-Payo, L. García-Fernández, 2021. Hollow fibre polymeric membranes for desalination by membrane distillation technology: A review of different morphological structures and key strategic improvements. *Desalination* 516, 115235.
- A. Subramani, J. Jacangelo, 2015. Emerging desalination technologies for water treatment: A critical review. *Water Res* 75, 164–187.
- Q. Zhao, T. Neveux, M. Mecheri, R. Privat, P. Guittard, J. Jaubert, 2018. Superstructure optimization (MINLP) within ProSimPlus Simulator. *Computer Aided Chemical Engineering*, 767–772.

A comparison of process synthesis approaches for multistage separation processes by gas permeation

Thibaut Neveux^{a,*}, Bernardetta Addis^b, Christophe Castel^c, Veronica Piccialli^d
and Eric Favre^c

^a*EDF R&D Chatou, 6 quai Watier, 78400 Chatou, France*

^b*Université de Lorraine, CNRS, LORIA, F-54000 Nancy, France*

^c*Université de Lorraine, CNRS, LRGP, F-54000 Nancy, France*

^d*University of Rome Tor Vergata, 00133 Rome, Italy*

thibaut.neveux@edf.fr

Abstract

Various optimization-based process synthesis approaches exist to propose a process structure and associated parameters, which have been applied on many applications. The goal of this work is to provide a comparison of methods on a common process synthesis problem.

Three methods are compared on an identical test problem, they differ in the synthesis approach (based on a superstructure or generative without predefined structures), process modelling (equation- or simulator-based), optimization algorithms (global NLP, MINLP, evolutionary) and implementation (dedicated or commercial programs). Results especially highlight the importance of the search space definition, and the need for accurate process models during the synthesis step.

Keywords: Process Synthesis, superstructure, generative approach, separation processes

1. Introduction

In the field of process synthesis, various methods have been proposed to select a set of equipment with their operational conditions and interconnection in a process flowsheet. These methods may vary from empirical approaches using expert judgment or heuristics to simulate a few configurations, towards optimization-based approaches where the process synthesis problem is posed as a mathematical problem, allowing to cover a wider search space than tedious manual iterations with process simulators. In the literature, numerous algorithms and application fields are reported. A common approach is to use a superstructure (Mencarelli et al., 2020), i.e. a postulated set of flowsheet alternatives, and to perform optimization among structural and design variables. Other approaches are *generative*, meaning that process structures are directly proposed from a set of available unit operations, using evolutionary or machine/deep learning algorithms (Nabil et al., 2022), instead of choosing among a predetermined set of structures.

This work is therefore a first attempt at comparing synthesis approaches on a common process synthesis problem. Three approaches are used on an identical test problem. At this stage, performance indicators are mainly: found process solutions, computational efficiency, and ease of use for a process engineer.

2. Synthesis approaches and their comparison

Three approaches (denoted here A, B and C), developed by the paper's authors, are selected for the comparison, summarized here. Main highlights of each approach are also reported in the Table 1, and will be discussed in §3 based on the case study.

Method A (superstructure - GO ; Ramírez-Santos et al., 2018) is superstructure-based using continuous splitter for flow alternatives. Modelling is equation-oriented, as all governing process equations are posed as constraints in the optimization problem. The continuous global optimization (GO) problem is solved with an algorithm combining multi-start strategy, monotonous Basin Hopping and local search. Implementation is performed in AMPL language with KNITRO commercial solver for local searches.

Method B (superstructure - MINLP ; Zhao et al., 2018) is also superstructure-based, using here integer *switches* for flow alternatives. Modelling is simulator-based using a commercial simulator, coupled to an optimization solver (*feasible path or unfeasible path*). The resulting Mixed Integer NonLinear Programming (MINLP) problem is solved by an Ant Colony Optimization algorithm. The graphical process simulator ProSimPlus is used, in communication with MIDACO solver.

Method C (ab-initio - EP ; Neveux, 2018) is a generative approach. Modelling is also simulator-based using an in-house simulator. The optimization problem relies on a two-level decomposition of the process synthesis problem, an upper level for designing the topology (i.e. the process flowsheet) using Evolutionary Programming (EP) and a lower level for evaluating the flowsheet (optimizing the degrees of freedom for a given structure). Implementation is performed in Fortran 2008, with calls to nonlinear optimization solvers for process evaluation (e.g. MIDACO, SLSQP).

While these methods do not represent the entirety of existing approaches, they cover various definitions of the search space (superstructure, generative), system modelling (equation- or simulator-based), optimization algorithms (continuous global optimization, mixed integer) and associated solvers, and software ergonomics (programming language, graphical interfaces).

The comparison procedure is summarized as follows: definition of the synthesis problem (common to all approaches), problem formulation to fit each method's requirements (approach-specific), optimization runs to isolate optimal solutions (approach-specific), validation of the solutions within a common process simulator (all approaches). The results are then compared on the basis of various criteria such as: found solutions and their exclusivity (process structure and parameters), computational time, ease of use for a process engineer, etc.

3. Case study: production of N₂ from air using membrane gas permeation

3.1. Definition of process synthesis problem

The production of N₂ from air using membrane gas permeation is chosen due to the numerous possible stages, recycling and design variables, with non-linear and non-convex economic objective and constraints. Industrial solutions are also known for commercial membrane materials, and prospective solutions can be assessed with other materials. The comparison is done for 90%, 95% and 99.9% N₂ purity, for given materials using fixed permeability, as well as variable materials using permeability as degrees of freedom. This case has already been treated using the method A, refer to Bozorg et al. (2019) for results and process assumptions.

The synthesis problem can be summarized as follows (see Figure 1): produce enriched nitrogen from air at the lowest levelized cost of production (objective function), with given N₂ purity and process outlet pressures (production constraints). Arising degrees of freedom occur at various level: at structural level, the choice of unit operations (membrane stages, compressors, vacuum pumps, mixers and splitters) and their connections (sequences, by-passes, recycling etc.); at unit operation level, each design or operational parameter (e.g. pressures, membrane areas); at 'product' level, the membrane permeability which could be kept fixed to given known values (e.g. commercial membranes) or be optimized within a known region for N₂/O₂ separation.

The basis for membrane calculation is the *MEMSIC* code (Bounaceur et al., 2017), a CAPE-OPEN unit operation software for multicomponent membrane gas separations, validated against a large set of chemical mixtures and membranes. All operations are assumed isothermal (i.e. cooling after compression, and Joule-Thompson effect neglected in membrane) and without pressure drop.

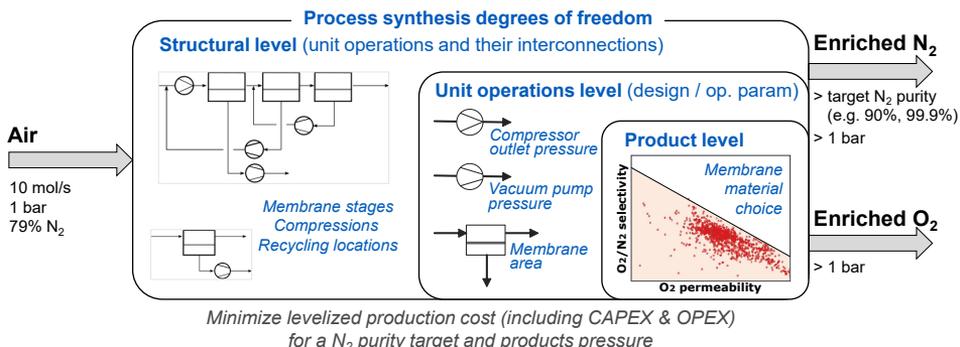


Figure 1: Overview of the process synthesis problem, and associated degrees of freedom

The following bounds are chosen for the design variables: vacuum levels (pressure on the permeate side of membrane) between 0.2 and 1 bar; compression ratio between 1 and 20 for gas compressor; membrane stage-cut (ratio of permeate to feed flowrates) between 0.05 and 0.95 (the required membrane area is calculated by the MEMSIC code knowing this design specification, inlet stream, membrane permeability and pressure ratio around the membrane).

3.2. Problem formulation: approach-specific implementation

Once the process synthesis problem is defined, it needs to be transposed for each approach to fit the method's requirements. We overview here the main particularities, please refer to the original papers for implementation details.

In terms of search space, methods A and B both require a superstructure, taken from Bozorg et al. (2019), allowing between 1 and 3 membrane stages, where all membrane stage outlets (retentate or permeate side) could be mixed and fed to the inlet of a module upstream (by-pass) or downstream (recycle) the current stage. In method C, only the process inlet (air) and outlets (enriched N_2 and O_2) are mandatory, the evolutionary algorithm can use unit operations from a given library (compressor, vacuum pump, membrane stage, mixer, splitter) to construct flowsheets.

In terms of system modelling, methods B and C are simulator-based (feasible path), meaning that the simulator is in charge of calling the membrane MEMSIC code and solving the flowsheet (tear streams) to provide required information to the optimizer (current objective function and process constraints, for the given set of action variables provided by the optimizer). Method A being equation-based, all the process equations are posed as optimization constraints, including flow conservation and unit operation equations. The membrane stage is therefore discretized into a finite number of 'cells' to get an algebraic system of equations from a differential one. This introduces a difference in physical modelling, whose influence will be discussed in §3.4.

In terms of optimization, the type and size of the problem will depend on the method. For method A, the size will depend on the discretization level. For a binary gas mixture (N_2/O_2) and 3 stages, the number of process variables and nonlinear equality constraints is between 748 and 18028 (for, respectively, 20 and 500 discretization cells per membrane stage), in addition to 17 continuous decision variables (splitters, pressures and membrane areas). For methods B and C, only the three production inequalities constraints are required (N_2 purity \geq target, and pressure of both outlets \geq 1 bar), as all process equations (governing unit operation equations and flowsheeting) are solved by the process simulator. For method B, 7 continuous (one design variable per unit operation with a degree of freedom) and 10 integer (switches for flow path) variables are required. For method C, the size will be different for each process evaluation as the topology is designed by the evolutionary algorithm, in this case 1 continuous variable per unit operation.

3.3. Optimization results: found solutions

In this section, we discuss the results obtained by fixing the permeability to 200 GPU (gas permeation unit) for O₂ and 44 GPU for N₂, i.e. values for commercial poly(p-phenylene oxide) membranes (PPO), in order to compare with industrial processes. Obtained results are presented in Figure 2. For each case, the best process structure is drawn as well as main process parameters (upstream and downstream pressure, P_{up} and P_{down} for each membrane stage) and overall process indicators (total membrane surface area A_{tot} , and levelized cost of production).

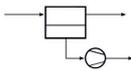
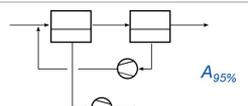
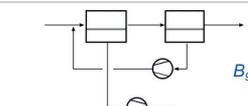
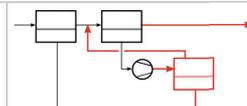
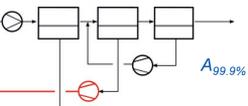
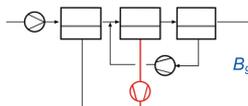
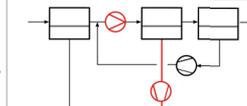
	Method A (superstructure – GO)	Method B (superstructure – MINLP)	Method C (ab-initio – EP)
90% N₂		$F_{up} = 1.0$ bar $F_{down} = 0.2$ bar $A = 2750$ m ²	13.8 €/t_{N₂} $A_{90\%}$ $B_{90\%}$ $C_{90\%}$
95% N₂	 $P_{up} = 1.0/1.0$ bar $P_{down} = 0.2/0.2$ bar $A_{tot} = 5290$ m ² 31.7 €/t_{N₂}	 $P_{up} = 1.0/1.0$ bar $P_{down} = 0.2/0.2$ bar $A_{tot} = 5112$ m ² 31.1 €/t_{N₂}	 $P_{up} = 1.0/1.0/1.0$ bar $P_{down} = 0.2/0.2/0.2$ bar $A_{tot} = 4915$ m ² 30.4 €/t_{N₂}
99.9% N₂	 $P_{up} = 2.9/2.9/2.9$ bar $P_{down} = 0.2/0.2/0.2$ bar $A_{tot} = 4335$ m ² 161.6 €/t_{N₂}	 $P_{up} = 2.7/2.7/2.7$ bar $P_{down} = 0.2/0.2/0.2$ bar $A_{tot} = 3668$ m ² 145.6 €/t_{N₂}	 $P_{up} = 1.0/2.4/2.4$ bar $P_{down} = 0.2/0.2/0.2$ bar $A_{tot} = 5611$ m ² 142.5 €/t_{N₂}

Figure 2: Found solutions for PPO membranes - $X_{yy\%}$ denotes the solution of the process synthesis approach X for the target of yy% N₂ purity

First, results exhibit high similarities with known industrial practices (Prasad et al., 1994), such as: an increasing number of membrane stages (and process complexity) with target N₂ purity, from one stage for 90% to three stages for 99.9%; an extraction of O₂ at the first stage(s) permeate; a recycling of other permeates at the 'N-1' stage; and an economic preference for vacuum operation when possible, without use of compressors up to 95% targeted purity.

More specifically, for 90% purity, all methods find the same simple solution (same flowsheet and design parameters). For 95% purity, both superstructure-based methods provide also identical solutions ($A_{95\%}$ and $B_{95\%}$) with a two-stages flowsheet and very closed design parameters. However, method C provides a better solution ($C_{95\%}$) with a three-stages flowsheet with same pressure levels, leading to a 4% lower membrane area and 2% lower production cost. It is worth mentioning that the flowsheet of solution ($C_{95\%}$) was indeed included in the superstructure of methods A and B, but not found by the optimization algorithms. This result will be validated and discussed in the validation section (§3.4).

For 99.9% purity, all methods propose three-stages flowsheets, with variations in flowsheets. Solution $A_{99.9\%}$ recycles the second stage permeate at the process inlet, whereas $B_{99.9\%}$ and $C_{99.9\%}$ mix with the O₂ process outlet. The best solution ($C_{99.9\%}$) is found by method C that was *not* included in the superstructure of methods A and B, with a flowsheet similar to solution $B_{99.9\%}$, simply consisting in pressurizing the retentate between stage 1 and 2, instead of at the process air inlet. This variant is comprehensible as no additional driving force (total pressure) may be required where the driving force is already high (O₂ partial pressure in retentate side). Putting the compressor after the first stage therefore lower the energy cost, further superstructure should therefore consider retentate recompression as options between each stage. Finally, solution ($A_{99.9\%}$) is significantly worse than $B_{99.9\%}$ and $C_{99.9\%}$, it will also be validated and discussed in the validation section (§3.4).

3.4. Validation: highlighting the importance of model accuracy and search space definition

The purpose of the validation step is to verify that the found solutions are actually retrieved within a common environment, different from the one used for optimization, here the COCO process simulator using the CAPE-OPEN MEMSIC code for membrane calculations.

From the simulations, it appears that all solutions found with various approaches and problem definition (N_2 target purity and membrane material) are validated, except the solution $A_{99.9\%}$ found by method A since the targeted 99.9% purity is not achieved. This is due to low deviation between rigorous and simplified models in the majority of the operating domain, but higher deviations in small regions with high-purity. This point is verified in Figure 3, where optimization runs are performed with increasing number of discretization cells for the simplified membrane model. As the discretization is more precise, both flowsheet and prediction of cost evolve until stabilizing to the same optimum as the one obtained with the rigorous membrane model.

For simplified models (e.g. physic, numeric, metamodel etc.), attention must therefore be placed in the regions of interests in addition to the overall domain. It highlights the need for accurate unit operation models, even for the exploration of the process structure (large) search space.

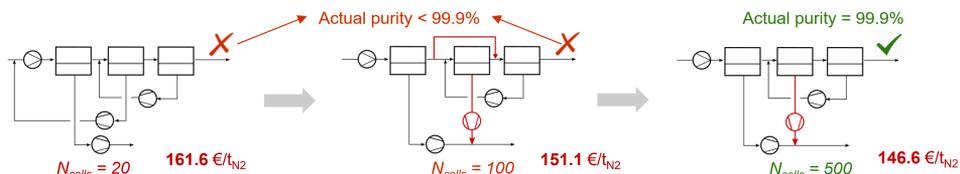


Figure 3: Effect of membrane discretization on synthesis results for method A (simplified model) and validation with the rigorous model under COCO simulator

Regarding solutions found by the method C, the fact that solution $C_{99.9\%}$ is validated confirms the interest of generative methods as this solution was not included in the superstructure of methods A and B. In addition, it is surprising to see that the solution $C_{95\%}$ is validated and that it was included in the superstructure, but was not found by any of the superstructure-based methods, despite various optimization approaches, system modelling and solvers. One possible explanation is on the way the various approaches explore the search space. In a superstructure-based approach, the entirety of the structural search space is explored at once as all foreseen alternatives are transposed into an optimization problem. Generative approaches usually work in two-levels, with an upper level being in charge of the flowsheet, and a lower level of the optimization of a (small) flowsheet. Each sub-optimization problem is therefore smaller and easier to solve.

3.5. Overall comparison

The Table 1 attempts to summarize found solutions, computational efficiency and ergonomics; in addition to details of the methods previously mentioned.

Regarding the found solutions, superstructure-based approaches (methods A and B) provide identical flowsheets and similar design parameters, if a sufficient discretization level is used when using the method A. The interest of the *ab-initio* method (generative) is highlighted as a better structure is found, which were not included in the superstructure of the other two methods, without the need to define a superstructure.

In terms of computational time, implementation differences makes it difficult to assess. It is here estimated based on the number of process evaluations (averaged over 10 random seeds, including constraint calculations for method A and flowsheet convergence for B and C) required to reach the reported solutions in Figure 2 for 99.9%. Superstructure based methods naturally need fewer evaluations than the *ab-initio* method, whose search space is wider.

Finally, in terms of ergonomics, implementations within commercial simulators allow chemical engineers to use familiar tools (graphical process simulators) instead of programming languages, which could promote acceptance of these methods.

Table 1: Overall comparison of synthesis approaches for the case study

	Method A (superstructure - GO)	Method B (superstructure - MINLP)	Method C (ab-initio - EP)
Synthesis approach / search space	Superstructure-based (choice among a postulated set of alternatives)	Superstructure-based (choice among a postulated set of alternatives)	Generative (available unit operations without postulate)
System modelling	Equation-oriented	Simulator-based	Simulator-based
Optimization	Continuous Global Optimization (Monotonic Basin Hopping + local NonLinear Programming)	Mixed Integer NonLinear Programming (solved by Ant Colony Optimization)	Evolutionary Programming (structure generation) + NonLinear Programming (evaluation)
Software tools	Programming (AMPL + KNITRO solver)	Graphical (commercial simulator ProSim + MIDACO solver)	Programming (in-house simulator + MIDACO/SLSQP solvers)
Membrane model	Simplified discretization of MEMSIC equation	Rigorous (direct call to MEMSIC)	Rigorous (direct call to MEMSIC)
Found solutions	+ (if adequate discretization in membrane model)	+ (similar as method A)	++ (similar as methods A and B + new ones not in superstruc.)
Computation time (N° of evaluations)	++ ($\simeq 2 \times 10^6$)	+ ($\simeq 3 \times 10^6$)	\simeq ($\simeq 10 \times 10^6$)
Ergonomics	\simeq (programming)	+ (graphical)	\simeq (programming)

4. Conclusion

From this case study, it is observed the importance of accurate models for unit operations, as simplified models lead to different optimal structures and associated parameters. By extension, the use of surrogate models for process synthesis should be done carefully, with efforts put during the validation of the surrogate in the whole operating domain.

Also, the definition of the structural search space is crucial, as postulating a superstructure restricts the search space. Meaning that if the optimal solutions are not in the superstructure, the optimizer cannot find it. *Generative* methods are therefore less biased and can explore a wider search space. As each approach exhibits various advantages, future works could include hybrid methods, for example using a generative approach for the definition of a restricted search space, then treated by a superstructure-based approach. This comparison can also be extended beyond the present case study to more constrained separation problems, meaning a smaller feasibility region, and more complex systems (multicomponent mixtures, multi-products, reactions, heat integration etc.).

References

- R. Bounaceur, E. Berger, M. Pfister, A. A. Ramirez Santos, E. Favre, 2017. Rigorous variable permeability modelling and process simulation for the design of polymeric membrane gas separation units: MEMSIC simulation tool. *Journal of Membrane Science* 523, 77–91.
- M. Bozorg, B. Addis, V. Piccialli, Álvaro A. Ramírez-Santos, C. Castel, I. Pinnau, E. Favre, 2019. Polymeric membrane materials for nitrogen production from air: A process synthesis study. *Chemical Engineering Science* 207, 1196–1213.
- L. Mencarelli, Q. Chen, A. Pagot, I. E. Grossmann, 2020. A review on superstructure optimization approaches in process system engineering. *Computers & Chemical Engineering* 136, 106808.
- T. Nabil, J.-M. Commenge, T. Neveux, 2022. Generative approaches for the synthesis of process structures. In: *Proceedings of the 14th International Symposium on Process Systems Engineering – PSE 2021+ June 19-23, 2022, Kyoto, Japan*. To appear in *Computer Aided Chemical Engineering*.
- T. Neveux, 2018. Ab-initio process synthesis using evolutionary programming. *Chemical Engineering Science* 185, 209–221.
- R. Prasad, F. Notaro, D. Thompson, 1994. Evolution of membranes in commercial air separation. *Journal of Membrane Science* 94 (1), 225–248.
- A. Ramírez-Santos, M. Bozorg, B. Addis, V. Piccialli, C. Castel, E. Favre, 2018. Optimization of multistage membrane gas separation processes. example of application to CO₂ capture from blast furnace gas. *Journal of Membrane Science* 566, 346–366.
- Q. Zhao, T. Neveux, M. Mecheri, R. Privat, P. Guittard, J. N. Jaubert, 2018. Superstructure optimization (MINLP) within ProSimPlus simulator. In: *28th European Symposium on Computer Aided Process Engineering*. Vol. 43 of *Computer Aided Chemical Engineering*. pp. 767–772.

Bifurcation analysis of combined agglomeration and layering granulation in fluidized bed spray processes

Eric Otto¹, Robert Dürr² and Achim Kienle^{1,3}

¹*Automation/Modelling, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*

²*Engineering Mathematics, Magdeburg-Stendal University of Applied Sciences, Breitscheidstraße 2, 39114 Magdeburg, Germany*

³*Process Synthesis and Process Dynamics, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany*

Abstract

Granulation in fluidized beds is an important unit operation in chemical engineering, finding widespread use in the production of food, fertilizer and pharmaceuticals in particulate form. The main particle size-enlargement mechanisms during this process are layering growth and agglomeration. Regarding the former, the possible occurrence of self-sustained oscillations during continuous operation with sieve-mill-cycle is well known. For agglomeration such phenomena have not been observed yet. The goal of this contribution is to investigate process stability for simultaneous layering growth and agglomeration. Therefore in this contribution a population balance based bifurcation analysis is conducted. It can be shown that for a wide range of process conditions an increased partition of agglomeration has a stabilizing effect on the process dynamics.

Keywords: fluidized bed, granulation, agglomeration, layering growth, bifurcation analysis

1. Introduction

Spray agglomeration and spray granulation (layering growth) in fluidized beds are size-enlargement unit operations, which are applied frequently during the production of goods in particulate form, e.g. in the chemical, food or pharmaceutical industry. Due to high mass and heat transfer rates in the fluidized bed both processes are highly efficient. Furthermore desired product particle properties can be influenced by multiple adjustable process conditions (Strenzke et al., 2020; Neugebauer et al., 2020). The mechanisms of size-enlargement are presented schematically in Fig. 1 for both processes, having in common that a liquid solution is sprayed onto the fluidized particle bed. During layering growth, the liquid is deposited on the surface and dried by the fluidization medium, resulting in continuous particle growth of an onion-like particle structure (Fig. 2). During an agglomeration process however the liquid solution on the

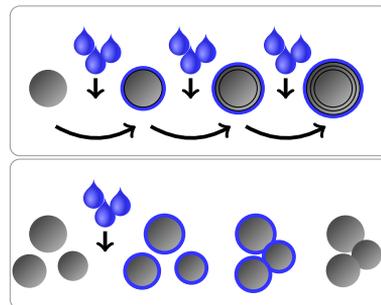


Figure 1: Schematic representation of the layering growth process (top) with continuous addition and drying of a liquid solution and agglomeration process (bottom) with solution addition, collision and drying.

particle, after collision with another particle and drying, acts as a binder, permanently connecting both particles, resulting in a blueberry-like cluster.

In general both layering growth and agglomeration can occur simultaneously in a spray fluidized bed, depending on the operation conditions. A variety of research contributions aims at identifying the dominant size enlargement mechanism under different process conditions (Villa et al., 2016; Rieck et al., 2020), which is important since the final particle properties depend on it. For example agglomerates usually are more porous than granules obtained by layering growth and therefore possess less mechanical strength. Besides of differences in the product particle properties, the dynamical behavior of the two processes differs drastically with respect to time constants and stability. From theoretical investigations (Radichkov et al., 2006) which were confirmed experimentally (Neugebauer et al., 2019) it is known that the particle size distribution in layering growth processes with sieve-mill-cycle as presented in Fig. 3 can exhibit self-sustained oscillations under certain process conditions. These oscillations are highly undesired from a production point of view and can be suppressed using suitable process control algorithms (Neugebauer et al., 2020; Palis and Kienle, 2014). To the author's best knowledge such behavior was not observed in agglomeration processes, a bifurcation analysis for a generic process did not reveal unstable regions (Bück et al., 2016). This leads to the question whether and under which circumstances oscillations can occur in a process where both agglomeration and granulation occur simultaneously. Vesjolaja et al. (2020) presented a process model for a drum granulation process, incorporating both growth mechanisms. They found oscillatory behavior, however a thorough bifurcation analysis was not conducted. This contribution aims at shedding light on the question of stability for the fluidized bed process by conducting a multi-parameter bifurcation analysis using a generic population balance process model containing a layering growth and an agglomeration term.

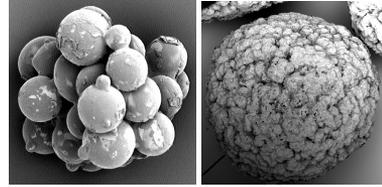


Figure 2: A blueberry-like agglomerate (left) and a spherical granule from layering growth (right).

2. Process Description and Model

In this contribution a generic, continuous fluidized bed spray process with external sieve-mill-cycle as presented in Fig. 3 is considered. Particles are fed to and withdrawn from the fluidization chamber, where the liquid solution is sprayed onto the particle surface and the size-enlargement occurs. The outlet particles are separated into three fractions by sieving according to the desired product size specifications. While the oversized particle fraction is ground up and recycled together with the undersized fraction, the product fraction is removed from the process.

For describing the evolution of a particle size distribution, the population balance equation is an established modeling framework (Ramkrishna, 2000). Here the evolution of a *volume*-based number density distribution $n(t, v)$ is described by the following partial differential equation

$$\frac{\partial n(t, v)}{\partial t} = \dot{n}_{\text{growth}}(t, v) + \dot{n}_{\text{agg}}(t, v) + \dot{n}_{\text{feed}}(t, v) - \dot{n}_{\text{out}}(t, v) + \dot{n}_{\text{mill}}(t, v) + \dot{n}_{\text{fines}}(t, v) \quad (1)$$

where the left-hand side describes the size-dependent accumulation of particles and the right-hand side accounts for the respective number fluxes. Typically the agglomeration term \dot{n}_{agg} is given in the following volume-based formulation (Hulburt and Katz, 1964)

$$\dot{n}_{\text{agg}} = \frac{1}{2} \int_0^v \beta(t, u, v-u) n(t, u) n(t, v-u) du - \int_0^\infty \beta(t, v, u) n(t, v) n(t, u) du. \quad (2)$$

The second-order agglomeration kinetics are described by the so-called agglomeration kernel $\beta(t, u, v)$ which describes the rate at which particles with volumes v and u collide and form a new particle. In this contribution we separate the kernel into a size-independent particle collision rate β_0 and a size-dependent coalescence probability $\beta_c(u, v)$ with values between 0 and 1, i.e. $\beta(u, v) = \beta_0 \beta_c(u, v)$. This kernel model is equally simple and flexible due to its generality. In a fluidized bed process the coalescence probability is a function depending on a variety of process and material specific conditions. To keep the model simple, we choose a generic two-parametric function $\beta_c = (u + v)^a / (uv)^b$ known as the Kapur kernel, which allows us to simulate different types of coalescence by varying the parameters (Otto et al., 2021). In order to keep values of β_c between 0 and 1, the kernel is normalized.

In contrast to this the layering growth term is introduced depending on the particle diameter (Mörl et al., 2007)

$$\dot{\tilde{n}}_{\text{growth}}(t, d) = -G \frac{\partial \tilde{n}(t, d)}{\partial d} = -\frac{\dot{m}_{\text{spray}}}{\rho A_{\text{tot}}} \frac{\partial \tilde{n}(t, d)}{\partial d}. \quad (3)$$

Here $\tilde{n}(t, d)$ describes the *diameter*-dependent number density distribution. In this contribution we assume that all particles are approximately spherical, i.e. $v = \pi/6d^3$, so that all *diameter*-based number density distributions $\tilde{n}(t, d)$ can be transformed into volume-based distributions $n(t, v)$ and vice-versa using the conservation of particle number in an infinitesimal class $n(t, v) dv = \tilde{n}(t, d) dd$.

The layering growth kinetics are described by the growth rate G which depends on the total particle surface A_{tot} and the liquid spray rate \dot{m}_{spray} as well as the liquid density ρ .

Previous studies have shown that the particle mill and especially the size of milled particles play a major role in the emergence of oscillations. Thus, the milling model is presented

$$\tilde{n}_{\text{mill}}(t, d) = \frac{\tilde{q}_{0, \text{mill}}(t, d)}{\mu_3(\tilde{q}_{0, \text{mill}}(t, d))} \mu_3(\tilde{n}_{\text{over}}(t, d)),$$

where $\tilde{q}_{0, \text{mill}}$ is the normalized number density distribution of particles after milling, \tilde{n}_{over} is the oversized particle size distribution and μ_3 is the third moment which is proportional to the total volume (Radichkov et al., 2006). For simplicity $\tilde{q}_{0, \text{mill}}$ is assumed to normalized Gaussian with the milling diameter μ_{mill} and standard deviation σ_{mill} . It is assumed that the milling diameter can be adjusted (Neugebauer et al., 2019).

The terms in Eq. (1), that are not described until now are part of the process model, they do however not play a major role with respect to this bifurcation analysis. Therefore we refrain from describing them explicitly in this contribution and refer the reader to Neugebauer (2020) for detailed information including parameters. We want to stress that the process model presented here is the same as presented in Neugebauer (2020) with exception of the additional agglomeration term. The volume-based formulation here can be transformed into the diameter-based formulation without any approximations and are therefore equivalent. Therefore the results are indeed comparable.

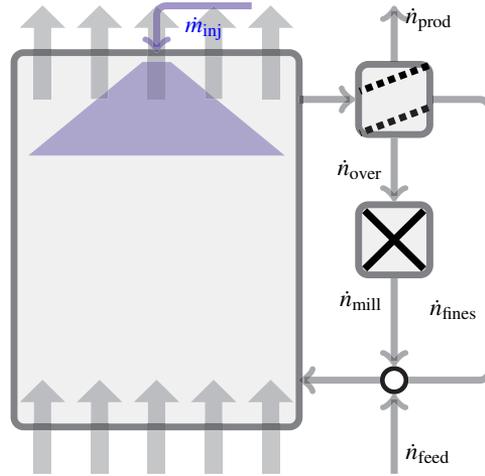


Figure 3: Schematic representation of the fluidization chamber and the sieve-mill-cycle.

3. Bifurcation Analysis

In order to generate stability maps for the variation of two general process parameters a simple prediction-correction scheme is used. To this end steady states are computed by optimization on a fixed, equidistant parameter grid. At every grid node (i, j) the minimization problem

$$n_{\text{cor}}^{(i,j)} = \arg \min_n \left\| \frac{\partial n}{\partial t} (v, \mathbf{p}^{(i,j)}) \right\|_1 \quad (4)$$

s.t. $n(v) \geq 0$

is solved (correction) using a steady state distribution from the adjacent grid point $(i-1, j)$ as initial guess (prediction).

The initial distribution predictions at $(0, j)$ in the stable area are generated by solving the PDE numerically over a long time horizon. The local stability is obtained by linearizing around the steady states and computing eigenvalues. In order to compute the minimization and for the numerical solution of the PDE the volume coordinate in the population balance is discretized using a finite volume method on a geometric grid with $N = 200$ classes. For the agglomeration term the discretization algorithm presented in Singh et al. (2015) is applied. The discretization of the other terms in Eq. 1 is straight forward. For the minimization of the L_1 -residuals of the time-differential, the Matlab built-in trust-region gradient-descent algorithm is utilized.

4. Results

Previous contributions (Radichkov et al., 2006; Neugebauer et al., 2020) have identified the milling grade μ_{mill} as primary influence on the stability of the layering growth process. Based on this we choose

μ_{mill} as the first bifurcation parameter. In order to investigate the influence of an increased rate of agglomeration events, we choose $\beta_0 \in [10^{-14}, 10^{-12}]$ as second bifurcation parameter, where the lower interval bound accounts for a negligible amount of agglomeration events compared to the growth by layering. Therefore the stability results for $\beta_0 = 10^{-14}$ are the same as in Neugebauer (2020). The influence of additional process parameters is investigated by conducting the parameter continuation for selected values.

The influence of an increased amount of injected solution \dot{m}_{inj} is investigated, since it influences the growth rate G proportionally. The results presented in Fig. 5 indicate an increased region of instability for larger values of \dot{m}_{inj} , which is in qualitative accordance with Radichkov et al. (2006). A change of the coalescence kernel β_c is modeled by varying the kernel parameter

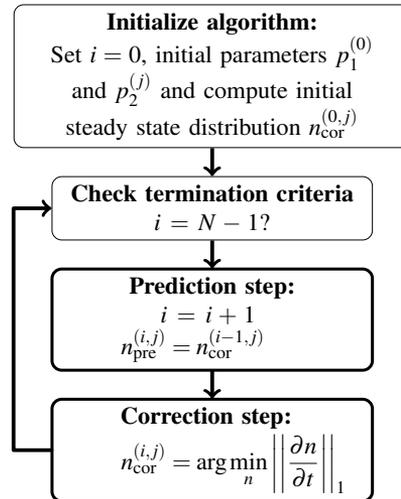


Figure 4: Parameter Continuation Scheme for one parameter.

Name	Symbol	Value	Unit
Bed mass	m_{bed}	100	kg
Mill grade	μ_{mill}	[0.1, 0.8]	mm
Collision rate	β_0	$[10^{-14}, 10^{-12}]$	s^{-1}
Injection rate	\dot{m}_{spray}	100	kg h^{-1}
Kapur parameters	a	0	
	b	0	
Feed rate	\dot{m}_{enuc}	72	kg h^{-1}

Table 1: Nominal process parameters.

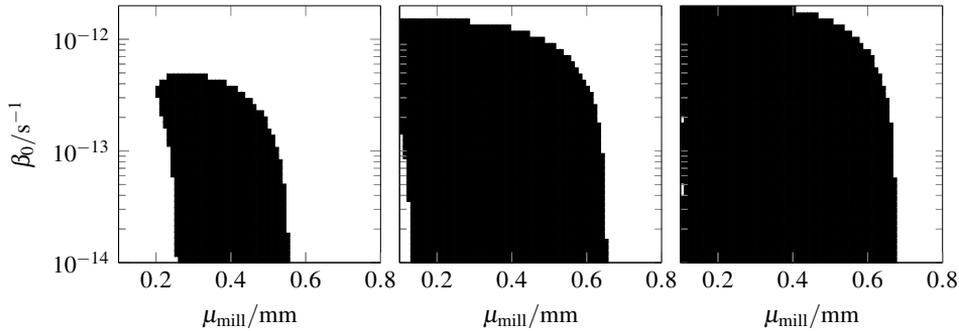


Figure 5: Stability map for a variation of the binder injection rate. Left: $\dot{m}_{\text{spray}} = 80 \text{ kg/h}$, middle: $\dot{m}_{\text{spray}} = 150 \text{ kg/h}$, right: $\dot{m}_{\text{spray}} = 200 \text{ kg/h}$. The region of instability is colored black.

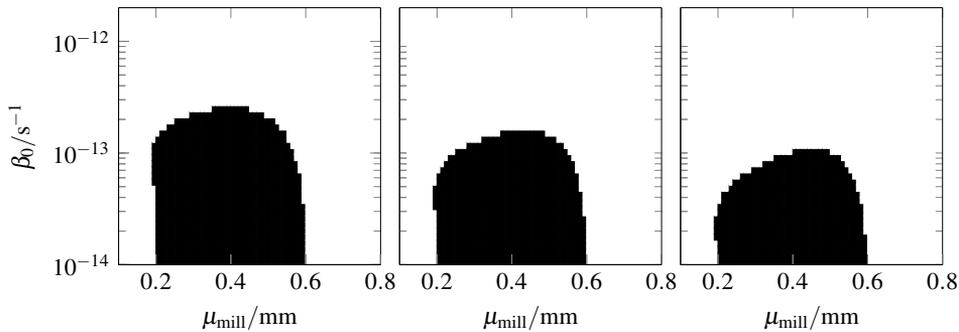


Figure 6: Stability map for a variation of kernel parameters: $b = 0.1$ (left), $b = 0.2$ (middle) and $b = 0.3$ (right). The region of instability is colored black

b . Increasing it results in preferential coalescence of smaller particles compared to coalescence of larger particles. As depicted in Fig. 6 increasing b from 0.1 to 0.2 to 0.3 reduces the region of instability slightly. Furthermore, the influence of the external nuclei feed rate is investigated and presented in Fig. 7. The region of instability is reduced significantly by increasing the feed rate, which is again in accordance with Radichkov et al. (2006).

The transitions from white to black areas in Figs. 5-7 represent supercritical Hopf bifurcation points, i.e. transition from a stable fixed point to a stable limit-cycle (Seydel, 1998).

5. Conclusion and Future Work

In this contribution a model based bifurcation analysis was presented for a fluidized bed spray process with layering growth and particle agglomeration. The process stability was investigated for different plant and kinetic parameters. The results generally indicate that self-sustained oscillations in the particle size distribution are dampened with an increasing number of agglomeration events taking place. This result is in agreement with previous observations on the stability of agglomeration processes.

Future contributions will be concerned with a more detailed bifurcation analysis for an extended process model including the thermal process conditions such as fluid temperature and moisture as

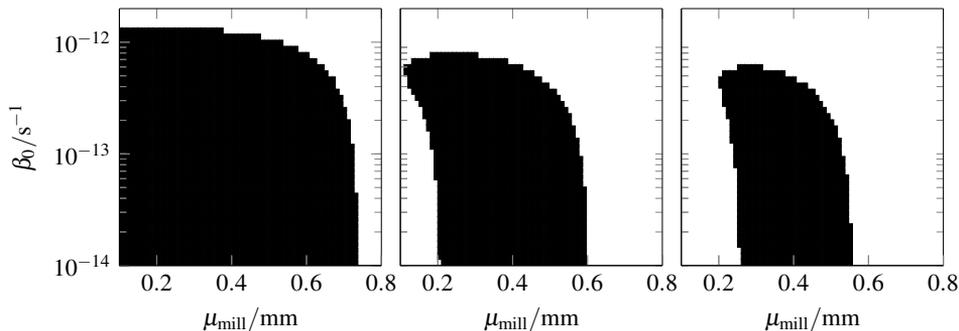


Figure 7: Stability map for variations of the external feed rate: $\dot{m}_{\text{enuc}} = 0 \text{ kg/min}$ (left), $\dot{m}_{\text{enuc}} = 1.2 \text{ kg/min}$ (middle) and $\dot{m}_{\text{enuc}} = 1.5 \text{ kg/min}$ (right). The region of instability is colored black

well as considering the particle porosity.

Acknowledgments

This work is funded by the European Regional Development Fund (ERDF) project "Center of Dynamic Systems". The financial support is hereby gratefully acknowledged.

References

- A. Bück, M. Wegner, C. Neugebauer, S. Palis, E. Tsotsas, 2016. Bifurcation analysis of process stability of continuous fluidized bed agglomeration with external product classification. In: Z. Kravanja, M. Bogataj (Eds.), 26th European Symposium on Computer Aided Process Engineering. Vol. 38 of Computer Aided Chemical Engineering. Elsevier, pp. 1881–1886.
- H. Hulbert, S. Katz, 1964. Some problems in particle technology: A statistical mechanical formulation. *Chemical Engineering Science* 19 (8), 555 – 574.
- L. Mörl, S. Heinrich, M. Peglow, 2007. Chapter 2 fluidized bed spray granulation. In: A. Salman, M. Hounslow, J. Seville (Eds.), *Granulation*. Vol. 11 of Handbook of Powder Technology. Elsevier Science B.V., pp. 21–188.
- C. Neugebauer, 2020. Model, dynamics and control of continuous fluidized bed layering granulation processes. Ph.D. thesis, Otto-von-Guericke-Universität Magdeburg.
- C. Neugebauer, A. Bück, A. Kienle, 2020. Control of particle size and porosity in continuous fluidized-bed layering granulation processes. *Chemical Engineering & Technology* 43 (5), 813–818.
- C. Neugebauer, E. Diez, A. Bück, S. Palis, S. Heinrich, A. Kienle, 2019. On the dynamics and control of continuous fluidized bed layering granulation with screen-mill-cycle. *Powder Technology* 354, 765 – 778.
- E. Otto, R. Dürr, G. Strenzke, S. Palis, A. Bück, E. Tsotsas, A. Kienle, 2021. Kernel identification in continuous fluidized bed spray agglomeration from steady state data. *Advanced Powder Technology*.
- S. Palis, A. Kienle, 2014. Discrepancy based control of particulate processes. *Journal of Process Control* 24 (3), 33–46.
- R. Radichkov, T. Müller, A. Kienle, S. Heinrich, M. Peglow, L. Mörl, 2006. A numerical bifurcation analysis of continuous fluidized bed spray granulation with external product classification. *Chemical Engineering and Processing: Process Intensification* 45 (10), 826–837, particulate Processes A Special Issue of Chemical Engineering and Processing.
- D. Ramkrishna, 2000. *Population Balances: Theory and Applications to Particulate Systems in Engineering*. Academic Press.
- C. Rieck, A. Bück, E. Tsotsas, 2020. Estimation of the dominant size enlargement mechanism in spray fluidized bed processes. *AIChE Journal* 66 (5), e16920.
- R. U. Seydel, 1998. *Practical Bifurcation and Stability Analysis*. Springer, New York, NY.
- M. Singh, J. Kumar, A. Bück, E. Tsotsas, 08 2015. A volume consistent discrete formulation of aggregation population balance equation. *Mathematical Methods in the Applied Sciences*.
- G. Strenzke, R. Dürr, A. Bück, E. Tsotsas, 2020. Influence of operating parameters on process behavior and product quality in continuous spray fluidized bed agglomeration. *Powder Technology* 375, 210 – 220.
- L. Vesjolaja, B. Glemmestad, B. Lie, 2020. Double-loop control structure for rotary drum granulation loop. *Processes* 8 (11).
- M. P. Villa, D. E. Bertín, I. M. Cotabarrén, J. Piña, V. Bucalá, 2016. Fluidized-bed melt granulation: Coating and agglomeration kinetics and growth regime prediction. *Powder Technology* 300, 61–72, 7th International Granulation Workshop 2015: Granulation across the length scales.

Enhancement of energy saving of reverse osmosis system via incorporating a photovoltaic system

Alanood A. Alsarayreh ^a, Mudhar A. Al-Obaidi ^b, Saad S. Alrwashdeh ^c, Raj Patel ^d and Iqbal M. Mujtaba ^d, *

^a *Chemical Engineering Department, Faculty of Engineering, Mu'tah University, Al Karak – Jordan*

^b *Middle Technical University, Technical Institute of Baquba, Baquba, Dayala – Iraq*

^c *Mechanical Engineering Department, Faculty of Engineering, Mu'tah University, , P.O Box 7, Al-Karak 61710 Jordan*

^d *Chemical Engineering Department, Faculty of Engineering and Informatics, University of Bradford*

Abstract

Reverse Osmosis (RO) desalination capacity has been increasing steadily over the years. However, the necessity of developing more efficient designs has become more crucial due the increased demands for freshwater. This study shows the development of a conceptual design of an efficient energy medium-scale RO brackish water desalination plant of the Arab Potash Company (APC) located in Jordan. In this regard, a photovoltaic (PV) solar farm has been combined with the original RO process to generate electricity in a sustainable way and decrease the overall energy consumption. To obtain an accurate estimation of the process performance, this work shows the utilisation of an efficient process model for the spiral wound RO process. Also, an accurate model for the photovoltaic system was obtained from the literature to conduct preliminary calculations of the proposed new configuration. The simulation results of the proposed RO design are compared to the original ones of APC and showed a considerable energy saving.

Keywords: Brackish water; Reverse osmosis; Photovoltaics; RO-PV system; energy consumption

1. Introduction

Reverse Osmosis (RO) membrane desalination systems are considered to be the most superior technology for desalinating brackish water and producing high-quality water with relatively low energy consumption (Alsarayreh et al., 2020). For medium sized RO desalination systems, the energy consumption is estimated to be about 3.5-4 kWh/m³ and 19 kWh/m³ for larger scale desalination system (Abdelkareem et al., 2018). However, there is an increased interest to promote the overall performance of RO system with an alleviated energy consumption. In this regard, the utilisation of renewable energy, especially solar energy, to power RO system desalination has proved to be one of the most viable and economic options to mitigate the overall energy consumption. The integration of solar energy systems with RO system is mainly to drive the high-pressure pumps (Nayar et al., 2017). An integration of RO systems and photovoltaic (PV) is one of the most successful and cost-effective desalination hybrid systems (Filippini et al., 2019). Several researchers have evaluated the feasibility of installing small and medium scales RO seawater and brackish water desalination systems powered by a PV system.

A small scale PV and brackish water two-stage configuration RO desalination system has been constructed by Alghoul et al. (2016). They concluded that RO system can produce 5.1 m^3 of freshwater during 10 hours a day at a specific energy consumption of 1.1 kWh/m^3 . The viability of using PV solar energy for brackish water RO desalination process in Jordan valley has been studied by Alsarayreh et al. (2017). They confirmed that employing PV systems for brackish water desalination is cost-effective with reduction of more than 20% in the initial cost.

While many renewable energy-powered RO units have been developed for seawater and brackish water desalination systems, only a few attempts have been made to investigate the feasibility of combining PV solar energy to an industrial medium scale- brackish water RO desalination system of multistage, multi pass configuration. Thus, it is feasible to test such integrated system and explore the maximum energy saving. Unarguably, this would introduce a reliable and cost-effective desalination system especially for coastal areas of intense sun radiation. Thus, this research focuses on installing PV solar farm to be integrated to the multistage multi-pass brackish water RO system of the Arab Potash Company (APC), located in Jordan.

The assessment of the hybrid system will be carried out based on intensive simulation based models developed for RO (developed by the same authors) and PV (available in the literature) systems. In this regard, the total energy consumption of RO desalination system will be firstly determined as a base case and the required number of PV panels and associated energy saving will be investigated for three different selected PV types.

2. Description of RO desalination plant powered by PV system

The layout of a brackish water RO powered by solar energy for the APC plant with capacity $1200 \text{ m}^3/\text{day}$ is shown in Fig. 1. It comprises of two passes with permeate and retentate reprocessing designs. The 1st pass has two stages of pressure vessels organized in the following order: (4:2). The 2nd pass contains three stages of pressure vessels organized in the following order: (2:1:1). The 1st pass's permeate is sent into the second pass for additional polishing. The 1st pass's high-concentration stream is drained. The low-concentration streams of the 2nd pass are gathered to make the high-quality water with salinity of 2 ppm. The 2nd pass's high-concentration stream is returned to the 1st pass's raw feed water. As shown in Fig. 1, a photovoltaic (PV) solar farm is integrated with a RO system to provide the necessary amount of power to drive the pumps.

3. Design of PV system

This study will determine the number of panels required to supply the RO system for each type of PV system with evaluating the generated power and associated construction area. The RO plant feed characteristics are 1098.62 ppm, $74 \text{ m}^3/\text{h}$, $25 \text{ }^\circ\text{C}$, and 9.22 atm of brackish water salinity, flow rate, temperature and pressure, respectively.

According to the simulation-based model developed by Al-Obaidi et al. (2018) based on the above feed conditions, the total energy consumption of the original RO plant was 0.837 kWh/m^3 for a water productivity rate of $48.6 \text{ m}^3/\text{h}$ ($1166.4 \text{ m}^3/\text{day}$). In this study, three type of solar panels were selected to be integrated to the RO plant. The first type is made by Monocrystalline, Sun-Power 315-SOLAR PAN, with an efficiency of 19.3% and 315 W, as maximum power (Al-Najideen and Alrwashdeh (2017). The second type is made by the TopSun Energy Ltd with an efficiency of 17.2% (Filippini et al. (2019). The third type is made by the FSM-155 W- Mono-crystalline, with an efficiency of 16% (Abd Elbar and Hassan, 2020). The inverter type GermanPV TRIO-27.6-TL-OUTD-S2X manufactured by GermanPV Inc has been used in this study and has a high value of CEC

efficiency of about 98% (Al-Najideen and Alrwashdeh, 2017). Tables 1, and 2 show the characteristics of selected types of solar panel and inverter device, respectively.

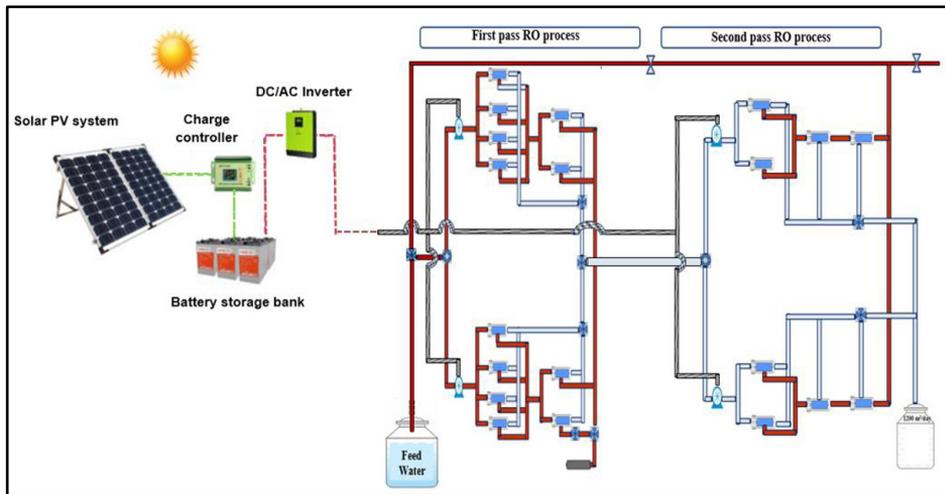


Fig. 1. A schematic representation of a PV-RO system of APC

Table 1. Technical features of the photovoltaic solar panel (Al-Najideen and Alrwashdeh, 2017, Filippini et al., 2019, Abd Elbar and Hassan, 2020)

Properties of the PV module	Sun power315-SOLAR PAN	TopSun Energy Ltd	FSM-155 W-Mono-crystalline
Maximum power output (P_m)	315 W	440 W	155 W
Maximum system voltage	1000 V	1000 V	1000 V
Module efficiency (η)	19.3%	17.2%	16%
Rated current at maximum power (I_m)	5.76 A	8.86 A	8.15 A
Rated voltage at maximum power (V_m)	54.7 V	49.67 V	30.7 V
Module length (L)	1.559 m	1.960 m	1.482 m
Module height (h)	1.046 m	1.308 m	0.676 m
Module depth (d)	0.046 m	0.040 m	0.035 m
solar radiation (G) at 25 °C	1000 W/m ²	1000 W/m ²	1000 W/m ²

Table 2. Inverter specifications (Al-Najideen and Alrwashdeh, 2017)

Properties for Input DC	Value	Properties for Output AC	Value
Max. recommended PV power	30 kW	Maximum efficiency	98 %
Max. DC voltage	1000 V		
Rated MPPT voltage range	252- 970 V	Nominal AC rated power	27.6 kW
Nominal DC rated power	28.6 kW	Maximum AC output power	30 kW
Number of MPP tracker inputs	2	Nominal AC rated line voltage	400 V
Max. input current / per MPP tracker input	64 A / 32 A	AC voltage rang	320–480 V
Number of DC input pairs (parallel connection)	5	AC maximum output current	45 A

4. Modelling of RO-PV system of APC

Al-Obaidi et al. (2018) developed a steady-state model for the APC medium-scale brackish water RO desalination system. The detailed model was validated against actual APC data before being used to study plant performance under variable operating conditions. For the convenience of the reader, it is recommended to revise Al-Obaidi et al. (2018) for the model's details of the RO system. This model was recently upgraded by including a specific sub model to measure the power generated by the PV system.

5. Modelling of Photovoltaic System

The required energy power of RO desalination $E_{Power Des.}$ is estimated using Eq. 1, (Filippini et al., 2019).

$$E_{Power Des.} = M_{RO} \times E_{Cons. RO} \quad (1)$$

where M_{RO} , and $E_{Cons. RO}$, are the productivity of freshwater of RO process (m³/day), and energy requirement for RO (evaluated in kWh/m³), respectively.

The on-grid type is the most basic PV installation system model in which the load is supplied by the user and the suggested inverter is simply equal to the nominal array power. However, the energy produced by the array equivalent to the energy available to the grid in MWh and can be calculated by the following equation

$$E_{grid} = E_{Power Des.} \times \eta_{inv.} \quad (2)$$

where $\eta_{inv.}$, is the inverter efficiency.

Based on the grid configuration, the grid may not be able to absorb all the supplied energy. In this regard, the actual amount of energy delivered can be calculated using Eq. 3 (Al-Najideen and Alrwashdeh, 2017).

$$E_{actual} = E_{grid} \times \eta_{absorbed} \quad (3)$$

where $\eta_{absorbed}$, is the efficiency grid absorption rate

The energy power for PV module $E_{Power Module}$, can be calculated using Eq. 4

$$E_{Power Module} = \eta_{Module} \times A_{Module} \times G \quad (4)$$

where η_{Module} , A_{Module} , and G are the solar panel module efficiency (%), the PV module area (m²) and the solar radiation in (kW/m²), respectively.

The required number of PV panels is calculated as follows:

$$NPV, Panels = \frac{E_{Power Des.}}{E_{Power Module}} \quad (5)$$

Eq. 6 can also be used to calculate the hourly total power generated by the PV system, measured in kW and based on the value of the PV array (Mokheimer et al., 2013);

$$Power_{PV} = N_{PV, Panels} \times V_{pv} \times I_{pv} \quad (6)$$

where $N_{PV, Panels}$, V_{pv} , and I_{pv} are the number of the PV modules, the voltage of each PV module and current of each PV module, respectively.

6. Evaluation of PV-RO system for various types of solar panels

As stated in section 3, the required power of RO desalination $E_{Power Des.}$ is estimated from Eq. 1 to be 976.28 kWh/day (40.67 kW), and annually about 356.34 MWh at the selected operating conditions. Using Eqs. 2 and 3, the annual energy available to the grid is estimated to be 349.2 MWh and the actual energy delivered annually is 345.7 MWh, with an inverter efficiency of 98% and grid absorption rate of 99%. The power of three PV modules, was estimated based on Eq. 4: (Sun power315-SOLAR PAN), (TopSun Energy Ltd), and (FSM-155 W-Mono-crystalline) and are 0.31 kW, 0.44 kW, 0.16 kW, respectively.

At standard test conditions for SOLAR PAN module, the rated voltage for each panel is 54.7 V and the rated current is 5.76A, whereas the maximum system voltage is 1000 V, according to the solar panel data sheet and Table 1. Dividing the maximum system voltage by the panel rated voltage gives the number of panels in series orientation; that is $18.3 = 18$ panels in series. Therefore, 7 parallel lines are required each with 18 panels connected in series (5.67 kW, 984.6 V, 5.76A). According to the panel data sheet in Table 1, the panel area with its frame is 1.63 m². Thus, the required area is 210.4 m².

For the TopSun Energy module, the rated voltage for each panel is 49.67 V and the rated current is 8.86 A, whereas the maximum system voltage is 1000 V. The number of panels in series orientation is $20.1 = 20$ panels. Therefore, 4.5 parallel lines are required each with 20 panels connected in series (8.8 kW, 993.4 V, 8.86 A). Thus, the panel area with its frame is 2.56 m² with required area of 235.8 m².

For the W-Mono-crystalline module, the rated voltage for each panel is 30.7 V and the rated current is 8.15 A, whereas the maximum system voltage is 1000 V. The number of panels in series orientation is $32.6 = 33$ panels, and therefore, 7.7 parallel lines are required each with 30 panels connected in series (5.12 kW, 1013.1 V, 8.15 A). Finally, the panel area with its frame is 1.002 m² and therefore 254.5 m² is required to construct this type of solar panel.

Using Eq. 5, the number of PV panels required for the Sun power315-SOLAR PAN, TopSun Energy Ltd and FSM-155 W-Mono-crystalline modules are 129, 92 and 254, respectively.

Table 3 shows the calculations of total power generated by the three selected PV systems in kW and based on the value of the PV array. The calculations of the generated power are also based on 100% energy saving. For the convenience of the plant operator, Table 3 shows the calculation of the power generated of three PV modules with 90%, 80%, 70%, 60% and 50% of energy saving. Also, the required construction area and the number of panels are calculated for each option. Logically, the reduction of energy saving implies lesser required construction area and panels. Using Eq. 6, the PV power required based on the number of panels for three selected types of PV modules are 40.674 kW, 40.677 kW, and 40.672 kW.

Table 3. The energy saving and construction area of three types of PV of different number of panels

Sun power315-SOLAR PAN				TopSun Energy Ltd Module type				FSM-155 W-Mono-crystalline			
No. panels	Power PV (kW)	Energy saving (%)	Area (m ²)	No. panels	Power PV (kW)	Energy saving (%)	Area (m ²)	No. panels	Power PV (kW)	Energy saving (%)	Area (m ²)
129	40.67	100	210.4	92	40.67	100	236.5	163	40.67	100	163
116	36.6	90	189.4	83	36.6	90	213	146	36.6	90	147
103	32.5	80	168.3	74	32.5	80	189	130	32.5	80	130
90	28.4	70	147.3	65	28.5	70	166	113	28.5	70	114
77	24.4	60	126.2	55	24.4	60	142	98	24.4	60	98
64	20.3	50	105.2	46	20.3	50	118	81	20	50	82

Most importantly, the PV module type TopSun Energy Ltd is the most feasible and favorable option to generate the required power of RO desalination process with the least number of panels compared to the other tested modules (Table 3). However, the

construction area required to install the Sun power 315-SOLAR PAN Module type is the lowest one compared to the other tested ones.

7. Conclusions

This paper presents the integration system of PV and RO brackish water desalination plant of Arab Potash Company. The simulation of the proposed PV-RO system is carried out based on a developed model of subsystems. The simulation results confirm the possibility of gaining full energy saving for three types of PV modules. The calculations have determined the number of panels and construction area required for each PV system. More specifically, the PV module of TopSun Energy Ltd has been selected as the most feasible option. This study therefore lays the foundation for further improvement to lowering the overall freshwater production cost of the RO desalination plant of APC.

References

- A.R. Abd Elbar, H. Hassan, 2020, Enhancement of hybrid solar desalination system composed of solar panel and solar still by using porous material and saline water preheating, *Solar Energy*, 204, 382-394.
- M.A. Abdelkareem, M.E.H. Assad, E.T. Sayed, B. Soudan, 2018, Recent progress in the use of renewable energy sources to power water desalination plants, *Desalination*, 435, 97-113.
- M.I. Al-najideen, S.S. Alrwashdeh, 2017, Design of a solar photovoltaic system to cover the electricity demand for the faculty of Engineering-Mu'tah University in Jordan, *Resource-Efficient Technologies*, 3, 440-445.
- M. Al-Obaidi, A. Alsarayreh, A. Al-hroub, S. Alsadaie, I.M. Mujtaba, 2018, Performance analysis of a medium-sized industrial reverse osmosis brackish water desalination plant, *Desalination*, 443, 272-284.
- M.A. Alghoul, P. Poovanaesvaran, M.H. Mohammed, A.M. Fadhil, A.F. Muftah, M.M. Alkilani, K. Sopian, 2016, Design and experimental performance of brackish water reverse osmosis desalination unit powered by 2 kW photovoltaic system, *Renewable Energy*, 93, 101-114.
- A. Alsarayreh, M. Majdalawi, R. Bhandari, 2017, Techno-economic study of PV powered brackish water reverse osmosis desalination plant in the Jordan Valley, *International Journal of Thermal and Environmental Engineering*, 14, 83-88.
- A.A. Alsarayreh, M.A. Al-Obaidi, R. Patel, I.M. Mujtaba, 2020, Scope and Limitations of Modelling, Simulation, and Optimisation of a Spiral Wound Reverse Osmosis Process-Based Water Desalination, *Processes*, 8, 573.
- G. Filippini, M. Al-Obaidi, F. Manenti, I.M. Mujtaba, 2019, Design and economic evaluation of solar-powered hybrid multi effect and reverse osmosis system for seawater desalination, *Desalination*, 465, 114-125.
- E.M. Mokheimer, A.Z. Sahin, A. Al-sharafi, A.I. Ali, 2013, Modeling and optimization of hybrid wind-solar-powered reverse osmosis water desalination system in Saudi Arabia, *Energy Conversion and Management*, 75, 86-97.
- K.G. Nayar, P. Sundararaman, C.L. O'connor, J.D. Schacherl, M.L. Heath, M.O. Gabriel, S.R. Shah, N.C. Wright, A.G. Winter, 2017, Feasibility study of an electrodialysis system for in-home water desalination in urban India, *Development Engineering*, 2, 38-46.

Design and simulation of methanol synthesis using heavy residues with carbon utilization

Siddig S. Khalafalla^a, Hussain A. Alibrahim^a, Fayez Nasir Al-Rowaili^b, and Umer Zahid^{a, c}

^a *Chemical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*

^b *Research and Development Center, Saudi Aramco, Dhahran 31311, Saudi Arabia.*

^c *Interdisciplinary Research Center for Membranes & Water Security, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.*
uzahid@kfupm.edu.sa

Abstract

Global carbon emissions are continuously increasing, clean and sustainable technologies are needed to utilize fossil fuels. Methanol is a clean fuel that can enable a smooth transition between fossil fuels and renewables. This work explored the production of methanol through two different fossil fuels, i.e., natural gas (case I) and vacuum residue (case II). Aspen Plus has been utilized to simulate the two processes for the methanol production capacity of 90 t/h with 99.9 wt. % purity. The two methanol production processes are compared in terms of energy, environmental and economic factors. The important reactive sections of the two processes are validated with the design data to ascertain the accuracy of the results. The results show that case II has a process efficiency of 49.5 % which is 1.6 % higher compared to the case I. From an economics point of view, case II design can produce methanol at a unit cost of \$317/t CH₃OH, which is 14% less than the case I design. As a result of the environmental analysis, the case I design releases fewer carbon emissions than case II. Despite this, the case II design offers CO₂ captured with high purity that can be used for another application which can further reduce the production cost of methanol.

Keywords: vacuum residue gasification; carbon capture and utilization; methanol; process simulation

1. Introduction

The high emission of greenhouse gases (GHGs) is attributed to the increase in energy demand. Carbon dioxide is considered as one of the main contributors to carbon emissions. Therefore, several technologies have been developed to reduce CO₂ emissions such as carbon capture, utilization and storage (CCUS) (Alibrahim et al., 2021). The utilization and conversion of CO₂ to clean fuels or other added-value chemicals is one of the potential approaches to reduce carbon emission in the chemical industry. Therefore, the conversion of CO₂ to methanol is one of the key routes. Methanol is also an important feedstock downstream to produce many valuable chemicals. Moreover, methanol blending with gasoline provides clean combustion characteristics. Besides, methanol has a wide-ranging of uses in painting, anti-freezing agents, production of biodiesel, and denitrification of wastewater (Khalafalla et al., 2020). Based on all these applications, methanol demand has rapidly increased in the global market. Currently, China is the largest consumer of methanol, accounting for around 58% of methanol usage. It is also

predicted to grow at a compound annual growth rate (CAGR) of 5.66 % from 2019 to 2027 due to the methanol-to-olefin industry (Inkwood Research, 2020). Most of the methanol is produced via the catalytic reaction of the synthesis gas. Recently, methanol production technologies utilize CO₂ as a feed source. Many studies have been conducted on CO₂ direct hydrogenations to methanol. Battaglia et al. (Battaglia et al., 2021) studied the techno-economics parameters for methanol synthesis through direct hydrogenation of CO₂ captured from power plants with hydrogen from water electrolysis. They also performed heat integration and utilization between different sections and reported savings of cooling and heating loads as high as 47% and 81% respectively. Ahmed (Ahmed, 2021) also performed a techno-economics analysis for the co-production of methanol and hydrogen. He reported an increase in efficiency compared to the standalone technologies.

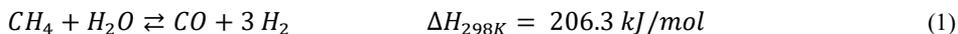
Gasification, another synthesis gas production technology, converts heavy carbon-based feeds to synthesis gas as an intermediate for many useful chemicals. Choi et al. (Choi et al., 2007) studied the composition of the synthesis gas produced through gasification of the vacuum residue. They concluded optimal operating conditions for the temperature and pressure are 1200 – 1250°C and around 1 bar respectively. Al-Attas et al. (Al-Attas et al., 2021) studied operational parameters of co-gasification of vacuum residue and biomass. They stated that oxygen content in the biomass increases the activation energy for the gasification reaction.

Saudi Arabia produces over 200,000 barrels per day (BPD) of vacuum residue used in asphalt production. Its massive production capacity rises environmental concerns in addition to environmental regulations of sulfur content in asphaltene (Al-Rowaili et al., 2021). In this study, the conversion of vacuum residue through gasification is proposed to reduce its carbon emissions. Both conventional steam methane reforming to methanol and vacuum residue gasification to methanol processes have been simulated using Aspen Plus® V11. Moreover, comparative analysis has been performed based on equal production rate and the purity of the methanol.

2. Process design

2.1. Methanol production from natural gas (Case I)

This process consists of five major units; sulfur removal unit, steam reforming, gas compression, methanol synthesis, and methanol purification unit. Toxic sulfur components harm the catalysts of both syngas reforming and methanol synthesis; thus, it is removed to a trace level less than 1 ppm. Natural gas (NG) is fed to the desulfurization unit which utilizes a fixed-bed reactor at high operation conditions of temperature and pressure, about 300 – 400°C and 30 – 130 bar respectively. The sweet NG stream is sent to the steam reforming unit where methane is converted to syngas as mentioned in Eq.1. The hot syngas is utilized to heat the feed of the reformer through a network of heat exchangers. The reactor effluent stream contains CO, CO₂, H₂, unreacted CH₄, and some amount of water.



The produced syngas is then compressed to operation conditions of the methanol synthesis block. Methanol synthesis is an exothermic process and happens in a commercial fixed-bed reactor operating at pressure and temperatures between 50 to 100 bar and 230 to 270°C, respectively. The crude methanol produced from the synthesis reactor is sent to the purification unit which consists of two distillation units to produce

90 t/h of methanol with a purity of 99.9 wt.%. Figure 1 shows a simplified block flow diagram of the steam reforming process of natural gas to methanol. This process has been simulated using Aspen Plus using the Peng-Robinson equation of state (PR-EOS).

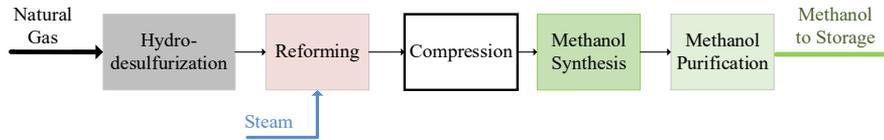


Figure 1: Block diagram of methanol production from natural gas (Case I)

2.2. Methanol production from vacuum residue (Case II)

Vacuum residue (VR) is utilized for syngas production via gasification technology. The composition and heating value for the VR is adapted from the literature (Al-Rowaili et al., 2021). PR-EOS has been selected as a method for the thermodynamic models to predict the thermodynamic properties. The block flow diagram for the case II design is illustrated in figure 2.

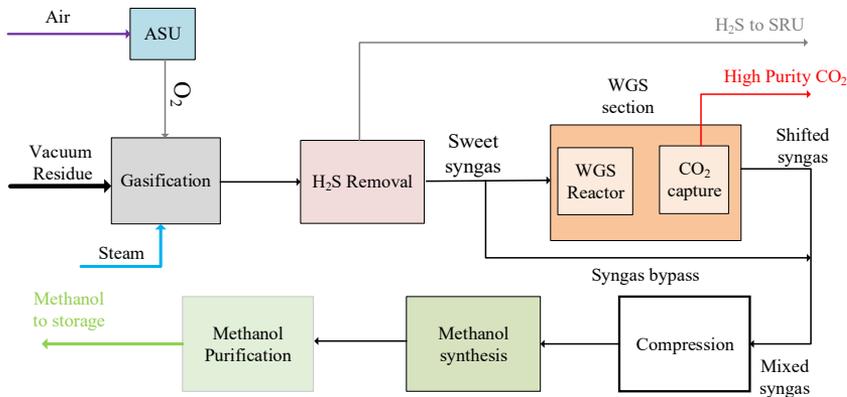
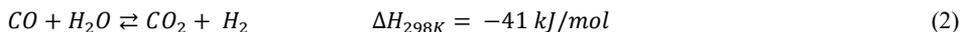


Figure 2: Block diagram of methanol production from vacuum residue (Case II) – sweet shift

The vacuum residue is assumed to be supplied from the refinery vacuum distillation unit (VDU). The vacuum residue along with oxygen and steam are fed to the gasifier. The raw syngas leaves the gasifier at 1300°C, which is utilized in the heat recovery steam generation unit (HRSG) to produce high-pressure steam and power. The raw syngas contains a mixture of gaseous products including CO, CO₂, H₂, H₂O, H₂S, and minor quantities of other gases. Methanol has been used as a physical solvent in the absorber for removing more than 99.9% of the sulfur components in the raw syngas. The sweet gas has a low syngas ratio that is not suitable for methanol production. Hence, 38% of the total syngas is sent to the water-gas shift reactor unit (WGS) where most of CO is converted to H₂ by the reaction with steam as Eq. 2. This step is key to meeting the H₂ deficiency in the sweet syngas ratio which requires a stoichiometric number (SN) from 2 as depicted in Eq. 3. Excess CO₂ is produced due to the water gas shift reaction and it is captured by methanol absorption to fix the stoichiometric number.



$$SN = \frac{[H_2] - [CO_2]}{[CO] + [CO_2]} \quad (3)$$

The mixed syngas is compressed and heated to 100 bar and 230°C respectively to meet the methanol synthesis reactor conditions as shown in figure 3. Methanol synthesis reactor is modelled as an adiabatic quenching reactor with five stages due to the exothermic nature of methanol synthesis as shown in Eq. 4 – Eq.6.

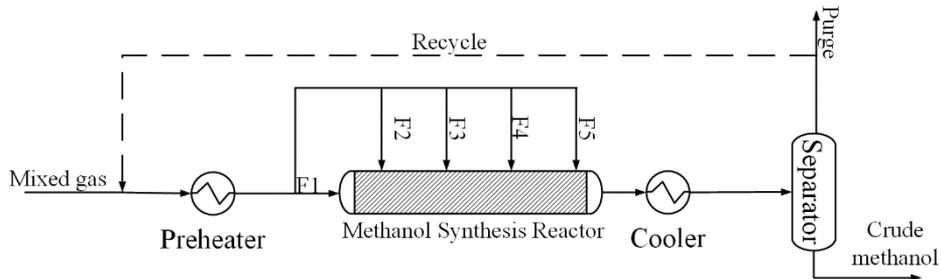
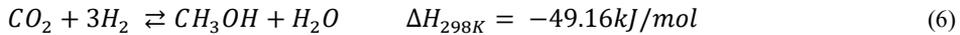
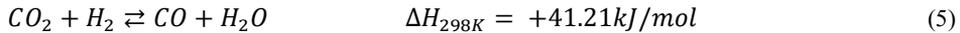
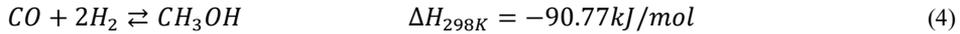


Figure 3: Process flow diagram for Methanol synthesis unit

The crude methanol produced from the synthesis reactor is sent to the purification unit that consists of two distillation units to produce 90 t/h of methanol with 99.9 wt.% purity.

3. Results

3.1. Model Validation

The simulation results have been validated for the main units such as the gasifier and the steam reformer as shown in figure 4-a. The nature of the feedstock used to produce syngas directly affected the quality of the syngas and their composition. The raw syngas produced from steam reforming technology has a higher H_2/CO ratio compared to VR gasification. Moreover, methanol synthesis is also validated with design data as illustrated in figure 4-b. The result shows well agreement between both the simulation results and design data.

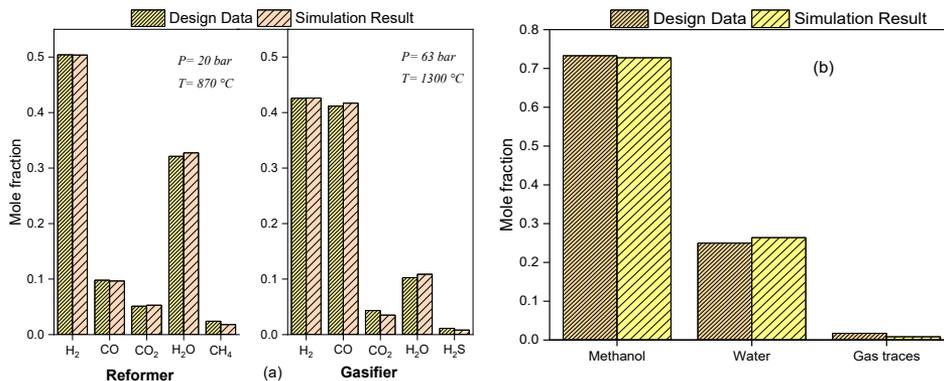


Figure 4: Model validation results: a) for reforming and gasification reactors respectively. b) for methanol reactor.

3.2. Techno-economic and Environmental Analysis

Energy analysis has been performed for both the proposed design and conventional steam reforming to estimate the energy consumption in these processes. Table 1 presents a summary of the energy consumption of various units in case I and case II designs.

Table 1: Energy consumption of various sections in case I and case II designs

Case I (MW)		Case II (MW)	
Steam reforming	208.9	VR Gasification	6.1
Compression	34.6	Cleaning unit	52.9
Methanol synthesis	225.4	WGS	190.8
Methanol Purification	149.9	Compression	31.1
		Methanol synthesis	188.4
		Methanol Purification	84.6

The findings illustrate that the case I process has a 14.9% higher total energy consumption compared to case II. Moreover, the energy efficiency has been calculated for the two processes and the result shows that the case I process has 47.9% which is 1.6% lower than the case II process. On the other hand, based on the environmental impact carbon emission has been calculated for both processes. Case I design is more environmentally sustainable compared to case II because it has a lower CO₂ emission per unit of methanol produced. Furthermore, economic analysis has been studied for both scenarios. The unit cost of methanol product from the case II process is \$317/tCH₃OH which is 14% lower compared to the case I process as shown in figure 5.

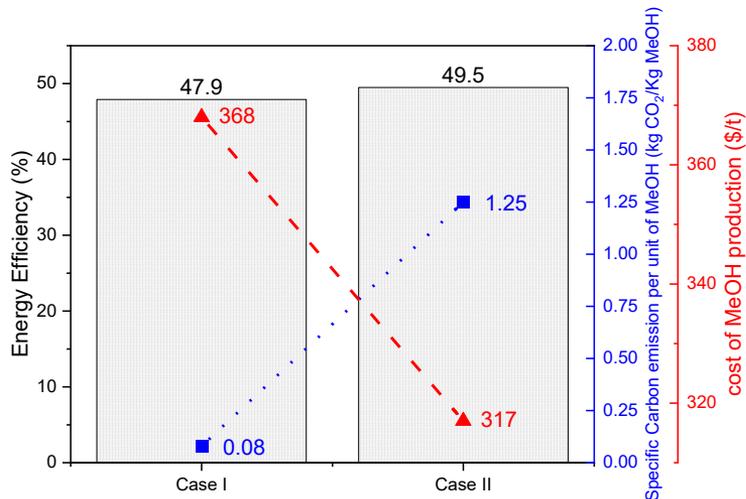


Figure 5: Techno-economic and environmental analysis.

4. Conclusion

In order to minimize the emissions of CO₂, the conversion of heavy vacuum residue to green fuel is proposed. In this study, comparative analysis between conventional

methanol production from natural gas and vacuum residue feed has been studied to produce 90 t/h with 99.9 wt. % purity. The overall results show that the proposed case II process has a lower energy consumption as compared to the conventional case I design. Moreover, case II is more economical compared to the natural gas-based case I since it has a lower cost of production. On the other hand, the proposed design has a higher specific carbon emission compared to the conventional steam reforming process.

Acknowledgment: The authors would like to acknowledge financial support from the Deanship of Research Oversight and Coordination (DROC) at King Fahd University of Petroleum & Minerals (KFUPM).

References

- Ahmed, U. (2021). Techno-economic analysis of dual methanol and hydrogen production using energy mix systems with CO₂ capture. *Energy Conversion and Management*, 228, 113663.
- Al-Rowaili, F. N., Khalafalla, S. S., Al-Yami, D. S., Jamal, A., Ahmed, U., Zahid, U., & Al-Mutairi, E. (2021). Techno-economic evaluation of methanol production via gasification of vacuum residue and conventional reforming routes. *Chemical Engineering Research and Design*.
- Al-Attas, T. A., Lucky, R. A., & Hossain, M. M. (2021). Apparent Kinetics of Co-Gasification of Biomass and Vacuum Gas Oil (VGO). *Chemistry – An Asian Journal*, 16(5), 507–520.
- Alibrahim, H. A., Khalafalla, S. S., Ahmed, U., Park, S., Lee, C.-J., & Zahid, U. (2021). Conceptual design of syngas production by the integration of gasification and dry-reforming technologies with CO₂ capture and utilization. *Energy Conversion and Management*, 244, 114485.
- Battaglia, P., Buffo, G., Ferrero, D., Santarelli, M., & Lanzini, A. (2021). Methanol synthesis through CO₂ capture and hydrogenation: Thermal integration, energy performance and techno-economic assessment. *Journal of CO₂ Utilization*, 44, 101407.
- Choi, Y.-C., Lee, J.-G., Yoon, S.-J., & Park, M.-H. (2007). Experimental and theoretical study on the characteristics of vacuum residue gasification in an entrained-flow gasifier. *Korean Journal of Chemical Engineering*, 24(1), 60–66.
- Inkwood Research. (2020). *Global Methanol Market Analysis, Stats, Trends, Size & Share 2019-2027*. <https://inkwoodresearch.com/reports/global-methanol-market/>
- Khalafalla, S. S., Zahid, U., Abdul Jameel, A. G., Ahmed, U., Alenazey, F. S., & Lee, C.-J. (2020). Conceptual Design Development of Coal-to-Methanol Process with Carbon Capture and Utilization. *Energies*, 13(23), 6421.

Integrated process synthesis and design of intensified/hybrid processes

Shuang Xu^a, Selen Cremaschi^a, Mario R. Eden^a and Anjan K. Tula^{b*}

^a*Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA*

^b*College of Control Science & Eng., Zhejiang University, Hangzhou 310027, China*
anjantula@zju.edu.cn

Abstract

Process intensification has been widely researched in recent years because of its possibility of reducing process capital/operating cost and carbon emissions. It has been shown that implementing intensified/hybrid unit operations into the process industry can significantly improve the process performance. This work aims to develop an integrated process synthesis framework, which includes reaction synthesis for the reaction pathways selection and process intensification in the early stages of process synthesis-design. So, given the product one needs to produce, this framework could identify the optimal reaction pathway with its best separation route, which includes both traditional and hybrid/intensified equipment. The framework was applied to a dimethyl carbonate (DMC) production case study. Eight potential reaction pathways are identified, and three of them are further analyzed in separation synthesis.

Keywords: process synthesis, intensification, reaction synthesis.

1. Introduction

Process intensification is defined as a set of innovative solutions, which aims to improve the whole process performance, such as lower capital/operating costs, minimizing waste, and carbon emissions. As a potential method for process improvement, identifying and applying process intensification will be critical. Lutze *et al.* (2013) proposed a phenomena-based methodology to identify and incorporate intensified equipment options into process synthesis. Tula *et al.* (2017) further integrated the phenomena-based methodology into a process synthesis framework to identify sustainable process flowsheets, including hybrid/intensified unit operations. Furthermore, researchers have evaluated different hybrid distillation configurations, and the results show potential energy savings up to 30% and capacity enhancement of 20%. Bhargava *et al.* (2019) indicated that the intensified equipment, e.g., dividing wall column, has 20-30% lower capital cost than conventional distillation systems. Applying intensification in the early stages of process design, e.g., separation synthesis, is one potential way to generate sustainable process flowsheets with higher energy efficiency and lower environmental impact. However, this method is usually applied for generating sustainable processes with a given reaction pathway without considering the possibilities of different raw materials. Generally, to produce a given target product, multiple reaction pathways are available, and each reaction pathway leads to different downstream processes. To identify the best process to produce the target product, not only process intensification but also different feasible reaction pathways should be considered. In this work, an integrated process synthesis framework including reaction synthesis and separation synthesis is proposed to identify the best processing route for chemical production.

2. Framework

The three-stage integrated process synthesis framework is shown in Figure 1. In stage 1, reaction synthesis is applied to identify novel reaction pathways with potential lower environmental impact and separation cost. Here, the input is either raw materials or target products so that the forward synthesis or retrosynthesis model can be used to generate multiple novel reaction pathways. In stage 2, with different identified novel reaction pathways, the downstream separation synthesis is applied to find its corresponding optimal process flowsheets. In stage 3, the identified process is designed and verified using rigorous models. A detailed description of reaction synthesis and separation synthesis is provided in the following sections.

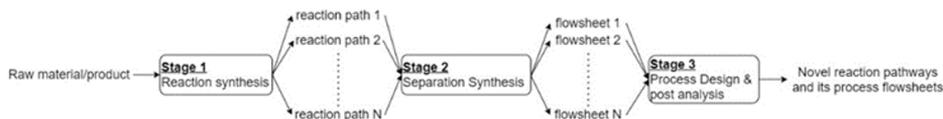


Figure 1. Integrated framework for process synthesis problem

2.1. Reaction Synthesis

Reaction synthesis was initially developed to assist organic chemistry and pharmaceutical research. The task of reaction synthesis is to identify an optimal reaction path to produce a product (retrosynthesis) or utilize the raw material (forward synthesis). Given the target product, the retrosynthesis model can reversely identify different starting materials. Generally, reaction synthesis methods can be classified into logic-centered approach, direct associative approach (Nishida *et al.*, 1981), and data-driven approach (Coley *et al.*, 2017). In this work, a data-driven approach-based reaction synthesis framework is used, which is shown in Figure 2. First, a specific target product or material is given to the reaction synthesis model. In step R1, a constructed retrosynthesis or forward synthesis tool is applied for reaction path prediction. The tool uses a neural network to find the reaction template where the product/reactant has the highest structural similarity to the input molecule. Then, these identified templates are used to predict the reaction pathways. In this way, multiple feasible reaction pathways, including the predicted and external literature pathways, are generated. In step R2, based on the identified reaction pathways, a rule-based screening method is applied to select the optimal reaction pathways. The screening method includes the following rules/conditions:

- 1) Viable reactants that have a good production pipeline.
- 2) Reactions that have lower absolute reaction enthalpy values are preferred.
- 3) Toxicity for reaction i : $LC50_i = \min\{LC50_{i,j}\}$, j represents component j . Reactions that have higher LC50 are preferred.
- 4) Separation driving force of reaction i : $DF_i = \prod_{j,k} DF_{i,jk}$, jk represent binary components j and k . Reactions with a higher DF value are preferred as they may be easier for separation.
- 5) Economic viability.

2.2. Separation Synthesis

The separation synthesis framework utilizes the thermodynamic insight-based process synthesis method by Tula *et al.* (2015). It has six steps (Figure 3): problem definition, problem analysis, feasible separation techniques identification, process alternative generation, process ranking, and design/verification. In step 1, the raw materials,

products, inlet conditions, and product purity specifications are specified. In step 2, azeotropic/eutectic analysis is performed, and binary ratios of pure component properties are calculated. In step 3, based on the data from the analysis step and predefined selection rules, all feasible separation techniques are identified. Here, conventional unit operations and recently developed hybrid/intensified separation techniques, including distillation-membrane, distillation-adsorption, distillation-crystallization, and membrane-crystallization, dividing wall column, are considered in identifying the feasible separation techniques. In step 4, based on the identified superstructure of separation techniques, all the feasible process alternatives are generated. In step 5, a ranking algorithm is applied to evaluate all the process alternatives. Here, all the processes are ranked based on their process energy indices, which is the summation of the estimated energy consumption of each unit operation. For example, the energy consumption for distillation and hybrid distillation is estimated based on first principle-based models (Lange, 2017) as shown in Equations (1) and (2). Energy consumption for the other unit operations is based on enthalpy balances (Equation (3)). After the evaluation step, the processes with lower process energy indices are selected. Finally, these top processes are further evaluated using rigorous models. In this way, the method identifies the best process alternatives and their optimized design/operating conditions.

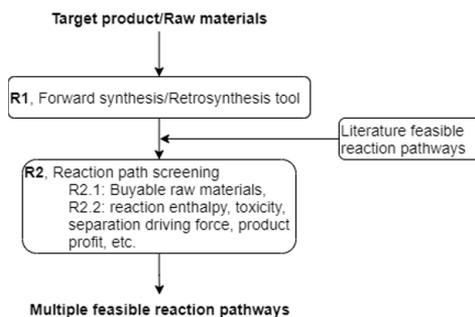


Figure 2. Framework for reaction synthesis.

$$Q_d = 57 \times \sum_i \frac{w_i}{\Delta T_i} \quad (1)$$

$$Q_{dhs} = \eta \times Q_d \quad (2)$$

$$Q_{hs} = H_{out}N_{out} - H_{in}N_{in} \quad (3)$$

Where Q_d is the reboiler duty; w_i is the inlet mass fraction of component i ; ΔT_i is the atmosphere boiling point different between component i and $i + 1$ (ranked based on boiling point); Q_{dhs} is the reboiler duty for distillation hybrid separation techniques; η is an constant coefficient; Q_{hs} is the heat duty for the other separation techniques; H_{out} and H_{in} are the enthalpy at outlet and inlet conditions; N_{out} and N_{in} are the outlet and inlet flowrates.

3. Case Study

The developed framework is applied to a DMC production case study. This case study aims to find the best reaction pathway and its separation route for DMC production. First, reaction synthesis (Section 3.1) is applied to generate and select the top reaction

pathways. Next, the selected top reaction routes are given to the separation synthesis (Section 3.2) step to determine the optimal separation flowsheet.

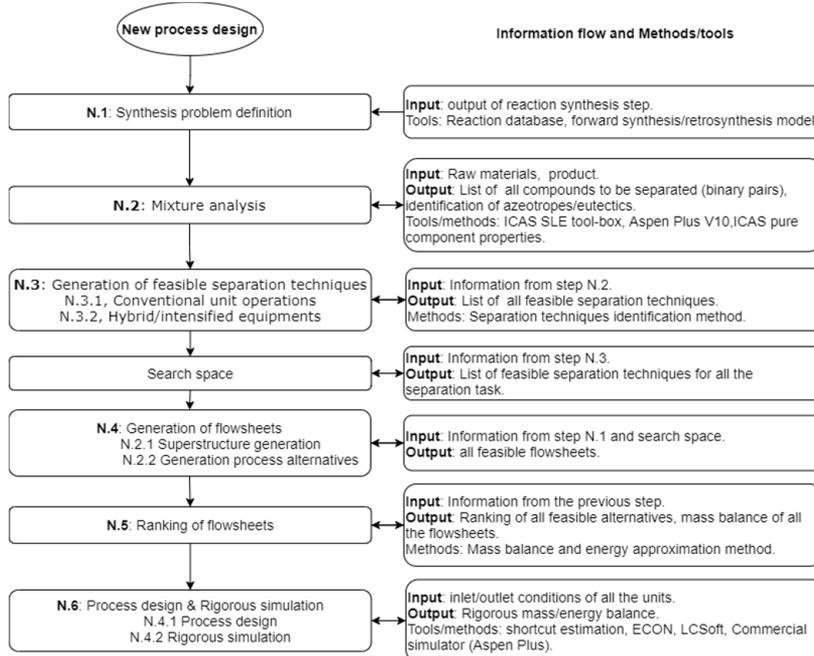


Figure 3. Framework for separation synthesis.

3.1. Reaction Synthesis

In this step, the retrosynthesis tool ASKCOS is applied to generate different feasible reaction pathways (github.com/ASKCOS). To maintain good prediction accuracy, default settings are used to generate these potential reaction pathways. A total of eleven reactions are generated, including eight reactions from ASKCOS and three from literature (Kongpanna *et al.*, 2015). After applying the rule-based approach, three of these reactions routes violate the first rule, which leads to eight reactions pathways for further screening. Table 1 shows the eight reactions and calculated/estimated reaction enthalpy, toxicity, and separation driving force. Based on the criteria in Section 2.1, the top three are reaction routes 3, 4, and 5, which are sent to separation synthesis for further analysis.

Table 1. Potential feasible reactions for DMC production.

No.	Reactants		Products		ΔH	LC50	DF
1	CH ₃ OH	C ₂ H ₃ ClO ₂	C ₃ H ₆ O ₃	HCl	0.88	47.46	0
2	CH ₃ OH	CH ₄ N ₂ O	C ₃ H ₆ O ₃	NH ₃	109.99	129.51	0
3	CH ₃ OH	CO ₂	C ₃ H ₆ O ₃	H ₂ O	-22.9	129.51	0
4	CH ₃ OH	C ₃ H ₄ O ₃	C ₃ H ₆ O ₃	C ₂ H ₆ O ₂	-22.05	129.51	0
5	CH ₃ OH	C ₃ H ₆ O ₃	C ₃ H ₆ O ₃	C ₃ H ₈ O ₂	26.67	129.51	0
6	CH ₃ OH	CO ₂	C ₃ H ₆ O ₃	H ₂ O	-305.7	129.51	0
7	CO	CH ₃ ONO	C ₃ H ₆ O ₃	NO	-189.6	63.67	1
8	CH ₃ OH	CCl ₂ O	C ₃ H ₆ O ₃	HCl	-70.75	128.14	0

3.2. Separation Synthesis

Separation synthesis is performed to produce 56.4 kmol/h of DMC using the identified three mechanisms. The raw materials requirements, inlet conditions, and product specifications are shown in Table 2. In the mixture analysis step, DMC/methanol and

DMC/water are identified as binary azeotropes. Based on pure component and mixture analysis, all feasible separation techniques are identified. Table 3 shows an example of all the feasible separation techniques for reaction 3.

Table 2. Inlet conditions for separation synthesis.

		Reaction 3	Reaction 4	Reaction 5
CH ₃ OH	kmol/h	56.4	56.4	56.4
C ₃ H ₆ O ₃	kmol/h	46.8	56.4	56.4
H ₂ O	kmol/h	46.8	-	-
C ₃ H ₄ O ₃	kmol/h	-	18.2	-
C ₂ H ₆ O ₂	kmol/h	-	46.8	-
C ₃ H ₆ O ₃	kmol/h	-	-	18.2
C ₃ H ₈ O ₂	kmol/h	-	-	46.8
Pressure	bar	10		
Purity specification	mol%	CH ₃ OH, C ₃ H ₆ O ₃ : 99.5 mol%		

Table 3. Feasible separation techniques for methanol/DMC and DMC/water.

Binary Mixture	Feasible Separation Techniques
Methanol/DMC	Membrane, adsorption, distillation-membrane, distillation-adsorption, extractive distillation
DMC/water	Azeotropic distillation, membrane, adsorption, distillation-membrane, distillation-adsorption

A total of 16, 127, and 166 process alternatives are generated for reactions 3, 4, and 5, respectively. All the alternatives are ranked based on their process energy indices. For each reaction pathway, the top two process alternatives are selected for rigorous simulation. Figures 4 and 5 show the flowsheets of these top selected process alternatives. Figure 6 compares the simulation results of all the top alternatives with respect to utility/capital cost and process environmental factors like carbon footprint and global warming potential (GWP). From Figure 6, the process alternative from reaction pathway 4 has the lowest utility, capital cost, and environmental impact compared to the top alternatives from other pathways. Reaction pathway 3 has two azeotropes, which leads to higher separation cost and carbon emissions.

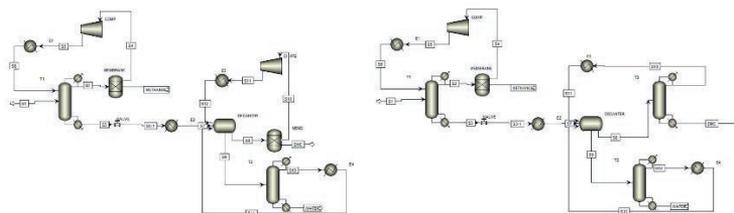


Figure 4. Top two separation configurations for reaction 3 (from left to right R3_P1, R3_P2).

4. Conclusions

An integrated process synthesis framework is developed to identify the best process flowsheet, including both reaction synthesis and separation synthesis. Given the product compound, the reaction synthesis generates multiple feasible reaction pathways, while the separation synthesis identifies the best downstream separation flowsheet, including intensified/hybrid separation techniques. The framework was applied to a DMC production case study. The results show that the best pathway for DMC production is using propylene carbonate and methanol as reactants with distillation-membrane and

dividing wall column for separation, which has 20% lower operating cost than the conventional extractive distillation process.

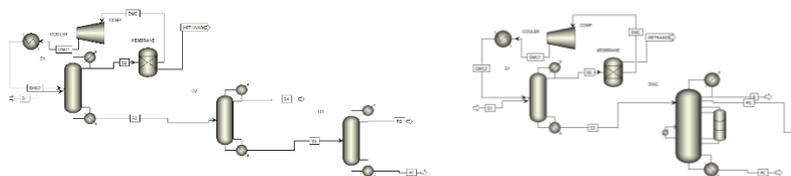


Figure 5. Top two separation configurations (from left to right R4(5)_P1, R4(5)_P2) for reaction 4 and 5 (similar reactants and byproducts, which leads to the same top two separation flowsheets).

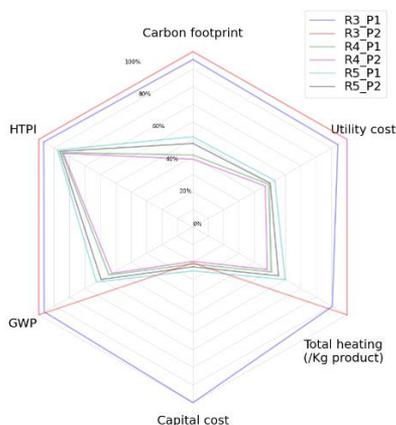


Figure 6. Comparison of six process alternatives.

References

- Lutze, P., Babi, D. K., Woodley, J. M., & Gani, R. (2013). Phenomena based methodology for process synthesis incorporating process intensification. *Industrial & Engineering Chemistry Research*, 52(22), 7127-7144.
- Tula, A. K., Babi, D. K., Bottlaender, J., Eden, M. R., & Gani, R. (2017). A computer-aided software-tool for sustainable process synthesis-intensification. *Computers & Chemical Engineering*, 105, 74-95.
- Bhargava, M., & Sharma, A. P. (2019). Introduction to Dividing-Wall Columns. *Chemical Engineering Progress*, 115(3), 30-34.
- Nishida, N., Stephanopoulos, G., & Westerberg, A. W. (1981). A review of process synthesis. *AIChE Journal*, 27(3), 321-351.
- Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., & Jensen, K. F. (2017). Prediction of organic reaction outcomes using machine learning. *ACS central science*, 3(5), 434-443.
- Tula, A. K., Eden, M. R., & Gani, R. (2015). Process synthesis, design and analysis using a process-group contribution method. *Computers & Chemical Engineering*, 81, 245-259.
- Lange, J. P. (2017). Don't forget product recovery in catalysis research—Check the distillation resistance. *ChemSusChem*, 10(1), 245-252.
- Kongpanna, P., Pavarajarn, V., Gani, R., & Assabumrungrat, S. (2015). Techno-economic evaluation of different CO₂-based processes for dimethyl carbonate production. *Chemical Engineering Research and Design*, 93, 496-510.

Macroalgae-based integrated biorefinery for hydrocolloids, chemicals and advanced biofuels production

Andrés I. Casoni, Alejandro H. Pedrozo, Fernando D. Ramos, Vanina Estrada,
M. Soledad Diaz*

Planta Piloto de Ingeniería Química- PLAPIQUI (UNS-CONICET), Camino la Carrindanga km. 7, Bahía Blanca (8000), Argentina

sdiaz@plapiqui.edu.ar

Abstract

In this work, we address the simultaneous design of an integrated macroalgae-based biorefinery and its heat exchanger network for the production of alginates, high added-value chemicals and advanced biofuels. We apply a superstructure approach, which is formulated as a mixed integer nonlinear programming (MINLP) model. The resulting MINLP model has more than 53000 constraints and continuous variables and 14000 discrete variables. Net present value is the objective function. Numerical results show that heat utility costs are reduced by 41 % due to heat integration, as compared to a base case, and the production of alginates and dimethyl-isosorbide (DMI) via sorbitol, as well as methane as biofuels are selected in the optimal scheme.

Keywords: MINLP, Heat Exchanger Network, Macroalgae biorefinery.

1. Introduction

The use of macroalgae in the biorefinery concept is known to positively contribute to the environmentally friendly production of bioproducts (Zapata-Boada et al., 2021). Despite recent contributions (González-Gloria et al., 2021; Kostas et al., 2021), there is still a gap in the understanding of using macroalgae species to obtain chemical products industrially (Sudhakar et al., 2019).

Alginates are a family of polysaccharides composed of mannuronic and guluronic which constitute the main commercial product from brown macroalgae with a wide range of applications in the food, cosmetic, pharmaceutical and nutraceutical industries. Currently, alginate is commercially produced in USA, Japan, China, France, and Norway (Gomez et al. 2009). The giant kelps *Macrocystis pyrifera* and *Lessonia* sp. (Laminariales, Phaeophyceae) are brown macroalgae widely distributed throughout the cold water of the Argentinean Patagonia, with great potential to be sustainably exploited (Zaixso and Boraso, 2015) for hydrocolloids, materials, chemicals and biofuels integrated production (Pedrozo et al., 2021; Ramos et al., 2021).

In this work, we propose a mixed-integer nonlinear programming (MINLP) model for the simultaneous plant design and heat exchanger network synthesis of a macroalgae-based integrated biorefinery for alginate, chemicals and advanced biofuels. In previous work, we formulated a biorefinery model (Casoni et al., 2020) for the production of the platform molecule isosorbide, biogas and fertilizers from the brown macroalgae. Isosorbide is used to obtain several final products that include a flame retardant, a biopolymer, a biosolvent

(dimethyl isosorbide) and a drug for heart disease (isosorbide dinitrate). In the present work, the potential production of alginate and the sequential production of hydrogen and methane through two-stage anaerobic digestion is included, to enhance macroalgae biomass valorisation and advanced biofuels production.

2. Process Description

We consider an integrated biorefinery potentially based on two species of macroalgae, alternatively, on corn starch that can be raw materials for sorbitol production by acid hydrolysis. Also, sorbitol can be directly bought from the market. This compound is subsequently transformed into isosorbide, a platform molecule. This molecule may be converted into different products: a biopolymer, a flame retardant, a drug for heart disease and a biosolvent (further details can be found in Casoni et al., 2020). Alginate can also be produced from macroalgae biomass. Figure 1 shows a simplified scheme with the different process alternatives embedded within the superstructure presented in this paper.

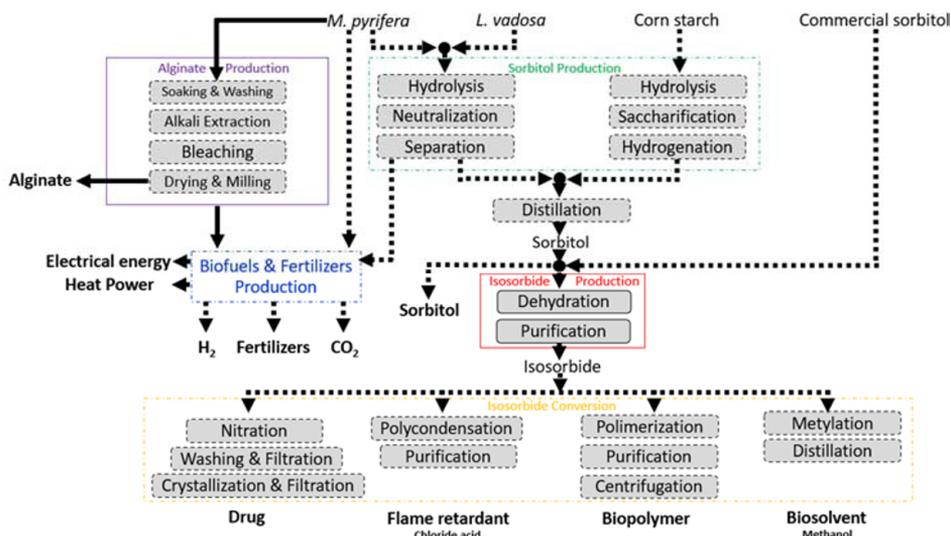


Figure 1. Simplified superstructure of process alternatives showing the five considered sections: Alginate Production, Sorbitol Production, Isosorbide Production, Isosorbide Conversion, and Biofuels and Fertilizers Production. Potential products of the integrated biorefinery are shown in bold. Chloride acid and methanol are byproducts.

A detailed description of Alginate Production Section is described hereunder. As a first step, macroalgae are treated with formaldehyde (FAL) 0.1% wt. in a 0.086 algae-to-FAL ratio in a soaking reactor. Then, the mixture is filtered to recover macroalgae biomass, which is washed using HCl 0.38 % wt., in a 0.038 HCl-to-algae ratio and subsequently filtered. The pre-processed biomass is mixed in an extraction reactor with a hot alkali solution (80 °C) of Na₂CO₃ 10 % wt. in a 0.91 algae-to-Na₂CO₃ ratio. A water stream is fed to this unit to dilute the solution to reach an 18.68 water-to-algae ratio. The stream containing sodium alginate (Na(alg)) is diluted in a reactor with water at 70 °C in a 141.15 water-to-Na(alg) ratio. The outlet stream is filtered in a rotating vibrating filter using expanded lava as filter aid in a 0.94 perlite-to-Na(alg) ratio. The residual algae biomass is sent to the Biofuels and Fertilizers Production Section, while the process stream composed of Na(alg) and water is fed to a CaCl₂ reactor. In this reactor, a solution of

CaCl₂ 10 % wt., is fed in a 1.81 Na(alg)-to-CaCl₂ ratio to transform sodium alginate into calcium alginate (Ca(alg)). This stream is then filtered to separate NaCl with water and the Ca(alg) is fed into a bleaching reactor to bleach it using NaClO 5 % wt. in a 0.0076 NaOCl-to-Ca(alg) ratio. Next, the stream is fed to an acid reactor where it is mixed with HCl 3.6 % wt. in a 0.195 HCl-to-Ca(alg) ratio. In this reactor, the calcium alginate is transformed into acid alginate (H(alg)), which is fed into a belt press unit to separate as much water as possible. The stream containing H(alg) and water is fed into a Na₂CO₃ reactor where it is mixed with Na₂CO₃ 10 % wt. in a 0.29 Na₂CO₃-to-H(alg) ratio. This treatment converts the H(alg) into Na(alg), which is dried at 50 °C and milled to obtain the final product: 90% sodium alginate.

3. Mathematical Model

We formulate a Mixed-Integer Nonlinear Programming (MINLP) model to address the simultaneous optimal design of the integrated macroalgal based biorefinery and its heat exchanger network through a superstructure representation as shown in Eqs. (1.1)-(1.2). The objective function is net present value maximization (NPV) (Eq. (1.1)), to consider a detailed economic evaluation. Equation (1.2) is the set of linear equations corresponding to process design and includes mass balances of reactive and nonreactive units, power consumption of process units, and Big-M constraints to model conditional units. Equation (1.3) refers to the set of nonlinear constraints of process synthesis, for instance, equipment design calculations and capital cost correlations. In this work, we have also embedded the potential matches between process streams and utilities to synthesize the heat exchanger network within the proposed superstructure, following Yee and Grossmann (1990), to simultaneously address the optimal heat exchanger network (HEN) design within the biorefinery design problem. Equation (1.4) is the set of linear equations, which includes global overall heat balances for hot and cold streams, temperature assignments, and logic constraints for stream matches and temperature differences. Equation (1.5) represents HEN nonlinear equations, such as heat balances at temperature stages, and capital cost correlations. Finally, Eq. (1.6) is the set of connection equations to link process design variables with HEN variables. It should be noted that the HEN model requires process flowrates from which determines utility and capital costs of heat exchanger units.

$$\max NPV = f(x) \tag{1.1}$$

$$\text{Process design} \left\{ \begin{array}{l} A_1x + A_2y \leq a \end{array} \right. \tag{1.2}$$

$$g(x, y) \leq 0 \tag{1.3}$$

$$\text{Heat exchanger network} \left\{ \begin{array}{l} B_1w + B_2z \leq b \end{array} \right. \tag{1.4}$$

$$h(w) \leq 0 \tag{1.5}$$

$$\text{Connection eqns.} \left\{ \begin{array}{l} C_1x = C_2w \end{array} \right. \tag{1.6}$$

$$x \in \mathbb{R}^n, y \in \{0,1\}^m$$

$$w \in \mathbb{R}^{n'}, z \in \{0,1\}^{m'}$$

where x and w are continuous variables associated with process and heat exchanger network design, respectively; y are binary variables related to the technology selection; and z are binary variables to model heat exchanger matches.

4. Numerical Results

A fixed production of 920 t/y of sorbitol and 300 t/y of alginate has been considered, taking into account current market demand. The resulting MINLP model is formulated in GAMS 35.2 (McCarl et al., 2017). The mathematical model for the simultaneous process and heat exchanger network design includes 53423 constraints, 53554 continuous variables, 13988 discrete variables. As initialization procedure, we first solve mass balances, and then, a model that includes heat balances, temperature assignments, and logic equations for stream matches and temperature differences. It should be noted that these equations are already linear, then no linearization procedures are required. This framework provides a suitable initial guess, which allows solving the MINLP problem in 26 seconds. We use DICOPT, with CONOPT and CPLEX as nonlinear and linear sub solvers, respectively (Grossmann et al., 2003). The optimal scheme includes the co-production of dimethyl-isosorbide (DMI) via sorbitol and alginate from the macroalgae *M. pyrifera*. Table 1 shows the main economic indicators for the optimal solution, and a base case where DMI is produced as sole product (no alginate production). The NPV of the optimal solution is 52.24 MM\$, representing an improvement of 17 %, with respect to the base case. Although the total capital cost increases by 57 % for the optimal design, it allows for producing alginates, improving the revenues by 28 %. These results show the benefits of co-producing alginates DMI and advanced biofuels.

Table 1. Main economic indicators.

	DMI and alginate	DMI
NPV (MM\$)	52.239	44.575
Investment (MM\$)	22.375	14.252
Revenues (MM\$/y)	22.232	17.409
Raw material cost (MM\$/y)	3.241	2.698

Regarding the heat exchanger network, numerical results indicate that utility costs are reduced by 41 % due to heat integration. The network includes 25 heat exchanger units. It is highlighted that the number of matches does not increase with respect to the case of using only utilities to satisfy energy requirements, and the optimal solution only requires cooling water as cooling utility. Furthermore, as the Alginate Production Section has a process stream that is fed to the Biofuels and Fertilizers Production Section, the energy production in this section increases in 133 %, with respect to the base case (no alginate production).

Due to uncertainties in alginate market demand, we performed a sensitivity analysis on the optimal technological route to assess the influence of the alginate plant capacity over the biorefinery NPV. In this sense, we compare the value of the economic objective function for the integrated biorefinery (alginate and DMI co-production) to a base case scenario in which alginate production is not included.

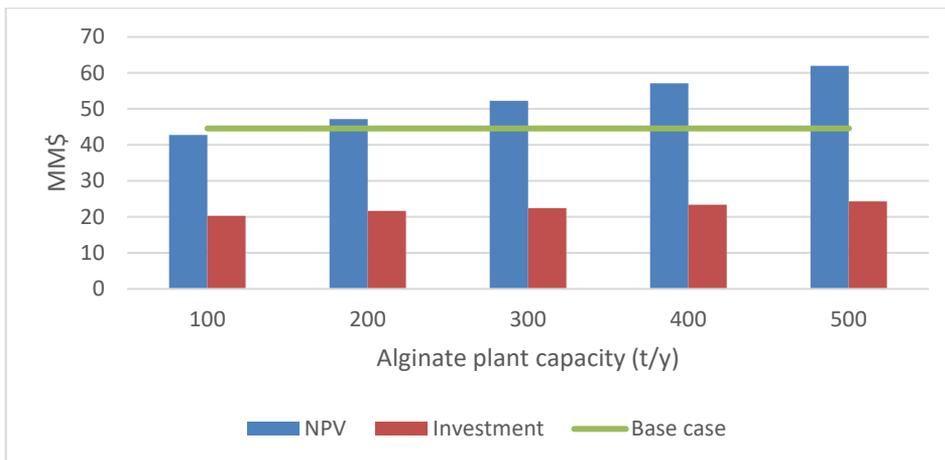


Figure 2. Sensitivity analysis of the alginate plant capacity.

As it can be seen in Fig. 2, alginate production is economically attractive if the alginate section capacity is higher than approximately 180 t/y due to the economies of scale (Roldán-San Antonio et al., 2021).

5. Conclusion

In the present work we propose an MINLP model for the sustainable production of alginates, chemicals and advanced biofuels based on macroalgae. The superstructure includes different process alternatives, as well as alternative process streams and utility streams matches for the simultaneous design of an integrated biorefinery and its heat exchanger network. The proposed model size increases more than two orders of magnitude with the inclusion of the design of its heat exchanger network and provides a decrease in energy consumption of 41% due to heat integration, with respect to a base case. The optimal configuration for NPV maximization includes alginates and DMI via sorbitol production, as well as energy production, which increases 133% due to the inclusion of the biofuels production section.

References

- S. Zapata-Boada, M. Gonzalez-Miquel, M. Jobson, R. Cuellar-Franca, 2021. Integrating process simulation and life cycle assessment to evaluate the economic and environmental performance of algae biodiesl. *Comp. Aided Chem. Eng.* 50, 1903-1908.
- K. D. González-Gloria, R. M. Rodríguez-Jasso, S. E. Aparicio, M. L. Chávez González, E. T. Kostas, H. A. Ruiz, 2021. Macroalgal biomass in terms of third-generation biorefinery concept: Current status and techno-economic analysis – A review. *Bioresour. Technol. Rep.* 16, 100863.
- E. T. Kostas, J. M. M. Adams, H. A. Ruiz, G. Durán-Jiménez, G. J. Lye, 2021. Macroalgal biorefinery concepts for the circular bioeconomy: A review on biotechnological developments and future perspectives. *Renew. Sust. Energ. Rev.* 151, 111553.
- M. P. Sudhakar, B. R. Kumar, T. Mathimani, K. Arunkumar, 2019. A review on bioenergy and bioactive compounds from microalgae and macroalgae-sustainable energy perspective. *J. Clean. Prod.* 228, 1320-1333.
- C. G. Gomez, M. V. Pérez Lambrecht, J. E. Lozano, M. Rinaudo, M. A. Villar, 2009. Influence of the extraction-purification conditions on final properties of alginates obtained from brown algae (*Macrocystis pyrifera*). *Int. J. Biol. Macromol.* 44, 365–371

- H. E. Zaixso, A. Boraso, 2015. *La Zona Costera Patagónica Argentina. Volumen I. Recursos Biológicos Bentónicos* Publisher: Editorial Universitaria de la Patagonia. Comodoro Rivadavia.
- H. A. Pedrozo, S. B. Rodríguez Reartes, A. R. Vecchietti, M. S. Díaz, I. E. Grossmann, 2021. Optimal design of ethylene and propylene coproduction plants with generalized disjunctive programming and state equipment network models. *Comp. Chem. Eng.* 149, 107295.
- F. D. Ramos, A. I. Casoni, V. Estrada, M. S. Díaz, 2021. Simultaneous optimization and heat integration of a macroalgae-based biorefinery. *Comp. Aided Chem. Eng.* 50, 1581-1586.
- A. I. Casoni, F. D. Ramos, V. Estrada, M. S. Díaz, 2020. Sustainable and Economic analysis of Marine Macroalgae Based Chemical Production – Process Design and Optimization. *J. Clean. Prod.* 276, 122792.
- T. F. Yee, I. E. Grossmann. 1990. Simultaneous optimization models for heat integration – II. Heat exchanger networks synthesis. *Comp. Chem. Eng.*, 28, 1165-1184.
- B. A. McCarl, A. Meeraus, P. van der Eijk, M. Bussieck, S. Dirkse, P. Steacy, F. Nelissen, 2017. *McCarl GAMS user guide*, Washington, DC, USA.
- I. E. Grossmann, J. Viswanathan, A. Vecchietti, R. Raman, E. Kalvelagen, 2003. *GAMS/DICOPT: A discrete continuous optimization package*. Washington, DC, USA.
- J. E. Roldán-San Antonio, E. Martín-Hernández, R. Briones, M. Martín, 2021. Process design and scale-up study for the production of polyol-based biopolymers from sawdust. *Sustain. Prod. Consum.* 27, 462-470.

Design Space Determination of Mesenchymal Stem Cell Cultivation by Dynamic Modeling under Uncertainty

Keita Hirono,^a Isuru A. Udugama,^a Yusuke Hayashi,^a Masahiro Kino-oka,^b
Hirokazu Sugiyama^{a,*}

^a*Department of Chemical System Engineering, The University of Tokyo, 7-3-1, Hongo Bunkyo-ku, Tokyo 113-8656, Japan*

^b*Department of Biotechnology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan*

**sugiyama@chemsys.t.u-tokyo.ac.jp*

Abstract

This work details the development of a five-step workflow for design space determination of mesenchymal stem cell (MSC) cultivation. The workflow employs mechanistic models to represent system dynamics of MSC cultivation processes while introducing stochastic simulations to capture uncertainties. The developed workflow was applied to an MSC cultivation case study where seeding density and medium change ratio were selected as critical process parameters (CPPs), while the number of cells cultivated was specified as the sole critical quality attribute (CQA). Monod kinetics based modelling and stochastic simulation were used to determine a dynamically evolving probabilistic design space. The results were then visualized on a 2D map with probability contours. In the presented calculations, the optimal batch termination time was identified as Day 10, given a probability preference of 90 %. The results illustrated the need to account for both raw material and process/operation uncertainties and dynamics in design space determination for MSC cultivation.

Keywords: Cell therapy, stochastic simulation, quality by design, decision support.

1. Introduction

Mesenchymal stem cells (MSCs) are a promising cell therapy candidate due to their multipotency, self-renewability, and ethical compliance (Ullah et al., 2015). Today, MSCs are used in several cell therapies, creating a need to produce MSCs at scale while guaranteeing product quality. Despite ongoing research and developments into producing MSCs at scale, many challenges remain in ensuring the quality manufacturing of MSCs. In Lipsitz et al. (2016), managing the influence of raw material and process variability was identified as the key aspect in ensuring quality MSC manufacturing. Figure 1 provides a schematic representation of MSC cultivation where dynamic cell expansion and metabolism are performed under raw material and process/operation uncertainties.

One strategy that can be employed to ensure product quality at scale is the use of Quality by Design (QbD) framework (European Medical Agency, 2017). The core concept in the QbD framework is developing a systematic and science-based understanding of how materials and processes affect the final product quality. Once all relationships are established, an operation envelop can be identified where the product is guaranteed to be of acceptable quality. This operation envelope is also referred to as the design space

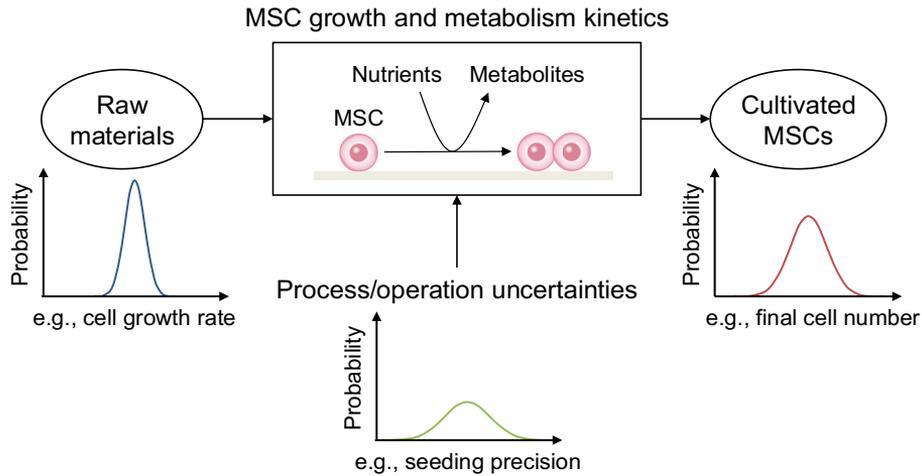


Figure 1. A schematic of an MSC cultivation process where cell expansion and metabolism occur under both raw material and process/operation uncertainties.

(García-Muñoz et al., 2015). In other words, the design space represents an operational envelope where critical process parameters (CPPs) satisfy all critical quality attributes (CQAs).

To apply the QbD framework for design space determination in MSC cultivation, the following aspects must be considered: 1) the time-dependent nature of the MSC cultivation, 2) the influence of raw material and process/operation uncertainties, 3) the lack of experimental data and the high cost of experiments. Inspiration can be drawn from Process Systems Engineering (PSE) to address these needs. For example, mechanistic models, often based on Monod kinetics, have been applied to describe the dynamics of cell growth and metabolism in therapeutic protein production (Badr et al., 2021). These models have also been used to describe MSC cultivation processes (Jossen et al., 2020). Combining stochastic simulations with mechanistic models can generate time-dependent data sets that capture raw material and operation uncertainties (Andersen et al., 2020). While the use of mechanistic models can also reduce the need for experimental data in determining a design space (García-Muñoz et al., 2015). To this end, a workflow that integrates these types of PSE tools with the QbD framework can address the specific needs for design space determination in MSC cultivation.

2. Proposed workflow

Figure 2 represents the proposed five-step workflow for design space determination in MSC cultivation.

2.1. STEP 1: CPP and CQA specification

This step specifies the CPPs and CQAs that are relevant to quality MSC cultivation. Ideally, the CPPs and CQAs should be specified based on fundamental relationships (e.g., stoichiometry). In practice, CPPs and CQAs will be based on expert opinion backed by available literature or experimental data. This is because the fundamental understanding that connects CPPs and CQAs in MSC cultivation is generally lacking, while literature data are sparse.

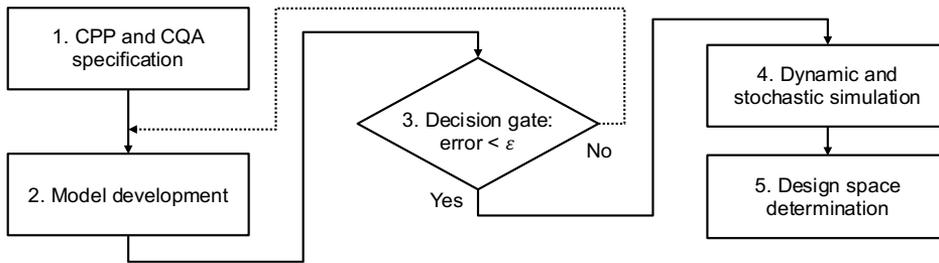


Figure 2. The applied workflow consisted of five steps to determine a design space and the optimal cultivation time.

2.2. STEP 2: Model development

In this step, the identified CPPs and CQAs are connected using a mechanistic model. In the context of cell cultivation, a mechanistic model will consist of a set of ordinary differential equations (ODEs), where empirical relationships such as Monod cell growth kinetics are used to describe state key relationships.

2.3. STEP 3: Decision gate

Through parameter estimation, the model is fitted to the available process data. The parameter estimation is performed by minimizing the error between the model output and the process data. The calculated error is then compared to a user-specified accuracy criterion (ϵ). If the error is less than ϵ , the workflow proceeds. Otherwise, the workflow reverts to Step 2.

2.4. STEP 4: Dynamic and stochastic simulation

Information required to determine a time-dependent probabilistic design space is produced in this step. First raw material and process/operation uncertainties are identified based on literature, experiments, or expert opinion. These uncertainties are then linked to specific parameters in the model developed in Step 2. Once the parameters are identified, the type of probability distribution they follow must be determined. This workflow assumes that all uncertain parameters follow a normal distribution, which is a typical assumption made in probabilistic design space determination (Ochoa et al., 2021). A stochastic simulation is then conducted by iteratively executing the model with the identified parameter sets. Each run will use a randomly sampled parameter set from the defined distribution, capturing the expected uncertainty. The time-dependent output generated by the stochastic simulations can then be compared against the CQAs to determine the probability that a given run satisfies all CQAs. Repeating this process for all combinations of CPPs enables the generation of time-dependent probability information required for design space determination.

2.5. STEP 5: Design space determination

The information from Step 4 is then used to determine and visualize a dynamic and probabilistic design space. If the number of CPPs is limited to two, then a contour map is used for design space visualization. In this case, the X-Y axes are the two given CPPs, and the contours represent the predicted probability that a given batch meets all CQAs. This visualization is performed for each day of cell cultivation. A decision-maker can make an informed decision about the ideal combination of CPPs and cultivation time based on their probability preference. If the number of CPPs is greater than two, geometric projection could be used to visualize the design space (García-Muñoz et al., 2015).

3. Case Study

The proposed workflow was applied to an MSC cultivation process described by Jossen et al. (2020). The source MSCs were derived from the abdomen of a healthy and post-chemotherapy donor and cryopreserved, respectively. The frozen MSCs were thawed and precultured before carrying out the main culture. The main cultivation was performed on a T₂₅-flask as a batch operation for 11 days. An initial culture media volume of 5 mL was used with two partial medium changes of 40 % and 60 % carried out on days 4 and 8, respectively.

3.1. STEP 1: CPP and CQA specification

Based on literature observations and expert opinion, the following CQAs and constraints were specified. The final viable cell number was specified as the sole CQA as this variable was linked to cell therapy product quality (Redaelli et al., 2012). The lower cell limit was specified as 1.4×10^6 cells to meet the product requirements, while the upper limit of 1.7×10^6 cells was set to prevent genome instability due to excessive cell expansion (Redaelli et al., 2012). In addition, a maximum ammonia concentration of 2.0 mmol/L was specified as a constraint (Schop et al., 2009). MSC seeding density was specified as a CPP based on previous experimental observations (Kino-oka, 2000). Medium change ratio was identified as the other CPP as medium change both replenished nutrient (e.g., glucose) needed for cell growth and removed metabolic products (e.g., lactate, ammonia).

3.2. STEP 2: Model development

An ODE based mechanistic model was developed to represent the MSC cultivation process. The model described the cell growth based on the Monod equation, which was modified to consider growth inhibition due to excessive cell-cell contact. The overall ODEs were formulated as follows:

$$\mu = \mu_{\max} \cdot \frac{C_S}{K_S + C_S} \cdot \frac{K_P}{K_P + C_P} \cdot \frac{X_{\max} - X_A}{X_{\max}} \quad (1)$$

$$t_{\text{lag}} = \beta \cdot \ln\left(\frac{X_0}{X_0^*}\right) + \gamma \quad (2)$$

$$\frac{dN_A}{dt} = \mu \cdot N_A \quad (t \geq t_{\text{lag}}) \quad (3)$$

$$V \frac{dC_X}{dt} = \frac{Y_{X/S}}{Y_{\text{cell}/S}} \cdot \mu \cdot N_A + p_X \cdot N_V \quad (4)$$

where β and γ , empirical parameters; μ , specific growth rate; μ_{\max} , maximum specific growth rate; C_X , metabolite concentration; K_X , Monod constant; N_A , attached cell number; N_V , viable cell number; P , metabolic product (i.e., lactate, ammonia); p_X , specific metabolite consumption/production rate; S , nutrient (i.e., glucose); t , time; t_{lag} , lag time; V , working medium volume; X_0 , seeding density; X_0^* , 1 cell/cm²; X_A , cell density on a bottom surface; X_{\max} , maximum cell density on the surface; $Y_{\text{cell}/S}$ and $Y_{X/S}$, nutrient consumption coefficients for cell mass production and metabolite consumption/production, respectively.

3.3. STEP 3: Decision gate

This step ensured the developed model adequately described cell number (i.e., CQA) and ammonia concentration (i.e., constraint). The model was fitted by parameter estimation, where the Nelder-Mead method was used to minimize the normalized root mean square error (NRMSE) between the predicted values and the literature observations (Jossen et al., 2020). The accuracy criterion (ϵ) was set as 10 % and 20 % for the cell number and

ammonia concentration, respectively. The NRMSE was calculated to be 5.8 % and 17 % for the cell number and ammonia concentration, respectively. As such, the workflow was able to proceed to the next step.

3.4. STEP 4: Dynamic and stochastic simulation

Based on literature, maximum specific growth rate (μ_{\max}) and maximum cell density (X_{\max}) were identified as model parameters that could vary due to raw material and process/operation uncertainties (Jossen et al., 2020; Schop et al., 2009). The maximum specific growth rate was determined as 2.71×10^{-2} /h (based on the parameter estimation in Step 3), while the standard deviation was set to 1.59×10^{-3} /h based on literature (Higuera et al., 2009). For the maximum cell density, the average (6.5×10^4 cells/cm²) and standard deviation (2.0×10^3 cells/cm²) were identified from the literature (Jossen et al., 2020).

The model developed in Step 2 was then executed 1,000 times using uncertain model parameters randomly sampled from their respective probability distributions. The model generated time-dependent profiles of the viable cell number (i.e., CQA) and the ammonia concentration (i.e., constraint). The time-dependent outputs of the 1,000 runs were then compared to the CQA and constraint requirements between days 8 and 11 to identify the time-dependent probability that a given run would meet all requirements (CQA and constraint). This process was repeated on 10,000 CPP combinations that uniformly covered the CPP ranges (i.e., the seeding density range: 6.0×10^3 – 1.0×10^4 cells/cm², medium change ratio range: 0.2–1.0).

The default *scipy.integrate.solve_ivp* solver with the BDF method was used for solving the ODEs, and the algorithm was implemented in Python 3.8. The total CPU time for the design space determination was ca. 15 h using an Apple M1 processor with 16 GB RAM.

3.5. STEP 5: Design space determination

Figure 3 visualizes the probabilistic design space between days 8 to 11 on a 2D map, where seeding density and medium change ratio (CPPs) are the X-Y axis, respectively, while contours are the predicted probability. The acceptable minimum probability of 90 % was adopted in this case study, in line with literature (García-Muñoz et al., 2015). Figure 3 shows that Day 10 provides the broadest design space for the adopted probability preference. Further investigation found that before Day 8, no CPP combination resulted in at least 90% of the batches producing 1.4×10^6 cells (lower bound of the CQA). In fact, most CPP combinations resulted in no batch reaching this criterion (visualized as a black region). On Day 9, CPP combinations with high seeding densities and medium change ratios resulted in over 90 % of the batches producing more than 1.4×10^6 cells (visualized as a white region). The underlying reason for this observation is that 1) high seeding densities provide a head start, and 2) a high change ratio accelerates cell growth due to the removal of inhibitors. However, as the cultivation progressed, the same high seeding

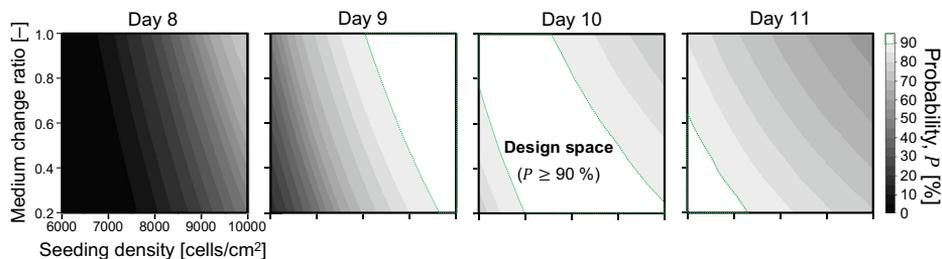


Figure 3. Visualized time evolving probabilistic design space between on days 8–11.

density and medium change ratio combinations resulted in batches exceeding 1.7×10^6 cells, which was the upper bound of the CQA requirement. This effect was observed as an enlarging grey area on the top righthand side of the design space on Days 10 and 11. To this end, Day 10 was the ideal cultivation termination date where the broadest set of CPP combinations produced at least 90 % of the batches between the 1.4 – 1.7×10^6 cells. In this design space determination, the ammonia constraint was found to be non-bounding.

4. Conclusions

This work detailed the development and application of a workflow that determined a dynamic and probabilistic design space for MSC cultivation. In the case study, the proposed workflow captured the dynamics of the MSC cultivation process by a Monod kinetic based mechanistic model. Raw material and process/operation uncertainties were then incorporated into the model by linking these uncertainties to the model parameters (maximum specific growth rate and maximum cell density). A dynamically evolving probabilistic design space was then determined using the QbD framework. The design space was then visualized over days 8–11 of the cultivation with the aid of an X-Y contour map. The contour map was then used as an early-stage decision support tool to determine the optimal cultivation time. In the future, this workflow can be further expanded to incorporate more realistic probability estimations and be developed into a detailed methodology applicable to a more complex case study.

Acknowledgements

This research was supported by AMED under Grant Number JP21be0704001.

References

- Andersen, E.B., et al., 2020, Big Data Generation for Time Dependent Process: The Tennessee Eastman Process for Generating Large Quantities of Process Data, *Comput. Aided Chem. Eng.*, 48, 1309–1314
- Badr, S., et al., 2021, Integrated design of biopharmaceutical manufacturing processes: Operation modes and process configurations for monoclonal antibody production, *Comput. Chem. Eng.*, 153, 107422
- European Medicines Agency, 2017, ICH guideline Q8 (R2) on pharmaceutical development
- García-Muñoz, S., et al., 2015, Definition of Design Spaces Using Mechanistic Models and Geometric Projections of Probability Maps, *Org. Process Res. Dev.*, 19, 1012–1023
- Higuera, G., et al., 2009, Quantifying in vitro growth and metabolism kinetics of human mesenchymal stem cells using a mathematical model, *Tissue Eng. Part A*, 15, 2653–2663
- Jossen, V., et al., 2020, An Approach towards a GMP Compliant In-Vitro Expansion of Human Adipose Stem Cells for Autologous Therapies, *Bioengineering*, 7, 77
- Kino-oka, M., et al., 2000, Valuation of Growth Parameters in Monolayer Keratinocyte Cultures Based on a Two-Dimensional Cell Placement Model, *J. Biosci. Bioeng.*, 89, 285–287
- Lipsitz, Y.Y., et al., 2016, Quality cell therapy manufacturing by design, *Nat. Biotechnol.*, 34, 393–400
- Ochoa, M.P., et al., 2021, Novel flexibility index formulations for the selection of the operating range within a design space, *Comput. Chem. Eng.*, 149, 107284
- Redaelli, S., et al., 2012, From cytogenomic to epigenomic profiles: Monitoring the biologic behavior of in vitro cultured human bone marrow mesenchymal stem cells, *Stem Cell Res. Ther.*, 3, 47
- Schop, D., et al., 2009, Growth, Metabolism, and Growth Inhibitors of Mesenchymal Stem Cells, *Tissue Eng. Part A*, 15, 8, 1877–1886
- Ullah, I., et al., 2015, Human mesenchymal stem cells - current trends and future prospective, *Biosci. Rep.*, 35, 2

Deterministic global optimization of multistage layer melt crystallization using surrogate models and reduced space formulations

Christian Kunde^a, Ronald Méndez^b and Achim Kienle^{a,b}

^a*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany*

^b*Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*
kunde@mpi-magdeburg.mpg.de

Abstract

This case study considers the deterministic global optimization of a multistage separation process using melt crystallization. A detailed model of the crystal growth process is replaced by a computationally more efficient data-based surrogate trained on simulation data from the detailed model. The resulting optimization problem is solved with the deterministic global solver MAiNGO using a reduced space formulation. The results demonstrate that this is a feasible approach for identifying globally optimal designs for this complex separation process.

Keywords: deterministic global optimization, layer melt crystallization, surrogate modeling, reduced space formulation, multistage crystallizer network

1. Introduction

Layer melt crystallization is a thermal separation process with very high theoretical selectivity in case of eutectic mixtures, lower required energy for the phase change than e.g. distillation, and the ability to treat heat sensitive substances (Ulrich et al., 1996). Crystal layers that grow on a cooled surface incorporate impurities from the melt if the growth rate is not extremely small (Wintermantel, 1986), which may necessitate additional recrystallization steps in a multistage process to achieve a desired purity. The layer growth process determines the separation efficiency of each stage and can be described by complex dynamic models. However, including differential equations in mathematical process optimization is computationally unfavorable, and using strongly simplified correlations for the separation efficiency could lead to significant prediction errors and suboptimal design decisions. Instead, an artificial neural network (ANN) that reproduces the separation efficiency accurately is used here as a surrogate for a dynamic model adapted from (Beierling et al., 2014). The surrogate model is trained on simulation data from the dynamic model and combined with mechanistic equations to create a hybrid overall process model that is used for the optimization of multistage crystallizer networks. In order to obtain certified globally optimal designs, the deterministic global solver MAiNGO (Bongartz et al., 2018) is used to solve a reduced space formulation of the optimization problem, i.e. a formulation in which the dimensionality of the problem is not increased by intermediate variables, in particular those appearing in ANNs. The structure of the following sections is as follows: The original model, the surrogate approach, and the optimization problem are presented in Section 2, results are summarized in Section 3, and a conclusion is given in Section 4.

2. Methods

2.1. Melt crystallization model

The layer melt crystallization model used here is adapted from the model presented in (Beierling et al., 2014). Beierling et al. (2014) consider a falling film crystallization process. In this process, the liquid mixture, i.e. the melt, continuously flows down the cooled inner surface of a tube and is pumped back to the top. A crystal layer grows on the cooled inner surface of the tube until a desired layer thickness is achieved and the remaining liquid is removed from the tube. In the next step, the crystal layer is melted by increasing the temperature of the tube and collected separately. Beierling et al. (2014) consider systems with eutectic solid-liquid equilibria, for which no impurities are embedded in the crystals themselves. However, if the growth rate is not extremely small, the crystal layer will incorporate impurities from the melt through liquid inclusions.

In a mixture containing components A and B, where A is the crystallizing component, the distribution of component B between melt and crystal layer is described by the following set of equations (Wellington and Wintermantel, 1991; Beierling et al., 2014).

$$\frac{\partial m_B^{\text{cr}}}{\partial m^{\text{cr}}} = k_{\text{diff}} w_B^{\text{m}} \quad (1)$$

$$k_{\text{diff}} = e \phi^f \quad (2)$$

$$\phi = \frac{w_B^{\text{m}}}{1 - w_B^{\text{m}}} \left(e^{\frac{v \rho^{\text{cr}}}{\beta \rho^{\text{m}}}} - 1 \right) \quad (3)$$

Here, $\frac{\partial m_B^{\text{cr}}}{\partial m^{\text{cr}}}$ is the differential increase of mass of the impurity B in the crystal layer m_B^{cr} per increase of total mass of the crystal layer m^{cr} , $k_{\text{diff}} \in [0, 1]$ is the differential distribution coefficient, and w_B^{m} is the mass fraction of B in the melt. For $k_{\text{diff}} = 0$ there are no liquid inclusions in the crystal layer and for the theoretical limit case $k_{\text{diff}} = 1$, the crystal layer contains only liquid inclusions. The dimensionless number ϕ is used to correlate the properties of the melt with the differential distribution coefficient k_{diff} , using parameters e and f . The crystal growth rate is denoted by v , the mass transport coefficient at the crystal surface by β , and the densities of the crystals and the melt by ρ^{cr} and ρ^{m} , respectively.

Equation (2) is a heuristic expression introduced in (Beierling et al., 2014) and should only be used in a range that is supported by experimental data. Otherwise, it may produce non-physical values, e.g. values of k_{diff} larger than one for large values of ϕ and positive parameters e, f .

Using dimensionless variables Y^{m} and Y_B^{m} relating to the liquid phase

$$m^{\text{m}} := Y^{\text{m}} m_0^{\text{m}}, \quad m_B^{\text{m}} := Y_B^{\text{m}} m_{B,0}^{\text{m}}, \quad (4)$$

with total initial melt mass m_0^{m} and mass of B in the initial melt $m_{B,0}^{\text{m}}$, the model Equations (1)-(3) are rewritten as

$$\frac{\partial Y_B^{\text{m}}}{\partial Y^{\text{m}}} = e \left(\frac{w_B^{\text{m}}}{1 - w_B^{\text{m}}} \left(e^{\frac{v \rho^{\text{cr}}}{\beta \rho^{\text{m}}}} - 1 \right) \right)^f \frac{Y_B^{\text{m}}}{Y^{\text{m}}}, \quad w_B^{\text{m}} = w_{B,0}^{\text{m}} \frac{Y_B^{\text{m}}}{Y^{\text{m}}}, \quad Y_B^{\text{m}}(Y^{\text{m}}=1) = 1, \quad (5)$$

where $w_{B,0}^{\text{m}}$ is the mass fraction of B in the initial melt. Note that the classical freezing rate is obtained by $Y^{\text{cr}} := \frac{m^{\text{cr}}}{m_0^{\text{m}}} = 1 - Y^{\text{m}}$.

In the present study, an operating mode with constant crystal growth rates v is considered. Additionally, the constant values for $e, f, \rho^{\text{cr}}, \rho^{\text{m}}$ and β provided in (Beierling et al., 2014) for a mixture of n-dodecanal/iso-dodecanal are used, see Table 1.

Instead of using the solution of Equation (5) directly as a melt crystallization model, we introduce an effective differential distribution coefficient $k_{\text{diff}}^{\text{eff}}$ with

$$Y_B^{\text{m}} := (Y^{\text{m}})^{k_{\text{diff}}^{\text{eff}}}, \quad (6)$$

in analogy to layer melt crystallization with a constant differential distribution coefficient. The effective differential distribution coefficient is then a function $k_{\text{diff}}^{\text{eff}} = f(Y^{\text{m}}, w_{\text{B},0}^{\text{m}}, v)$ of the final liquid yield, the initial mass fraction of B, and the crystal growth speed. This function is approximated by a data-based surrogate $k_{\text{diff}}^{\text{eff}} \approx \hat{f}(Y^{\text{m}}, w_{\text{B},0}^{\text{m}}, v)$ as described in the subsequent sections. This strategy removes the differential equation (5) from the model in order to be able to apply deterministic global optimization for a multi-stage separation process, at the cost of introducing some approximation error.

2.2. Crystallizer network model

With the results from the previous section, the model for each melt crystallization stage $n = 1, \dots, N$ reads

$$\dot{m}_n^{\text{m}} = Y_n^{\text{m}} \dot{m}_{0,n}^{\text{m}}, \quad \dot{m}_{\text{B},n}^{\text{m}} = Y_{\text{B},n}^{\text{m}} \dot{m}_{\text{B},0,n}^{\text{m}}, \quad \dot{m}_n^{\text{cr}} = \dot{m}_{0,n}^{\text{m}} - \dot{m}_n^{\text{m}}, \quad \dot{m}_{\text{B},n}^{\text{cr}} = \dot{m}_{\text{B},0,n}^{\text{m}} - \dot{m}_{\text{B},n}^{\text{m}}, \quad (7)$$

$$Y_{\text{B},n}^{\text{m}} = (Y_n^{\text{m}})^{k_{\text{diff},n}^{\text{eff}}}, \quad k_{\text{diff},n}^{\text{eff}} = \hat{f}\left(Y_n^{\text{m}}, \frac{\dot{m}_{\text{B},0,n}^{\text{m}}}{\dot{m}_{0,n}^{\text{m}}}, v_n\right) \quad (8)$$

where \dot{m} -variables are time-averaged mass flows for the batch-operated stages. Each stage has two output flows: one for the remaining melt after crystallization $\dot{m}_n^{\text{m}}, \dot{m}_{\text{B},n}^{\text{m}}$ and one for the melted crystal layer $\dot{m}_n^{\text{cr}}, \dot{m}_{\text{B},n}^{\text{cr}}$.

The feed mass flows $\dot{m}_{0,n}^{\text{m}}, \dot{m}_{\text{B},0,n}^{\text{m}}$ for each stage $n = 1, \dots, N$ are calculated as

$$\dot{m}_{0,n}^{\text{m}} = \dot{m}_0 \beta_n^{\text{feed}} + \sum_{l=1}^N \dot{m}_l^{\text{m}} \beta_{l,n}^{\text{m}} + \sum_{l=1}^N \dot{m}_l^{\text{cr}} \beta_{l,n}^{\text{cr}}, \quad (9)$$

$$\dot{m}_{\text{B},0,n}^{\text{m}} = \dot{m}_{\text{B},0} \beta_n^{\text{feed}} + \sum_{l=1}^N \dot{m}_{\text{B},l}^{\text{m}} \beta_{l,n}^{\text{m}} + \sum_{l=1}^N \dot{m}_{\text{B},l}^{\text{cr}} \beta_{l,n}^{\text{cr}}, \quad (10)$$

and the overall network product flows $\dot{m}^{\text{m}}, \dot{m}^{\text{cr}}$ as

$$\dot{m}^{\text{m}} = \sum_{n=1}^N \dot{m}_n^{\text{m}} \beta_n^{\text{m,out}}, \quad \dot{m}^{\text{cr}} = \sum_{n=1}^N \dot{m}_n^{\text{cr}} \beta_n^{\text{cr,out}}, \quad (11)$$

$$\dot{m}_{\text{B}}^{\text{m}} = \sum_{n=1}^N \dot{m}_{\text{B},n}^{\text{m}} \beta_n^{\text{m,out}}, \quad \dot{m}_{\text{B}}^{\text{cr}} = \sum_{n=1}^N \dot{m}_{\text{B},n}^{\text{cr}} \beta_n^{\text{cr,out}}. \quad (12)$$

Here, $\beta_n, \beta_{l,n}$ are binary parameters that determine mass flow connections in the crystallizer network and $\dot{m}_0, \dot{m}_{\text{B},0}$ is the overall feed of the network.

Individual crystallizer network configurations, i.e. parameter sets for $\beta_n, \beta_{l,n}$, are selected from the superstructure depicted for $N=3$ stages in Figure 1. This results in 4 unique configurations for $N=2$ stages and 31 unique configurations for $N=3$ stages.

The product specifications are given as lower and upper bounds of the mass fraction of B in each product flow.

$$\dot{m}_{\text{B}}^{\text{m}} \geq w_{\text{B},\text{lo}}^{\text{m}} \dot{m}^{\text{m}} \quad (13)$$

$$\dot{m}_{\text{B}}^{\text{cr}} \leq w_{\text{B},\text{up}}^{\text{cr}} \dot{m}^{\text{cr}} \quad (14)$$

Table 1: Parameter specifications for the crystallizer model.

parameter	value	unit
ρ^{cr}	938.6 ^a	kg m ⁻³
ρ^{m}	844.24 ^a	kg m ⁻³
e	0.4903 ^a	1
f	0.3493 ^a	1
β	3.67 · 10 ⁻⁶ ^a	ms ⁻¹

^a(Beierling et al., 2014)

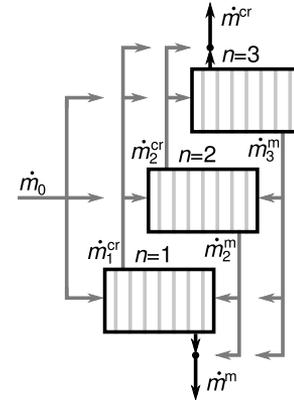


Figure 1: Superstructure for the crystallizer network for $N=3$ stages.

The crystallization model is valid only in a limited range for ϕ , i.e. in the range of measurement data for the original model in (Beierling et al., 2014), and for the mass fraction of B in the melt, i.e. for values sufficiently far away from the eutectic composition. These additional constraints are added to the model for each stage $n = 1, \dots, N$ as

$$\dot{m}_{B,n}^m \leq w_{B,\text{up}}^m \dot{m}_n^m \quad (15)$$

$$\dot{m}_{B,n}^m \left(e^{\frac{v_n \rho^{\text{cr}}}{\beta \rho^{\text{m}}}} - 1 \right) \leq \phi_{\text{up}} (\dot{m}_n^m - \dot{m}_{B,n}^m), \quad \dot{m}_{B,n}^m \left(e^{\frac{v_n \rho^{\text{cr}}}{\beta \rho^{\text{m}}}} - 1 \right) \geq \phi_{\text{lo}} (\dot{m}_n^m - \dot{m}_{B,n}^m) \quad (16)$$

Additionally, the crystal growth speeds on each stage $n = 1, \dots, N$ are limited by

$$v_n \in [v_{\text{lo}}, v_{\text{up}}]. \quad (17)$$

The objective function, taken from (Kunde et al., 2016), represents total annualized cost comprising storage tanks for each crystallization stage, a heat exchanger as the actual crystallizer vessel, as well as energy costs for cooling. Parameters are updated to the current case study and the objective function is adapted for mass flows and a variable crystal growth speed. The resulting objective function can be written in a condensed form as

$$J = \left(a_1 J_1^{0.65} + a_2 N J_1^{0.32} + a_3 J_2 \right) \text{€ year}^{-1}, \quad (18)$$

$$J_1 = \sum_{n=1}^N \dot{m}_n^{\text{cr}} v_n^{-1} \text{kg}^{-1} \text{m}, \quad J_2 = \sum_{n=1}^N \dot{m}_n^{\text{cr}} \text{kg}^{-1} \text{s}. \quad (19)$$

The model parameters used in the following case study are listed in Table 2.

Table 2: Additional parameter and domain specifications for the case study, $n = 1, \dots, N$.

parameter	value	unit	variable	domain	unit
$(\dot{m}_0, \dot{m}_{B,0})$	(0.3, 0.0234)	kg s^{-1}	Y_n^m	[0.05, 0.99]	1
$(w_{B,\text{up}}^{\text{cr}}, w_{B,\text{lo}}^{\text{m}})$	(0.01, 0.4)	1	$\dot{m}_{0,n}^m, \dot{m}_n^{\text{cr}}$	$[10^{-9}, 10]$	kg s^{-1}
$w_{B,\text{up}}^m$	0.90	1	$\dot{m}_{B,0,n}^m, \dot{m}_{B,n}^{\text{cr}}$	[0, 10]	kg s^{-1}
$(\phi_{\text{lo}}, \phi_{\text{up}})$	(0.01, 0.45)	1	v_n	$[0.01\beta, 0.5\beta]$	m s^{-1}
a_1	$1.030049 \cdot 10^1$	1			
a_2	$4.744830 \cdot 10^1$	1			
a_3	$1.273072 \cdot 10^5$	1			

2.3. Data generation and surrogate training

The data required for the training of the surrogate model $k_{\text{diff}}^{\text{eff}} \approx \hat{f}(Y^m, w_{B,0}^m, v)$ is obtained by numerical solution of Equation (5) using “ode45” in MATLAB 2019b and Equation (6). Data points are calculated on a grid for $Y^m \in \{0.01, 0.012, \dots, 0.99, 1\}$, $v \in \{0.005\beta, 0.01\beta, \dots, 0.5\beta, 0.501\beta\}$, and $w_{B,0}^m \in \{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 0.01, 0.02, \dots, 0.99\}$. In the crystallization model, feasible input values of f are constrained by bounds on $w_{B,0}^m$, ϕ , and v , see also Equations (15)-(17). In order to improve the data support on the boundary of this feasible region, an extended feasible region is defined by relaxing the corresponding constraints by a small amount. All data points that lie inside the extended feasible region are used for the surrogate training.

An ANN $\hat{f}(Y^m, w_{B,0}^m, v) = b^{[2]} + W^{[2]} \sigma(b^{[1]} + W^{[1]}(Y^m, w_{B,0}^m, v)^\top)$ comprising a single hidden layer with 20 neurons is used as a surrogate, where the activation function σ is the element-wise hyperbolic tangent. The surrogate is trained using the function “train” in MATLAB 2019b, with Levenberg-Marquardt backpropagation minimizing the mean squared error (MSE) and default options except for the minimum gradient set to 10^{-16} and the regularization parameter set to 10^{-11} .

The final MSE for all data points inside the feasible region is $2.0603 \cdot 10^{-08}$. Additionally, the corresponding regression results for the mass fraction of B in the melted crystal layer are shown in Figure 2. The relative error of predictions \hat{w}_B^{cr} for w_B^{cr} using the trained ANN is less than 1% for data points in the feasible region, including for values close to zero.

2.4. Optimization

In order to minimize the cost J of the process such that the model constraints Equations (7)-(19) are fulfilled, the optimization problem is solved as a reduced space formulation in the deterministic global solver MAiNGO (Bongartz et al., 2018). The general structure of the optimization problem is

$$\begin{aligned} \min_{x \in X = [x_{\text{lo}}, x_{\text{up}}]} \quad & J(x, y) \\ \text{s. t.} \quad & h(x, y) = 0, \\ & g(x, y) \leq 0, \\ & y = h_y(x), \end{aligned} \quad (20)$$

where $h_y(x)$ are explicit functions or sequences of functions of the selected optimization variables x that are used to substitute dependent variables y . The resulting optimization problem has a much smaller dimension compared to a problem where y are included in the optimization variables. In particular, this avoids introducing a large number of intermediate variables for the activations of the neurons in an ANN. MAiNGO solves reduced space formulations using automated methods for calculating relaxations for $h_y(x)$ in a branch-and-bound framework, as well as tailored relaxations for specific expressions such as the hyperbolic tangent appearing in ANNs (Schweidtmann and Mitsos, 2019). The optimization variables and bounds that are used in this case study are listed in Table 2. The optimization problem is solved as a non-linear program for each individual configuration selected from the superstructure in Figure 1. The problem is solved in MAiNGO v0.4.0 with subsolvers IPOPT and CLP with default options except for the absolute optimality gap, which was reduced to 10^{-6} such that the algorithm terminates at the default relative optimality gap of 1%. Additionally, for all configurations with $N=3$ stages, the target lower bound is set to the best known upper bound for configurations with $N=2$ stages rounded up to two significant digits in order to save computation time, since the latter case is much faster to solve.

3. Results

The optimal crystallizer network configuration is shown in Figure 3. This configuration comprises a pre-treatment stage ($n=2$) that separates the mixture to below specification for the melted crystal product and two purification stages ($n=1, 3$) that are connected in series and separate the remaining mixture to above product specification. By mixing both product flows, \hat{m}^{cr} is exactly at specification. For the considered case study, the mixture has to be processed by at least two stages in a countercurrent fashion to achieve the product specifications. By including the pre-treatment stage, a smaller amount of the mixture has to be processed in that way, thereby reducing overall cost. The optimal cost for the process is $1.22680 \cdot 10^5 \text{€} \cdot \text{year}^{-1}$. This is compared to the optimal cost of $1.26504 \cdot 10^5 \text{€} \cdot \text{year}^{-1}$ when limiting possible configurations to countercurrent cascades of crystallizers. Note that, in order to allow for comparability, the parameter values were chosen such that constraints due to model validity (15)-(17) were not active and no variable was at their bound

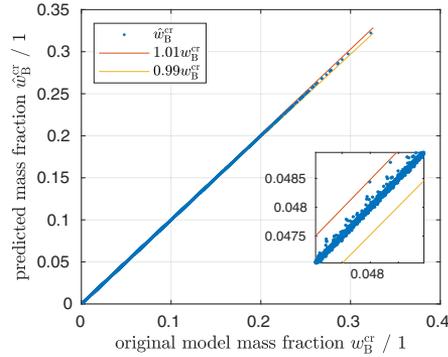


Figure 2: Regression results for the mass fraction of B in the crystal layer $w_B^{\text{cr}} = \frac{1-(y^m)^{\text{diff}}}{1-y^m} w_{B,0}^m$. The inlay shows a magnification of the value with the largest relative error.

at both optimal solution points.

We observed computation times of up to 1300s for configurations with three stages on a standard desktop computer. The overall computation time for all configurations with two or three stages was 6500s. The additional effort for data generation and surrogate training was in the order of 200s.

4. Conclusions

This case study demonstrates the applicability of deterministic global optimization to multistage melt crystallization processes by replacing the computationally expensive dynamic crystal growth model with a static surrogate. The approach can be extended to more detailed crystallization models provided sufficient input-output data is available. This includes operating modes other than constant crystal growth rates as long as the operating mode can be described by a limited number of parameters.

Local optimization approaches might also benefit from surrogates for crystallizer models by trading some model accuracy for computationally more efficient models.

The same optimal configuration as in the current case study was identified before for a simplified crystallization model with constant differential distribution coefficient and equal crystal growth rate on each stage in (Kunde et al., 2016), with cost reductions up to 25% compared to countercurrent cascades in a large parameter region. It was also found that countercurrent cascades are optimal in most of the remaining parameter space. An analogous parameter study could show whether those results can be transferred to the current model and how much improvement over the countercurrent cascade can be achieved.

Acknowledgements: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – TRR 63 “Integrierte chemische Prozesse in flüssigen Mehrphasensystemen” (Teilprojekt B9) – 56091768.

References

- T. Beierling, J. Micovic, P. Lutze, G. Sadowski, 2014. Using complex layer melt crystallization models for the optimization of hybrid distillation/melt crystallization processes. *Chemical Engineering and Processing: Process Intensification* 85, 10–23.
- D. Bongartz, J. Najman, S. Sass, A. Mitsos, 2018. MAiNGO - McCormick-based Algorithm for mixed-integer Nonlinear Global Optimization. Tech. rep., Process Systems Engineering (AVT.SVT), RWTH Aachen University. URL <http://permalink.avt.rwth-aachen.de/?id=729717>
- C. Kunde, D. Michaels, J. Micovic, P. Lutze, A. Górák, A. Kienle, 2016. Deterministic global optimization in conceptual process design of distillation and melt crystallization. *Chemical Engineering and Processing: Process Intensification* 99, 132–142.
- A. M. Schweidtmann, A. Mitsos, 2019. Deterministic global optimization with artificial neural networks embedded. *Journal of Optimization Theory and Applications* 180 (3), 925–948.
- J. Ulrich, J. Bierwirth, S. Henning, 1996. Solid layer melt crystallization. *Separation & Purification Reviews* 25 (1), 1–45.
- G. Wellinghoff, K. Wintermantel, 1991. Schmelzkristallisation - theoretische Voraussetzungen und technische Grenzen [Melt crystallization - theoretical presumptions and technical limitations]. *Chemie-Ingenieur-Technik* 63 (9), 881–888, 89.
- K. Wintermantel, 1986. Die effektive Trennwirkung beim Ausfrieren von Kristallschichten aus Schmelzen und Lösungen - eine einheitliche Darstellung. *Chemie Ingenieur Technik* 58 (6), 498–499.

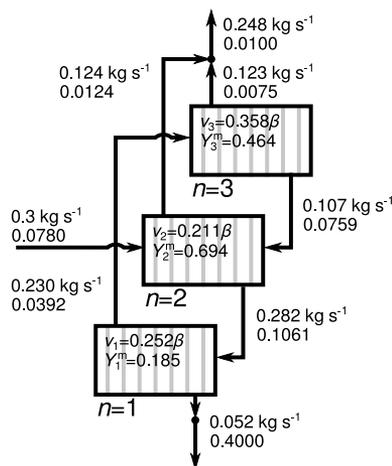


Figure 3: Optimal design for the considered case study for crystallizer networks with up to three stages. The number pairs are total mass flow and mass fraction of B at each location.

Model-Based Design Space for Protein A Chromatography Resin Selection

Steven Sachio, Cleo Kontoravdi and Maria M. Papathanasiou

Centre for Process Systems Engineering, Imperial College London, London SW7 2AZ, United Kingdom

maria.papathanasiou11@imperial.ac.uk

Abstract

As demand for biopharmaceuticals rises, manufacturers are required to meet multiple competing key performance indicators (KPIs) such as process sustainability, efficiency and product efficacy and quality. Advanced process optimisation and control in biopharmaceutical manufacturing is challenged by the lack of online Process Analytical Technologies (PAT). This results in processes relying heavily on wet-lab experimentation, which may be costly and inefficient. In this work, a novel methodology for evaluating process robustness and alternative operating strategies using design space identification is proposed to accelerate process design and optimisation. The focus in this work is on the initial separation step for the purification of monoclonal antibodies (mAbs) separating the majority of process impurities generated upstream using affinity (protein A) chromatography. A high fidelity process model is used to computationally explore the multi-dimensional design space. The performance and robustness of the process under three different resin properties and a variety of input conditions are evaluated using the framework. Three scenarios for each of the resins are considered resulting in a total of nine design spaces. The results indicate that using a higher column protein A density resin can increase operational flexibility.

Keywords: computing and systems engineering, process design and development, process optimisation, biopharmaceutical manufacturing

1. Introduction

Biopharmaceuticals is the fastest growing sector of the pharmaceutical industry (Walsh, 2018; Philippidis, 2019). As demand for biopharmaceuticals rises and there is a drive towards sustainable processes, now manufacturing needs to meet multiple competing key performance indicators (KPIs) (Nasr et al., 2017). The Quality by Design (QbD) initiative was started to address inefficient, out-of-date and wasteful practices to process development and manufacturing. QbD encourages systematic decision making and deep understanding of pharmaceutical development and manufacturing in contrast to the traditional statistical experimentation approach (Rathore and Winkle, 2009). This initiative encourages manufacturers to design quality into the product from the early phases of drug development instead of treating it as a secondary objective. Monoclonal antibodies (mAbs) are the most popular biologics as they can be manufactured to target almost any substance. However, the implementation of QbD for biopharmaceuticals, specifically therapeutic proteins, faces three main challenges: product heterogeneity, limited measurements for process monitoring and lack of online process control strategies due to lack of online Process Analytical Technologies (PAT) (Papathanasiou and Kontoravdi, 2020). In this work, a novel methodology for evaluating process robustness and process optimisation using model-based design space is proposed to accelerate process design and optimisation. The methodology is intended to be applied

during the process design phase within the QbD framework to act as a virtual experimentation platform to reduce reliance on resource-intensive experimentation. The capture step of the mAb production process is the main focus in the case study of this work.

2. The Framework

Evaluating robustness and performance of a process under different parameters and conditions is vital for designing optimal and robust processes. In this section, the framework for model-based design space identification is presented. A schematic of the framework is shown in Figure 1.

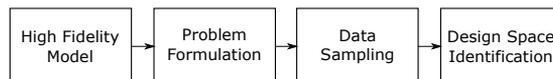


Figure 1: Model-based design space identification framework.

Step 1. High Fidelity Model. First, a high fidelity model is developed and validated using experimental data to ensure accurate description of the system. The validity of the findings observed from the framework is highly determined by the quality of the model. Therefore, ensuring that the model is validated across the range of parameters which will be sampled from is important.

Step 2. Problem Formulation. Next, the problem is formulated starting with identifying what are the objectives of this study. What are the key performance indicators (KPIs) of interest and what are the parameters/variables (e.g., design variables, process parameters) that the design space need to cover. It is also possible to introduce constraints to the problem such as target minimum yield.

Step 3. Data Sampling. The choice of the sampling method used can affect the computation time, accuracy and the coverage of the space. In this work, the quasi-random Sobol sequence (Sobol and Shukman, 1993) is used to sample the multi-dimensional design space with good coverage. For application in process design and development within the QbD paradigm, it is crucial to ensure good accuracy, hence, the high fidelity model is used directly for data sampling.

Step 4. Design Space Identification. Finally the design space is visualised and identified from the collected computational data samples. The constraints are applied to find out which sample points do not satisfy the desired performance constraints.

The multi-dimensional design space can be used to evaluate the robustness of the process by identifying acceptable disturbance ranges. This information can also be used to identify the optimal design given a range of disturbance expected. The case study in the next section presents how the framework can be applied to the capture process of mAb production.

3. Case Study: Monoclonal Antibody Capture Process

The capture process is the first step in downstream separation of mAbs. The objective of this process is to concentrate the Harvest cell culture fluid (HCCF) from upstream reactor removing the bulk impurities. In this process, the mAb concentration from upstream is treated as a disturbance as there can be fluctuations. This variation in feed mAb concentration can impact the separation performance of the process. Model-based design space identification (DSI) can be used to design a process which is able to handle the variation in feed mAb concentration.

3.1. The System

For the purposes of this work, the two-column capture process model as presented in Steinebach et al. (2016) is adapted. The two-column system consist of three major steps: A – interconnected

loading, B – interconnected wash, and C – batch operation (washing, elution, and regeneration). As shown in Figure 2, in the first step, column 1 is loaded beyond its dynamic binding capacity and the breakthrough is fed onto column 2. In step B, column 1 is washed and the breakthrough which may contain some mAb is mixed with the fresh feed and loaded onto column 2. In the last step the operation is in batch mode, column 1 is washed further and then the products are eluted and the column regenerated while column 2 is being loaded. Then the column positions are swapped, and the process is repeated.

3.2. Problem Formulation & Data Sampling

The objective for the DSI study is to investigate the performance and robustness of the process under different resin properties and conditions to find the optimum resin which satisfy the performance constraints of $Yield \geq 99\%$ and $Productivity \geq 15 \text{ mg mL}^{-1} \text{ h}^{-1}$. $Yield$ is the percentage of mAb recovered to the amount fed, while $Productivity$ is amount of mAb produced per time taken per resin volume used. The constraints are chosen based on the nominal performance reported by Steinebach et al. (2016). The resins used in this study are summarised in Table 1. The column protein A density is the density of protein A ligands per volume of the column. It was calculated based on the reported values of the protein A ligand density in the solid phase of the resin and both the bed and particle porosity (Hahn et al., 2005; Liu et al., 2015; Steinebach et al., 2016). Three scenarios of design spaces for each of the three different resins are considered and summarised in Table 2.

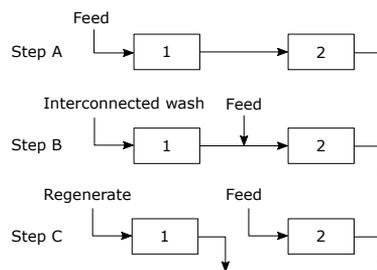


Figure 2: Schematic of the two-column capture process

Table 1: Column protein A density of resins.

Resin	Column Protein A Resin Density (mg mL^{-1})
MabSelect SuRe	34.41
MabSelect SuRe LX	43.82
CaptivA PriMAB	31.25

Table 2: Design space problem scenarios.

Scenario	Type	Parameters Varied
1	Column Dimension	D_c, HD_c
2	Switching Time	$T_{load}, T_{wash C}$
3	Process Parameters	$c_{feed}, Q_{feed}, T_{load}, T_{wash C}$

In the first scenario, only the column diameter, D_c , and column height to diameter ratio, HD_c , are varied. The height to diameter ratio is varied to ensure the shape remains like a column. In this scenario, the design space can be used to size the column for process design. In scenario two, only the loading time, T_{load} , and the interconnected washing time, $T_{wash|C}$, are varied (the times combined together is the switching time of the column). The switching time is one of the most important design variable which determines how long the column is loaded before the products are eluted and collected. In this scenario, the design space can help to choose the appropriate switching time which satisfies our production constraints. Finally, in the last scenario, the design space is a four-dimensional design space where the feed mAb concentration, c_{feed} , feed volumetric flowrate, Q_{feed} and the switching time parameters are varied simultaneously. The volumetric flowrate is a control variable of the process. When a disturbance in the form of fluctuations in mAb feed

concentration occurs, the volumetric flowrate can be adjusted to keep the mass flowrate of mAb going into the column within the designed range. The design space in the last scenario can identify ranges of c_{feed} and Q_{feed} which satisfies the performance constraints. Lastly, the quasi-random Sobol sequence (Sobol and Shukman, 1993) was used to sample the parameter space directly on the high fidelity model.

4. Results and Discussion

First, the design spaces of the resin MabSelect SuRe is shown and discussed in detail. The resin MabSelect SuRe is the nominal resin used in the study by Steinebach et al. (2016).

4.1. MabSelect SuRe Design Space

The design space of all three scenarios for the MabSelect SuRe resin is shown in Figure 3. A total of 128 samples are used to construct each design space.

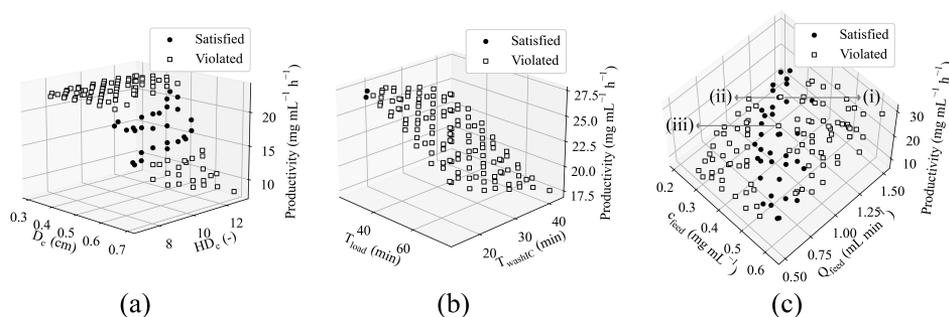


Figure 3: Design space of all scenarios for the MabSelect SuRe resin. Scenario 1, 2, and 3 are shown in (a), (b), and (c) respectively. Solid black circle markers indicate that the constraints are satisfied and white square that they are violated.

The design space for the first scenario is shown in Figure 3 (a) plotted in 3D (diameter, height to diameter ratio and productivity). A region of samples which satisfied both constraints can be seen. This information can be used to design a column which is able to satisfy both performance constraints. From Figure 3 (a), it can be seen that the majority of the samples satisfied the productivity constraint ($\geq 15 \text{ mg mL}^{-1} \text{ h}^{-1}$), while only a subset of samples shown in solid black circles satisfied both constraints. The yield constraint proves to be the limiting one for this design space.

Shown in Figure 3 (b) is the design space of scenario two. From Figure 3 (b) it can be seen that the yield is again the limiting constraint as all of the sample were able to satisfy the productivity constraint. Furthermore, the design space here is very narrow with only two samples which satisfied both constraints. This results aligns with the work by Steinebach et al. (2016) as the process was optimised using the MabSelect SuRe resin. The results also indicate that at lower switching times, higher productivity and yield is achieved. This is because when the switching time is reduced, the column is prevented from becoming saturated with mAb. Therefore, by lowering the switching time, the driving force for separation can be maintained at a higher value. Although lower switching time can give better performance in terms of yield and productivity, a lower loading time can result in a higher buffer consumption. Therefore, there is a trade-off between economic and environmental performance.

The four-dimensional design space from the last scenario is shown in Figure 3 (c). Looking at the design space in Figure 3 (c) it is clear that there are samples, marked (i) and (iii), which

violated one or more of the constraints despite the fact that it is very near to another sample, marked (ii), which satisfied both constraints. This is due to the fact that the design space here is a four-dimensional space represented in a two-dimensional space.

Table 3: Parameter and KPI values of samples (i), (ii), and (iii) in the design space of scenario three.

Sample	c_{feed}	Q_{feed}	T_{load}	T_{washIC}	Yield	Productivity
(i)	0.32	1.25	62.50	35.50	98.84	18.58
(ii)	0.31	1.25	43.36	28.62	99.17	18.33
(iii)	0.39	1.05	56.23	30.40	98.51	19.37

The labelled samples from Figure 3 (c) are shown in detail in Table 3. The samples (i) and (iii) did not satisfy the constraint due to violation in the yield constraint. On the other hand, sample (ii) satisfied both constraints. From Table 3, it can be seen that the value of T_{load} for both violated samples are significantly higher than that of sample (ii). From the design space in scenario two it was found out that to achieve higher yield, lower switching time is preferable. It is further evident here where samples (i) and (ii) are very close in the values of c_{feed} and Q_{feed} but having a larger T_{load} resulting in violation of the yield constraint. Using a lower switching time can reduce loss of products achieving higher yield. However, by doing so more buffer is consumed during washing, elution and regeneration as they are performed more often. There is, therefore, a trade-off which needs to be considered when designing the process.

4.2. Comparison of Resin Performance

Shown in Figure 4 is the design spaces of all three resins.

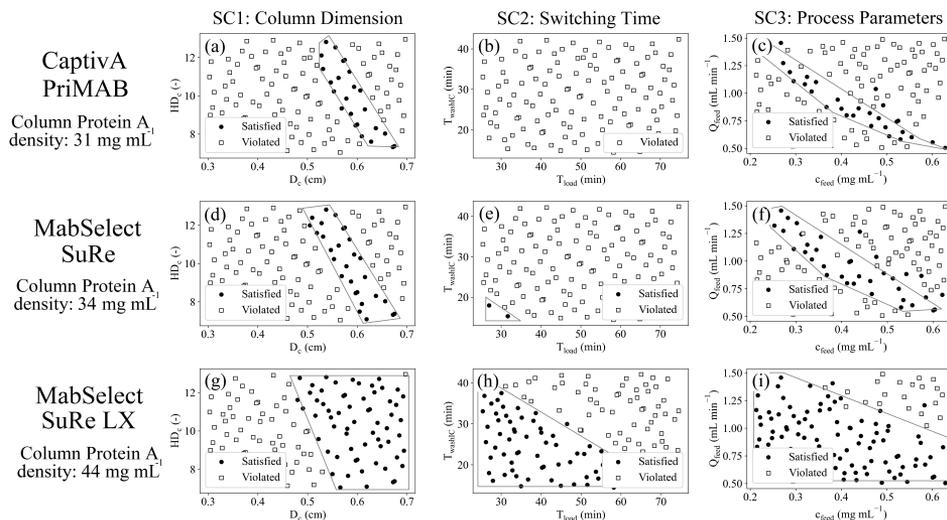


Figure 4: Comparison of the design spaces (as columns: scenarios 1, 2, and 3) for three different resins (as rows: Captiva PriMAB, MabSelect SuRe and MabSelect SuRe LX)

The Captiva PriMAB resin has the lowest column protein A density out of the three resins. From Figure 4 (c), it is shown that it does not have a single sample which satisfies both of the constraints. The column protein A density is directly correlated with how much mAb can be adsorbed onto the column. Hence, the lower the column protein A density the lower the yield is. From Section 4.1, it

was shown that the yield constraint is harder to satisfy than the productivity constraint. This leads to wider design spaces seen in resins with larger column protein A density shown in Figure 4.

Using the design spaces, the robustness of the process under different resin properties can be evaluated. For example, given a case where it is known that the upstream process produces a feed stream with mAb concentrations of between $0.35 - 0.45 \text{ mg mL}^{-1}$. Then, using the design spaces from Figure 4 (c), (f), and (i) it is possible to figure out what value of Q_{feed} that can satisfy the performance constraints. In this case, the CaptivA PriMAB resin can operate at between $0.85 - 0.90 \text{ mL min}^{-1}$, the MabSelect SuRe resin can operate between $0.85 - 1.00 \text{ mL min}^{-1}$, and the MabSelect SuRe LX resin can operate between $0.55 - 1.30 \text{ mL min}^{-1}$. This information can also be used in the opposite way. For instance, if the process has a nominal volumetric flowrate of 1.00 mL min^{-1} , the design spaces from Figure 4 (c), (f), and (i) shows that the CaptivA PriMAB resin can process feed streams containing c_{feed} of about $0.32 - 0.43 \text{ mg mL}^{-1}$ satisfying both constraints. On the other hand, the MabSelect SuRe and MabSelect SuRe LX resin can operate in the range of $0.31 - 0.45 \text{ mL min}^{-1}$ and $0.20 - 0.52 \text{ mL min}^{-1}$ respectively. Therefore, using the design spaces, an optimal resin that can satisfy the performance constraints and be robust against disturbance from the upstream process can be selected.

5. Conclusion

In this work, the framework for DSI within the QbD paradigm has been presented and applied in a case study. As shown in the case study, the proposed framework offers the opportunity to evaluate process robustness which is crucial for manufacture of biopharmaceuticals with stringent quality standards. Model-based DSI can be used to identify the range of operation conditions which satisfies a set of constraints. The case study has shown how model-based DSI can evaluate process robustness of protein A affinity chromatography for mAb capture under different resin properties and identify the optimum resin for a robust process. It was found out that resin with higher column protein A density has greater operational flexibility. This information is also important for process integration where it is vital to ensure compatibility between units to design a robust and optimal process. The approach can be transferred to other types of process operations (e.g., tangential flow filtration, bioreactor). It can be tailored to meet different types of study objectives (e.g., glycosylation in mAbs, buffer consumption) and to address other parameters (e.g., temperature, pH). Current and future work is focused on combining more advanced modelling techniques such as multi-parametric programming (mp-P) and data driven/hybrid models to accelerate data sampling in an effort to increase the framework's capabilities for online use.

References

- R. Hahn, P. Bauerhansl, K. Shimahara, C. Wizniewski, A. Tscheliessnig, A. Jungbauer, 2005. Comparison of protein A affinity sorbents ii. mass transfer properties. *J Chromatogr A* 1093 (1-2), 98–110.
- Z. Liu, S. S. Mostafa, A. A. Shukla, 2015. A comparison of protein A chromatographic stationary phases: performance characteristics for monoclonal antibody purification. *Biotechnol Appl Biochem* 62 (1), 37–47.
- M. M. Nasr, M. Krumme, Y. Matsuda, B. L. Trout, C. Badman, S. Mascia, C. L. Cooney, K. D. Jensen, A. Florence, C. Johnston, K. Konstantinov, S. L. Lee, 2017. Regulatory perspectives on continuous pharmaceutical manufacturing: Moving from theory to practice: September 26-27, 2016, international symposium on the continuous manufacturing of pharmaceuticals. *J Pharm Sci* 106 (11), 3199–3206.
- M. M. Papathanasiou, C. Kontoravdi, 2020. Engineering challenges in therapeutic protein product and process design. *Current Opinion in Chemical Engineering* 27, 81–88.
- A. Philippidis, 2019. Top 15 best-selling drugs of 2018. *Genetic Engineering & Biotechnology News* 39 (4), 16–17.
- A. S. Rathore, H. Winkle, 2009. Quality by design for biopharmaceuticals. *Nat Biotechnol* 27 (1), 26–34.
- I. M. Sobol, B. V. Shukman, 1993. Random and quasirandom sequences: Numerical estimates of uniformity of distribution. *Mathematical and Computer Modelling* 18 (8), 39–45.
- F. Steinebach, M. Angarita, D. J. Karst, T. Muller-Spath, M. Morbidelli, 2016. Model based adaptive control of a continuous capture process for monoclonal antibodies production. *J Chromatogr A* 1444, 50–56.
- G. Walsh, 2018. Biopharmaceutical benchmarks 2018. *Nat Biotechnol* 36 (12), 1136–1145.

Sustainable Energy Systems Design with Heat Pump Technologies for Meeting Peak Heating Demand

Xueyu Tian,^a Fengqi You^a

^a *Cornell University, Ithaca, New York, 14853, USA*
xt93@cornell.edu

Abstract

This study tackles the sustainable design of carbon-neutral energy systems with electrified peak heating systems, namely heat pump technologies for meeting electricity, heat, and cooling demand. We first propose a superstructure of the carbon-neutral energy systems for electricity, heat, and cooling generation. A multi-period optimization model with a predefined time horizon and temporal resolution is developed based on the proposed superstructure for total annualized cost minimization. The objective is to determine the optimal design of the carbon-neutral energy systems in the investigated region, seasonal operations, energy mix, and corresponding capacity of each technology while meeting the electricity, heat, and cooling demand. The applicability of the proposed modeling framework is illustrated through a case study using Cornell University as the living laboratory, explicitly considering the electrification of the peak heating systems.

Keywords: carbon neutrality; energy systems; decarbonization; electrification; mixed-integer nonlinear programming.

1. Introduction

The Paris Agreement sets a goal to curb global greenhouse gas (GHG) emissions, driving vast penetration of renewable energy worldwide. Extensive research on deep decarbonization of energy systems is conducted at the community-level (Tian et al., 2019), city-level (Wiryadinata et al., 2019), state-level (Zhao and You, 2020), and country-level (Vaillancourt et al., 2017). Electrification of heat and cooling generation and decarbonization of electricity generation is identified as a promising lever to address the ambitious climate goals (de Chalendar et al., 2019). However, heat and cooling generation stand a chance to affect the stability of the power system due to the surge in electric load involved (Sun et al., 2021), if they are electrified in an uncontrolled way. Therefore, it seems to be a reliable and promising decarbonization option by exploring renewable heat and cooling generation technologies rather than simply using electrified counterparts (Sánchez-Bautista et al., 2017). Among the vast array of renewable heat and cooling generation technologies (Gong et al., 2015), geothermal energy and deep water source cooling system show great potentials for the decarbonization transition of energy systems (Lee et al., 2019). Recent research efforts have also identified the values of green hydrogen (Dodds et al., 2015), large-scale heat pumps (Bach et al., 2016), biomass and biogas (Kassem et al., 2020), and thermal energy storage (Ochs et al., 2020) for decarbonizing the heating system. There is a lack of studies addressing the sustainable design of energy systems toward carbon neutrality by simultaneously exploring renewable electricity, heat and cooling generation, and electrified heating and cooling

options in the region with a humid continental climate, such as New York State (NYS) (Zhao and You, 2021). To bridge the knowledge gap, this paper aims to develop a modeling and optimization framework to handle the optimal design of carbon-neutral energy systems. In this paper, a multi-period optimization model, given a time horizon and a temporal resolution for total annualized cost (TAC) minimization, is built. The aim is to determine the optimal design of the carbon-neutral energy systems in the investigated region, seasonal operations, energy mix, and corresponding capacity of each base-load and peak-load technology involved. A tailored global optimization algorithm is applied to circumvent the computational challenges induced by separable concave terms in the objective function. The applicability of the proposed modeling framework is illustrated through a case study developed using the real data from the main campus of Cornell University, located in Ithaca, NYS.

2. Problem Statement and Model Summary

We are given a superstructure of carbon-neutral energy systems, including a set of renewable electricity generation technologies and a set of renewable and electrified heating and cooling options, as shown in Figure 1. To capture the optimal design, seasonal operations, energy mix, and corresponding capacity of each base-load and peak-load technology in the carbon-neutral energy systems, a time horizon and a set of time periods are specified to improve the temporal resolution of the model. The multi-period optimization problem of the proposed carbon-neutral energy system with earth source heat, lake source cooling (LSC), on-site electricity generation, and peak heating options for the total annualized cost minimization is formally defined in this section. The aim is to determine the optimal design of the carbon-neutral energy systems in the target region, seasonal operations, energy mix, and the corresponding capacity of each base-load and peak-load technology involved while fulfilling the seasonal demand for electricity, heat, and cooling.

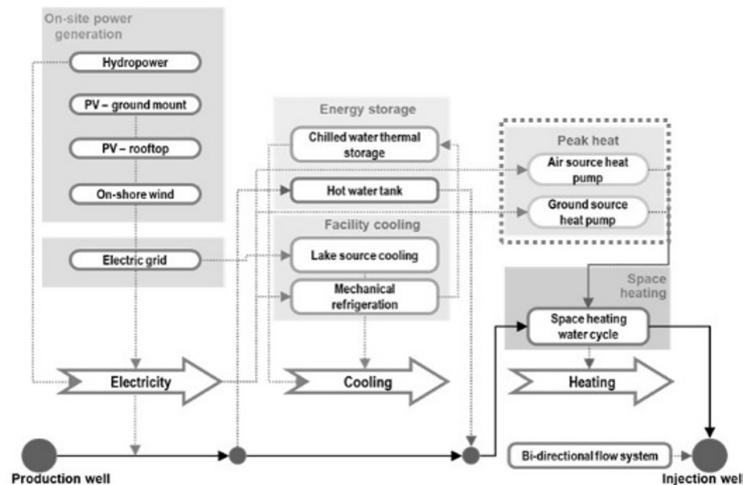


Figure 1. Overview of the superstructure of the proposed carbon-neutral energy systems with earth source heat, lake source cooling, on-site electricity generation, and peak heating technologies.

The general multi-period optimization model is subjected to the mass balance and configuration constraints, energy balance constraints, logic constraints, and techno-economic evaluation constraints (Gong and You, 2018). The integer decision variables

represent the selection of technologies. The number of geothermal well-pairs is an integer decision variable. Other essential decision variables such as the mass flow rates, energy flows, and capacities are continuous variables. The objective function, total annualized cost, includes integer variables such as the numbers of production wells and injection wells and thus is a mixed-integer function. The nonlinear terms mainly come from the separable concave terms induced by the economy of scale. Therefore, the resulting problem is a mixed-integer nonlinear programming (MINLP) problem. The general form of this MINLP problem is summarized as follows.

$$\begin{aligned} \min \quad & TAC = AIC + AOC + RE & (1) \\ \text{s.t.} \quad & \text{mass balance and configuration constraints;} \\ & \text{energy balance constraints;} \\ & \text{logic constraints;} \\ & \text{techno-economic evaluation constraints;} \end{aligned}$$

where AIC , AOC , and RE refer to the annualized investment cost, annual operating cost, and replacement cost, respectively.

The resulting MINLP problem includes integer and continuous variables embedded in nonlinear functions. For the global optimization of this problem, general-purpose global optimization solvers could be intractable. A global optimization strategy is developed to solve this non-convex MINLP problem. Specifically, we substitute the separable concave functions for capital cost model with successive piecewise linear functions for a relaxation of the original MINLP. The resulting MINLP problem is solved iteratively following the branch-and-refine algorithm (You and Grossmann, 2011).

3. Application to Sustainable Design of Campus Energy Systems

The proposed multi-period optimization modeling framework for energy systems decarbonization is applied to address the optimal design of the carbon-neutral energy systems using the main campus of Cornell University located in Ithaca as the living laboratory. Based on the optimization results, the optimal configuration of the carbon-neutral energy system in the target region, seasonal operations, energy mix, and corresponding capacity of each base-load and peak-load technology involved are determined while accommodating the seasonal demand for electricity, heat, and cooling across the main campus of Cornell University located in Ithaca, NYS. The case study aims to obtain the global optimal solution of the multi-period optimization problem with a monthly model resolution for the proposed carbon-neutral energy system with earth source heat, LSC, on-site electricity generation, and peak heating options, namely heat pump technologies. The maximum monthly average demand for cooling and heating are 46.7 MW and 68.0 MW, respectively. In 2020, there are 23,094 students, 9,907 non-faculty staffs, and 1,530 faculties on campus.

3.1. Systems Configuration

Herein, we demonstrate the peak heating options with different base-load capacities, where only ground source heat pumps are employed with no thermal energy storage. Specifically, 2,037, 1,265, and 493 ground source heat pumps with a typical capacity of 19 kW for each (Tian et al., 2020) in the North American region are deployed to handle the peak-load heat demand for the cases with two, three, and four base-load earth source heat pumps, respectively. We note that deploying thousands of geothermal heat pumps

across the campus could be a significant challenge in practice. This result helps inform the technology limitations and practice needs.

The selection of cooling technologies, peak heating systems, and the corresponding operating levels vary from month to month. The block flow diagram in Figure 3 reflects the technology selection in cold winter days when the heat demand is large and exceeds the capacity of base-load earth source heat systems. The column chart shows the breakdowns of electricity used to accommodate the building demand across the campus and drive all the electrical appliances for cooling and heating purposes, including the LSC, auxiliary chiller, air source heat pump, and ground source heat pump. The ground source heat pump provides extra heat as the peak heating option on cold winter days when the base-load earth source heat cannot fulfill all the heat demand. On the contrary, the base-load LSC occasionally fails to satisfy the peak cooling load on hot summer days. Thus mechanical chillers are utilized to generate additional chilled water during the night, which is warmed by the campus cooling load during the day, as shown in Figure 4.

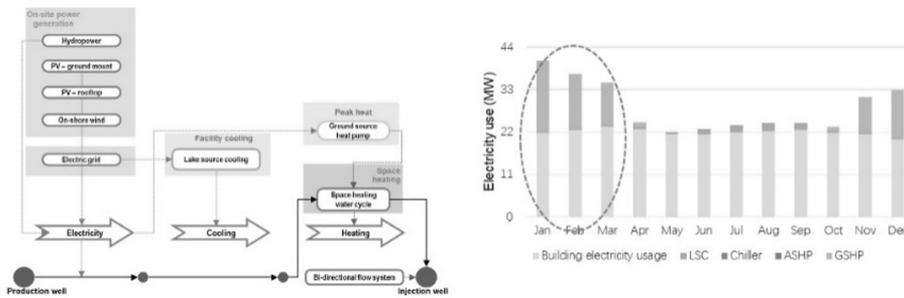


Figure 3. Technology/process configuration on cold winter days.

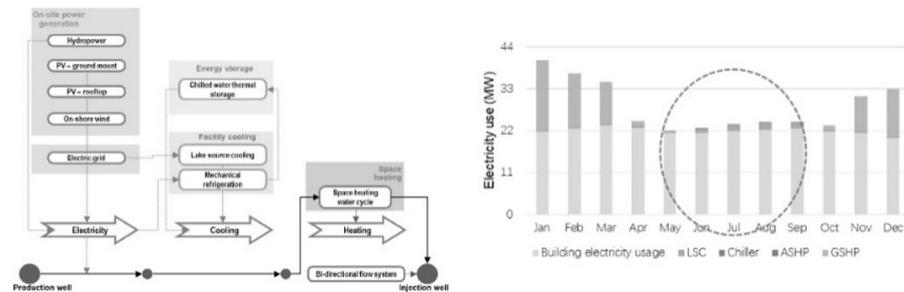


Figure 4. Technology/process configuration on hot summer days.

3.2. Techno-economic analysis

For a better understanding of the optimal design from an economic perspective, the breakdowns of the total annualized cost are given in Figure 6. The total annualized cost sums up the annualized investment cost and annual operating cost, calculated with a 15-year project lifetime (from 2021 to 2035) and a 5% discount rate. Because the capacity of base-load earth source heat systems has significant impacts on the selection of peak-load technologies and the corresponding capacity, the sensitivity of the base-load capacity is systematically analyzed. The number of geothermal well-pairs is chosen as the investigated input parameter, ranging from two to five. This case study considers at most five geothermal well-pairs because five well-pairs are sufficient to satisfy base-load heat

demand for the main campus of Cornell University located in Ithaca. We note that as the number of geothermal well-pairs increases from two to five, the annualized investment cost increases from \$93.62 MM/yr to \$95.64 MM/yr, while the annual operating cost decreases from \$31.56 MM/yr to \$28.20 MM/yr. In terms of capital investment, solar panels (58%-59%), wind turbines (23%), and hydroelectric power plant (8%-9%) are the major contributors. When the base-load earth source heat capacity is low, the operating cost associated with the ground source heat pump is more pronounced. The remaining annual operating cost is mainly sourced from the operations of solar and wind farms. We find that the number of base-load geothermal well-pairs shows little impact on the total annualized cost.

We note that all the techno-economic analysis results are obtained based on the current electricity price, which is volatile in the future. We perform breakeven analysis to gain insights and justify the need for earth source heat over the heat pumps. For the earth source heat-based systems, we consider five well-pairs based on the optimization results, and we consider all the electricity driving the heat pumps is sourced from the local grid. The results show that if the electricity price is higher than \$0.0202/kWh, the earth source heat system outperforms the ground source heat pumps from an economic perspective.

3.3. Energy Flow

The Sankey chart in Figure 5 shows the annual energy flow profile for the case with only two geothermal well-pairs as the base-load heat supplier. We find that 65.3% of the total electricity is sourced from the local grid, with an annual consumption of 160,812 MWh during the on-peak time and 393 MWh during off-peak hours. 0.13% of the total electricity is consumed to drive the geothermal pumping systems, the ground source heat pump utilizes 19%, and 3% are attributed to LSC assisted by mechanical chillers. It is also found that the ground source heat pumps provide approximately one-third of the building heat demand across the campus, and the base-load geothermal well-pairs fulfill the remaining heat demand. Finally, 99% of the cooling is provided by the LSC (with COP of around 30) and only 1% from the mechanical chiller.

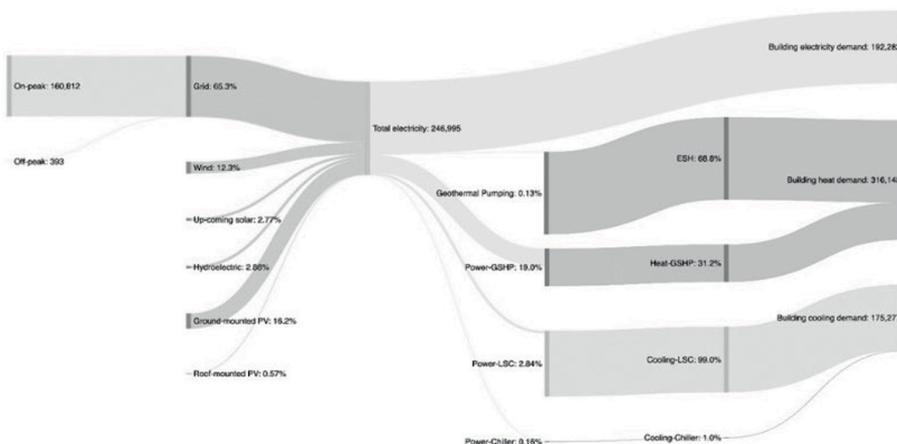


Figure 5. Annual energy flow profile with two base-load geothermal well-pairs (unit: MWh/yr).

4. Conclusion

A multi-period optimization model, given a time horizon and a temporal resolution for total annualized cost minimization, was built. The aim was to simultaneously determine

the optimal design of the carbon-neutral energy systems in the investigated region, seasonal operations, energy mix, and corresponding capacity of each base-load and peak-load technology involved while fulfilling the seasonal demand for electricity, heat, and cooling. The applicability of the proposed modeling framework was illustrated through a case study considering electrified peak heating options, namely heat pump technologies, which was developed by leveraging the real-world data from the main campus of Cornell University, located in Ithaca, New York.

References

- B. Bach, J. Werling, T. Ommen, M. Münster, J.M. Morales, B. Elmegaard, 2016, Integration of large-scale heat pumps in the district heating systems of Greater Copenhagen. *Energy*, 107, 321-334.
- J.A. de Chalendar, P.W. Glynn, S.M. Benson, 2019, City-scale decarbonization experiments with integrated energy systems. *Energy & Environmental Science*, 12(5), 1695-1707.
- P.E. Dodds, I. Staffell, A.D. Hawkes, F. Li, P. Grünewald, W. McDowall, P. Ekins, 2015, Hydrogen and fuel cell technologies for heating: A review. *International Journal of Hydrogen Energy*, 40(5), 2065-2083.
- J. Gong, F. You, 2015, Sustainable design and synthesis of energy systems. *Current Opinion in Chemical Engineering*, 10, 77-86.
- J. Gong, F. You, 2018, A new superstructure optimization paradigm for process synthesis with product distribution optimization: Application to an integrated shale gas processing and chemical manufacturing process. *AIChE Journal*, 64, 123-143.
- N. Kassem, J. Hockey, C. Lopez, L. Lardon, L.T. Angenent, J.W. Tester, 2020, Integrating anaerobic digestion, hydrothermal liquefaction, and biomethanation within a power-to-gas framework for dairy waste management and grid decarbonization: a techno-economic assessment. *Sustainable Energy & Fuels*, 4(9), 4644-4661.
- I. Lee, J.W. Tester, F. You, 2019, Systems analysis, design, and optimization of geothermal energy systems for power production and polygeneration: State-of-the-art and future challenges. *Renewable and Sustainable Energy Reviews*, 109, 551-577.
- F. Ochs, A. Dahash, A. Tosatto, M.B. Janetti, 2020, Techno-economic planning and construction of cost-effective large-scale hot water thermal energy storage for Renewable District heating systems. *Renewable Energy*, 150, 1165-1177.
- A.d.F. Sánchez-Bautista, J.E. Santibañez-Aguilar, et al., 2017, Optimal Design of Energy Systems Involving Pollution Trading through Forest Plantations. *ACS Sustainable Chemistry & Engineering*, 5, 2585-2604.
- L. Sun, F. You, 2021, Machine learning and data-driven techniques for the control of smart power generation systems: An uncertainty handling perspective. *Engineering*, 7 (9), 1239-1247.
- X. Tian, T. Meyer, H. Lee, et al., 2020, Sustainable design of geothermal energy systems for electric power generation using life cycle optimization. *AIChE Journal*, 66, e16898.
- X. Tian, F. You, 2019, Carbon-neutral hybrid energy systems with deep water source cooling, biomass heating, and geothermal heat and power. *Applied Energy*, 250, 413-432.
- K. Vaillancourt, O. Bahn, E. Frenette, O. Sigvaldason, 2017, Exploring deep decarbonization pathways to 2050 for Canada using an optimization energy model framework. *Applied Energy*, 195, 774-785.
- S. Wiryadinata, J. Morejohn, K. Kornbluth, 2019, Pathways to carbon neutral energy systems at the University of California, Davis. *Renewable Energy*, 130, 853-866.
- F. You, I.E. Grossmann, 2011, Stochastic inventory management for tactical process planning under uncertainties: MINLP models and algorithms. *AIChE Journal*, 57(5), 1250-1277.
- N. Zhao, F. You, 2020, Can renewable generation, energy storage and energy efficient technologies enable carbon neutral energy transition? *Applied Energy*, 279, 115889.
- N. Zhao, F. You, 2021, New York State's 100% renewable electricity transition planning under uncertainty using a data-driven multistage adaptive robust optimization approach with machine-learning. *Advances in Applied Energy*, 2, 100019.

Isopropanol/ n -Butanol/ Ethanol separation from diluted fermentation broth by distillation. Process optimization using MILP techniques

Eszter Toth^a, Gianluca Mancino^a, Ludovic Raynal^a

^a*IFPEN, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*

Abstract

This work has been performed in the framework of the development of the bio-based fermentation process to produce IBE (isopropanol/ n-butanol/ ethanol). The so-called IBE fermentation is indeed an interesting sustainable alternative to produce fossil-based products. However, product inhibition in fermentation leads to dilute fermentation broths in water (15-25 g IBE/L) which implies high energy demand for products/water separation. In the present study focus has been put on ready-to-industrialize downstream processes for Isopropanol/ n-Butanol/ Ethanol separation from dilute fermentation broth, using conventional distillation and shell-and-tube heat-exchanger technologies. A reference process scheme, using 5 distillation columns, as well as an IFPEN patented distillation sequence with 3 columns only were optimized using an in-house tool. The tool allows to simultaneously optimize the heat exchanger network configuration and the distillation columns' operating pressures for a given material balance, using Mixed Integer Linear Programming (MILP) optimization techniques. When comparing to the reference process, the optimized heat exchanger network leads to significant vapor consumption reduction and to also significant total separation cost reduction, when both investments and utilities costs are considered. The IFPEN patented scheme even without optimization is found to be more interesting than the reference scheme and is shown to be even more interesting after optimization. The tool can be applied to any distillation process, leading to significant cost savings.

Keywords: Isopropanol/ n-Butanol/ Ethanol fermentation, Distillation, Optimization, Mixed Integer Linear Programming

1. Introduction

N-butanol is used as a solvent and could be used as a fuel. It reached 4.1 million tons consumption in 2015. Isopropanol, a 1.9 million tons market in 2015, can also be used as a solvent and has also shown its interest as an additive in fuels. Finally, isopropanol can also be transformed into propylene, a major chemical intermediate today.

N-butanol and isopropanol are currently produced by petrochemical routes, but new technological routes are under development including bio-based technologies. Their production can be carried out by fermentation, e.g. ABE production (acetone, n-butanol, and ethanol) or IBE production (isopropanol, n-butanol, ethanol). The ABE process using *Clostridium*-type bacteria was one of the first large-scale industrial microbial process for chemical production, as well as the largest fermentation process under sterile conditions (Ni, 2009; Köpke, 2011). Initially, acetone was the main compound of interest for its use in the production of cordite during World War I and its use in the production of other chemicals. Nowadays, acetone is sometimes considered as an undesirable by-product due

to its low properties as a biofuel or chemical. The reduction of acetone into isopropanol is possible using different bacteria belonging to the genus *Clostridium*. The production process of IBE is therefore based on the ABE fermentation process and uses strains of the genus *Clostridium* capable of fermenting simple sugars and reducing a large part of the acetone into isopropanol. However, a small part of acetone is still produced during IBE fermentations.

N-butanol production by *Clostridium*-type strains, for both ABE and IBE fermentations, is limited by its impact towards microbial growth, typically when its concentration is between 7 and 15 g/L. This greatly limits the final alcohol concentration of the fermentation broth, to approximately 10 to 30 IBE g/L. Let's note at this point that some aspects of ABE process can easily be used for IBE process: for example, n-butanol inhibition in fermentation leads to the same issues of diluted broth separation.

Such low concentrations imply high energy demand for products/water separation and therefore high separation costs. Several solutions have been reviewed in literature to achieve cost reduction (Kujawska, 2015; Vane, 2008). A part of them concerns the downstream process, i.e. alcohol/water and alcohol/alcohol separations. Hybrid distillation/liquid-liquid extraction processes using solvents have been proposed, mesitylene has been considered by Kraemer (2010, 2011)–, and ethylene glycol by Zhang (2020). Mesitylene appears to be an interesting solution as far as energy demand is concerned, but capital cost investment is not given. Moreover, after alcohol separation, the vinasses are meant to be recycled to the fermenter, and the residual mesitylene's toxicity to micro-organisms must be investigated. For ethylene glycol, the authors use n-butanol in higher concentration of about 4.46 wt %, which further implies low energy demand for separation but is hard to achieve with ABE/ IBE fermentation. Another way to achieve separation cost reduction is to increase alcohol concentration in broth. In order to limit n-butanol inhibition on bacteria, In Situ Product Recovery Techniques (ISPR), consisting in removing n-butanol from the fermenter during its production, have been considered (Outram, 2016). A part of these techniques, such as perstraction, are at research level and would need significant development before industrial level. Gas stripping (Xue, 2012) or two-phase fermentation (González-Peñas, 2020) in the presence of a biocompatible solvent in the fermenter are two types of techniques nearer to industrial scale. However, those techniques are relatively expensive and there is no evidence that the cost of ISPR techniques is offset by the decrease in separation costs as several parameters must be considered. It would then be important to properly estimate the cost/benefit ratio of ISPR techniques.

In the present study focus has been put on ready-to-industrialize downstream processes for isopropanol/ n-butanol/ ethanol separation from dilute fermentation broth, using conventional distillation and shell-and-tube heat-exchanger technologies. The originality of the work consists here in a thorough process study including both process flow scheme optimization and an associated techno-economic estimation.

2. Methods

As aforementioned, the IBE process is close to the ABE process. For the downstream separation part, the main difference is that, unlike acetone, isopropanol forms an azeotrope with water which makes infeasible the complete dehydration of isopropanol by conventional distillation: for example, at atmospheric pressure, the isopropanol-water azeotrope contains 12 wt % of water. . Moreover, that little amount (approx. 0.2 - 0.6 g/L)

of acetone is still being produced, which represents an undesirable impurity to be eliminated from the main products.

2.1. Process description. 5 columns distillation process – reference case.

The reference distillation process was adapted from the ABE literature (van der Merwe, 2013) and is shown in Fig. 1. It involves five distillation columns to perform the separation of the various components. The Beer column recovers IBE at the top and eliminates 98.7 % of water at the bottom. The Acetone column allows the elimination of undesired acetone. With the IPA column, it is possible to separate the isopropanol/ water azeotrope and the small amount of ethanol at the top. The Water and Butanol columns are the last two columns used to break the water/n-butanol heterogenous azeotrope and recover the n-butanol at the bottom of the Butanol column.

2.2. Process description. 3 columns distillation process – base case

Another way to achieve the required separation is to use the sequence adapted from an IFPEN patent (Mikitenko, 1983) discussed by Pucci (1986) and shown in Fig. 2. This sequence consists only of a Beer column and a second column to separate the butanol from the other components. These are finally sent to a third column which separates isopropanol/water azeotrope and ethanol from the undesired acetone. The particularity of this scheme is in the second column. This column provides a three-phase area (two liquid and one vapor) on some trays. The aqueous phase is subtracted from the three-phase zone and is recycled and mixed with the feed of the first column.

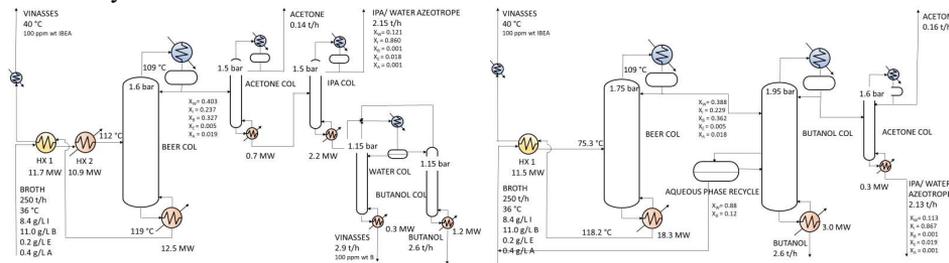


Fig 1. Flowsheet of the 5 columns reference case

Fig 2. Flowsheet of the 3 columns base case

2.3. Flowsheet simulation

The flowsheets were simulated using SIMSCI’s PRO/II v 10.2 software. The base thermodynamic method used was SRK-Simsci (SRKS). The unary and binary, both for vapor/liquid and liquid/liquid equilibria, were adjusted based on literature and in-house experimental data.

2.4. Process optimization using Mixed Integer Linear Programming (MILP) techniques

To achieve separation cost reduction, including both investments and utilities, the two process schemes were optimized with an in-house optimization tool, based on Mixed Integer Linear Programming (MILP) techniques. Fig 3. shows the general optimization procedure.

The global material balances, as shown in Fig. 1. and 2. are kept constant for the whole study. First, data are generated using PRO/II v 10.2 process simulation software. A linear equation set, including temperature vs. enthalpy curves for process streams at different pressures, energy consumption vs pressure with different number of trays for columns and utility consumption is coded into GAMS Studio V 25.1.2 after equation parameter fitting using Microsoft Excel. The novelties of the model stay in the fact that on the one

hand streams can either give or receive heat, i.e. they are not mandatorily defined as hot or cold streams, and on the other hand the column's operating pressures are simultaneously optimized with the heat exchanger network, i.e., the pressures of some of them are adapted in order to allow reboiler/condenser integration between columns. Some extra features, such as steam and electricity generation are also enabled by the model but not used in the current study. The goal is to maximize process/process energy integration, thus minimizing utility consumption while considering penalties for extra investment costs due extra heat exchangers or extra number of trays for columns. The objective function thus corresponds to a global optimum expressed in €/h, taking into account operational and capital expenditures. Calculations are based on an IFPEN internal data base with utility and investments costs, 2019 being the reference year. The objective function is then minimized using CPLEX 12.10 solver provided by GAMS Studio V 25.1.2. As a result, we get an optimized heat exchanger network and column's operating pressures and an idea of optimum number of trays. To end with, the resulting heat exchanger network is integrated in the flowsheet and used to determine real utility consumption, equipment are sized, and investment cost determined to end up with total separation cost, in €/t of IBE.

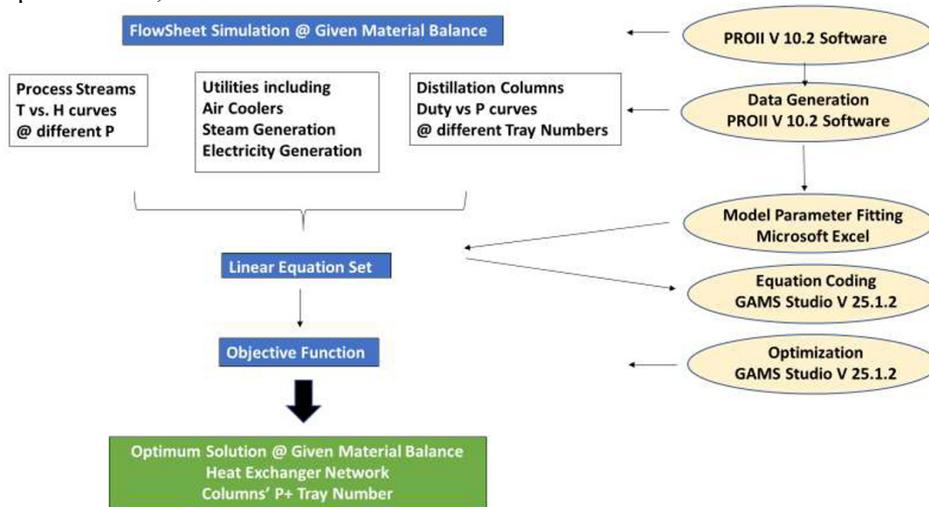


Fig 3. Optimization procedure

3. Results and discussion

3.1. Optimized process with 5 distillation columns. Energy demand.

The resulting flowsheet is shown in Fig. 4. The optimized scheme consists in an integration between IPA and Butanol columns. The operating pressure of the first is raised to 5.5 bars, further increasing the condenser temperature, and allowing its heat of condensation to be recovered for the Butanol column's reboiler. Moreover, liquid-liquid demixing is no longer present at this level of pressure in the IPA column, leading to cheaper column internals and easier operation.

Optimization mainly impacts the pre-heating of the fermentation broth and the Beer column's reboiler duty. Indeed, heat can be recovered not only from the bottom effluent of the column (the so-called vinasses), but also from the top, requiring two additional

exchangers. The broth is divided into two streams, one stream recovers heat from the vinasses and the other from the top stream. After mixing those two streams, extra 1.9 MW can be recovered from the top. In the reference case a total of 23.4 MW of low-pressure steam was necessary to achieve the required separation (10.9 MW and 12.5 MW respectively for the HX 2- pre-heater and the Beer column reboiler, as shown in Fig.1). In the optimized case 13.9 MW are sufficient. The Beer column operating pressure was slightly increased compared to the reference case to compensate extra pressure drop from the extra heat exchangers.

3.2. Optimized process with 3 distillation columns. Energy demand.

The resulting flowsheet is shown in Fig. 5. The main idea is roughly the same than for the 5 columns process. Extra heat can be recovered from the top of the Beer Column. The vinasses transfer heat to the broth in two heat exchangers (HX 6 and HX 1) and recover heat from the top in heat-exchanger HX7. Low-pressure steam consumption falls from 18.3 MW down to 15.4 MW, requiring two extra heat exchangers.

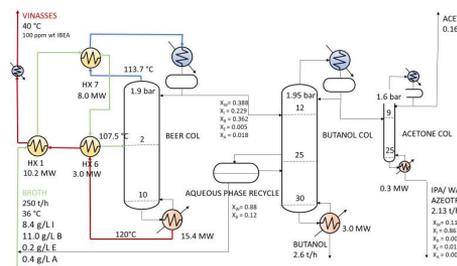


Fig 4. Optimized 5 columns flowsheet

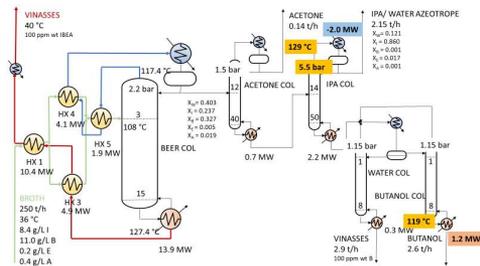


Fig 5. Optimized 3 columns flowsheet

3.3. Total separation cost

For both distillation sequences, hot utility demand remarkably decreases when optimized. The total separation cost, including total utilities consumption and investment costs has further been calculated. The results, given in normalized to 100 base are shown in Fig 6. The 3 columns optimized process ends up being 34% cheaper than the reference case. Because of the liquid-liquid demixtion zone in the Butanol column of this process, it may however be hard to operate. In that view, the more classical, optimized 5 columns process, with 28 % lower separation cost compared to the reference case can be considered as the best solution.

3.4. Alternative solutions

Vapor recompression of the Beer column's top stream to provide heat to the bottom could be an option, considering the 10 °C temperature difference between the top and bottom tray temperatures. Nevertheless, after total separation cost calculation, even if this solution is interesting from utility consumption point of view, the investment cost of the compressor appears to be too high, as it is shown in Fig. 6.

To end with, aforementioned ISPR processes could lead to broth concentrations above 25 g IBE /L. Fig 7. shows that the relative gain on separation tends to decrease in this concentration region.

4. Conclusions

The IFPEN process optimization tool allowed to identify process flow scheme with heat exchanger integration with significant decrease in the IBE fermentation downstream

process cost. It is shown here that up to 30% reduction in production cost can be achieved. Such approach and tool can be applied to any distillation process to lead to economically interesting distillation processes which is especially of great importance for bio-based processes.

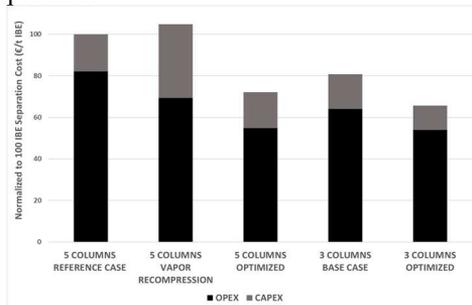


Fig 7. Influence of broth concentration on separation cost

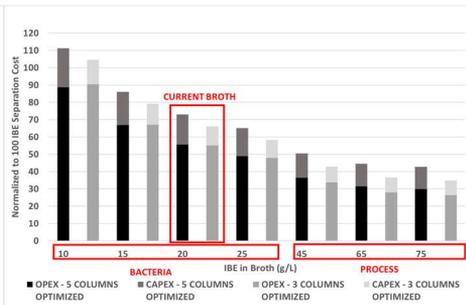


Fig 6. Optimization results @ 20 g/L IBE in broth

References

- H. González-Peñas, Th. A. Lu-Chao, G. Eibes, J. M. Lema, 2020, Energy requirements and economics of acetone–butanol–ethanol (ABE) extractive fermentation: a solvent-based comparative assessment, *Bioprocess and Biosystems Engineering*, 43, 2269–2281
- M. Köpke, P. Dürre, 2011, Biochemical production of biobutanol, *Handbook of biofuels production*, Chapter 10, Woodhead Publishing Limited, p 221–256
- K. Kraemer, A. Harwardt, R. Bronneberg, W. Marquardt, 2011, Separation of butanol from acetone–butanol–ethanol fermentation by a hybrid extraction–distillation process, *Computers and Chemical Engineering*, 35, 949–963
- A. Kujawska, J. Kujawski, M. Bryjak, W. Kujawski, 2015, ABE fermentation products recovery methods—A review, *Renewable and Sustainable Energy Reviews*, 48, 648–661
- Y. Ni, Z. Sun, 2009, Recent progress on industrial fermentative production of acetone-butanol-ethanol by *Clostridium acetobutylicum* in China. *Applied Microbiology and Biotechnology*, 83(3), 415–423.
- A.B. van der Merwe, H. Cheng, J.F. Görgens, J.H. Knoetze, 2013, Comparison of energy efficiency and economics of process designs for biobutanol production from sugarcane molasses, *Fuel*, 105, 451–458
- P. Mikitenko, L. Asselineau, A. Hurpin, 1983, Procédé de fractionnement d'une solution aqueuse de butanol et d'acétone, *Demande de brevet d'invention No. FR2549043 A1*
- V. Outram, C.A. Lalander, J.G.M. Lee, E.T. Davis, A.P. Harvey, 2016, A comparison of the energy use of in situ product recovery techniques for the Acetone Butanol Ethanol fermentation, *Bioresource Technology*, 220, 590–600
- A. Pucci, P. Mikitenko, L. Asselineau, 1986, Three-phase distillation. Simulation and application to the separation of fermentation products. *Chemical Engineering Science*, 41, 3, 485–494
- L.M. Vane, 2008, Separation technologies for the recovery and dehydration of alcohols from fermentation broths, *Biofuels, Bioprod. Bioref.*, 2, 553–588
- C. Xue, J. Zhao, C. Lu, S.T. Yang, F. Bai, I.C. Tang, 2012, High-Titer n-Butanol Production by *Clostridium acetobutylicum* JB200 in Fed-Batch Fermentation with Intermittent Gas Stripping. *Biotechnology and Bioengineering*, 109, 11, 2746–2756
- H. Zhang, Q. Ye, L. Chen, N. Wang, Y. Xu, Y. Li, 2020, Purification of isopropanol-butanol-ethanol (IBE) from fermentation broth: process intensification and evaluation, *Chemical Engineering & Processing: Process Intensification* 158, 108182

Automatic synthesis of hybrid processes using distillation and liquid-liquid extraction for the separation of azeotropic mixtures

Thulasi Sasi^a, Mirko Skiborowski^{a,b}

^a*TU Dortmund University, Laboratory of Fluid Separations, Emil-Figge-Straße 70, Dortmund 44227, Germany*

^b*Hamburg University of Technology, Institute of Process Systems Engineering, Am Schwarzenberg-Campus 4 (C), 21073 Hamburg, Germany*

Abstract

Solvent-based separation processes, such as liquid-liquid extraction are important options for the separation of aqueous mixtures with low relative volatilities, tangent pinches or azeotropes, for which distillation is either infeasible or energy intensive. As the choice of solvent is of utmost importance, various methods for solvent screening have been proposed, focusing either on properties like solubility and capacity, or the minimum solvent requirement. While some do evaluate the energy for solvent recovery, they assume a fixed process structure, based on the integration of an extraction and a distillation column. Therefore, additional constraints restrict feasible solvents to those that do not form azeotropes and show specific boiling points. While such constraints are oftentimes reasonable, overcoming this limitation requires the identification of solvent-specific process configurations, which is a complex task usually performed by graphical analysis or trial and error. For this purpose, the current work introduces a novel algorithmic synthesis approach and illustrates its application for the separation of a ternary mixture of acetone, acetic acid and water using chloroform as solvent.

Keywords: process synthesis, azeotropic multi-component mixtures, extraction-distillation hybrid processes.

1. Introduction

The separation of aqueous non-ideal mixtures continuously gains importance, while low relative volatilities or azeotropes limit a direct and energy-efficient application of distillation for these systems. Solvent-based separation processes, such as liquid-liquid extraction are of special interest, as the addition of a suitable mass separating agent (MSA) that introduces immiscibility in the system, allow for the separation of the target product even from complex mixtures. Thereby, depending on the choice of a suitable MSA, liquid-liquid extraction can be a very efficient means of separation with very little energy requirements, which primarily result from the recovery of the MSA. For this reason, various methods for solvent screening and computer-aided molecular design have been proposed (Gertig et al., 2020).

Solvent recovery is required to keep the solvent in the loop. Although regeneration strategies such as back-extraction are of interest for non-volatile solvents, this is usually accomplished using distillation or simply stripping and accounts for the major energy requirement in the process (Sprakel & Schuur, 2019). While many studies screen solvents

based on distribution coefficients (at infinite dilution) or the minimum required solvent flowrate, which correlates with the energy required to regenerate it, the latter also depends on the complexity of the individual separation. Solvent screening is therefore usually focusing on the selection of a suitable solvent for a specific process configuration, usually limited to homogeneous distillation, posing constraints on the solvent (e.g. Gertig et al., 2020). Such constraints can only be dropped if process synthesis is performed, and for each solvent the most promising flowsheet is selected. This is especially necessary for the separation of multi-component mixtures, where multiple separation steps are required to recover the solvent and constituent components from the extract and raffinate phases. Kraemer et al. (2011) have demonstrated the combined approach of solvent screening and process synthesis for the separation of butanol from fermentation broth, whereas each process configuration was manually designed. Kaul et al. (2018) also combined solvent screening and flowsheet synthesis, building on expert knowledge for the generation of process variants. For a systematic approach, it is necessary to automate the synthesis method. Expert systems developed in the 90s such as SPLIT by Wahnschafft et al. (1992) and PROSYN by Schembecker & Simmrock (1997) demonstrated the synthesis of separation processes including extraction and distillation. The former is a blackboard-based synthesis method where the separation problem is divided into so called *binary separation tasks*, which effectively oversee the non-idealities in the multi-component space, the latter is a combined heuristic-numeric approach where the separation process is designed using heuristic rules and the physical properties are estimated numerically. While such expert systems can systematically guide process synthesis, they require subsequent simulation to validate the feasibility of the developed process configurations and cannot easily be automated for an effective evaluation of different solvent candidates. To overcome this limitation, the current approach proposes a thermodynamically-sound method for an algorithmic process synthesis of hybrid extraction-distillation processes. The method generates process alternatives solely based on a thermodynamic model of the chemical system. It extends our previously developed algorithmic framework, for the synthesis of homogeneous and heterogeneous distillation processes (Sasi & Skiborowski 2020a, Sasi et al. 2020b) introducing liquid-liquid extraction as an alternative unit operation that is evaluated by means of the pinch-based shortcut method of Redepenning et al. (2017), which determines the minimum required solvent flowrate for a pre-defined recovery. The synthesis approach can be applied to azeotropic mixtures as well, thereby allowing the use of azeotrope-forming solvents for extraction, which require a more complex process configuration for solvent recovery. The subsequent sections summarize the extended synthesis method and illustrate the application of the method for the separation of a ternary mixture of acetone, acetic acid and water using chloroform as solvent.

2. Methodology

The overall process synthesis framework extends our previously presented algorithmic framework (Sasi & Skiborowski 2020a, Sasi et al. 2020b), which can broadly be divided in the three consecutive steps shown in Figure 1. The method builds on the topological analysis of the mixture to identify the distillation regions and compartments as well as miscibility gaps. Based on this initial knowledge potential process sequences are generated in a tree structure before they are finally converted to process configurations with closed recycles. The application of the method only requires the specification of a thermodynamic model for VLE and LLE computations, as well as a problem definition

with feed composition and flowrate, desired products and the definition of a potential solvent. The method is implemented in MATLAB, whereas property data can be extracted from literature or directly from the property database of a process simulator, such as Aspen Plus.

2.1. Mixture topology

In order to evaluate the feasibility of individual splits by means of distillation, first the topology of the mixture, which includes the determination of homogeneous and/or heterogeneous azeotropes (Skiborowski et al. 2016) and the identification of distillation regions and compartments (Rooks et al. 1998) is performed. For liquid-liquid extraction, a necessary phase separation is evaluated by means of the homotopy continuation method of Bausa & Marquardt. (2000).

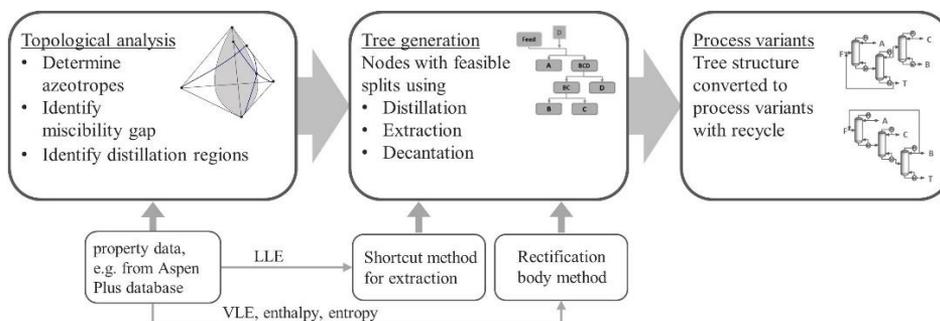


Figure 1: Illustration of the framework of the synthesis methodology

2.2. Generation of process sequences in a tree structure

Process sequences are generated through the evaluation of individual splits, creating a tree structure. The process starts from the initial feed mixture and adds feasible splits as new branches of the tree. Each node is further split until all desired products are reached. For each node, distillation, decantation and extraction are considered, with a potential subcooling to a predefined operating temperature being considered for the latter options. In case the product specification cannot be reached immediately, the program evaluates if recycling of individual products can improve the recovery or shift a feed stream to a different distillation region.

2.2.1. Feasibility test for distillation

For a separation by means of distillation the maximum recovery of individual splits is determined by means of a general feasibility criteria, which evaluates if both product compositions are located in the same distillation region. This is achieved by evaluating the terminals of product residue curves and pinch lines with respect to the topological distillation regions. Thus, if a sharp split is not feasible due to the intersection of the mass balance line with a distillation boundary, the highest possible purity of the limiting product is determined by means of a simple bisection. Through application of the rectification body method (Bausa et al. 1998) it is furthermore checked if both products are located in the same compartment, furthermore providing an estimate of the minimum energy demand (MED) for the specific separation. Thus, this combined procedure evaluates split feasibility without an explicit computation of the distillation boundaries, considers potential boundary crossing for finite reflux distillation and provides an estimate of the MED.

2.2.2. Feasibility test for extraction

For the evaluation of a separation by means of extraction, a pinch-based shortcut method for an isothermal counter-current liquid-liquid extraction column (Redepenning et al. 2017) is applied to determine the feasibility of extraction and the minimum solvent flowrate for the separation of a given feed composition. For the application the solvent composition and the desired purity of the key component in the raffinate phase must be defined a-priori. The shortcut method builds on the identification of the saddle-node pinch that governs the minimum solvent operation. The computations build on a rigorous thermodynamic model for LLE computation and exploit continuation methods for a robust and computationally efficient evaluation of the minimum solvent flowrate.

2.3. Generation of variants including recycle

While potential recycle streams are already introduced during the generation of the tree structure, the recycle streams are not closed in this initial synthesis phase. In case such recycle streams are used the recycles are closed in a subsequent step. Furthermore, individual nodes that are not desired products or pure components are also recycled to suitable destinations. If the recycle cannot be converged due to topological restrictions, an additional pure component stream is introduced with the recycle.

3. Case Study

For illustration of the methodology the separation of a ternary mixture of acetone, acetic acid and water using chloroform as solvent is considered. Although the ternary mixture is zeotropic at ambient pressure, separation by means of distillation is energy-intensive due to a rather low relative volatility of acetic acid and water and a tangent pinch near pure water. Acetone and water also exhibit low relative volatility at high acetone purity. Previous studies have shown the use of chloroform as a suitable solvent for liquid-liquid extraction (Minotti et

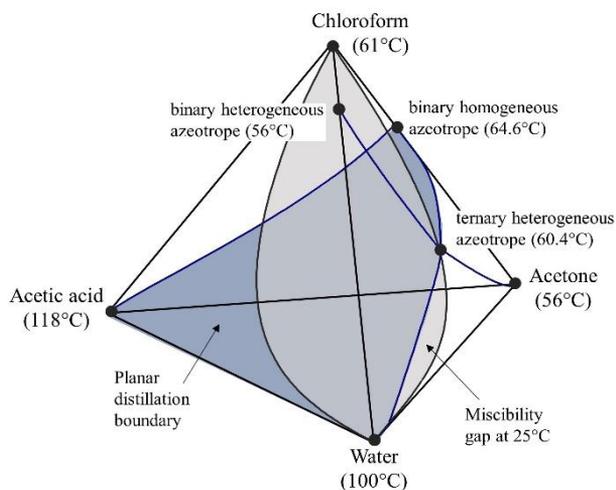


Figure 2: Topology of the quaternary mixture of acetone, acetic acid, water and chloroform at 1 atm including the miscibility gap at 25°C

al. 1996, Caballero & Grossmann 2009, Redepenning et al. 2013). Figure 2 illustrates the topology of the quaternary mixture, including chloroform, which introduces two binary and a ternary azeotrope at 1 atm pressure, and a miscibility gap between water and chloroform, which extends into the quaternary region. The thermodynamic model builds on the UNIQUAC parameters obtained from Aspen Plus. The azeotropes induce a distillation boundary, which divides the quaternary composition space into two distillation regions with acetone and chloroform being the unstable nodes of the individual distillation regions. Note that the ternary sub-space of acetone-chloroform-water is even more complex, with four individual distillation regions.

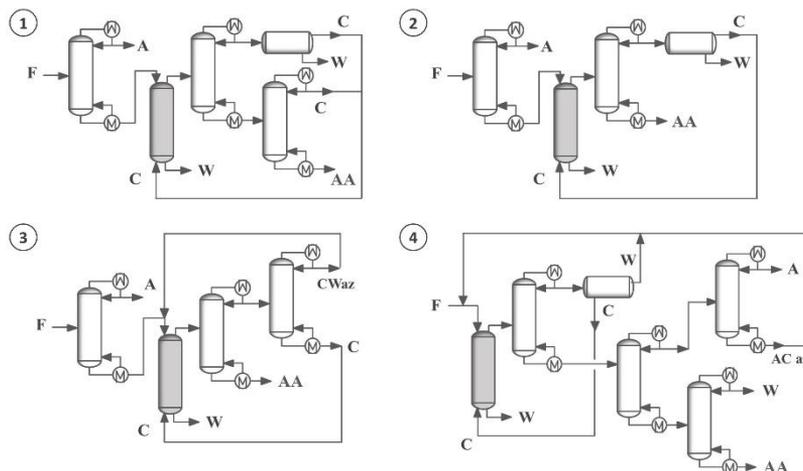


Figure 3: Process variants for the separation of a ternary mixture of acetone (A), acetic acid (AA) and water (W) using chloroform (C).

For process synthesis a feed stream of 10 mol/s with 70 mol-% water and 15 mol-% each of acetone and acetic acid is to be separated in the individual components. For decantation and extraction an operating temperature of 25°C is assumed. Application of the methodology results in overall six feasible process variants, including the direct split (DS) and indirect split (IS) sequence for the separation of the initially zeotropic ternary mixture. The computations for the generation of these variants take less than 5 min of computational time on a desktop PC. Figure 3 illustrates the four variants that utilize chloroform as a solvent and incorporate an extraction column with a closed solvent recycle. Due to the azeotropes formed with chloroform, solvent recovery is more complex. Variants 1-3 first remove acetone from the feed in a distillation column, followed by extraction of the binary acetic acid-water mixture using chloroform. Variant 4 starts with liquid-liquid extraction and requires the most unit operations. The MED of all process variants is illustrated in Figure 4. For streams leaving an extraction column or decanter at subcooled conditions, additional heating duty is accounted for in the MED if such streams are further processed using distillation. Compared with the DS and IS sequences only variant 1 and 2 offer the potential for energy savings. These process variants require 22% and 26% less energy than the DS sequence respectively. Variant 2 offers the lowest MED and requires one distillation column less than variant 1. The two streams from the decanter have a purity of 99.4 mol-% chloroform and 99.88 mol-% water respectively. Based on the comparison of variant 2 and 3 it becomes apparent that the chloroform water separation is much more efficient by decantation than distillation. Variant 4 does not only require the most number of equipment, but also the highest MED, which is also caused by a large water recycle stream that needs to be mixed with the acetone-chloroform azeotrope prior to recycling to the extraction column.

4. Conclusion

Given the selection of a suitable solvent, hybrid extraction-distillation processes present an important option for the energy-efficient separation of complex non-ideal mixtures. Yet, depending on the choice of solvent a more complex process design may be required, which so far had to be manually designed. In order to overcome this limitation and allow

for an algorithmic generation of integrated process flowsheets the algorithmic framework for the synthesis of homogeneous and heterogeneous azeotropic distillation processes has been extended to include liquid-liquid extraction. The application of the extended methodology to the separation of acetone, acetic acid and water with chloroform illustrates the capability of the method to synthesize hybrid separation processes for a solvent that introduces additional azeotropes and distillation boundaries. The results further illustrate the potential energy savings of such solvent-based processes, even for close-boiling systems that do not experience azeotropes. The extended synthesis method provides the basis for a solvent screening that considers tailored process configurations for each individual solvent. Such evaluations will be the focus of future work, which will also consider processes with alternative unit operations such as heteroazeotropic distillation and extractive distillation.

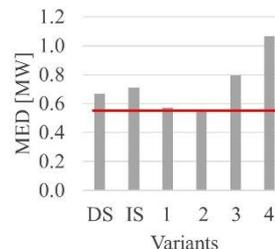


Figure 4 MED of the process variants shown in Figure 3.

Reference

- J. Bausa et al., 1998, Shortcut methods for nonideal multicomponent distillation: I. Simple columns, *AIChE J* 44 (10), 2181–2198.
- J. Bausa & W. Marquardt, 2000, Quick and reliable phase stability test in VLE flash calculations by homotopy continuation, *Comput. & Chem. Eng.*, 24 (11), 2447–2456.
- J. A. Caballero & I. E. Grossmann, 2009, Rigorous design of complex liquid-liquid multi-staged extractors combining mathematical programming and process simulators, *Comput. Aided Chem. Eng.*, 27, 981–986.
- C. Gertig et al., 2020, Computer-aided molecular and processes design based on quantum chemistry: current status and future prospects, *Current Opinion in Chem. Eng.*, 27, 89–97.
- M. Kaul et al., 2018, Hierarchical design of extraction-distillation processes using shortcut apparatus models with piece-wise linearized thermodynamics, *Chem. Eng. Sci.*, 192, 422–433.
- K. Kraemer et al., 2011, Separation of butanol from acetone–butanol–ethanol fermentation by a hybrid extraction–distillation process, *Comput. & Chem. Eng.*, 35(5), 949–963.
- M. Minotti et al., 1996, A Geometric Method for the Design of Liquid Extractors, *Ind. Eng. Chem. Res.*, 35(8), 2672–2681
- C. Redepenning et al., 2013, Shortcut method for the design of extraction columns for multicomponent mixture separations, *Comput. Aided Chem. Eng.*, 32, 1039–1044.
- C. Redepenning et al., 2017, Pinch-based shortcut method for the conceptual design of isothermal extraction columns, *AIChE J.*, 63, 1236–1245.
- R. Rooks et al., 1998, Structure of Distillation Regions for Multicomponent Azeotropic Mixtures, *AIChE J.*, 44 (6), 1382–1391.
- G. Schembecker & K.H. Simmrock, 1997, Heuristic-numeric design of separation processes for azeotropic mixtures, *Comput. & Chem. Eng.*, 21, S231–S236.
- T. Sasi & M. Skiborowski, 2020a, Automatic synthesis of distillation processes for the separation of homogeneous azeotropic multicomponent systems, *Ind. Eng. Chem. Res.*, 59, 20816–20835.
- T. Sasi et al., 2020b, Automatic synthesis of distillation processes for the separation of heterogeneous azeotropic multi-component mixtures, *Comput. Aided Chem. Eng.*, 48, 1009–1014.
- M. Skiborowski et al., 2016, A Unifying Approach for the Calculation of Azeotropes & Pinch Points in Homogeneous & Heterogeneous Mixtures, *Ind. Eng. Chem. Res.*, 55(24), 6815–6834.
- L.M.J. Sprakel & B. Schuur, 2019, Solvent developments for liquid-liquid extraction of carboxylic acids in perspective, *Sep. Purif. Technol.*, 211, 935–957.
- O.M. Wahnschafft et al., 1992, Split: II. Automated synthesis of hybrid liquid separation systems, *Comput. & Chem. Eng.*, 16, S305–S312.

A framework for optimal and flexible schemes design under uncertainty & sustainable aspects

Jaime D. Ponce-Rocha^{a,b}, Martín Picón-Núñez^a, and Ricardo Morales-Rodriguez^{a*}

^a*Departamento de Ingeniería química, Univeridad de Guanajuato, Guanajuato, 36050, México.*

^b*Dirección de Ingeniería y Construcción de Plantas, CIATEQ A. C., Querétaro, Querétaro, 76150, México.*

ricardo.morales@ugto.mx

Abstract

The modern design of chemical processes should involve holistic objective functions as well as the consideration of uncertainty sources. This work presents an optimization framework for flexible design combining stochastic programming with metaheuristic algorithm as solver, the 4E analysis as objective functions, and the integration of a decision-making criterion based on statistical fundamentals, applied in optimal distribution designs under uncertainty (80 % value of cumulative distribution functions (CFD)). The framework was applied to a separation and purification scheme (SPS) based on hybrid configurations and evaluating four extracting agents to separate an ABE mixture. The results show that the use of 80 % value of CFD is an adequate indicator for decision-making in multi-objective optimization under uncertainty. The similarities of the final designs using different extracting agents allows to conclude that a unique flexible and adaptable SPS design can be selected.

Keywords: Optimization under uncertainty, MOGA, Flexible design, 4E analysis.

1. Introduction

In the past decades, the objectives for optimal designs were focused on simple targets, because the standard metrics for chemical plants consisted of the assessment of the operational aspects and comparison of their performance through common targets (*e.g.* economy, heat duty, CO₂ emissions, etc.). These optimal designs are strictly related to steady-state operating conditions, and therefore, they are not naturally capable to respond to external perturbations and internal unexpected behaviors. In this sense, the introduction of the flexibility concept and multi-objective optimization are some of the principal and attractive challenges in Process System Engineering (Pretoro *et al.*, 2019).

Particularly, the design of chemical processes that involves multiple conflicting objectives and uncertainties, are identified as relevant and interesting optimization challenges, where the combination of two mathematical and solution strategies can be applied to tackle those issues: multi-objective optimization and optimization under uncertainty. Both strategies allow to obtain a final reliable and robust design (Zhang & Taflanidis, 2019). Regarding to flexible design, this generally involves system oversizing with respect to the nominal design, which represents major investment and operational costs, as well as better energy manipulation to compensate the variability and keep the final product specifications (Pretoro *et al.*, 2021).

Based on the actual state of the art, the metrics of modern process design must gradually adapt and consider internal and external perturbations, besides the fast-changing market demand, and the resources availability, without neglecting sustainability aspects and optimal operating conditions. Therefore, the objective of this work is to develop and implement a conceptual framework for optimal and flexible schemes design, associated to uncertainty sources using a computational platform based on the use of statistical indicators as decision-making criteria for the selection of the final design.

2. Case study

The production of biofuels and bio-bulk commodities is still an attractive path to satisfy a partial global market demand, in this sense, the acetone-butanol-ethanol production via *Clostridia* microorganism is an interesting bioprocess to contribute to this goal, which naturally has diverse sources of uncertainties; but due to the nonidealities in the mixture, high dilution ratio, multiple products and variable yields, the design of ABE separation and purification schemes (SPS) is a complex system to model and optimize (Grisales & Olivar, 2017). In previous works, liquid-liquid extraction (LLE) and conventional distillation were employed as separation and purification techniques for the ABE mixture, getting satisfactory energy results, and showing the relevance of suitable selection of extracting agents (EA) through computer-aided molecular design (Sanchez-Ramirez *et al.*, 2018). Based on the previous statements, the framework was evaluated using the hybrid SPS for ABE mixture.

3. Methodology

The framework for optimal and flexible schemes design under uncertainty & sustainable aspects involves the main idea to combine the stochastic programming fundamentals (wait-and-see strategy) but replaces the conventional solvers by a metaheuristic algorithm to solve the problem statement. This framework includes four steps illustrated in Figure 1 and are described below.

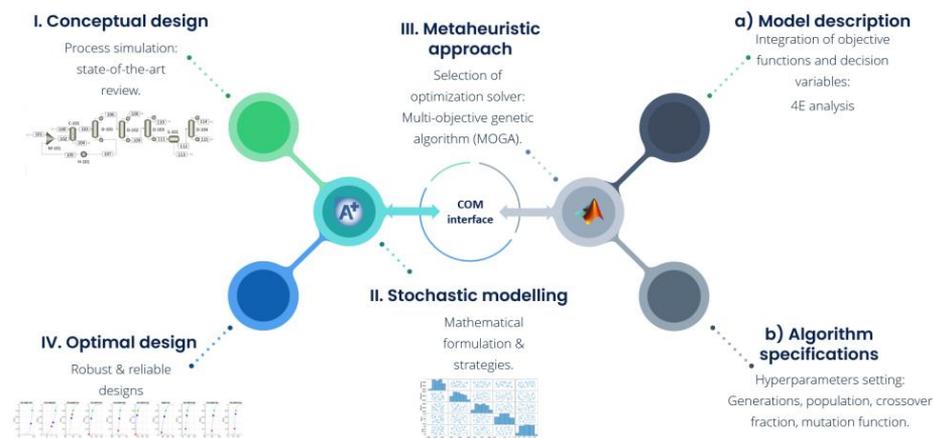


Figure 1. Framework for optimal and flexible design under uncertainty & sustainable aspects.

3.1. Conceptual design: Thermodynamic description

Aspen plus was employed for the conceptual design of the SPS (illustrated in Figure 2), including one LLE column (C-101) to remove ABE from the aqueous phase, evaluating four EA: heptyl acetate (HEP), hexyl acetate (HEX), heptyl-hexyl acetate mixture (HEP-

HEX), and 2-ethyl-1-hexanol (2-ETH-1-OL). Four conventional distillation columns (D-101, D-102, D-103, and D-104) were used to recover and purify the extracting agent, acetone, ethanol, and butanol, respectively. Phase equilibria was calculated employing the NRTL thermodynamic model and the equation of state Hayden-O'Connell (Ponce Rocha, 2021).

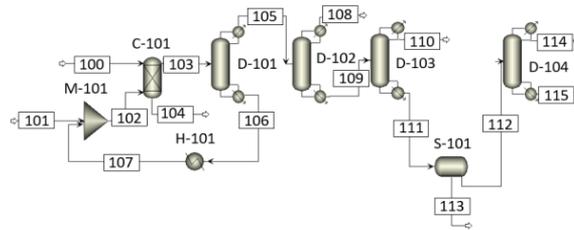


Figure 2. Case study employed to evaluate the MOOUU framework.

3.2. Stochastic modelling: MOOUU

The multi-objective optimization under uncertainty (MOOUU) employed in this work was formulated as following (1):

$$\underset{x}{\text{Min}} \quad \underline{Z} = [Z_k(\underline{x}, \theta^*)] \quad \text{s. t.} \quad \mathbf{h}(x, \theta^*) = 0 \quad \mathbf{g}(x, \theta^*) \leq 0 \quad x \in \mathbf{X}, \theta^* \in \Theta^* \quad (1)$$

Where \underline{Z} , describes the vector of k objective functions under uncertainty (Exergy, energy, economy, and environment), \underline{x} represents the vector of n-dimensional states associated with structural and operational specifications (FMAEX: extracting agent mass flowrate, NS: number of stages, FS: feed location, RR: reflux ratio, DF: distillate flowrate, or BF: bottom flowrate), θ^* is the vector of s-dimensional uncertainties present in the selected variables. \mathbf{h} & \mathbf{g} represent the vectors of equality and inequality constraints associated with operational and structural limitations.

3.2.1 Objective function definition: 4E-analysis

The combination of energetic, exergetic, economic, and environmental analysis (4E analysis) has been reported as a useful tool to design and optimize energetic intensive processes (Ponce Rocha, 2021). In the present work, this 4E functions are calculated as follows:

a) Exergetic analysis

The exergy losses (σT_0) can be obtained as illustrated in Equation (2), where $\Delta \dot{E}x_{In}$ represents the exergy input, and $\Delta \dot{E}x_{Out}$ is the exergy output, of each stream in the analyzed system, which includes the heating, cooling, and concentration changes effects.

$$\sigma T_0 = \Delta \dot{E}x_{In} - \Delta \dot{E}x_{Out} \quad [kW] \quad (2)$$

b) Energetic analysis

A common metric to validate the feasibility of biofuels separation is the ratio of the energy requirements and product throughput, as illustrated in equation (3).

$$ER = \frac{\text{Energy consumption [MJ]}}{\text{Product [kg]}} \quad (3)$$

c) *Economic analysis*

The total annual cost (TAC) is one of the more important indicators that includes the capital cost, operating cost, and the payback period (4).

$$TAC = \frac{\text{Capital cost}}{\text{Payback period}} + \text{Utilities} \quad (4)$$

d) *Environmental analysis*

A simple metric to evaluate the environmental impact is the use of the E factor, this indicator is the ratio between waste and products generated in the process.

$$E_{factor} = \frac{\text{Waste [kg]}}{\text{Product [kg]}} \quad (5)$$

3.2.2 *Uncertainty identification and propagation*

The feed flow stream (100) was identified as one of the principal uncertainties sources in the system since this is coming from the fermentation section. Therefore, the propagation of the uncertainty in the feed composition was generated using the following values: acetone (A) 2,275±112 kg/h, butanol (B) 5,931±276 kg/h, ethanol (E) 1,592±135 kg/h, water (W) 10,479±498 kg/h, carbon dioxide (CO₂) 82±5 kg/h. The Latin hypercube sampling (LSH) method was used for the generation of seventy-five random scenarios using a uniform distribution function.

3.3 *Metaheuristic algorithm: simulation-optimization approach*

In this section, a computational interface (COM) was used to integrate Aspen Plus and Matlab. This interface involves three main steps: 1) Matlab generates random vector values using the LSH, 2) The vector values are used by Aspen Plus to evaluate each scenario under uncertainty, 3) The optimal operating and design results are returned to matlab to calculate the objective function values. This step is performed according to the hyperparameters definition. The optimal values are saved to be used in the next step for graphical, statistical and feasibility analysis. The Matlab suite: gamultiobj, a multi-objective genetic algorithm (MOGA) and a variant of NSGA-II were used for metaheuristic optimization. Each optimization employed the following hyperparameters: 50 generations, 200 individuals, and a crossover fraction of 0.8 and an adaptative feasible mutation function. The mass purity (P) and mass recovery (R) constraints are specified for equations (6) and (7).

$$[P_{HEP}, P_A, P_B, P_E] \geq [0.999, 0.950, 0.995, 0.900] \quad (6)$$

$$[R_{HEP}, R_A, R_B, R_E] \geq [0.999, 0.995, 0.000, 0.992] \quad (7)$$

3.4 *Optimal distribution designs: decision making for flexible design*

After the evaluation of each uncertain scenario employing the evaluated EA, a set of optimal designs was obtained, which can be analyzed employing statistical indicators as tool for establishing a robust and reliable design in uncertainty scenarios. Three criteria were evaluated: mean, mode, and the 80 % of the cumulative distribution function (CDF).

4. Results

Figure 3 shows the distribution designs for each EA, using the mean (blue dot), mode (red dot) and 80 % of CFD (green dot). Then, the minimum, average, and maximum flowrate vectors were evaluated in the final optimal design obtained for each statistical indicator for each extracting agent. In this sense, the 80 % of CFD indicator showed better results

assuring the robustness of the design through the convergence analysis of the average value and the lower and upper boundaries of the feed flowrate vectors.

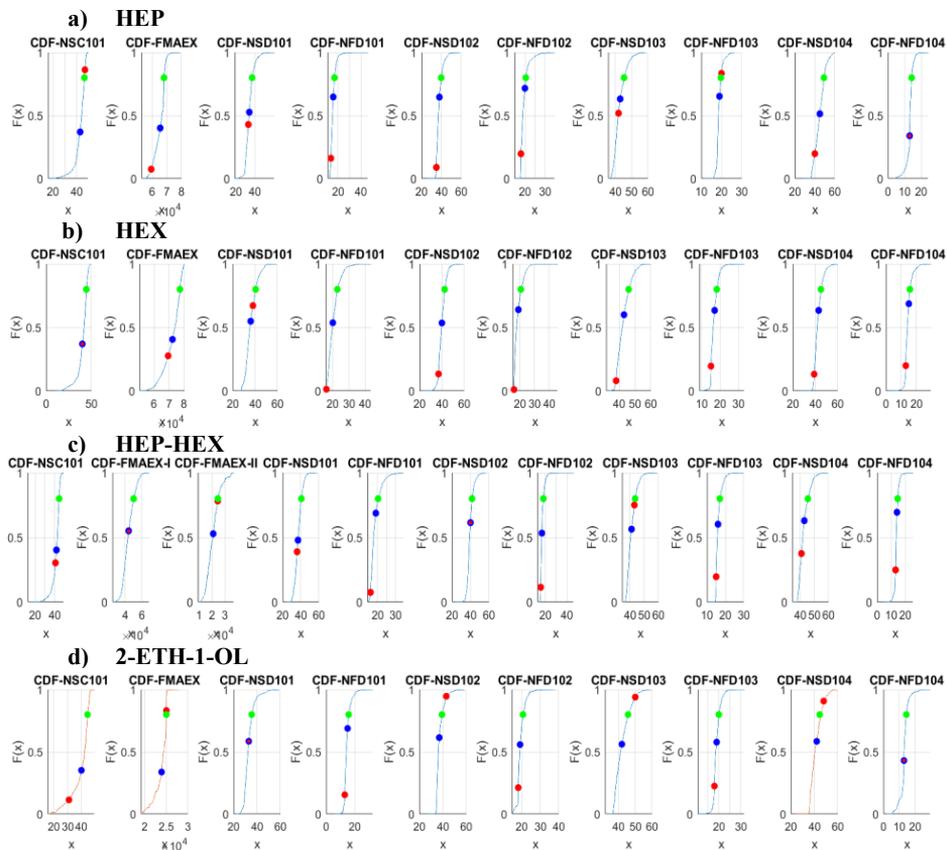


Figure 3. CFD for optimal distribution structural designs for each EA. FMAEX: extracting agent mass flowrate, NS: number of stages, FS: feed location.

Table 1 illustrates the optimal design results summary obtained for the 80 % of CFD indicator, employing the evaluated EA. In Table 1 is possible to observe similar structural designs specification such as, number of stages and feed locations for each equipment.

Table 1. Comparative structural designs for each design and operating variables.

Variable	N _S	F _{MAEX}	N _S	N _F	N _S	N _F	N _S	N _F	N _S	N _F
	C ₁₀₁	kg/h	D ₁₀₁	D ₁₀₁	D ₁₀₂	D ₁₀₂	D ₁₀₃	D ₁₀₃	D ₁₀₄	D ₁₀₄
HEP	47	68,080	37	17	40	20	46	20	49	14
HEX	44	77,292	41	23	43	21	46	18	45	16
HEP-HEX	45	24,753-50,418	41	20	42	18	44	17	46	15
2-ETH-1-OL	45	25,000	36	16	39	21	46	20	46	15

Table 2 shows the average results and standard deviations of the objective functions values for each extracting agent, employing the average values and the lower and upper boundaries of the uncertain values of the feed flowrate vectors. The use of HEP as EA gives the best results in three criteria (exergy, energy and environmental). The second option was the use of the HEP-HEX mixture as EA, followed by 2-ETH-1-OL and HEX.

Table 2. 4E objective functions values

	σT_0 [MW]	ER [MJ/kg ABE]	TAC [MUSD/year]	E _{factor}
HEP	2.00±0.06	12.19±0.42	8.34±0.22	0.24±0.01
HEX	2.28±0.09	17.33±0.35	10.96±0.29	0.44±0.01
HEP-HEX	2.29±0.07	14.88±0.49	9.86±0.33	0.32±0.02
2-ETH-1-OL	2.03±0.07	16.50±0.45	7.87±0.29	0.80±0.02

5. Conclusions

In this framework, a novel strategy to optimize under uncertainty a complex SPS was applied. The implementation of 4E analysis and MOGA showed satisfactory results in the stochastic problem solution. The 80 % of the CFD indicator was validated as a satisfactory indicator to select reliable and robust designs. Due to the identical targets in separation and purification tasks, each optimal final design for each EA has similar structural specifications, in this sense, a single design can be proposed and identified with adaptable and flexible characteristic, where EA flowrate is variable due to their physicochemical properties. Regarding to the EA, HEP showed the best performance results, but the idea of a flexible and adaptable design would allow to use other EA considering market availability.

6. Acknowledgments

Jaime D. Ponce-Rocha acknowledges the National Council of Science and Technology (CONACyT) for the scholarship received to develop this project (731399).

References

- A. D. Pretoro, L. Montastruc, F. Manenti, & X. Joulia, 2019, Flexibility analysis of a distillation column: indexes comparison and economic assessment, *Computers & Chemical Engineering*, 124, 93-108.
- J. Zhang, & A. Taflanidis, 2019, Evolutionary multi-objective optimization under uncertainty through adaptive Kriging in augmented input space, *Journal of Mechanical Design*, 142, 011404.
- A. D. Pretoro, L. Montastruc, X. Joulia, & F. Manenti, 2021, Accounting for dynamics in flexible process design: A switchability index. *Computers & Chemical Engineering*, 145, 107149.
- V. H. Grisales Diaz, & G. Olivar Tost, 2017, Energy efficiency of acetone, butanol, and ethanol (ABE) recovery by heat-integrated distillation, *Bioprocess and Biosystems Engineering*, 41(3), 395-405.
- E. Sánchez-Ramírez, J. D. Ponce-Rocha, J. G. Segovia-Hernández, F. I. Gómez-Castro, & R. Morales-Rodríguez, 2018, A framework for optimised sustainable solvent mixture and separation process design, *Computer Aided Process Engineering*, 43, 755-760.
- J. D. Ponce-Rocha, M. Picón-Núñez, A. Carvalho, A. Santos, F. I. Gómez-Castro, & R. Morales-Rodríguez, 2021, A design and optimization framework for (bio-) chemical process based on exergo-economic and environmental aspects, *Computer Aided Chemical Engineering*, 50, 1897-1902.

Optimization-based assessment framework for CO₂ utilization to fuels strategies

Thai Ngan Do,^a Chanhee You,^a Minseong Park,^a Changsu Kim,^a Seongju Lee,^b
and Jiyong Kim^{a,*}

^a*School of Chemical Engineering, Sungkyunkwan University, 16419, Republic of Korea*

^b*Department of Energy and Chemical Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon, 22012, South Korea*

*Corresponding Author's E-mail: jiyongkim@skku.edu

Abstract

Carbon capture and utilization for fuel production is one of the attractive and effective solutions addressing climate change and energy security. In which captured CO₂ is considered as raw materials for high-energy-density products (e.g., methanol, dimethyl ether, Fischer-Tropsch fuel, gasoline) via different technologies (catalytic conversion, thermochemical energizing, electrochemical reduction). This study developed an optimization-based framework to analyze and assess CO₂ utilization strategies for fuel products regarding technical, economic, and environmental performance. To achieve this goal, we first generated a superstructure involving a series of technologies (carbon conversion and separation) to produce value-added fuels from captured CO₂ as a feedstock. We then simulated all the involved processes and estimated the technical and economic parameters (mass and energy flow, and sizing and costing data) that were further adopted into the optimization model. The optimization models were developed to identify the optimal CO₂ utilization strategies with different criteria: energy efficiency, production cost, profit, and CO₂ reduction. As a result, we can determine the best CO₂ utilization strategy over various technological pathways to produce different targeted fuels, which makes CO₂-based fuels economically and/or environmentally viable.

Keywords: Process design; CO₂ utilization superstructure; Optimization.

1. Introduction

Carbon capture and utilization is one of the attractive and effective solutions addressing global warming and energy security. This action supports reducing CO₂ emission by capturing and utilizing CO₂ as a raw material for high-value fuels and chemicals such as methanol (MeOH), dimethyl ether (DME), Fischer-Tropsch fuel (FT fuel), and gasoline. Besides mitigating emission, the CO₂-based fuels can share the burden on fossil fuels by replacing (e.g., FT fuels, gasoline) or blending (e.g., MeOH, DME) with conventional fuels.

Recently, the literature has found much research on CO₂ utilization technologies such as catalytic conversion, electrochemical reduction, and thermochemical energizing (Kim et al, 2012, and Mevawala et al., 2017). While some work on the high performed catalyst or thermochemical/electrochemical materials, others study process design and optimization, techno-economic and environmental analysis of CO₂ utilization to fuels (Do et al., 2019 and 2020). Also, many studies conduct a technological superstructure and optimization-based framework for fuels and chemicals from various feedstock materials. Kim et al.

developed a superstructure of fuels from biomass and assessed the conversion strategies with different evaluation criteria, the uncertainty, and sensitivity analysis (Kim et al., 2013). Han et al. developed an assessment framework for optimal strategies to utilize industrial residue gas (Han et al., 2019).

This study developed an optimization-based assessment framework to analyze and assess CO₂-to-fuel utilization strategies with different evaluation criteria: energy efficiency, economic, and environmental friendliness. The problem statement and methodology are presented in Section 2. The generation of the technological superstructure, including various carbon conversion and separation technologies, which aim to produce CO₂-based fuels: MeOH, DME, FT fuels, and gasoline, are discussed in Section 3. Then, the process simulation of possible CO₂-to-fuels pathways among the superstructure was developed and the techno-economic parameters (e.g., mass and energy flow, and sizing and costing data) were further obtained and presented in Section 4. In section 5, the optimization models were developed to identify the optimal CO₂ utilization strategies with different criteria: energy efficiency, production cost, and CO₂ reduction. Finally, the best CO₂ utilization strategy over various technological pathways for different targeted fuels is discussed in Section 6.

2. Methodology

This study aims to propose and develop an optimization-based assessment framework of CO₂ utilization to fuels, which supports identifying the optimal strategies for specific CO₂-based fuels, as shown in Figure 1. We first define the problem by selecting captured CO₂ as the main feedstock, selecting final products and technologies, and developing the superstructure by integrated technologies to convert CO₂ forward to final products. The process simulation of all possible production pathways is then developed and the technical (mass and energy flow), and economic data (sizing and costing data) are obtained. The optimization model is developed based on the generated superstructure framework and the associated techno-economic data. The model identifies the optimal CO₂-to-fuel strategy for a given problem and given final products, such as: maximizing energy efficiency (EEF), minimum unit production cost (UPC), and net CO₂ emission (NCE).

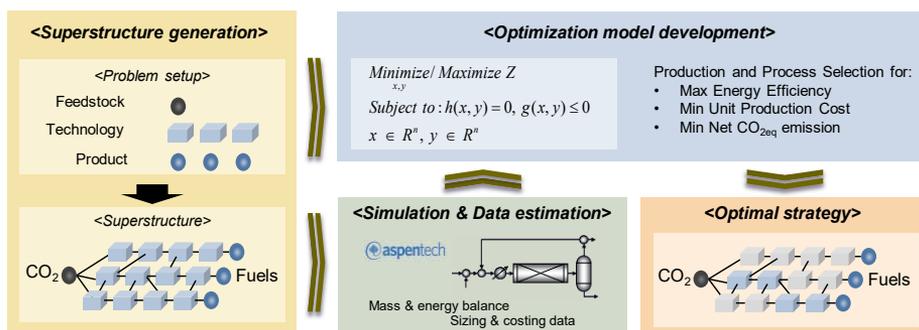


Figure 1. Methodology for the assessment of CO₂ utilization strategy for liquid fuels

3. Superstructure

In this study, the compounds were classified into three groups: feedstock (i.e., captured CO₂), intermediates (e.g., syngas, raw fuels), and final products (e.g., MeOH, FT fuel, DME, and gasoline). The technologies including reaction/conversion technology (e.g.,

direct CO₂ hydrogenation, reverse water-gas shift, thermochemical CO₂ energizing, electrochemical CO₂ reduction, methanol/DME/FT synthesis from syngas, methanol-to-gasoline, methanol dehydration) component separation (e.g., CO₂ absorption, CO adsorption), and product purification (e.g., flash drum, distillation). Note that a single technology includes various operating equipment.

The CO₂ utilization superstructure was developed by connecting compounds and multiple techniques for targeted products, as presented in Figure 2. Here, there are many production pathways for a specific product. For example, FT fuels can be produced via direct CO₂ hydrogenation or FT synthesis from syngas (indirect pathway), in which, the intermediate syngas can be converted through reverse water-gas shift or thermochemical splitting, or electrochemical reduction.

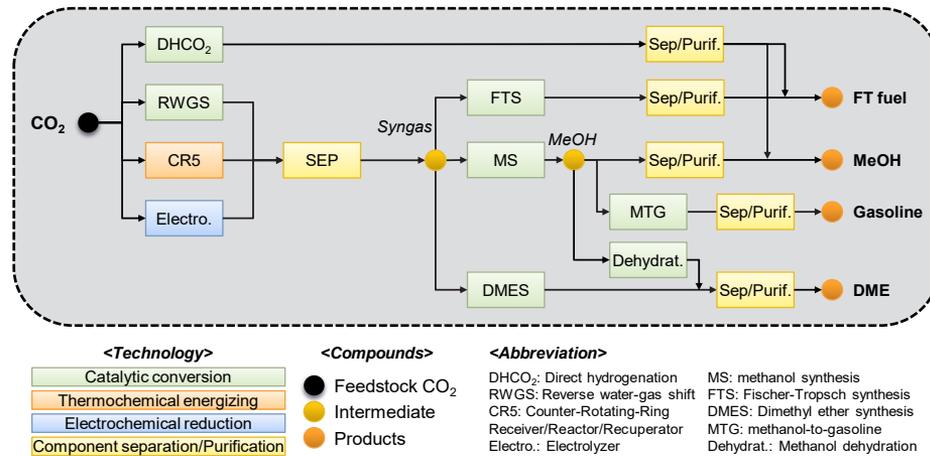


Figure 2. Superstructure of CO₂ utilization to fuels

4. Simulation and parameter estimation

In this study, CO₂ feedstock is assumed to be captured from the flue gas of coal-powered electric utility (500 MW), which is about 3.1 million ton of captured CO₂ per year. Each process was simulated using Aspen Plus V.11. More detailed process simulation and modeling could be found in the literature (Do et al., 2022). The mass flow, energy flow, sizing and costing data were then obtained for further model input and analysis. Table 1 summarizes major technical and economic parameters.

Table 1. Technical and economic parameters of the CO₂-to-fuels.

No.	Technological pathway	Production	Input (ton/h)	Output (ton/h)	Yield ^a	TCI ^b (M\$)	TOC ^c (M\$/y)	ER ^d (MW)	DCE ^e (ton CO _{2eq} /h)	EO ^f (10 ³ GGE/h)
1	RWGS-MS-SEP	MeOH_1	437.9	249.9	0.57	664	2,093	759	36.7	47.4
2	CR5-MS-SEP	MeOH_2	433.3	253.5	0.59	10,934	1,875	679	32.8	48.1
3	ELECZ-MS-SEP	MeOH_3	431.7	255.5	0.59	3,243	1,900	990	30.4	48.4
4	RWGS-COSEP-MS-SEP	MeOH_4	439.1	260.0	0.59	3,639	2,462	708	26.5	49.2
5	CR5-COSEP-MS-SEP	MeOH_5	425.3	257.2	0.60	35,241	2,156	567	29.2	48.6
6	ELECZ-COSEP-MS-SEP	MeOH_6	426.2	264.6	0.62	7,291	2,091	1,169	18.8	50.1
7	RWGS-CO2SEP-MS-SEP	MeOH_7	437.2	253.1	0.58	692	2,204	1,699	33.8	48.0
8	CR5-CO2SEP-MS-SEP	MeOH_8	424.2	256.6	0.60	31,874	1,677	1,322	29.2	48.7
9	ELECZ-CO2SEP-MS-SEP	MeOH_9	425.6	263.9	0.62	5,430	1,953	1,873	18.9	50.0
10	DHCO2MEOH-SEP	MeOH_10	440.1	270.4	0.61	734	2,068	383	14.6	51.3
11	RWGS-COSEP-FTS-SEP	FT_fuel_1	443.5	82.5	0.19	4,048	2,805	1,217	6.3	30.9
12	CR5-COSEP-FTS-SEP	FT_fuel_2	429.5	81.9	0.19	35,644	2,486	1,032	7.5	30.6
13	ELECZ-COSEP-FTS-SEP	FT_fuel_3	429.6	83.7	0.19	7,701	2,399	1,691	4.5	31.4
14	RWGS-CO2SEP-FTS-SEP	FT_fuel_4	442.2	79.9	0.18	1,016	2,561	2,146	9.8	30.0

15 CR5-CO2SEP-FTS-SEP	FT_fuel_5	428.4	81.7	0.19	32,229	2,000	1,776	7.5	30.6
16 ELECZ-CO2SEP-FTS-SEP	FT_fuel_6	429.0	83.6	0.19	5,798	2,248	2,364	4.5	31.4
17 DHCO2FT-SEP	FT_fuel_7	442.8	119.2	0.27	880	2,203	438	145.9	44.7
18 RWGS-MS-DEH.-SEP	DME_1	437.9	179.5	0.41	1,250	2,149	992	36.7	46.9
19 CR5-MS-DEH.-SEP	DME_2	433.3	182.1	0.42	11,526	1,932	915	32.8	47.6
20 ELECZ-MS-DEH.-SEP	DME_3	431.7	183.6	0.43	3,837	1,957	1,228	30.4	48.0
21 RWGS-COSEP-MS-DEH.-SEP	DME_4	439.1	186.8	0.43	4,240	2,520	949	26.5	48.7
22 CR5-COSEP-MS-DEH.-SEP	DME_5	425.3	184.7	0.43	35,837	2,213	805	29.2	48.2
23 ELECZ-COSEP-MS-DEH.-SEP	DME_6	426.2	190.1	0.45	7,899	2,151	1,415	18.8	49.6
24 RWGS-CO2SEP-MS-DEH.-SEP	DME_7	437.2	181.9	0.42	1,283	2,261	1,935	33.8	47.5
25 CR5-CO2SEP-MS-DEH.-SEP	DME_8	424.2	184.3	0.43	32,470	1,734	1,561	29.2	48.2
26 ELECZ-CO2SEP-MS-DEH.-SEP	DME_9	425.6	189.6	0.45	6,038	2,012	2,118	18.9	49.5
27 DHCO2MEOH-DEH.-SEP	DME_10	440.1	194.3	0.44	1,351	2,129	635	14.6	50.8
28 RWGS-COSEP-DMES-SEP	DME_11	417.6	179.9	0.43	3,573	1,679	511	182.2	46.9
29 CR5-COSEP-DMES-SEP	DME_12	402.2	95.9	0.24	35,170	1,307	315	183.7	25.0
30 ELECZ-COSEP-DMES-SEP	DME_13	407.0	114.0	0.28	7,250	1,405	1,041	159.1	29.7
31 RWGS-CO2SEP-DMES-SEP	DME_14	410.4	80.5	0.20	434	1,200	1,355	215.9	21.0
32 CR5-CO2SEP-DMES-SEP	DME_15	401.1	95.7	0.24	31,770	813	1,049	195.7	25.0
33 ELECZ-CO2SEP-DMES-SEP	DME_16	406.3	113.9	0.28	5,348	1,253	1,713	159.1	29.8
34 RWGS-MS-MTG-SEP	Gasoline_1	437.9	120.0	0.27	1,520	2,054	527	11.4	45.9
35 CR5-MS-MTG-SEP	Gasoline_2	433.3	125.8	0.29	11,814	1,834	435	6.4	48.1
36 ELECZ-MS-MTG-SEP	Gasoline_3	431.7	113.7	0.26	4,064	1,863	769	4.7	43.5
37 RWGS-COSEP-MS-MTG-SEP	Gasoline_4	438.1	97.5	0.22	4,381	2,045	537	5.4	37.3
38 CR5-COSEP-MS-MTG-SEP	Gasoline_5	424.1	99.7	0.23	35,993	1,673	363	6.7	38.1
39 ELECZ-COSEP-MS-MTG-SEP	Gasoline_6	425.6	93.9	0.22	8,016	1,819	1,019	2.9	35.9
40 RWGS-CO2SEP-MS-MTG-SEP	Gasoline_7	437.2	104.7	0.24	1,453	2,165	1,464	9.7	40.1
41 CR5-CO2SEP-MS-MTG-SEP	Gasoline_8	424.2	99.7	0.24	32,627	1,642	1,119	6.7	38.1
42 ELECZ-CO2SEP-MS-MTG-SEP	Gasoline_9	425.6	94.0	0.22	6,157	1,933	1,754	2.9	35.9
43 DHCO2MEOH-MTG-SEP	Gasoline_10	440.1	139.9	0.32	1,701	2,034	176	6.0	53.5

^a Yield: calculated based on the material input and production output.

^b TCI (M\$): total capital investment cost.

^c TOC (M\$/y): total annual operating cost

^d ER (MWh): energy requirement for process.

^e DCE (ton/h): direct CO_{2eq} emission (vent-out gas, purge gas) from process.

^f EO (10³ GGE/h): Energy flow of products.

Abbreviations: RWGS: Reverse water-gas shift, DH: Direct CO₂ hydrogenation, CR5: Counter-Rotating-Ring Receiver/Reactor/ Recuperator, ELECZ: Electrochemical reduction, MS: Methanol synthesis, FTS: Fischer–Tropsch synthesis, DMES: dimethylether synthesis, DEH.: Methanol dehydration, MTG: methanol-to-gasoline, COSEP: CO separation by adsorption, CO2SEP: CO₂ separation by absorption, SEP: separation and purification.

5. Optimization model

To identify the optimal strategy for CO₂ utilization to fuels, we develop optimization models using a mixed-integer linear programming (MILP) technique to identifies the optimal CO₂-to-fuel pathway for each CO₂-based fuel type. Eq. (1) uses to identify the maximum energy efficiency strategy, which is maximum energy stored in product with a certain input energy to process. Eq. (2) seeks for the most economic strategy, which is minimum unit production cost for a fixed amount of product. Other, minimum net CO₂ emission strategy is identified vis Eq. (3), which emits as producing a certain amount of product.

$$MaxEEF_s = \sum_j \rho_i P_{ij} \quad (1)$$

$$MinUPC_s = \sum_j \alpha_j + \sum_j \psi_j + \sum_{i \in I^F} \xi_i F_{ij} + \sum_{i \in I^U} \varpi_i U_{ij} \quad (2)$$

$$MinNCE_s = \sum_j \lambda_j + \sum_{i \in I^F} \varepsilon_i U_{ij} - \sum_{i \in I^P} \tau_i F_{ij} \quad (3)$$

Where ρ_i is the heating value of final product $i \in I^P$. F_{ij} , U_{ij} , P_{ij} , is the the amount of feestock $i \in I^F$, utilities $i \in I^U$ and product of pathway j , respectively. α_j is the

total capital investment cost of pathway j , ψ_j is the fixed operating cost factor of pathway j . ϖ_i and ξ_i are the unit costs for utilities $i \in I^U$ and feedstock $i \in I^F$, respectively. λ_j is the amount of CO₂ directly emitted by technology j . ε_i is the amount of indirect CO₂ emission by using utility $i \in I^U$, and τ_i is the CO₂ inventory for feedstock $i \in I^F$.

The optimization model was constrained by the demand satisfaction and feed availability and minimum feed purchase, as expressed in Eqs. (4) – (5).

$$P_i \geq \beta_i \tag{4}$$

$$\gamma_i \leq F_i \leq \delta_i \tag{5}$$

Where β_i is the demand of product. γ_i and δ_i are the minimum purchase and feed's availability, respectively.

6. Results and discussion

With the proposed models, we can identify the optimal strategy to utilize CO₂ to different fuels objecting to max EEF, min UPC and min NCE. Figure 3 presented the optimal strategies, including CO₂-to-fuel technological pathways and three main evaluated criteria for the production of MeOH, FT fuel, DME and gasoline.

Target Fuels	Problem	Optimal Strategy	Description
Methanol (MeOH)	(a) Max EEF	CO ₂ → CR5 → SEP → MS → Sep/Purif. → MeOH	<ul style="list-style-type: none"> • EEF: 82.7% • UPC: 18.13 \$/GGE • NCE: -0.93 kgCO₂/GGE
	(c) Min UPC	CO ₂ → DHCO ₂ → Sep/Purif. → MeOH	<ul style="list-style-type: none"> • EEF: 71.2% • UPC: 3.60 \$/GGE • NCE: -1.01 kgCO₂/GGE
	(d) Min NCE	CO ₂ → DHCO ₂ → Sep/Purif. → MeOH	<ul style="list-style-type: none"> • EEF: 71.2% • UPC: 3.60 \$/GGE • NCE: -1.01 kgCO₂/GGE
Fischer-Tropsch (FT fuel)	(a) Max EEF	CO ₂ → DHCO ₂ → Sep/Purif. → FT fuel	<ul style="list-style-type: none"> • EEF: 58.2% • UPC: 4.46 \$/GGE • NCE: 3.22 kgCO₂/GGE
	(c) Min UPC	CO ₂ → DHCO ₂ → Sep/Purif. → FT fuel	<ul style="list-style-type: none"> • EEF: 58.2% • UPC: 4.46 \$/GGE • NCE: 3.22 kgCO₂/GGE
	(d) Min NCE	CO ₂ → DHCO ₂ → Sep/Purif. → FT fuel	<ul style="list-style-type: none"> • EEF: 58.2% • UPC: 4.46 \$/GGE • NCE: 3.22 kgCO₂/GGE
Dimethyl ether (DME)	(a) Max EEF	CO ₂ → CR5 → SEP → DMES → Sep/Purif. → DME	<ul style="list-style-type: none"> • EEF: 93.9% • UPC: 32.56 \$/GGE • NCE: 2.47 kgCO₂/GGE
	(c) Min UPC	CO ₂ → DHCO ₂ → Dehydrat. → Sep/Purif. → DME	<ul style="list-style-type: none"> • EEF: 58.4% • UPC: 4.01 \$/GGE • NCE: 0.18 kgCO₂/GGE
	(d) Min NCE	CO ₂ → DHCO ₂ → Dehydrat. → Sep/Purif. → DME	<ul style="list-style-type: none"> • EEF: 58.4% • UPC: 4.01 \$/GGE • NCE: 0.18 kgCO₂/GGE
Gasoline	(a) Max EEF	CO ₂ → DHCO ₂ → MTG → Sep/Purif. → Gasoline	<ul style="list-style-type: none"> • EEF: 81.3% • UPC: 3.71 \$/GGE • NCE: -1.86 kgCO₂/GGE
	(c) Min UPC	CO ₂ → DHCO ₂ → MTG → Sep/Purif. → Gasoline	<ul style="list-style-type: none"> • EEF: 81.3% • UPC: 3.71 \$/GGE • NCE: -1.86 kgCO₂/GGE
	(d) Min NCE	CO ₂ → CR5 → SEP → MS → MTG → Gasoline	<ul style="list-style-type: none"> • EEF: 72.3% • UPC: 21.92 \$/GGE • NCE: -2.78 kgCO₂/GGE

Figure 3. Optimal strategy for CO₂-based fuel production using different criteria

The optimal strategy with the highest process energy efficiency for MeOH is through thermochemical CO₂ energizing subsequent methanol synthesis, at 82.7%. However, it is challenged with extremely high production cost at 18.13 \$/GGE (GGE – gallon gasoline equivalent). The direct CO₂ hydrogenation to MeOH is the most economic and eco-friendly pathway at 3.6 \$/GGE and reduces 1.01 kgCO₂/GGE (NCE<0), respectively.

For production of FT fuel, the direct hydrogenation of CO₂ is also the optimal strategy over other pathways and over three examined problems. Compared to MeOH, FT fuels is less energy efficient (58.2 %), higher cost (4.46 \$/GGE), and emit 3.22 kgCO₂/GGE (NCE>0). Here, FT is indicated as an unfavorable choice for eco-hydrocarbon synthesis. Similar to MeOH, the CO₂-to-DME pathway of thermochemical CO₂ energizing subsequent DME synthesis is the optimal strategy for maximum EEF. The optimal strategy for cost and environmental problems is through direct CO₂ hydrogenation to MeOH subsequent MeOH dehydration to DME, which results in 4.01 \$/GGE and 0.18 kgCO₂/GGE emission.

The direct CO₂ hydrogenation to MeOH followed methanol-to-gasoline is the optimal gasoline production strategy regarding max EEF (81.3%) and min UPC (3.71 \$/GGE). This strategy also performs good environmental performance, which reduces CO₂_{eq} at NCE of -1.86 kgCO₂/GGE. However, the optimal eco-friendly strategy is through thermochemical energizing technology for methanol synthesis subsequent methanol-to-gasoline, at -2.78 kg CO₂/GGE with the trade-off of extremely high production cost.

7. Conclusion

In this study, we proposed and developed an optimization-based assessment framework for CO₂ utilization to fuels strategies. The framework is based on the CO₂-to-fuels superstructure includes a number of conversion and separation technologies for desired fuel production. Then, the optimization model enables us to assess CO₂ utilization strategies with evaluation criteria regarding energy efficiency, unit production cost, and process eco-friendliness. Thereby, the framework determines the best CO₂ utilization strategy over various technological pathways to produce different targeted fuels, which makes CO₂-based fuels economically and/or environmentally viable. It can also support the policymaker or company identify a suite of strategic solutions to utilize captured CO₂ for fuels.

References

- J. Kim, T. A. Johnson, J. E. Miller, E. B. Stechel, C.T. Maravelias, 2012, Fuel production from CO₂ using solar-thermal energy: System level analysis. *Energy Environ. Sci.* 5, 8417–8429.
- C. Mevawala, Y. Jiang, D. Bhattacharyya, 2017, Plant-wide modeling and analysis of the shale gas to dimethyl ether (DME) process via direct and indirect synthesis routes. *Appl. Energy* 204, 163–180
- TN. Do, J. Kim, 2019, Process development and techno-economic evaluation of methanol production by direct CO₂ hydrogenation using solar-thermal energy. *J CO₂ Util*, 33, 461–72
- TN. Do, J. Kim, 2020, Green C2-C4 hydrocarbon production through direct CO₂ hydrogenation with renewable hydrogen: Process development and techno-economic analysis. *Energy Convers Manag*, 214, 112866.
- J. Kim, SM. Sen, CT. Maravelias, 2013, An optimization-based assessment framework for biomass-to-fuel conversion strategies. *Energy Environ Sci*, 1093–104
- S. Han, S. Kim, Kim, YT., Kwak, G., Kim, J., 2019, Optimization-based assessment framework for carbon utilization strategies: Energy production from coke oven gas. *Energy Convers Manag*, 187, 1–14
- TN. Do, C. You, J. Kim, 2022, A CO₂ utilization framework for liquid fuels and chemical production: techno-economic and environmental analysis. *Energy Environ Sci.*, 15, 169–184

A novel approach to continuous extraction of active ingredients from essential oils through combined chromatography

Constantin Jurischka^a, Bogdan Dorneanu^a, Christian Stollberg^b, Harvey Arellano-Garcia^a

^a*LS- Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus D-03044, Germany*

^b*Hochschule Wismar, Verfahrenstechnik Biogener Rohstoffe, Malchow D-23999, Germany*

Abstract

This contribution introduces a combined liquid chromatography purification designed for a continuous and resource-efficient process, integrating the rotating columns and the simulated bed principles. The approach is demonstrated and validated based on bisabolol oxides A and B, which are effective ingredients with anti-inflammatory and spasmolytic effects, prepared from chamomile essential oil. The results show superior efficiency to the traditional selective methods for isolating ingredients from multicomponent mixtures, as well as reduction in resources and costs.

Keywords: Essential oil; Preparative chromatography; Simulated moving bed (SMB).

1. Introduction

Bioactive ingredients from plant extracts are the basis of many innovative products in the food, cosmetics, or the pharmaceutical industries. However, at present, mainly diluted products are extracted as complex mixtures, as a large amount of by-products accumulates as waste at the end of the process, and can only be recovered with increased effort. Essential oils are complex volatile substances composed of many ingredients with broad spectrum of bioactivity and, therefore, subject to extensive research (Chouhan et al. 2017). Various processes have been used to extract these ingredients: harvesting as plant materials, washing, extraction of phytonutrients, or purification of the components of interest (Belwal et al. 2020). Traditional techniques consist of liquid-liquid extraction, solid-liquid extraction, or solid phase micro-extraction. Modern techniques include ultrasound-assisted extraction, pressurised liquid extraction, supercritical fluid extraction, microwave-assisted extraction, or instant controlled pressure drop extraction (Yahya et al. 2018). Often, for the same plant, extraction conditions may vary in terms of extraction time or amount of solvent used. Conventional methods suffer from some serious disadvantages, such as excessive solvent, time and energy requirements, which demand the development of new techniques that are more efficient, less expensive and least hazardous (Abbas et al. 2021). Modern techniques show significant improvement over the traditional ones in terms of less extraction time, greater extraction yield and lower solvent consumption, but depend on parameters such as solvent composition, amount of solvent, and plant material loading (Bagade et al. 2021). Moreover, green extraction technologies come with their own shortcomings: high capital investment, high running

cost, complex configuration, training, maintenance cost, which limit their large scale application (Picot-Allain et al. 2021).

Using the example of bisabolol oxides A and B (BOA, BOB), effective ingredients with anti-inflammatory and spasmolytic effects, prepared from the essential oil of a bisabolol oxide-rich chamomile (Schilcher et al. 1987), a combined liquid chromatographic purification is designed and implemented in the following sections up to their complete separation.

2. Materials and methods

The developed procedure is implemented using a system for continuous liquid chromatography (Wissenschaftliche Gerätebau Dr.-Ing. Knauer GmbH, Berlin; Type: CSEP®9116) with a 64-port multifunction valve (100 bar pressure-resistant). Up to 16 separation units (columns) can be combined in parallel or in series. In the present work, (semi-)preparative columns (C18H, 10 μ m 8x120mm) are used. The active ingredients were detected using a photo diode array (PDA) detector (Knauer Smartline S2800).

2.1. Materials

The chamomile oil produced by steam distillation is separated from the paraffins (long-chain, branched and unbranched n-alkanes) contained in it using a simple crystallization method. Subsequently, the oil could be prepared by distillation in such a way that the low-volatility components (β -farnesene, germacrene D, elemene) are largely depleted and only present in traces (remaning low-volatility compounds < 2 wt. %). The concentration of the valuable active ingredients increases during this step, and a pre-fractionated oil is produced with the composition presented in Table 1.

Table 1: Chamomile oil content

Ingredient	Mass fraction
β -farnesene	1.68 wt.%
Bisabolol oxide A	45.52 wt.%
Bisabolol oxide B	28.51 wt.%
Chamazulene	9.55 wt.%
Rest	14.74 wt.%

To avoid policy hurdles during a possible product development, the process design is limited to a solvent mixture of ethanol/water (mobile phase – MP), which is pharmacologically harmless.

The modular structure of the simulation toolbox offers flexibility in adding new modules for improved representation of the internal geometries as well as of the different phenomena (*e.g.*, holdup, fluid flow, kinetics, etc.) taking place inside the TBR, as well as expansion to energy balance to move away from the isothermal assumptions.

2.2. Analysis

Gas chromatography systems are used to analyse the starting, intermediate, and end products. Their structure is elucidated and identified using a GC-MS system (Shimadzu QP2010S), and then transferred to a GC-FID system for quantification (Shimadzu 2010Plus).

2.3. The rotating columns principle

Bauer et al. (1996) introduced a continuous chromatographic process based on a rotating multifunctional valve in which individual fractions can be separated spatially from a multifunctional mixture. In principle, the process can be compared with annular chromatography (Brozio et al. 2004), whereas the annular gap built up with the stationary phase (SP) is divided into discrete units and, thus, replaced by traditional (semi-)

preparative chromatography columns. The columns connected in parallel are filled with the MP in the direction of the flow, whereby the feed is introduced at a fixed point. A significant disadvantage of the process is the mandatory isocratic loading of the MP across all pillar positions. In the case of a heterogeneous feed consisting of components with different adsorption-desorption mechanisms, this leads to long processes, high cycle frequency and high consumption. To overcome this, in the first stage of the extraction process, the task of the MP is divided into position segments, which allow the gradient elution and ultimately a cleaning of difficult-to-elute components with equilibration of the columns for the subsequent cycle. The two components considered, BOA and BOB, are difficult to separate by liquid chromatography using the ethanol-water combination, and sometimes elute at the same time under the desired high feed concentrations. To circumvent this problem, the two eluents are first fractionated together.

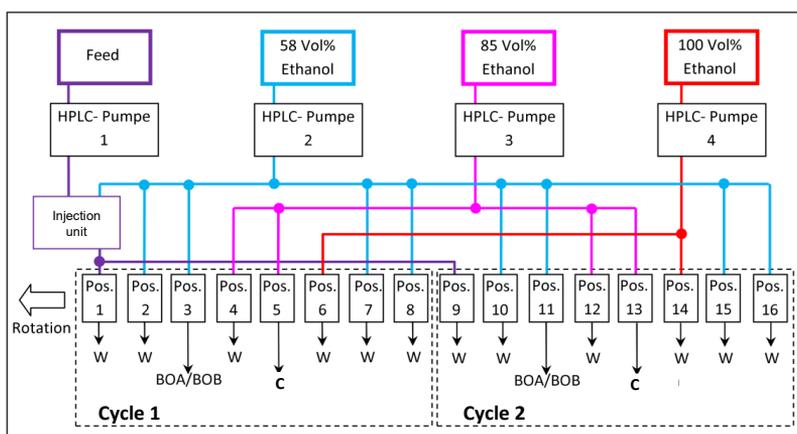


Figure 1: Fractionation system for BOA and BOB

Figure 1 shows the schematic interconnection of the system for the continuous fractionation of the bisabolol oxides as a binary mixture.

2.4. The simulated moving bed (SMB)

Counter current chromatography also offers the possibility of a continuous process management and is characterised by high efficiency compared to other alternatives, but it is limited to binary mixture. For the procedure developed here, both the SP and MP are moved in counter current to one another. The structure of the SMB used in the separation considers a four-zone configuration. A true moving bed (TMB) model (Charton & Nicoud, 1995) is used to define the parameters for the SMB. The residence time of the compounds in a chromatographic system results from the distribution equilibrium. The distribution coefficient is determined as the concentration quotient between the MP and the SP and is to be established via the retention time (t_R), taking into account the dead time (t_0) and the porosity (ϵ) of the packed column, based on the following equation:

$$K_i = \frac{t_R - t_0}{t_0 \frac{1 - \epsilon}{\epsilon}} \quad (1)$$

The selectivity between substances is defined as:

$$\alpha = \frac{t_{R(A)} - t_0}{t_{R(B)} - t_0} \quad (2)$$

The triangular theory developed by Ruthven & Ching (1989) for linear adsorption isotherms are used for the design of the simple SMB process. A system of equations is set up for the respective volume flows in each of the zones as follows:

$$\dot{V}_{MP(I)} = \dot{V}_{Elu} + \dot{V}_{MP(IV)} \quad (3)$$

$$\dot{V}_{MP(II)} = \dot{V}_{MP(I)} - \dot{V}_{Ex} \quad (4)$$

$$\dot{V}_{MP(III)} = \dot{V}_{MP(II)} + \dot{V}_F = \dot{V}_{MP(I)} - \dot{V}_{Ex} + \dot{V}_F \quad (5)$$

$$\dot{V}_{MP(IV)} = \dot{V}_{MP(III)} - \dot{V}_{Ra} = \dot{V}_{MP(I)} - \dot{V}_{Ex} + \dot{V}_F - \dot{V}_{Ra} \quad (6)$$

Where \dot{V} is the volume flow, in ml/min.

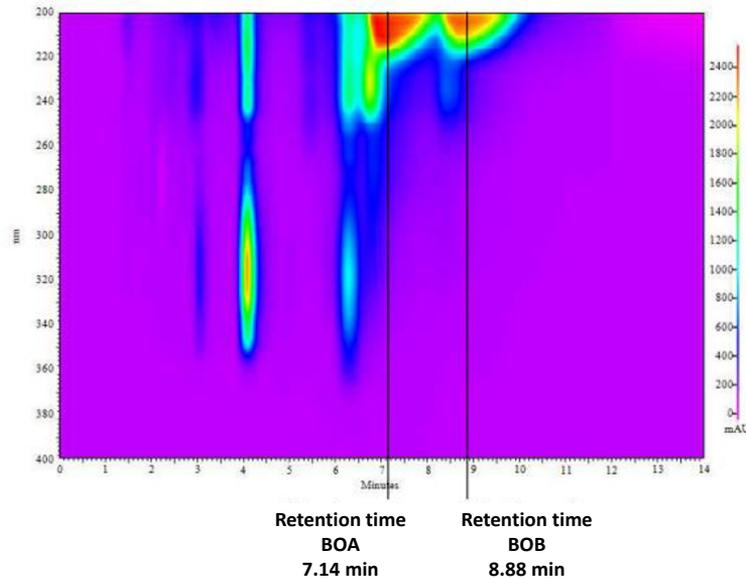


Figure 2: Retention times for the bisabolol oxides

The substance-specific retention times for the bisabolol oxides to be separated using the SMB method are determined in an optimized batch method. For this purpose, the most efficient separation of the components, with the lowest possible MP consumption is implemented. Due to the ethanol-water mixture used as MP, a relatively low maximum selectivity is set. Figure 2 shows the underlying chromatogram for the determination of the retention times as the basis for establishing the operating point for the TMB model. The following user-specified constraints are defined further:

- a) The permissible minimum flowrate ratio in Zone I must be increased by approximately 25% to ensure complete elution in the extract. Furthermore, if the flowrate increase, the eluent consumption increases as well

- b) The permissible maximum flowrate ratio in Zone IV must be reduced by approximately 25% in order to prevent carryover into Zone I. However, too great reduction is equivalent with an inefficient utilization of the SP
- c) For Zones I and II, an operating point within the range from Points 1 and 2 must be fixed ($m_{II} \leq m_{III}$)

For the operating conditions of the TMB model, the flowrate ratios in Zones I and IV were thus fixed, whereby the separation area for a practical operating point is spanned (triangle). From this, the flowrate for Zones II and III can be determined.

Table 2: Parameters for the SMB stage

Stationary phase (SP)	
Flowrate	0.1944 ml/min
Cycle rate	14.26 min
Mobile phase (MP) zone flow rates	
MP Zone I (Fixed)	1.500 ml/min
MP Zone II	1.084 ml/min
MP Zone III	1.162 ml/min
MP Zone IV	0.808 ml/min
Input and Output flows	
Eluent flow (Zone I – Zone IV)	0.692 ml/min
Feed flow (Zone III-Zone II)	0.078 ml/min
Extract flow (Zone I – Zone II)	0.416 ml/min
Raffinate flow (Zone III – Zone IV)	0.354 ml/min

The parameters determined for the SMB are summarised in Table 2. When assessing the test with a fixed operating point from this table, the concentration of the extract (BOB) at the extraction point is assessed as too low. The MP flowrate in Zone I should therefore be corrected downwards to 1.3 ml/min (instead of the value of 1.5 ml/min resulted from the

modelling) in order to cause a shift in the concentration profile of this component in the direction of Zones I and II (increase in the concentration at the extraction point). The eluent flowrate is consequently reduced to 0.492 ml/min, while the extract flow is only 0.216 ml/min.

3. Results and discussion

The rotating column stage is compared with a discontinuous batch application. According to the GC-FID analysis, BOA is detected to 91.3 wt. %, while BOB to 93.8 wt. % in their respective fractions. The losses are carried out via the respective column Positions 2 and 10 (BOA) and 4 and 12 (BOB) via the waste (W).

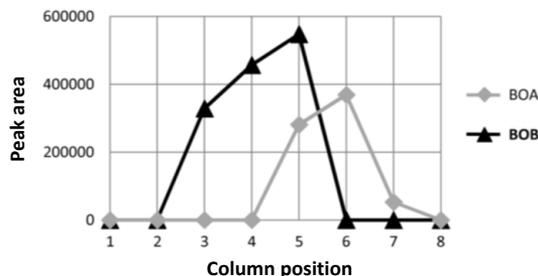


Figure 3: Profile concentration of the SMB application for the bisabolol oxides separation

The SMB-method for the subsequent separation of the binary fraction is developed based on linear adsorption isotherms. Figure 3 shows the concentration profile (GC-FID) after the iterative optimisation at steady state (Cycle 7). The subsequent SP overloading and the system approaching the efficiency limits took place by continuously increasing the feed concentration. The decrease in

the raffinate (BOA) purity due to increasing contamination by BOB is used for the evaluation of the process. Under an essential oil feed of 5µl/ml·cycle, a purity of 98.1% is achieved for BOA, and > 99.0% for BOB, according to the GC-FID analysis.

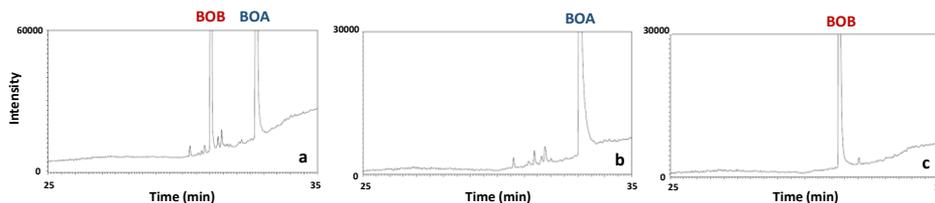


Figure 4: GC-FID of: a) Feed; b) Raffinate as BOA; c) Extract as BOB

The GC-FID for determining the fractions' purity also show that the impurities contained in the feed fraction are almost completely separated via the Raffinate, whereas the Extract almost exclusively contains the target component (Figure 4).

4. Conclusions

In the present work, a coupled method strategy for the isolation of valuable active ingredients is developed based on the example of an essential oil as a starting mixture. The approach integrates the rotating columns and the simulated moving bed principles. During the first stage (the rotating column stage), a binary mixture target (bisabolol A and B oxides) are fractionated at the same time. The oxides are then separated in the second stage (the SMB stage). The combination of the two processes shows superior results in terms of efficiency to the traditional selective methods for isolating different ingredients from multicomponent mixtures (e.g., batch chromatography), with the additional reduction in resources and costs.

The proposed procedure is intended to provide practical solutions for the continuous chromatographic extraction of ingredients from heterogeneous multicomponent mixtures. Through subsequent processing steps, possible co-elutions of valuable active ingredients can be achieved, which have to be further purified.

References

- M. Abbas et al., 2021, Optimization of ultrasound-assisted, microwave-assisted and Soxhlet extraction of bioactive compounds from *Lagenaria siceraria*: A comparative analysis, *Bioresource Technology Reports* 15, 100746
- S. Bagade et al., 2021, Recent advances in microwave assisted extraction of bioactive compounds from herbal samples: A review, *Critical Reviews in Analytical Chemistry* 51, pp. 138-149
- J. Bauer et al., 1997, *GIT Labor-Fachz.* 41, pp. 6-9
- T. Belwal et al., 2020, Optimization of extraction methodologies and purification technologies to recover phytonutrients from food, *Phytonutrients in Food*, pp. 217-235
- J. Brozio et al., 2004, A rigorous model for annular chromatography, *Chemical Engineering & Technology* 27 (9), pp. 962-970
- F. Charton & R.M. Nicoud, 1995, Complete design of a simulated moving bed, *Journal of Chromatography A* 702 (1-2), pp. 97-112
- S. Chouhan et al., 2017, Antimicrobial activity of some essential oils – Present status and future perspectives, *Medicines* 4 (3), 58
- C. Picot-Allain et al., 2012, Conventional versus green extraction techniques – a comparative perspective, *Current Opinion in Food Science* 40, pp. 144-156
- D.M. Ruthven & C.B. Ching, 1989, Counter-current and simulated counter-current adsorption separation processes, *Chemical Engineering Science* 44 (5), pp. 1011-1038
- H. Schilcher, 1987, *Die Kamille-Handbuch für Ärzte, Apotheker und andere Naturwissenschaftler*, Wissenschaftliche Verlagsgesellschaft, Stuttgart
- N.A. Yahya et al., 2018, An overview of cosmeceutically relevant plant extracts and strategies for extraction of plant-based bioactive compounds, *Food and Bioproducts Processing* 112, pp. 69-85

Digital design and optimization of an integrated reaction-extraction-crystallization-filtration continuous pharmaceutical process

Timothy J. S. Campbell^a, Chris D. Rielly^a, Brahim Benyahia^a

*a*Loughborough University, Department of Chemical Engineering, Epinal Way,
Loughborough, Leicestershire, LE11 1TU, UK
B.Benyahia@lboro.ac.uk

Abstract

This work investigates the digital design of a continuous pharmaceutical plant comprising a continuous three stage reaction, liquid-liquid extraction, multistage cooling and antisolvent crystallization, and wash-filtration. Firstly, the mathematical models were developed and validated in conjunction with the available experimental data obtained from the literature and research partners. The resulting digital twin was used for steady state optimization to deliver optimal options for plant design and operation, including process capacities and number of crystallization stages. After the identification of the optimal design and optimal steady state operation, the digital twin was used to perform uncertainty propagation and global sensitivity analysis to identify the Critical Process Parameters (CPP) and Critical Material Attributes (CMA) and deliver robust and cost-effective methods for a systematic implementation of Quality-by-Design (QbD). This approach is aimed at demonstrating that the plant can be operated within the robust quality bounds which provide a built-in quality assurance for the final product. Several Critical Quality Attributes (CQA) which impact drug safety and efficacy were considered which includes the average crystal size, crystal size distribution, coefficient of variation and product purity were considered as the CQA.

Keywords: Integrated continuous pharmaceutical plant, Continuous multistage crystallization, Global Sensitivity Analysis, Uncertainty Analysis, Digital Twin, Quality-by-Design.

1. Introduction

Continuous manufacturing has been identified as a pivotal technology in overcoming the inherent limitations of current batch manufacturing methods used within the pharmaceutical sphere. New research in process analytical technologies and real-time control strategies have made continuous pharmaceutical manufacturing (CPM) feasible for wider use alongside its capacity to surpass quality compliance of former batch processing methods (Mascia et al., 2013). Despite the progress in research, technical hurdles remain before the wider adoption of CPM. One such hurdle is developing systematic and economic methods to design, optimize and operate integrated continuous pharmaceutical processes (ICPP) due to the complexity of integrating the synthesis and purification steps to produce the active pharmaceutical ingredient (API), formulations and final dosage forms (Benyahia, 2018). The advent of digital transformation opened new opportunities for computer aided optimal design, operation and decision making using digital twins (Ramin et al., 2018). However, successful use of digital twins requires the development of more robust mathematical models and systematic methodologies to understand and quantify the impact of the process and design parameters, material attributes and technology selection on drug product safety and efficacy. Besides the traditional objectives above, an effective digital twin should allow reliable identification of the critical process parameters (CPP) and critical material attributes (CMA) which influence the critical quality attributes (CQA) of the drug product. The last decade has seen a wider adoption of a range of process simulators in the pharmaceutical sector, including high fidelity models, surrogate models, flow sheet models, and plant-wide dynamic models (Benyahia et al., 2012; Maloney et al. 2020). There is clear evidence of the benefits of these process models and digital twins in the development and optimal operation and control of the world's first ICPP

and the kilogram scale continuous upstream plant for Prexasertib Monolactate Monohydrate (Mascia et al., 2013; Cole et al., 2017; Lakerveld et al., 2015). More recently, dynamic models of ICPPs were used for the design of a carfilzomib process (Maloney et al. 2020) and optimize steady state and dynamic performance such as start-up and shut down (Liu and Benyahia, 2021). The impact of recycles have not yet been investigated along with the impact of uncertainties that these present. A significant challenge in the implementation of Quality-by-Design (QbD) is the identification of the CPP and CMA which are commonly addressed by combining human expertise and expensive designs of experiments (e.g., Factorial DoE). The development of predictable and reliable mathematical models or digital twins may provide a systematic tool to understand the influence of the different factors and design parameters on the CQA's of the final product. In this paper, a dynamic simulator of an integrated upstream process comprising reactions, liquid liquid extraction, crystallization and wash-filtration was developed and used to optimally design and operate an ICPP. The digital twin was then used to perform uncertainty propagation and global sensitivity analysis (Fysikopoulos et al., 2019) to help identify more reliably the CPPs and CMAs which are essential in the subsequent implementation of QbD. This approach is aimed at demonstrating that the plant can operate with a built-in quality and safety assurance for the final product. The average crystal size, crystal size distribution, coefficient of variation and purity were considered as the CQA's, and all simulations were performed using gPROMS formulated products 2.0.1 which specializes in process simulation and optimization.

2. Methodology

2.1. Process description

The digital twin of the upstream continuous pharmaceutical process for the synthesis and purification of ibuprofen was created in gPROMS 2.0.1 and consisted of a sequence of three tubular reactors, a continuous liquid-liquid extractor, a three-stage continuous crystallization and finally a filtration step (Figure 1). The parameter estimation of the mathematical model of the reaction step was based on a continuous flow synthesis performed by Bogdan (Bogdan, 2009) which consists of three plug flow reactors where a sequence of three reactions takes place: a Freidel-Crafts acylation, a mediated aryl migration and a saponification reaction which finally produces ibuprofen as a potassium salt. The reaction is followed by a liquid liquid extraction (LLE) used as a preliminary purification step followed by a three-stage continuous crystallization process. The solubility and kinetic data of growth and nucleation for the antisolvent crystallization of ibuprofen were obtained from the literature (Afrose. 2017). In addition, a combined wash and wash-filtration step was considered to separate the crystals from the mother liquor and reduce the impurities and residual solvents in the final product. The intensified plant was designed for a productivity of 6.58 g/hr at an assumed efficiency of 70%.

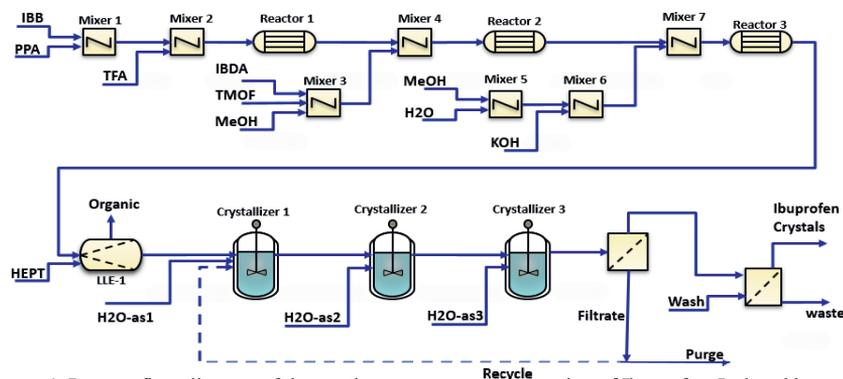


Figure 1. Process flow diagram of the continuous upstream processing of Ibuprofen. Isobutyl benzene (IBB), Propanoic acid (PPA), triflic acid (TFA), Iodobenzene diacetate (IBDA), Trimethyl orthoformate (TMOF), Methanol (MeOH), Potassium Hydroxide (KOH), Heptane (HEPT).

2.2. Optimization

The objective is to optimize the design features and operating conditions to maximize productivity of the final product at steady state, here the crystals obtained after wash-filtration, under a set of quality and efficiency constraints (e.g., Mean Crystal Size, purity, yields) and operation constraints (e.g., Maximum process capacity, maximum flow rates etc.). The mathematical formulation of the optimization problem is described below.

$$\begin{aligned}
 & \underset{F_{as,i}, T_i, V_i, \tau_i, F_{org}, V_{LLE}, P_{filt}}{\text{Max}} \quad \text{Pr} \\
 \text{s.t.} \quad & \dot{x} = f(x, y, u, p, t) \\
 & 0 = g(x, y, u, p, t) \\
 & \bar{d} \geq 200 \mu\text{m} \\
 & w_{T,TA} \leq 0.01 \text{ kg/kg} \\
 & w_{T,KOH} \leq 0.01 \text{ kg/kg} \\
 & 0 \leq F_{AS,i} \leq 200 \text{ mL h}^{-1} \\
 & 283 \text{ K} \leq T_i \leq 313 \text{ K} \\
 & 20 \text{ min} \leq \tau_i \leq 100 \text{ min} \\
 & 0 \text{ mL h}^{-1} \leq F_{org} \leq 100 \text{ mL h}^{-1} \\
 & 0.5 \text{ Bar} \leq P_{filt} \leq 2.0 \text{ Bar} \\
 & S_i \geq 1 \\
 & i = 1, 2, 3
 \end{aligned}$$

Where \bar{d} is the average crystal size (D [4,3]), $w_{T,TA}$ and $w_{T,KOH}$ are the residual mass fraction of triflic acid and potassium hydroxide impurities in the product, left after the wash-filtration stage. $F_{AS,i}$ is the flowrate of the anti-solvent fed to each mixed suspension mixed product removal crystallizer (MSMPR), T_i, τ_i, S_i are respectively the temperature, the residence time and the relative supersaturation in each MSMPR, F_{org} is the flowrate of the organic solvent to the LLE stage, and P_{filt} is the pressure of the filter. The optimization problems were solved using a Control Vector Parameterization (CVP) technique which converts the dynamic optimization problems into nonlinear programming problems solved using an efficient SQP (Successive Quadratic Programming) solver built into gPROMS.

2.3. Global Sensitivity Analysis

Upon completion of the optimization, the optimal results were used as nominal values to develop global system analysis which included a variance based global sensitivity analysis, using a quasi-random (Sobol) sampling technique. The objective is to identify the most influential CPPs and CMAs. The CQAs, used here as the outputs or responses were the mean crystal size, crystal purity or total mass fraction of the impurities and productivity. Process and material parameters include the reactant flowrates, the operating temperatures of the MSMPRs, the antisolvent streams for the MSMPRs, the residence times of the MSMPRs, and the wash flowrate.

3. Results and discussion

3.1. Optimization

The optimization results correspond to the set of optimal decision variables including the operating conditions and design options along with their upper and lower bounds and are summarised in table 1. Two scenarios are considered, the base case with no recycle and the second with a 50% recycle after the filtration stage. Table 1 shows that the organic solvent flowrate remains the same because the LLE is prior to the crystallization where the recycle stream is being fed. However, the steady state crystallization temperatures dropped at each crystallizer to help address the dilution effect inherent to the recycle stream which contains significant amounts of solvent and antisolvent. The optimal residence times associated with the scenario at 50% recycle has also significantly increased as shown in table 1 which suggests that larger capacities are required to achieve effective crystallization under the recycle conditions. The key performance indicators of each case are summarized in table 2 including the attainable maximum production mass flow rates in both scenarios, considered here as the

objective functions, as well as the CQAs and key performance indicators. Overall, the scenario considered in presence of 50% recycle shows improved productivity and yields. Both cases met the required impurity contents however the recycle scenario exhibits as expected a slightly higher impurity content due the accumulation of the impurities in the system. To better control the level of impurities our future work will consider the recycle ratio as a decision variable. The dynamic simulations of the integrated process using the optimal steady state conditions are presented in Figure 2.

Table 1. The optimal values of the input factors from the optimization simulation for both non-recycle and the 50% recycle case.

Variables	Units	Optimal values		Lower Bound	Upper Bound
		Base case	50% Recycle		
Organic Solvent flowrate	mL/h	95.96	95.96	0	400
MSMPR 1 temperature	K	309	305.330	283	313
MSMPR 2 Temperature	K	304	309.088	283	313
MSMPR 3 Temperature	K	305	283.615	283	313
Anti-solvent flowrate MSMPR 1	mL/h	170.4	188.0	0	100
Anti-solvent flowrate MSMPR 2	mL/h	11.2	5.0	0	150
Anti-solvent flowrate MSMPR 2	mL/h	3.6	6.0	0	200
Mean residence time MSMPR 1	Min	16.0	76.4	0	100
Mean residence time MSMPR 2	Min	35.5	120	0	120
Mean residence time MSMPR 3	Min	19.7	120	0	120
Pressure filter flowrate set point	kg/hr	0.0078	0.0088	0	1
Pressure drop filter	bar	1.6	1.7	0.5	2.0

Table 2. CQA and key performance indicators for the base case and the recycle scenario.

	Units	Base case	50% Recycle
Mass flowrate	g/hr	7.41	8.4
Overall Process yield	%	65%	73%
Impurity content- Triflic acid	kg/kg	0.0025	0.0032
Impurity content- KOH	kg/kg	0.0036	0.0046

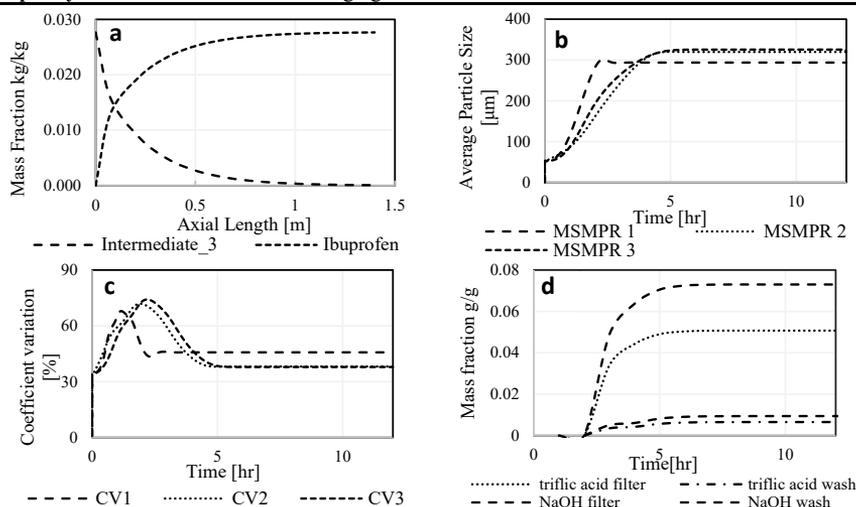


Figure 2. (a) Mass fraction of Ibuprofen (API) in the final PFR showing the conversion of intermediate 3 to Ibuprofen. (b) Dynamic profile of the average particle size in each MSMPR. (c) Coefficient of variation in each of the MSMPRs. (d) Mass fraction of impurities in the final crystal product after filtration then after the wash-filtration.

Table 3. The global sensitivity analysis total order factor sensitivities table of the CQA's and CPP's.

		Mass Flowrate Crystal	Average Particle size D43	Coefficient of variation	Impurity Triflic acid KOH	
Reactant Mass flowrate	Prop acid	0.007	0.008	0.008	0.000	0.000
	IBB	0.007	0.009	0.009	0.000	0.000
	IBD	0.007	0.008	0.007	0.001	0.000
	KOH	0.027	0.022	0.021	0.006	0.913
	TMOF	0.003	0.003	0.003	0.001	0.000
RDT	MSMPR_1	0.000	0.000	0.000	0.000	0.000
	MSMPR_2	0.000	0.000	0.000	0.000	0.000
	MSMPR_3	0.000	0.000	0.000	0.000	0.000
Temp	MSMPR_1	0.000	0.047	0.047	0.000	0.000
	MSMPR_2	0.010	0.121	0.124	0.001	0.000
	MSMPR_3	0.199	0.050	0.051	0.008	0.001
Anti-Solvent	MSMPR_1	0.647	0.573	0.571	1.112	0.109
	MSMPR_2	0.000	0.000	0.000	0.001	0.000
	MSMPR_3	0.000	0.000	0.000	0.001	0.000
Flowrate Wash		0.000	0.000	0.000	0.000	0.000

At the optimal production, the final production rate of the reaction step is 9.39 g/hr. After the first purification step which occurs in the LLE, a three-stage crystallization process is used to purify the API and recover ibuprofen as a crystalline form. The manipulation of temperature and anti-solvent flowrate help achieve the required supersaturation levels across the three MSMPRs which results in increased mean crystal size from one stage to the other.

As shown in Figure 2b, the final mean crystal size obtained in the last MSMPR stage is 325 μ m. The coefficient of variation has a sharp increase in all MSMPRs due to the higher initial rates of nucleation. When steady state is reached, all coefficients of variation stabilize. The slurry leaving the final crystallization stage, goes through a filtration then a wash-filtration step to reduce the residual solvent and impurities content in the final crystal product. The wash stage reduces the mass fraction of the main impurities, triflic acid and potassium hydroxide, to below 0.005g/g as indicated in Figure 2d. The dynamic simulations of the integrated process using the optimal steady state conditions are presented in Figure 2. At the optimal production, the final production rate of the reaction step is 9.39g/hr. After the first purification step which occurs in the LLE, a three-stage crystallization process is used to purify the API and recover ibuprofen as a crystalline form. The manipulation of temperature and anti-solvent flowrate help achieve the required supersaturation levels across the three MSMPRs which results in increased mean crystal size from one stage to the other. As shown in Figure 2b, the final mean crystal size obtained in the last MSMPR stage is 325 μ m. The coefficient of variation has a sharp increase in all MSMPRs due to the higher initial rates of nucleation. When steady state is reached, all coefficients of variation stabilize. The slurry leaving the final crystallization stage, goes through a filtration then a wash-filtration step to reduce the residual solvent and impurities content in the final crystal product. The wash stage reduces the mass fraction of the main impurities, triflic acid and potassium hydroxide, to below 0.005g/g as indicated in Figure 2d.

3.2. Global Sensitivity Analysis

The results of the global sensitivity analysis are used to identify the hierarchy of CPPs for the integrated process. This is a critical step in the implementation of QbD for process operation. The total order sensitivities factors can be seen in table 3. The most influential three parameters of each CQA are highlighted in grey with the most influential being the darkest. The crystal quality was most influenced by the antisolvent flowrate into MSMPR 1, temperature of MSMPR 3 and the mass flowrate of the reactant potassium hydroxide. The Anti-solvent flowrate into MSMPR 1 had a dominating effect on the solubility curve in the following crystallizers. Of the three antisolvent streams it is over 10 times the amount of the other factors thus the comparative impact is respective of this. Regarding the triflic acid impurity in the

residue on the crystals, the antisolvent stream again had the highest impact due to its diluting effect reducing the overall amount within the residue of the intermediate from the filtration stage. With respect to the KOH impurity, the reactant mass flowrate is the most influential factor as unreacted KOH will remain in the liquid phase. More unreacted KOH will result in a higher proportion left in the residue. Further investigation is required on the wash phase as this has zero effect on any of the impurity factors. This could be due to the low comparative impact to the other CPPs or other unknown factors.

4. Conclusions

A dynamic mathematical model of a continuous pharmaceutical plant comprising a continuous three stage reaction, liquid-liquid extraction, multistage cooling and antisolvent crystallization, and wash-filtration was developed using data from the literature. This resulting process simulator used a digital twin for steady state optimization to deliver optimal options for plant design and operation and maximize productivity. To enhance cost-efficiency and environmental performance of the plant, an additional optimization scenario was conducted under 50% recycle of the mother liquor obtained from the filtration stage without further purification. The results showed that the scenario associated with 50% recycle increases the startup time and impurity content, which in turn required reoptimized crystallization and wash-filtration conditions. However, the 50% recycle was also shown to improve the overall productivity of the process. The GSA identified that the antisolvent flowrate into MSMPR 1, the temperature of MSMPR 3 and the mass flowrate of the reactant KOH had the most impact on the CPP's. The future work will be focused on the optimization of recycle to purge ratio and the incorporation of downstream processes in the design and optimization.

This work was conducted as part of the EPSRC CMAC HUB (EPSRC EP/P006965/1). The authors would like to thank PSE Enterprise for their support and for providing the licenses for gPROMs Formulated Products.

References

- Afrose, A., 2017. Crystallisation techniques for developing a novel dry powder inhaler formulation for ibuprofen, s.l.: Queensland University of Technology.
- Benyahia, B. 2018. Applications of a plant-wide dynamic model of an integrated continuous pharmaceutical plant: Design of the recycle in the case of multiple impurities. In *Computer Aided Chemical Engineering* (1st ed., Vol. 41). Elsevier B.V.
- Benyahia, B., Lakerveld, R., & Barton, P. I. 2012. A plant-wide dynamic model of a continuous pharmaceutical process. *Industrial and Engineering Chemistry Research*, 51(47), 15393-15412.
- Cole et al., 2017. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow CGMP conditions. *Science*, 356(6343), 1144-1150.
- Fysikopoulos, D., Benyahia, B., Borsos, A., Nagy, Z. K., & Rielly, C. D. 2019. A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers and Chemical Engineering*, 122, 275-292.
- Liu, J., & Benyahia, B. 2021. Systematic model-based dynamic optimization of a combined cooling and antisolvent multistage continuous crystallization process. *Computer Aided Chemical Engineering*, 50, 1221-1227.
- Maloney, A. J., İçten, E., Capellades, G., Beaver, M. G., Zhu, X., Graham, L. R., Brown, D. B., Griffin, D.J., Sangodkar, R., Allian, A., Huggins, S., Hart, R., Rolandi, P., Walker, S. D., Braatz, R. D. 2020. A Virtual Plant for Integrated Continuous Manufacturing of a Carfilzomib Drug Substance Intermediate, Part 1-310, *Org. Process Res. Dev.*, Vol. 24, pp. 1861-1908.
- Mascia, S., Heider, P. L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P. I., Braatz, R. D., Cooney, C. L., Evans, J. M. B., Jamison, T. F., Jensen, K. F., Myerson, A. S., & Trout, B. L. 2013. End-to-end continuous manufacturing of pharmaceuticals: Integrated synthesis, purification, and final dosage formation. *Angewandte Chemie - International Edition*, 52(47), 12359-12363.
- Ramin, P., Mansouri, S.S., Udugama, IA, Benyahia, B., Gernaey K.V. 2018. Modelling continuous pharmaceutical and bio-based processes at plant-wide level: A roadmap towards efficient decision-making. *Chemistry Today* 36 (2), 26-30.

Multi-objective optimization of natural gas liquefaction process simulation via kriging surrogate model

Lucas F. Santos,^{a,b} Caliane B. B. Costa,^a Jose A. Caballero,^b Mauro A. S. S. Ravagnani^a

^a*Department of Chemical Engineering, State University of Maringá, Av. Colombo 5790, Maringá - 87020900, Brazil*

^b*Institute of Chemical Process Engineering, University of Alicante, Ap. Correos 99, Alicante – 03080, Spain*
lfs.francisco.95@gmail.com, pg54347@uem.br

Abstract

A multi-objective optimization framework is proposed to design single-mixed refrigerant natural gas liquefaction processes considering the conflicting goals of minimizing both power consumption and heat transfer area utilization. In the present approach, the black-box process simulation is embedded into a nonlinear programming (NLP) problem via kriging surrogate model. To deal with the conflicting objectives, the ε -constraint methodology is applied. The surrogate NLP problems with fixed ε are solved in GAMS using CONOPT to determine a non-dominated solution candidate of the original multi-objective problem. The Pareto Front achieved with the present framework dominates recent single-objective result from the literature. The non-dominated solutions have power consumption that ranges from 0.2800 to 0.4060 kW/(kg NG) and heat transfer area utilization multiplied by overall heat transfer coefficient varying from 0.0699 to 0.1852 kW/°C. A trade-off solution can be achieved by increasing 5.29 % of power consumption to save 31.8 % of heat transfer area.

Keywords: Multi-objective optimization, Kriging surrogate model, Natural gas liquefaction, Simulation optimization, Process simulation.

1. Introduction

The natural gas liquefaction process design presents a clear trade-off between energy consumption and equipment size (Khan et al., 2016). The concern of achieving high energy efficiency and diminished equipment size is further emphasized in offshore processes, where the plant site size is significantly restricted. These liquefaction processes consist of cryogenic refrigeration cycles to cool down the natural gas to about -160 °C to liquefy, store, transport, and commercialize it safely as liquefied natural gas (LNG). Optimization has been successfully used to determine refrigerant composition and thermodynamic cycle conditions in natural gas liquefaction processes to improve power consumption, exergy efficiency, or total annual cost (Austbø et al., 2014). Despite the vast literature on single-objective optimization, the evaluation of the natural gas liquefaction processes optimal trade-offs has been timidly addressed in the literature. Khan et al. (2016) investigated the optimization of the dual-mixed refrigerant LNG process under cold and warm ambient conditions. The authors used NSGA-II algorithm to trade-off between the minimization of specific compression energy and the area of heat exchangers in terms of overall heat transfer coefficient times the area (UA). Ghorbani et

The final pre-refrigeration temperature is considered to be $-24.0\text{ }^{\circ}\text{C}$. The pressure drop in the multi-stream heat exchangers (MSHEs) is 100 kPa for hot streams, 10 kPa for cold streams, and 50 kPa for water inter-stage coolers. The intermediate pressures are determined to guarantee a constant compression ratio in all four compressors, and the intermediate cooling temperature is $30\text{ }^{\circ}\text{C}$. The temperature of hot streams leaving the MSHEs is considered to be the same, $-24.0\text{ }^{\circ}\text{C}$ for the first and $-157.2\text{ }^{\circ}\text{C}$ for the second to achieve the desired LNG condition of $-160.0\text{ }^{\circ}\text{C}$, 110 kPa, and 3.7 % of vapor fraction. The process design problem is formulated as a multi-objective optimization model, considering the minimization of specific power consumption and UA. The process constraint is that a minimum temperature approach of $3\text{ }^{\circ}\text{C}$ must be assured throughout the MSHEs. To overcome the issue of this constraint not being smooth for kriging modelling (Santos, 2021), Santos et al. (2022) proposed to discretize the MSHEs 1 and 2 in K sections. The optimal process design can be formulated as the following multi-objective optimization problem

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{D}} \quad & \mathbf{f}(\mathbf{x}) = \left[\sum_{p \in \text{PM}} \frac{W_p(\mathbf{x})}{\dot{m}_{NG}}; \sum_{ex \in \text{HE}} UA_{ex}(\mathbf{x}) \right] \\ \text{s. t.} \quad & \mathbf{g}_{\kappa}(\mathbf{x}) = 1 - \frac{\min_{k \in \Omega_{\kappa}} \{Th_{\kappa,k}(\mathbf{x}) - Tc_{\kappa,k}(\mathbf{x})\}}{3} \leq 0, \kappa = 1, \dots, (2K) \end{aligned} \quad (1)$$

In Eq. (1) $W_p(\mathbf{x})$ is the work consumption of the pressure manipulator unit p in the set of compressors and pumps PM , and $UA_{ex}(\mathbf{x})$ is the global heat transfer coefficient multiplied by the area of the heat exchanger ex in the set of all heat exchangers HE . $Th_{\kappa,k}(\mathbf{x})$ and $Tc_{\kappa,k}(\mathbf{x})$ are the temperature of hot and cold composite curves in the κ section of the MSHEs, and Ω_{κ} is the set of the k points from composite curves calculations that belongs to section κ (Santos et al., 2022). $\mathcal{D} = [\mathbf{x}^{lb}, \mathbf{x}^{ub}]$ is a box constraint for the decision variables bounded by \mathbf{x}^{lb} and \mathbf{x}^{ub} as proposed in Qyyum et al. (2020), and \dot{m}_{NG} is the mass flow rate of the natural gas stream.

3. Kriging-based multi-objective optimization

To solve the multi-objective simulation optimization problems as in Eq. (1), the proposed framework is presented in Algorithm 1. The computer code to compute \mathbf{f} and \mathbf{g} at given $\mathbf{x} \in \mathcal{D}$ is a function written in MATLAB connected to the process simulator Aspen HYSYS. First, m_0 samples of \mathbf{X} is generated by a Latin Hypercube algorithm to maximize the minimum distance between points in the search space \mathcal{D} . The value of \mathbf{f} and \mathbf{g} are calculated in the simulation for each $\mathbf{x} \in \mathbf{X}$. The initial data is defined as $D_0 = [\mathbf{X} \ \mathbf{Y}]$. Given D_0 , m_f , and D , the single-objective optimization for each \mathbf{f}_i objective is solved using the framework for constrained black-box optimization proposed in Santos et al. (2022). In the case of Eq. (1), the objective function \mathbf{f}_1 is the specific power consumption and \mathbf{f}_2 is the overall UA. The above-mentioned approach consists of using the data in D_0 to fit kriging models for the objectives and constraints. These models are implemented in an NLP problem and solved in GAMS with multi-start local optimization with CONOPT local solver. The solution, which is promising a minimizer candidate, is evaluated in the simulation and the values are appended to the data. This process is repeated until convergence or simulation evaluation budget m_f . The values of \mathbf{f}_1 at each single-objective solution, $\mathbf{f}_1(\mathbf{x}_1^*)$ to $\mathbf{f}_1(\mathbf{x}_2^*)$, bound the ε vector. The ε -constraint vector is defined as the equally distributed n_{ε} -vector from $\mathbf{f}_1(\mathbf{x}_1^*)$ to $\mathbf{f}_1(\mathbf{x}_2^*)$. The single-objective solutions are appended to the data, D_2 . For the other entries of ε , the ε -constrained single-objective optimization is solved using the same framework for constrained black-box

optimization proposed in Santos et al. (2022). Each optimization solution is appended to the data so that $D_{ite} = D_{ite-1} \cup [\mathbf{x}_{ite}^*, \mathbf{f}(\mathbf{x}_{ite}^*), \mathbf{g}(\mathbf{x}_{ite}^*)]$. The last step of the proposed framework is to eliminate dominated and infeasible solutions from D_{n_ε} to form the set of Pareto solution, P , found by the algorithm.

Algorithm 1: Kriging-based multi-objective optimization framework

Input: Initial sample size $m_0 \in \mathbb{N}_+$, maximum number of samples $m_f \in \mathbb{N}_+$, box-constrained design space $\mathcal{D} \in \mathbb{R}^{2 \times n}$, the computer code to compute \mathbf{f} and \mathbf{g} at given $\mathbf{x} \in \mathcal{D} \subseteq \mathbb{R}^n$, and the number of ε -constrait values $n_\varepsilon \in \mathbb{N}_+$

- 1) Generate m_0 samples \mathbf{X} with respective simulation output values $\mathbf{Y} = [\mathbf{f}; \mathbf{g}]^\top$ and define $D_0 = [\mathbf{X} \ \mathbf{Y}]$;
- 2) Solve the single-objective optimization problem for each \mathbf{f}_i using the framework proposed in Santos et al. (2022) to determine \mathbf{x}_1^* and \mathbf{x}_2^* , with D_0, m_f, \mathcal{D} ;
- 3) Define ε as the equally spaced n_ε -vector from $\mathbf{f}_1(\mathbf{x}_1^*)$ to $\mathbf{f}_1(\mathbf{x}_2^*)$;
- 4) Append to the data: $D_2 = D_0 \cup [\mathbf{x}_1^*, \mathbf{f}(\mathbf{x}_1^*), \mathbf{g}(\mathbf{x}_1^*)] \cup [\mathbf{x}_2^*, \mathbf{f}(\mathbf{x}_2^*), \mathbf{g}(\mathbf{x}_2^*)]$;

for $ite = 3, \dots, n_\varepsilon$ **do**

- 5) Solve the ε -constrained single-objective optimization

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{D}} \quad & \mathbf{f}_2(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{f}_1(\mathbf{x}) \leq \varepsilon_{ite-1} \\ & \mathbf{g}(\mathbf{x}) \leq 0, \end{aligned}$$

using the framework proposed in Santos et al. (2022) to determine \mathbf{x}_{ite}^* , with $D_{ite-1}, m_f, \mathcal{D}$;

- 6) Append the solution to the data: $D_{ite} = D_{ite-1} \cup [\mathbf{x}_{ite}^*, \mathbf{f}(\mathbf{x}_{ite}^*), \mathbf{g}(\mathbf{x}_{ite}^*)]$;

end

- 7) Eliminate dominated and infeasible solutions from D_{n_ε} to form P , the set of Pareto solutions;

Output: Set of non-dominated solutions, P .

4. Results

The proposed multi-objective, kriging-based optimization framework is applied to the design of SMR natural gas liquefaction processes considering power consumption and heat exchanger area utilization objectives. The considered parameters of the optimization approach are initial sample size $m_0 = 10n$, function evaluation budget $m_f = 20n$, number of decision variables $n = 8$, number of sections into which each MSHE is divided $K = 10$, and number of non-dominated solution $n_\varepsilon = 18$.

Figure 2 presents the Pareto Front determined by the non-dominated solutions of the present approach. It also includes the best result from the literature to this liquefaction process with the given specifications (Qyyum et al., 2020). This figure shows that the present approach was able to determine an energy-optimal solution that is better than the literature, with energy saving of 1.51 %, from 0.2843 to 0.2800 kW/kg of natural gas. Also, given the present Pareto Front, the literature solution is a dominated one. It means that, for the same specific power consumption, the present approach would design a liquefaction process with reduced UA. The proposed Pareto Front ranges from 0.2800 to 0.4046 kW/(kg NG) for power consumption and 0.0699 to 0.1852 kW/°C for UA. The shape of this Pareto curve shows that the increase in heat exchanger area is more pronounced as the power consumption approaches low values. Based on this insight, one can choose non-dominated solutions that better balances high energy efficiency and heat exchanger area utilization as the solutions that presents the smallest ℓ^2 and ℓ^1 norm. These solutions are highlighted in Figure 2 in the blue square and magenta hexagram, respectively. The best trade-off based on the ℓ^2 norm presents an increase in power consumption of 5.29 % and 31.80 % UA decrease compared to the energetically-optimum solution. The ℓ^1 norm-based solution presents an increase in power consumption of 10.61 % and 44.65 % UA decrease compared to the energy-optimum solution.

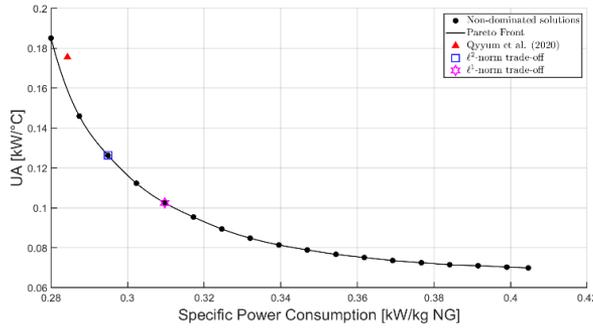


Figure 2: Pareto Front and comparison with Qyyum et al. (2020).

Table 1 presents the values of the decision variables and objective functions for the base case (Qyyum et al., 2020), energy-optimum, UA-optimum, ℓ^2 -norm trade-off, and ℓ^1 -norm trade-off solutions. The energy-optimum result presents a refrigerant composition that is reduced in propane and i-pentane in comparison with the literature result. The UA-optimum result shows that for reducing the heat exchangers area utilization high compression ratio and refrigerant with more contents of light-components are preferred. Differently from what one would expect, the trade-off solutions are complex and unique, i.e., not mere interpolations of the single-objective ones. Therefore, the use of multi-objective optimization approach is justified as a means to derive complex solutions that trades-off optimally between conflicting objectives to the process design.

The computational time is dominated by simulation evaluation, which takes around 4.7 s. The mean elapsed time for the multi-objective optimization algorithm was 143 min. The advantage of the present approach over the well-established population-based metaheuristics is the reduced budget of time-consuming simulation calculation, in the present case $10n \times n_\varepsilon = 1440$. The circumstances that the present approach would no longer be efficient is when the surrogate models do not capture the true functions behavior. That is often the case for high-dimensional problems or stiff functions.

Table 1: Multi-objective optimization results for SMR natural gas liquefaction processes

Optimization Results	Qyyum et al. (2020)	Energy-optimum	UA-optimum	ℓ^2 -norm trade-off	ℓ^1 -norm trade-off
m_N [kg/h]	0.2210	0.2426	0.2620	0.1871	0.1500
m_{C1} [kg/h]	0.4525	0.4344	0.5500	0.4132	0.4244
m_{C2} [kg/h]	0.9420	0.9504	0.9976	0.8878	0.8090
m_{C3} [kg/h]	0.9600	0.8000	0.8322	0.8001	0.8904
m_{iC4} [kg/h]	0.6525	0.6450	0.8000	0.6235	0.5775
m_{iC5} [kg/h]	0.8250	0.7108	0.6585	0.6838	0.6871
P_{suc} [kPa]	237.0	263.0	150.0	205.8	172.1
P_{dis} [kPa]	4080	5047	5500	5486	5500
Net work consumption [kW/(kg NG)]	0.2840	0.2800	0.4064	0.2948	0.3097
UA [kW/°C]	0.1759	0.1852	0.0699	0.1263	0.1025

5. Conclusions

This paper presented a multi-objective optimization framework to the SMR natural gas liquefaction process design problem, considering the trade-off between energy efficiency and heat exchanger area utilization. Kriging surrogate model, ϵ -constraint methodology, and gradient-based solver in GAMS are used in the proposed approach to determine candidates of non-dominated solutions of the original black-box optimization problem. The Pareto Front determined by the present approach shows that the recent literature result of energy-efficient SMR process is a dominated solution. The objective functions of non-dominated solutions range from 0.2800 and 0.4046 kW/(kg NG) for power consumption and 0.0699 and 0.1852 kW/°C for UA. Two trade-off solutions were analyzed and the results showed that 5.29 % and 10.61 % increase in the power consumption can lead to 31.80 % and 44.65 % heat exchanger area decrease, respectively. The trade-off solutions are complex process configurations instead of interpolations of the single-objective optima, which justifies the use of multi-objective optimization approach to tackle competing objectives.

Acknowledgments

The authors LFS, CBBC, and MASSR acknowledge the National Council for Scientific and Technological Development – CNPq (Brazil), processes 148184/2019-7, 440047/2019-6, 311807/2018-6, 428650/2018-0, and Coordination for the Improvement of Higher Education Personnel – CAPES (Brazil) for the financial support. The author JAC acknowledge financial support from the “Generalitat Valenciana” under project PROMETEO 2020/064.

References

- B. Austbø, S. W. Løvseth, T. Gundersen, 2014, Annotated bibliography—Use of optimization in LNG process design and operation, *Computers & Chemical Engineering*, 71, 12, 391–414.
- B. Ghorbani, M.-H. Hamed, R. Shirmohammadi, M. Hamed, M. Mehrpooya, 2016, Exergoeconomic analysis and multi-objective Pareto optimization of the C3MR liquefaction process, *Sustainable Energy Technologies and Assessments*, 17, 56–67.
- M. S. Khan, I. Karimi, M. Lee, 2016, Evolution and optimization of the dual mixed refrigerant process of natural gas liquefaction, *Applied Thermal Engineering*, 96, 320–329.
- H. Mofid, H. Jazayeri-Rad, M. Shahbazian, A. Fetanat, 2019, Enhancing the performance of a parallel nitrogen expansion liquefaction process (NELP) using the multi-objective particle swarm optimization (MOPSO) algorithm, *Energy*, 172, 286–303.
- T. V. Nguyen, E. D. Rothuizen, W. B. Markussen, B. Elmegaard, 2018, Thermodynamic comparison of three small-scale gas liquefaction systems, *Applied Thermal Engineering*, 128, 712–724.
- M. A. Qyyum, T. He, K. Qadeer, N. Mao, S. Lee, M. Lee, 2020, Dual-effect single-mixed refrigeration cycle: An innovative alternative process for energy-efficient and cost-effective natural gas liquefaction, *Applied Energy*, 268, 115022.
- L. F. Santos, C. B. B. Costa, J. A. Caballero, M. A. S. S. Ravagnani, 2022, Framework for embedding black-box simulation into gams via kriging surrogate model applied to natural gas liquefaction process optimization, *Applied Energy*, 310, 118537.
- L. F. Santos, C. B. B. Costa, J. A. Caballero, M. A. S. S. Ravagnani, 2021, Kriging-assisted constrained optimization of single-mixed refrigerant natural gas liquefaction process, *Chemical Engineering Science*, 116699.
- R. Song, M. Cui, J. Liu, 2017, Single and multiple objective optimization of a natural gas liquefaction process, *Energy* 124, 19–28.

Integrated synthesis, modeling, and assessment of waste-to-resource alternatives

Adrián Pacheco-López^a, Ana Somoza-Tornos^{a,b}, Moisès Graells^a, Antonio Espuña^{a*}

^a*Chemical Engineering Department, Universitat Politècnica de Catalunya, Escola d'Enginyeria de Barcelona Est, C/ Eduard Maristany 16, Barcelona 08019, Spain*

^b*Renewable and Sustainable Energy Institute, University of Colorado Boulder, Boulder, Colorado 80303, United States of America*

**antonio.espuna@upc.edu*

Abstract

To address the urgent need to transform the current economic paradigm towards a more circular one, this contribution proposes a framework for assisting the selection of the most promising waste-to-resource alternatives. An ontology-based approach is used to manage and centralize the knowledge across the different stages of this methodology. Feasible combinations of operations lead to processing routes connecting waste with products and the paths are assessed to pre-select the most suitable ones. Then, the network is optimized, and a set of configurations are obtained in order to maximize economic profit and minimize environmental impact. The performance of the proposed framework is illustrated through a case study for the treatment of mixed plastic waste. Results show how pyrolysis-based treatments lead to suitable options according to some specified objectives but also reveal the main drawbacks and conditions that these treatments should improve. The work also demonstrates how to efficiently assess a large number of options, and discusses how to build and improve new processing networks.

Keywords: circular economy, pyrolysis, plastic waste, integrated modeling, sustainable development

1. Introduction

Numerous processes/product design strategies are being proposed towards the new circular economy paradigm. Among them, it should be emphasized the line based on improved design of products to reduce waste generation and maintain resources within the material cycles as long as possible, as well as promoting the use of renewable or better-performing resources to reduce environmental impacts as much as possible (Ellen MacArthur Foundation, 2015) and improve overall sustainability.

A considerable number of technologies are now being developed to achieve such objectives. Consequently, the number of available alternatives is set to grow exponentially. Decision-makers need to assess these alternatives and select the best ones for each kind of waste or material cycle and optimize them attending to all three main pillars of sustainability: maximizing economic performance, minimizing environmental impacts, and promoting social benefits (Sillanpää and Ncibi, 2019). However, traditional optimization methods are very complex and not efficient enough when the number of structural alternatives is very large (Yang et al., 2013).

To address this challenge, this contribution proposes a way to integrate several models and procedures to generate, assess, and optimize the most promising alternatives for the treatment and revalorization of a generic waste stream, providing a systematic and fully

integrated tool to support decision-making towards the circular economy. The resulting framework (Figure 1) encompasses interoperable modules, which exchange the necessary information to perform a comprehensive route generation and assessment, as well as the optimization of a waste-to-resource network considering the best options.

1.1. Problem Statement

- Given an ontology filled with relevant information of processes suitable to treat a specific waste source.
- Obtain a set of promising treatment pathways and an optimized process network attending to environmental and economic objectives.

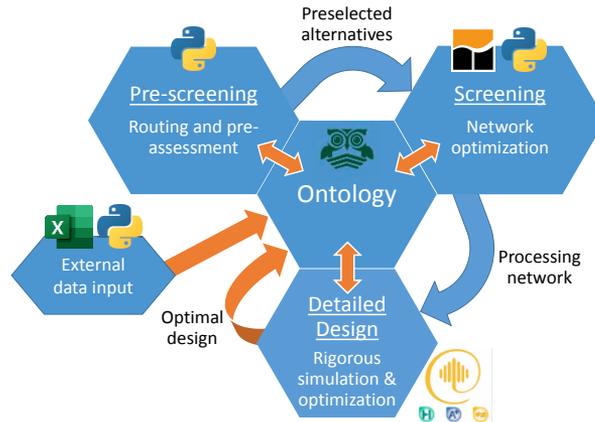


Figure 1. Schematic representation of the proposed framework.

2. Methodology

The proposed framework is structured into four main stages:

- The first one corresponds to knowledge management, where many single process steps are characterized as specific instances, including all the corresponding attributes that hold specifications such as properties, process parameters, and relationships with other instances such as process states and other process steps. Such information is managed through an ontology-based system. Ontologies have proven to be an efficient repository of information that allows establishing versatile relationships between different entities and storing parameters flexibly.
- The second one uses the information from the ontology to implicitly generate all the feasible routes (e.g. using short-path algorithms). Different criteria may be used to limit this generation based on usual branch and cut procedures. Once the routes are generated, they are sorted according to a proposed Global Performance Indicator (GPI) and the most promising ones are selected (Pacheco-López et al., 2021).
- The third one starts with this pre-selection of alternatives, where a more extensive process screening is performed. To do so, a superstructure with the most promising alternatives is assessed and optimized to select the most suitable processing networks based on different objectives, such as maximizing economic profit and minimizing the three endpoint life cycle assessment indicators (Somoza-Tornos et al., 2021).
- The fourth stage corresponds with the design and simulation of the process steps that are introduced in the ontology to be used by the other modules. It can be used as a preliminary step to fill up the ontology with new alternatives or afterward to further optimize selected alternatives and therefore enrich the ontology.

3. Case Study

The integrated decision-making framework performance is illustrated through a case study that consists of mixed plastic waste (MPW) treatment to obtain valuable products and reduce its environmental footprint. The chosen feedstock corresponds to a simulated MPW sample from sorting plants where packaging plastic is separated from municipal solid waste. It is composed of 40% of polyethylene (PE), 35% of polypropylene (PP), 18% of polystyrene (PS), 4% of polyethylene terephthalate (PET), and 3% of polyvinyl chloride (PVC) as proposed by Adrados et al., 2012.

The used ontology is an extended version from the one proposed previously (Pacheco-López et al., 2021), which was adapted from OntoCAPE (Marquardt et al., 2010) and filled with extra processes found in the literature to increase the number of alternatives available. For the pre-screening stage, all estimations were made and unified as presented in previous works, such as monetization of environmental impacts, maturity evaluation, and unitary costs calculation. The separation steps were simulated to split the gas and oil phases into all their components as commercial grade products. Pyrolyses were simulated and validated with experimental data from different studies. Market prices were updated to 2019 from the Prodcom database (Eurostat - European Commission and Eurostat, 2021). For the screening stage, all the estimations and assumptions for the needed parameters are as presented by Somoza-Tornos et al., 2021, such as the gate-to-gate approach for the life cycle assessments. The volume of plastic waste to treat is the amount collected in the EU28 (9.4 MT; PlasticsEurope, 2020) and escalated to a city of around 3 million inhabitants such as the Metropolitan Area of Barcelona, leading to a throughput of 5.72 tons of MPW per hour. Specific products demand satisfaction has been obviated in this case and all obtained products are assumed to be potentially sold at market price. For simplicity, only a deterministic model is used, although a stochastic model can be implemented in future developments to deal with cost estimation uncertainties.

4. Results and Discussion

Once the ontology is filled with the corresponding process steps and state specifications, the algorithm is run selecting the chosen starting material as mentioned above. The results for the pre-screening and screening stages are then obtained systematically. Figure 2 shows the graph obtained as a previous step to the route generation and assessment in the pre-screening stage, where all possible connections among the processes available in the ontology are shown. When the routing is implicitly performed, a total of 45 tentative processes are considered and 136 alternative paths are generated, assessed, and sorted according to their GPI. For simplicity, Table 1 shows a partial list of the sorted paths where it can be appreciated that the pyrolysis at 500°C with and without catalysts appear as the most promising, as opposed to landfilling and incineration. Once a representative set of the most promising processes are selected, the network optimization is performed in the screening stage, where a multi-objective optimization is performed using the ϵ -constraint method. By doing so, different configurations and results are obtained depending on the objectives trade-off. Figure 3 shows the optimal configuration when the objective is to maximize profit. Applying the ϵ -constraint for each one of the three environmental endpoint indicators, the corresponding Pareto optimal solution curves are obtained, as illustrated in Figure 5. The points in these plots are color-coded according to the different configurations, which are shown in Figure 4.

As envisaged in the previous stage, pyrolysis reactors that operate at 500°C with and without catalysts appear as the most profitable options from economic and environmental points of view in the screening stage as well. Involving these processes, the procedure

obtained four different configurations. The gray configuration (Figure 3) is obtained for the maximized economic performance and the green one when the objective is to minimize the impacts on the three endpoints. The first configuration consists of the pyrolysis at 500°C without catalyst and separation of the liquid and gas phase into their pure components to be sold. When minimizing environmental impacts, the configuration consists only of the pyrolysis using zeolite as catalyst (green configuration) and selling the pyrolytic gas and oil fractions for other uses (fuels for instance). In between the two-abovementioned anchor points, two more configurations could serve as a trade-off between the four objectives, blue and orange, which entail several pyrolyses with and without catalysts along with separation of some of their products (see Figure 4). As observed in Figure 5, the more separations are included in the configurations, the higher the environmental impact on all categories, and the higher the profit. This can be explained by the higher added value of the products when separated against the separation cost, which is the opposite case for environmental indicators due to energy requirements.

Table 1. Sorted paths obtained in the pre-screening stage according to the proposed GPI.

N°	Path composition	GPI
1	Pyrolysis MPW 500°C + Separation	779
2	Pyrolysis MPW 500°C /red mud/ + Separation	707
3	Pyrolysis MPW 500°C /ZSM5/ + Separation	509
7	Sorting + Pyrolysis PE 740°C + Separation	158
...
132	Sorting + Pyrolysis PS 425°C + Separation	28
134	Sorting + Pyrolysis PE 450°C	19
135	Incineration	0
136	Landfill	0

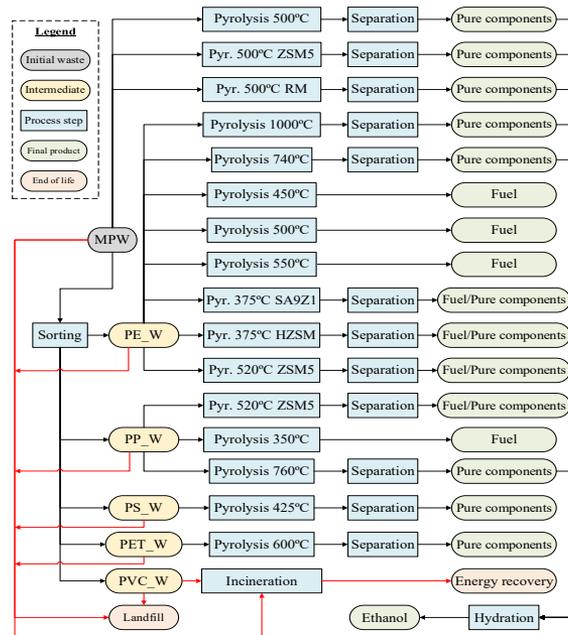


Figure 2. Implicitly generated graph in the pre-screening stage with tentative connections.

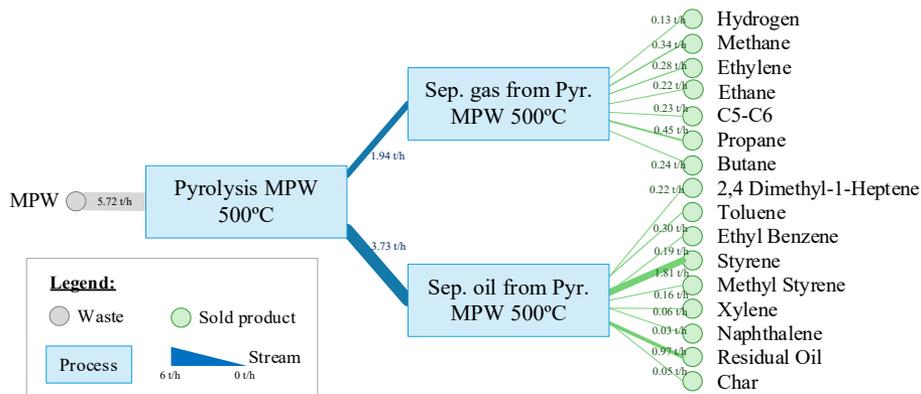


Figure 3. Configuration obtained for maximized profit. Color-coded as gray in Figure 5.

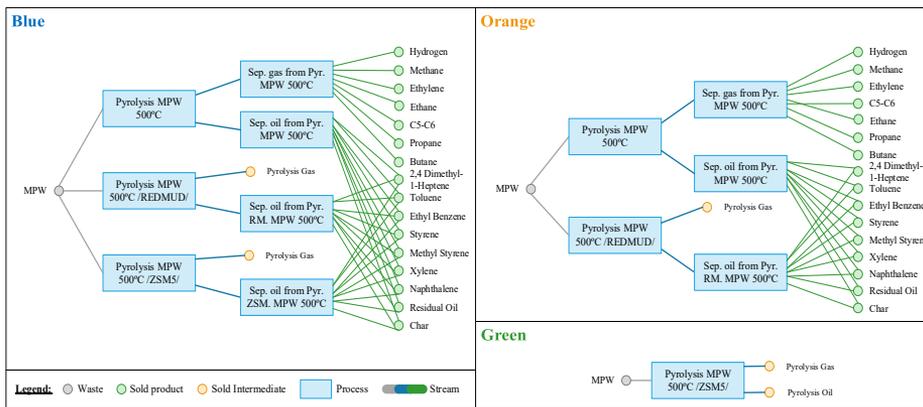


Figure 4. All possible configurations are color-named to ease their identification in the Pareto fronts in Figure 5.

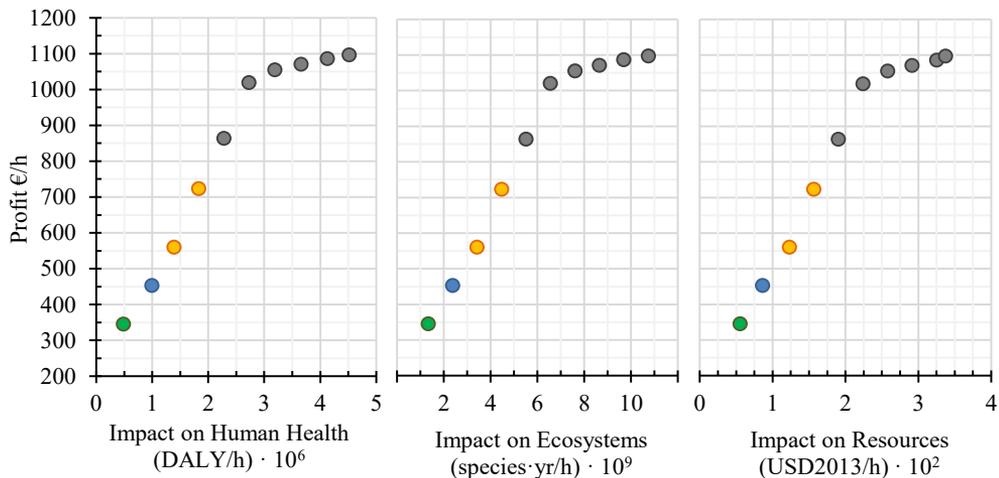


Figure 5. Pareto points for the trade-off between profit and the three environmental endpoint indicators. The points are color-coded to represent different configurations, as shown in Figure 4.

5. Conclusions

This work presents an integrated and systematic tool to generate feasible processing routes using short-path algorithms and assess them with different performance indicators, including a newly proposed one, that simultaneously accounts for economic, environmental, and maturity objectives. The best routes obtained at this stage are then stored back and enriched through their integration in an ontology and passed forward to a third stage where a network multi-objective optimization is performed and different network configurations are obtained and classified according to their performance, attending to different Pareto fronts. The presented case study illustrates the identification and assessment of opportunities in the recycling of plastic wastes. These opportunities are shown to be potentially profitable not only environmentally but also from the economic point of view. Additional developments should entail a more in-depth design of the selected configuration, along with energy integration, and optimization of equipment parameters (e.g. sizing, operating conditions...), including results, feedback to enrich the ontology. Future work will focus on this part as well as on further integrating other useful tools, such as automated information extraction and flowsheet development.

Acknowledgments

Financial support received from the "Ministerio de Ciencia e Innovación" and the European Regional Development Fund, both funding the research Project CEPI (PID2020-116051RB-I00) is fully acknowledged. Adrian Pacheco-Lopez thanks the financial support received from the "Ministerio de Ciencia e Innovación" (grant ref. PRE2018-087135).

References

- Adrados, A., de Marco, I., Caballero, B.M., López, A., Laresgoiti, M.F., Torres, A., 2012. Pyrolysis of plastic packaging waste: A comparison of plastic residuals from material recovery facilities with simulated plastic waste. *Waste Manag.* 32, 826–832. <https://doi.org/10.1016/j.wasman.2011.06.016>
- Ellen MacArthur Foundation, 2015. *Towards a Circular Economy: Business Rationale for an Accelerated Transition*. Ellen MacArthur Found. 20.
- Eurostat - European Commission, Eurostat, 2021. *Statistical Classification of Economy Activity in the European Union (NACE 2). 2019 Annual detailed data by PRODCOM list*. <http://ec.europa.eu/eurostat/web/prodcom/data/excel-files-nace-rev.2> [WWW Document]. URL <http://ec.europa.eu/eurostat/web/prodcom/data/excel-files-nace-rev.2> (accessed 9.21.21).
- Marquardt, W., Morbach, J., Wiesner, A., Yang, A., 2010. *OntoCAPE-A Re-Usable Ontology for Chemical Process Engineering*, RWTHedition. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-04655-1>
- Pacheco-López, A., Somoza-Tornos, A., Graells, M., Espuña, A., 2021. Synthesis and assessment of waste-to-resource routes for circular economy. *Comput. Chem. Eng.* 153. <https://doi.org/10.1016/j.compchemeng.2021.107439>
- PlasticsEurope, 2020. *Plastics – the Facts 2020*. PlasticEurope 1–64.
- Sillanpää, M., Ncibi, C., 2019. Circular economy and sustainable development, *The Circular Economy*. <https://doi.org/10.1016/b978-0-12-815267-6.00006-2>
- Somoza-Tornos, A., Pozo, C., Graells, M., Espuña, A., Puigjaner, L., 2021. Process screening framework for the synthesis of process networks from a circular economy perspective. *Resour. Conserv. Recycl.* 164, 105147. <https://doi.org/10.1016/j.resconrec.2020.105147>
- Yang, X.S., Koziel, S., Leifsson, L., 2013. Computational optimization, modelling and simulation: Recent trends and challenges. *Procedia Comput. Sci.* 18, 855–860. <https://doi.org/10.1016/j.procs.2013.05.250>

MINLP model for work and heat exchange networks synthesis considering unclassified streams

Lucas F. Santos,^{a,b} Caliane B. B. Costa,^a Jose A. Caballero,^b Mauro A. S. S. Ravagnani^a

^aDepartment of Chemical Engineering, State University of Maringá, Av. Colombo 5790, Maringá - 87020900, Brazil

^bInstitute of Chemical Process Engineering, University of Alicante, Ap. Correos 99, Alicante – 03080, Spain
lfs.francisco.95@gmail.com, pg54347@uem.br

Abstract

The optimal synthesis of work and heat exchange networks (WHENs) is deeply important to achieve simultaneously high energy efficiency and low costs in chemical processes via work and heat integration of process streams. This paper presents an efficient MINLP model for optimal WHENs synthesis derived from a superstructure that considers unclassified streams. The derived model is solved using BARON global optimization solver. The superstructure considers multi-staged heat integration with isothermal mixing, temperature adjustment with hot or cold utility, and work exchange network for streams that are not classified *a priori*. The leading advantage of the present optimization model is the capability of defining the temperature and pressure route, *i.e.* heating up, cooling down, expanding, or compressing, of a process stream entirely during optimization while still being eligible for global optimization. The present approach is tested to a small-scale WHEN problem and the result surpassed the ones from the literature.

Keywords: Work and heat exchange networks, Mixed-integer nonlinear programming, Unclassified streams, Process synthesis, Global optimization.

1. Introduction

Recovering work and heat in chemical processes is fundamental for achieving high energy efficiencies, and it can be performed through the optimal synthesis of work and heat exchange networks (WHENs). This synthesis problem can be approached with superstructure-derived mixed-integer nonlinear programming (MINLP). The main challenge of WHENs synthesis is the lack of predefined pressure and temperature change routes of process streams, which hampers *a priori* stream classification. In other words, differently from heat integration and because of temperature variation in compression and expansion, a stream may change its thermal identity from hot to cold or vice-versa, making it difficult to target energy demands. The resulting MINLP models are intrinsically difficult to solve due to nonconvexity and combinatorial complexity that scales up quickly with the size of the problem, *i.e.* with the number of streams and superstructure stages (Santos et al., 2020b).

This problem has been addressed via either thermodynamic analysis (pinch variations to account for work and heat) or mathematical programming. The latter approach has shown promising results in the literature. Wechsung et al. (2011) proposed a superstructure with fixed thermodynamic routes for process streams that were classified as hot or cold, and

with fixed or variable pressure. To further simplify the MINLP problem, heuristics were considered for the appropriate placement of pressure manipulators. The optimization problem aimed to minimize the WHEN irreversibility.

Onishi et al. (2014a) proposed a WHEN superstructure based on Wechsung et al. (2011), except that the well-known heat exchanger network (HEN) superstructure from Yee and Grossmann (1990) was introduced for heat integration, and mathematical programming was used for placing the pressure manipulators. The MINLP formulation was developed to minimize the network total annualized cost (TAC). Onishi et al. (2014b) elaborated a multi-stage superstructure considering high-pressure streams as cold streams and low-pressure as hot streams. Huang and Karimi (2016) presented some modifications to the work of Onishi et al. (2014b), like the decision of the final device to adjust stream temperature based on the stream need instead of its identity.

Onishi et al. (2018) developed an innovative WHEN superstructure to deal with streams that are not classified *a priori*. Generalized disjunctive programming (GDP) was used to deal with the pressure manipulator selection and the classification of the stream. The problem had an MINLP formulation that was a convex hull reformulation of the GDP. It was incorporated with a Pinch-based optimization model for heat integration to minimize the TAC. Nair et al. (2018) added to the model of Onishi et al. (2018) the possibility of phase change, variable heat capacity, as well as compression and expansion of streams with no net pressure change (cycles).

Differently from previous authors that relied on mathematical programming, some interesting results used meta-heuristics to deal with the WHEN synthesis problem. Pavão et al. (2019) proposed a new approach to WHEN synthesis considering non-isothermal mixing and utilities in every HEN stage in parallel with the other heat transfer devices. The solution approach comprised Simulated Annealing (SA) for the combinatorial level and Rocket Fireworks Optimization for the continuous one. Santos et al. (2020a) proposed a new superstructure-derived MINLP model with a reduced number of decision variables. Change of variables and inner-level optimization were considered to diminish the combinatorial size of the optimization problem. The solution approach consisted of a two-level meta-heuristic optimization, using SA in the combinatorial problem and Particle Swarm Optimization in the nonlinear problem. Lin et al. (2021) proposed a two-piece WHEN synthesis framework composed of a targeting phase followed by the detailed HEN synthesis. The former was performed by optimizing a model of thermodynamic paths of process streams using hybridization of genetic algorithm and golden section method. The latter was performed using mathematical programming to minimize TAC.

A major challenge in this synthesis problem is dealing with streams classification. Some authors considered energy targeting or fixed temperature and pressure routes based on thermodynamic insights. Other authors included the classification of the stream as binary decision variables in the optimization problem. The objective of this paper is to present a WHEN superstructure and a derived MINLP model that deals with streams classification without binary decision variables or thermodynamic heuristics. That is achieved by introducing the novel reciprocal heat exchangers in the heat integration superstructure.

2. WHEN model

2.1. Problem definition

The WHEN synthesis problem can be defined as determining a set of electric turbines and compressors, single-shaft turbine-compressors, helper motors, electric generators, heat exchangers, heaters, and coolers that perform the required temperature and pressure

changes of process streams with minimum operating and capital costs. Therefore, for the set of S streams ($s = 1, \dots, S$), it is given their initial and final states (T_{in} , P_{in} , T_{out} , and P_{out}), heat capacity flow rates (CP) and heat exchange coefficients (h). Hot and cold utilities are available with known inlet and outlet temperatures (TS_{in} , TW_{in} , TS_{out} , and TW_{out}), individual heat exchange coefficients (h_s and h_w) and costs (CHU and CCU). The prices of purchase and selling electricity (CE and PE) are given as well as economic capital cost equations ($capC$, $capT$, and $capA$), polytropic coefficient (κ); and compression and expansion efficiencies (η_c and η_e).

2.2. Superstructure

The superstructure is based on Santos et al. (2020a), but without a classification section (Figure 1). The proposed superstructure has three sections: heat integration, temperature adjustment, and work exchange network. The heat integration section is based on Yee and Grossmann (1990), updated to deal with unclassified streams with reciprocal heat exchangers. This superstructure artifact consists of using the notation of s, n, ss, nn, k for a heat exchanger between hot stream (s, n) and cold stream (ss, nn) in heat integration stage k for the hot stream and $K - k - 1$ for the cold one, where s is the stream number, n is the WHEN superstructure stage and K is the total number of heat integration stages. The temperature adjustment section allows using a hot or cold utility to achieve the desired cooling or heating task. The work exchange network section is based on Onishi et al. (2018), which uses convex hull reformulation of a GDP model to deal with the selection of compressors and turbines.

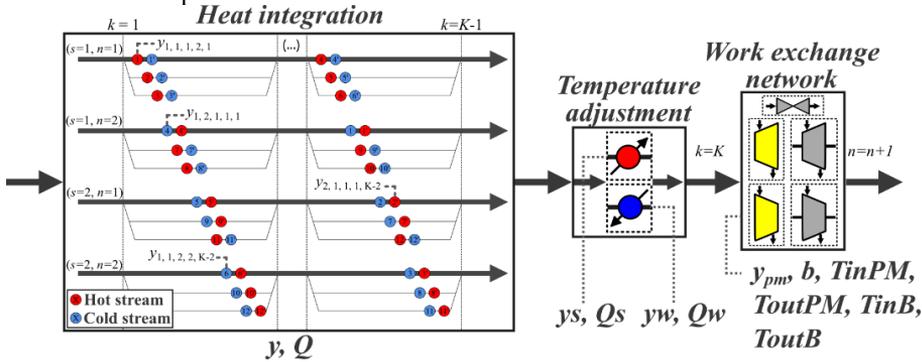


Figure 1. Proposed WHEN superstructure.

2.2.1. MINLP Model

In the heat integration section, Eqs. (1) – (6), the binary variable $y_{s,n,ss,nn,k}$ stands for the existence (1) or inexistence (0) of a heat exchanger between hot stream (s, n) and cold stream (ss, nn) in heat integration stage k for the hot stream and $K - k - 1$ for the cold one. In addition, Q , T , dTh , dTc , and A are the heat load, temperature, temperature difference at the hot and cold end, and area of a heat exchanger, respectively, considering their subscript index. CP , h , and the upper and lower limits of temperature (T^{up} and T^{lo}) are parameters given by the problem statement. Note that to deal with pre-classified streams one might just fix to zero the heat load and binary variable of the heat exchangers that regard streams with opposite thermal identity.

$$\forall s, n, ss, nn, k < K - 2: \quad T_{s,n,k+1} = T_{s,n,k} - \sum_{ss,nn \in he} \frac{Q_{s,n,ss,nn,k}}{CP_s} + \sum_{ss,nn \in he} \frac{Q_{ss,nn,s,n,K-k-1}}{CP_s} \quad (1)$$

$$Q_{s,n,ss,nn,k} \leq y_{s,n,ss,nn,k} Q^{up} \quad (2)$$

$$dT h_{s,n,ss,nn,k} \leq T_{s,n,k} - T_{ss,nn,K-k} + (T^{up} - T^{lo})(1 - y_{s,n,ss,nn,k}) \quad (3)$$

$$dT c_{s,n,ss,nn,k} \leq T_{s,n,k+1} - T_{ss,nn,K-k-1} + (T^{up} - T^{lo})(1 - y_{s,n,ss,nn,k}) \quad (4)$$

$$y_{s,n,ss,nn,k} + y_{ss,nn,s,n,K-k-1} \leq 1 \quad (5)$$

$$A_{s,n,ss,nn,k} \left(dT h_{s,n,ss,nn,k} dT c_{s,n,ss,nn,k} \frac{dT h_{s,n,ss,nn,k} + dT c_{s,n,ss,nn,k}}{2} \right)^{\frac{1}{3}} \geq Q_{s,n,ss,nn,k} \frac{h_s + h_{ss}}{h_s h_{ss}} \quad (6)$$

In the temperature adjustment section, Eqs. (7) – (16), it is possible to place a heater ($y_{s,n} = 1$), a cooler ($y_{w,n} = 1$), or none ($y_{w,n} + y_{s,n} = 0$) at stream s in stage n . Q_s , Q_w , $dTsh$, $dTsc$, $dTwh$, $dTwc$, A_s , and A_w are the heat load of the heater and cooler, temperature difference at the hot and cold end for the heater and cooler, and heat exchange area of the heater and cooler, respectively, considering the subscript index referred to streams and stages. The upper limit of heat load (Q^{up}), $TSout$, $TSin$, $TWout$, $TWin$, h_s , and h_w are parameters given by the problem statement.

$$\forall s, n, k = K - 2:$$

$$T_{s,n,k+1} = T_{s,n,k} + Q_{s,n}/CP_s - Q_{w,n}/CP_s \quad (7)$$

$$Q_{s,n} \leq y_{s,n} Q^{up} \quad (8)$$

$$Q_{w,n} \leq y_{w,n} Q^{up} \quad (9)$$

$$y_{w,n} + y_{s,n} \leq 1 \quad (10)$$

$$dTsh_{s,n} \leq TSout - T_{s,n,k} + (TSin - T^{lo})(1 - y_{s,n}) \quad (11)$$

$$dTsc_{s,n} \leq TSin - T_{s,n,k+1} + (TSin - T^{lo})(1 - y_{s,n}) \quad (12)$$

$$dTwh_{s,n} \leq T_{s,n,k} - TWout + (T^{up} - TWin)(1 - y_{w,n}) \quad (13)$$

$$dTwc_{s,n} \leq T_{s,n,k+1} - TWin + (T^{up} - TWin)(1 - y_{w,n}) \quad (14)$$

$$A_{s,n} \left(dTsh_{s,n} dTsc_{s,n} \frac{dTsh_{s,n} + dTsc_{s,n}}{2} \right)^{\frac{1}{3}} \geq Q_{s,n} \frac{h_s + h_s}{h_s h_s} \quad (15)$$

$$A_{w,n} \left(dTwh_{s,n} dTwc_{s,n} \frac{dTwh_{s,n} + dTwc_{s,n}}{2} \right)^{\frac{1}{3}} \geq Q_{w,n} \frac{h_s + h_w}{h_s h_w} \quad (16)$$

The work exchange network (WEN) section is based on Onishi et al. (2018). The cost calculation is based on Santos et al. (2020a), in which the TAC is the sum of operating and capital costs of WEN and HEN. The resulting MINLP model is implemented in GAMS 37.1.0 and solved with the global optimization solver BARON 21.1.13 (Tawarmalani & Sahinidis, 2005) to minimize the TAC.

3. Case study

This two-stream problem was proposed by Onishi et al. (2014a) and later approached by Lin et al. (2021). Table 1 and Table 2 present the stream and cost data, in which Wc , Wt , Qs , and Qw are given in kW, and A is given in m^2 . Notice that the problem statement contains several considerations to simplify mathematically the synthesis task. Some other parameters from the problem statement include T^{lo} and T^{up} of 350 and 750 K, minimum temperature approach ($dTmin$) of 5 K, polytropic coefficient of 1.352, compressor and turbine efficiencies of 100 %, and annualization factor of 0.18. Notice that $dTmin$ is the lower limit of all dT variables. For this case study, the authors disregarded the use of single-shaft turbine-compressors. The model size for this small-scale case study was 397 equations, 397 variables (80 discrete ones), and the best solution is found in about 5 s.

The results of solving this WHEN synthesis problem with the proposed MINLP model are illustrated in Figure 2.

Table 1. Stream data for the case study.

Stream	T_{in} [K]	T_{out} [K]	P_{in} [MPa]	P_{out} [MPa]	CP [kW/K]	h [kW/m ² K]
s1	650	370	0.1	0.5	3.0	0.1
s2	410	650	0.5	0.1	2.0	0.1
HU	680	680				1.0
CU	300	300				1.0

Table 2. Cost data for the case study.

Operation	Capital cost	Unit	Operating cost	Unit
Compressors	$capC = 30.317(Wc)^{0.62}$	(k\$)	$CE = 0.45504$	(k\$/kW _y)
Turbines	$capT = 1.5338(Wt)^{0.81}$	(k\$)	$PE = 0.0$	(k\$/kW _y)
Heat exchangers	$capA = 21.778 + 3.4467(A)$	(k\$)	$UT = 0.377Q_s + 0.1Q_w$	(k\$/y)

The TAC of this solution is 834.7 k\$/year, a value that surpassed the literature results of 1207 in Onishi et al. (2014a) and 837 k\$/year in Lin et al. (2021). This shows that strategies of energy targeting and fixing thermodynamic routes can simplify big-size WHEN problem at the price of potentially disregarding the global optimum.

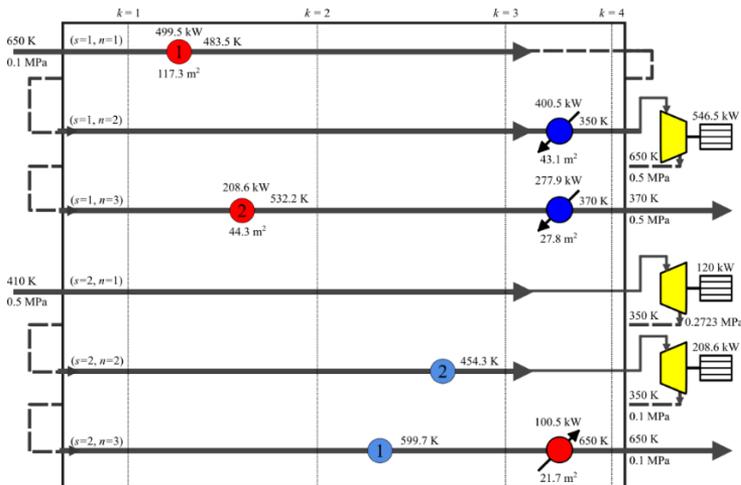


Figure 2. WHEN result for the case study.

Compared to the result of Lin et al. (2021), the proposed solution for this small-scale WHEN synthesis problem has one less heat exchanger recovers 2.2 kW more heat, consumes 2.1 kW less net work, and consumes 2.2 kW less cold utility. The main advantages of the present WHEN are achieving a higher inlet temperature to the second turbine, producing more work, and the increased heat integration, which saved utility. For future work, one might consider using the idea of reciprocal heat exchangers and the present MINLP model for medium or big-sized WHEN problems. For solving such problems efficiently with global optimization solver, such as BARON, improving the model to provide tighter bounds might be required.

4. Conclusions

An efficient MINLP model for optimal WHENs synthesis was proposed from a superstructure that considers unclassified streams via reciprocal heat exchangers. A small-scale WHEN case study from the literature was used to test the developed model. It was solved using the BARON global optimization solver in GAMS. The leading advantage of the present optimization model is the capability of defining the temperature and pressure route of a process stream entirely during optimization, instead of relying on energy targeting or fixed thermodynamic routes. Results surpassed the ones from the literature from 837.0 to 834.7 k\$/year of total annualized cost. For future work, the present MINLP model should be tested to medium or big-sized WHEN problems and be modified to tighten the bounds for global optimization.

Acknowledgments

The authors LFS, CBBC, and MASSR acknowledge the National Council for Scientific and Technological Development – CNPq (Brazil), processes 148184/2019-7, 440047/2019-6, 311807/2018-6, 428650/2018-0, and Coordination for the Improvement of Higher Education Personnel – CAPES (Brazil) for the financial support. The author JAC acknowledge financial support from the “Generalitat Valenciana” under project PROMETEO 2020/064.

References

- K. Huang, & I. A. Karimi, 2016, Work-heat exchanger network synthesis (WHENS), *Energy*, 113, 1006–1017.
- Q. Lin, C. Chang, Z. Liao, J. Sun, B. Jiang, J. Wang, & Y. Yang, 2021, Efficient Strategy for the Synthesis of Work and Heat Exchange Networks, *Industrial & Engineering Chemistry Research*, 60, 4, 1756–1773.
- S. K. Nair, H. Nagesh Rao, & I. A. Karimi, 2018, Framework for work-heat exchange network synthesis (WHENS), *AIChE Journal*, 64, 7, 2472–2485.
- V. C. Onishi, N. Quirante, M. A. S. S. Ravagnani, & J. A. Caballero, 2018, Optimal synthesis of work and heat exchangers networks considering unclassified process streams at sub and above-ambient conditions, *Applied Energy*, 224, 567–581.
- V. C. Onishi, M. A. S. S. Ravagnani, & J. A. Caballero, 2014a, Simultaneous synthesis of heat exchanger networks with pressure recovery, Optimal integration between heat and work. *AIChE Journal*, 60, 3, 893–908.
- V. C. Onishi, M. A. S. S. Ravagnani, & J. A. Caballero, 2014b, Simultaneous synthesis of work exchange networks with heat integration, *Chemical Engineering Science*, 112, 87–107.
- L. V. Pavão, C. B. B. Costa, & M. A. S. S. Ravagnani, 2019, A new framework for work and heat exchange network synthesis and optimization, *Energy Conversion and Management*, 183, 617–632.
- L. F. Santos, C. B. B. Costa, J. A. Caballero, & M. A. S. S. Ravagnani, 2020a, Synthesis and optimization of work and heat exchange networks using an MINLP model with a reduced number of decision variables, *Applied Energy*, 262, 114441.
- L. F. Santos, C. B. B. Costa, J. A. Caballero, & M. A. S. S. Ravagnani, 2020b, Search Space Analysis in Work and Heat Exchange Networks Synthesis using MINLP Models, *Computer Aided Chemical Engineering*, 48.
- M. Tawarmalani, & N. V. Sahinidis, 2005, A polyhedral branch-and-cut approach to global optimization, *Mathematical Programming*, 103, 2, 225–249.
- A. Wechsung, A. Aspelund, T. Gundersen, & P. I. Barton, 2011, Synthesis of heat exchanger networks at subambient conditions with compression and expansion of process streams, *AIChE Journal*, 57, 8, 2090–2108.
- T. F. Yee, & I. E. Grossmann, 1990, Simultaneous optimization models for heat integration—II Heat exchanger network synthesis, *Computers & Chemical Engineering*, 14, 10, 1165–1184.

Planetary Boundaries Assessment of Flue Gas Valorization into Ammonia and Methane

Sebastiano C. D'Angelo,^a Julian Mache,^a Gonzalo Guillén-Gosálbez^a

*^aInstitute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zürich, Vladimir-Prelog-Weg 1, 8093 Zürich, Switzerland
gonzalo.guillen.gosalbez@chem.ethz.ch*

Abstract

Carbon capture and utilization (CCU) has attracted growing interest to potentially curb CO₂ emissions while generating valuable chemicals. These emerging technologies will coexist with their fossil analogs soon, creating opportunities to combine both. In this context, flue gas valorization from power plants can play a role in this transition. The success of CCU technologies will ultimately depend on its environmental sustainability, which should be evaluated based on global indicators such as the Planetary Boundaries (PBs). Here we assessed the absolute sustainability level of an integrated facility producing ammonia (NH₃) and synthetic natural gas (SNG) using hydrogen (H₂) from water electrolysis, and nitrogen (N₂) and CO₂ from flue gas from a combined-cycle natural gas power plant. The LCA-PBs assessment showed that the alternative technologies could contribute to operating the Planet safely by significantly reducing the impact on the climate change and ocean acidification Earth-system processes, compared to the fossil Haber-Bosch (HB) and conventional steam reforming processes. Overall, the investigated process could smoothen the transition towards low-carbon technologies. More broadly, the application of the PBs to quantify the environmental performance of the integrated system opens up new avenues for the absolute sustainability assessment of emerging low-carbon technologies within the chemical sector and beyond.

Keywords: flue gas valorization, planetary boundaries, life cycle assessment, circular economy

1. Introduction

The phasing out of fossil fuels for power generation is a key step in the sustainable transition of the energy and chemical sectors. However, during this transition, greenhouse gases (GHG) emissions from power plants still under operation will have to be reduced, possibly following a circular economy approach. This could be accomplished by using natural resources more efficiently and combining them with the valorization of side outputs into valuable materials that can re-enter the economy. In this context, carbon capture and utilization (CCU) is gaining increasing traction (Kätelhön et al., 2019). Recent work by Castellani et al. (2018) proposed an integrated process that separates flue gas from power plants into its main components: carbon dioxide (CO₂) and nitrogen (N₂). These streams are then upgraded into synthetic natural gas (SNG) and ammonia (NH₃) through the Sabatier and the Haber-Bosch (HB) process, respectively, using electrolytic hydrogen (H₂). This concept shows several advantages. First, what was originally a waste stream could be upgraded into promising low-carbon energy vectors. Second, the produced SNG could reduce the consumption of natural gas power plants, in line with circular thinking. Third, by producing H₂ from intermittent energy sources such as wind

or solar power, part of the excess electricity during peak power generation times could be absorbed, avoiding curtailment (Staffell et al., 2019).

The environmental impact of these emerging technologies has been estimated through life cycle assessment (LCA), evaluating their sustainability along the whole supply chain. Several works applied LCA on NH_3 and SNG production through many alternative pathways to the business as usual (BAU) (Castellani et al., 2018). However, these studies often focus on carbon footprint solely while neglecting other impact categories.

In recent times, absolute sustainability assessments gained broad interest. Among these, the recently proposed planetary boundaries (PBs) concept (Rockström et al., 2009) establishes a set of critical thresholds on key Earth-system processes that enable comprehensive absolute sustainability assessments. Transgressing the PBs could shift the Planet's current state, challenging the Earth's resilience. These limits, all together, define a safe operating space (SOS) within which anthropogenic activities should lie. Despite their relevance in sustainable development, studies incorporating the PBs in chemicals and fuels assessments are scarce.

In this work, we apply a PB-LCA methodology to evaluate valorization pathways of flue gas from a natural gas power plant to produce NH_3 and SNG using renewable H_2 . In particular, we quantify the transgression levels relative to the SOS, focusing on a plant located in Germany and comparing the results with the fossil-based analogs. To the best of the authors' knowledge, this is the first time that such a novel methodology of environmental assessment is applied to the proposed process.

2. Methodology

2.1. Process description

The block flow diagram of the process is presented in Figure 1. It encompasses four main stages: membrane separation, water electrolysis, the Sabatier process, and the HB process. The plant is designed to use 10% of the electricity produced by wind in Germany for electrolytic H_2 production. Given an annual wind energy production of 131.7 TWh in 2020 (Burger, 2021) and an annual plant operating time of 8000 h, this represents an average consumption of 1.65 GW, corresponding to about 136 kt h^{-1} of flue gas (see also Table 1). Flue gas was assumed to contain exclusively N_2 (80 mol%) and CO_2 (20 mol%), approximating the average flue gas composition for a combined-cycle natural gas power plant (Al Hashmi et al., 2018; Castellani et al., 2018).

H_2 production is modeled in Aspen Custom Modeler[®] using the proton exchange membrane (PEM) technology (Ni et al., 2006) since it is currently considered the most suitable for intermittent operation (Staffell et al., 2019).

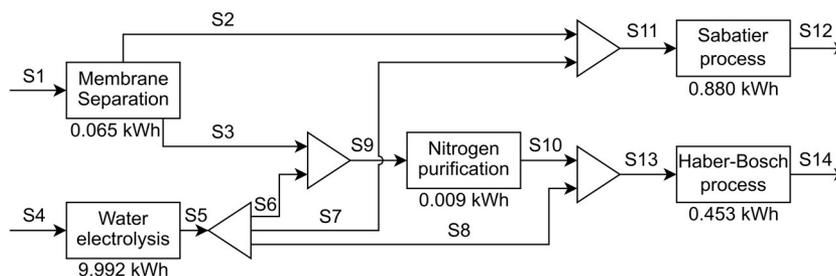


Figure 1. Block diagram of the case study. See Table 1 for the corresponding streams table. The electricity consumption, per kg flue gas, is shown below each process step.

Table 1. Selection of streams from the studied process. All the values are mass flows, in kt h^{-1} . The values denoted with a * correspond to water that is assumed to be immediately flashed.

Comp.\Stream	S1	S2	S3	S4	S6	S7	S8	S10	S12	S14
N ₂	97.67	1.16	96.52	0	0	0	0	96.52	1.16	0
CO ₂	38.36	34.53	3.84	0	0	0	0	0	0.19	0
H ₂	0	0	0	0	0.07	6.63	21.07	0	0.33	0
CH ₄	0	0	0	0	0	0	0	1.40	12.52	1.40
NH ₃	0	0	0	0	0	0	0	0	0	115.77
H ₂ O	0	0	0	248.12	0	0	0	3.14*	28.11*	0

The flue gas (S1 in Figure 1) is fed to a two-stage membrane separation, which provides a CO₂-rich stream (S2) and a N₂-rich stream (S3). A CO₂ recovery of 90% and a CO₂ purity of 95 mol% can be achieved in S2 (Castellani et al., 2018). The Sabatier process is tolerant to N₂ impurities, since N₂ behaves as an inert gas. In contrast, it is not possible to feed CO₂ impurities to the HB process. Consequently, as for the industrial standard, a small purification step that converts all the CO₂ to methane in the N₂-rich stream through methanation is assumed (D'Angelo et al., 2021). Accordingly, H₂ is fed in stoichiometric ratio to CO₂ to consume it entirely.

For the Sabatier reaction step (S11 to S12 in Figure 1), the process was modeled here assuming a CO₂ conversion of 99.5% (Castellani et al., 2018). To reach industrial grade, the SNG needs to be further purified from the N₂ impurity, which can be easily performed with a pressure swing adsorption unit using zeolites, for instance, at a low energy expense (Jayaraman et al., 2004). In the current study, this step was omitted for simplicity.

Furthermore, the water generated from the N₂ purification and Sabatier reaction (streams S10 and S12) is assumed to be separated using a flash at ambient conditions.

Finally, for NH₃ (S13 to S14 in Figure 1), a detailed model in Aspen HYSYS[®] was designed, based on D'Angelo et al. (2021) and considering an N₂ conversion of 98.6%.

2.2. Environmental Assessment

The LCA was performed following the ISO 14040 (ISO, 2014).

Three different scenarios were considered differing in the electricity source. All three scenarios assume that electricity powering the membrane separation, N₂ purification, Sabatier process, and HB process must be non-intermittent to ensure a smooth operation of the compressors. Accordingly, these steps are powered by the German 2020 power grid mix (Burger, 2021). At the same time, the first scenario (Sc1) assumes that PEM electrolysis is powered by the same mix, while the second (Sc2) uses offshore wind energy for the water-splitting step, and the last (Sc3) onshore wind instead.

A cradle-to-gate study was adopted to quantify the absolute sustainability level of three alternative flue gas valorization routes and an equivalent functional unit for the BAU. Further conversion of the products SNG and NH₃ is considered out of the work's scope. The selected functional unit was 1 kg of flue gas valorized for the three alternative scenarios. For the BAU, a system expansion approach was used. Notably, we considered the direct emissions from venting the flue gas, plus the impact of the BAU (fossil analogs) associated with the equivalent amount of natural gas and NH₃ produced from flue gas valorization. An attributional approach was selected.

In the second LCA phase, the life cycle inventories (LCIs) are modeled. The foreground system includes all the subprocesses depicted in Figure 1 and Table 1. At the same time,

the underlying energy and raw materials suppliers belong to the background system, here modeled with Ecoinvent v3.5, accessed through SimaPro v9.2. All the inventories, wherever possible, were regionalized for the German or European (RER) region. The H₂ inventory was obtained by combining results from Aspen Custom Modeler[®] with literature data, assuming H₂ storage in salt caverns (D'Angelo et al., 2021). The oxygen by-product obtained from water splitting was considered vented.

In the third phase, we used the characterization factors proposed by Ryberg et al. (2018) to estimate the impact on the control variables of the PBs. Nine Earth-system processes characterized by 11 control variables were considered, but two of them, atmospheric aerosol loading and novel entities, are yet to be quantified so they were omitted from this analysis. Since any of the PBs, if trespassed, could lead to catastrophic events, their joint ensemble defines the SOS for human anthropogenic activities. The same methodology described in D'Angelo et al. (2021) was used to quantify the levels of transgression (LTs). However, a different downscaling method was applied. Notably, the SOS was scaled down to the German market using a non-egalitarian approach involving the share of global GDP in 2019 covered by Germany. At the same time, the production plant was scaled up to cover the valorization of all the flue gas produced from natural gas power plants in the country (BEDW, 2020).

Finally, in step 4 of the LCA methodology, the results are interpreted, and potential recommendations are drawn. Here, we analyzed the impacts of the indicators exceeding the SOS to identify the main hotspots.

3. Results

The results of the selected scenarios (Sc1, Sc2, Sc3, and BAU) are shown in Figure 2.

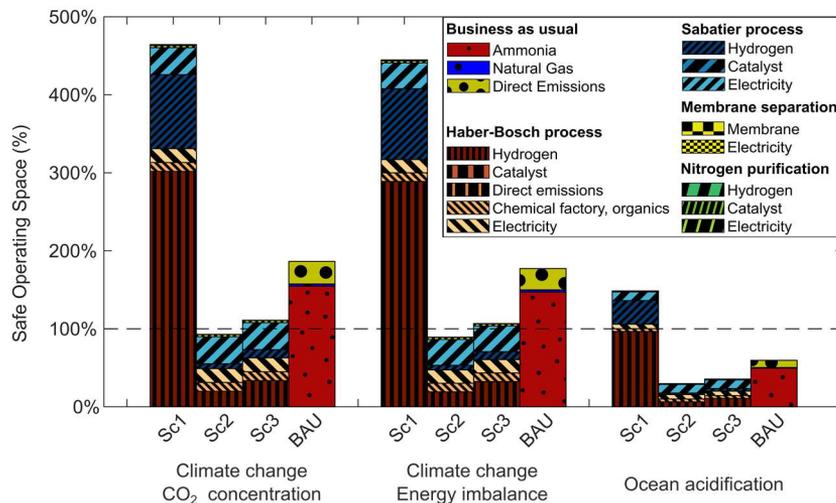


Figure 2. Impacts on the PBs control variables, and breakdown, for the four scenarios considered, quantified in terms of share of the SOS. Only the impacts exceeding the limit of 100% SOS are here shown.

Only the indicators exceeding the full SOS in at least one scenario are here presented, namely, climate change – CO₂ concentration (CC-CO₂), climate change – energy imbalance (CC-EI), and ocean acidification (OA). This is not surprising since these three

indicators are strictly associated with GHG concentration in different Earth compartments, i.e., the atmosphere for the first two and marine waters for OA.

The first key result is that the sustainability ranking witnesses Sc2 dominating for all the three selected control variables (29.7%-92.8% among the three indicators in focus), followed by Sc3 (35.5%-110.9%), BAU (59.6%-186.4%), and finally Sc1 (148.5%-464.4%). Sc2 is the only scenario that does not transgress any PB.

Moreover, the performance of the three valorization scenarios is strictly linked to the sustainability of the electricity employed for H₂ production. In fact, H₂ causes most of the impact in Sc1, with around 85.6% of the total contribution for all three indicators. Conversely, the H₂ contribution to the total impact shrinks down to around 28.1-28.2% for Sc2, where the electricity from the grid powering the rest of the facility plays the greatest role (49.6-49.9%). This fact is due to the remarkably high electricity consumption of water splitting (see also Figure 1) and the carbon intensity of the German power grid. In fact, more than a third (36.5%) of this energy mix relies on coal and natural gas power plants (Burger, 2021). Moreover, H₂ consumption explains the different magnitude of the impact of the various sub-processes in the valorization scenarios. The dominating role of the NH₃ production section is clearly highlighted, with a contribution to the overall alternative scenarios' impacts in the range of 52.76%-71.28%. Continuing along with the ranking, the second-highest contribution to the total impacts is SNG production, with a share of 27.9%-44.0% of the total impact.

On the contrary, the impacts associated with the very energy-efficient membrane separation are negligible. As anticipated, the power consumption associated with each part of the process dictates the overall performance. As highlighted in Table 1, the HB process requires the highest amount of H₂, almost 3.2 times as much as for SNG production. This is due to the fact that the N₂ feed to the HB process is much larger than the CO₂ feed that is meant to react through the Sabatier reaction. However, the latter needs more direct electric energy input, excluding H₂, than the HB process (see Figure 1), and a higher stoichiometric ratio H₂-to-substrate is required to convert CO₂ (4) with respect to N₂ (3). Finally, shifting the focus on the BAU, NH₃ production dominates the impact with about 83.0% of the share, followed by the emissions from venting the flue gas (15.4%-15.5%) and natural gas extraction and transport (about 1.6%). The lower relevance of the natural gas extraction compared to the SNG production can be explained by the much higher energy intensity of the Sabatier reaction and the very optimized extraction and transport of natural gas.

4. Conclusions

This work assessed an integrated process valorizing flue gas from natural gas power plants into NH₃ and SNG through using electrolytic H₂. The absolute environmental impact exerted on the Earth's biophysical limits was quantified through the PBs framework. For the first time, this novel LCA methodology is applied to such a process, to the best of the author's knowledge.

We found that the proposed process could perform more sustainably than the current fossil analog when using H₂ from renewable power sources such as wind. This fact can be true even in a country with energy mixes yet to be decarbonized, such as Germany, which still relies heavily on fossil coal and natural gas. However, the final feasibility of the proposed process will be strongly dependent on its economic affordability, which in the long run should be favored by the energy transition to renewables and lower capital costs for the electrolyzers.

For future studies, a comparison with an approach using the same membrane module to separate CO₂ from the flue gas, but assuming at a later stage the use of carbon capture and storage, could shed further light on the competitiveness of such a process relative to other environmentally promising approaches.

Acknowledgments

This publication was created as part of NCCR Catalysis, a National Centre of Competence in Research funded by the Swiss National Science Foundation.

References

- A.B. Al Hashmi, A.A.A. Mohamed, Z.E. Dadach, 2018, Process Simulation of a 620 Mw-Natural Gas Combined Cycle Power Plant with Optimum Flue Gas Recirculation, *Open Journal of Energy Efficiency*, 7, 33–52
- Bundesverband der Energie- und Wasserwirtschaft Deutschland (BEWD), 2020, Energiemarkt Deutschland 2020, URL <https://www.bdew.de/service/publikationen/bdew-energiemarkt-deutschland-2020/>
- B. Burger, 2021, Net Public Electricity Generation in Germany in 2020.
- B. Castellani, S. Rinaldi, E. Morini, B. Nastasi, F. Rossi, 2018, Flue gas treatment by power-to-gas integration for methane and ammonia synthesis – Energy and environmental analysis. *Energy Conversion and Management*, 171, 626–634
- S. C. D'Angelo, S. Cobo, V. Tulus, A. Nabera, A.J. Martín, J. Pérez-Ramírez, G. Guillén-Gosálbez, 2021, Planetary Boundaries Analysis of Low-Carbon Ammonia Production Routes. *ACS Sustainable Chemistry and Engineering*, 9, 9740–9749
- International Organization for Standardization (ISO), 2014. ISO 14040:2006 - Environmental management - Life Cycle Assessment: Principles and framework. ISO.
- A. Jayaraman, A.J. Hernandez-Maldonado, R.T. Yang, D. Chinn, C.L. Munson, D.H. Mohr, 2004, Clinoptilolites for nitrogen/methane separation, *Chemical Engineering Science*, 59, 2407–2417
- A. Kätelhön, R. Meys, S. Deutz, S. Suh, A. Bardow, 2019, Climate change mitigation potential of carbon capture and utilization in the chemical industry, *Proceedings of the National Academy of Sciences of the United States of America*, 166, 11187–11194
- M. Ni, M.K.H. Leung, D.Y.C. Leung, 2006, Electrochemistry Modeling of Proton Exchange Membrane (PEM) Water Electrolysis for Hydrogen Production, in: *World Hydrogen Energy Conference 16*, Lyon
- J. Rockström, W. Steffen, K. Noone, Å. Persson, F.S. Chapin, E. Lambin, T.M. Lenton, M. Scheffer, C. Folke, H.J. Schellnhuber, B. Nykvist, C.A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P.K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R.W. Corell, V.J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, J. Foley, 2009, Planetary Boundaries: Exploring the Safe Operating Space for Humanity. *Ecology and Society*, 14, 2
- M.W. Ryberg, M. Owsianiak, K. Richardson, M.Z. Hauschild, 2018, Development of a life-cycle impact assessment methodology linked to the Planetary Boundaries framework. *Ecological Indicators*, 88, 250–262
- I. Staffell, D. Scamman, A. Velazquez Abad, P. Balcombe, P.E. Dodds, P. Ekins, N. Shah, K.R. Ward, 2019, The role of hydrogen and fuel cells in the global energy system. *Energy & Environmental Science*, 12, 463–491

A techno-economic assessment of biochar production from date pits in the MENA region

Aisha Al-Rumaihi, Mohammad Alherbawi, Prakash Parthasarathy, Hamish R. Mackey, Gordon McKay and Tareq Al-Ansari*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar.

*talansari@hbku.edu.qa

Abstract

The Middle East and North Africa region (MENA) is the largest global producer of dates, with an annual production of more than 8 million tonnes. The date palm has been an iconic part of this region's culture since 3000 BC and played a key role in its development over the centuries. Incidentally, date pits account for approximately 10% of the fruit, representing a vital biomass resource due to its high carbon and low ash contents. Several studies investigated the valorisation of date pits into biochar via pyrolysis and the product applications as solid fuel, adsorbents, or soil enhancers. However, few have evaluated the economics of pyrolysis using date pits as feedstock. As such, this study presents a techno-economic assessment of biochar production from date pits. Proximate and elemental analyses of date pits samples are conducted in the laboratory, while a 500 tonne/day biorefinery is designed and evaluated using adapted pyrolysis prediction models. In addition, the process is evaluated in two modes: slow pyrolysis at 250 °C and fast pyrolysis at 400 °C. The obtained results are promising, where slow pyrolysis yielded 50 wt.% of biochar with 62 % elemental carbon content. Besides, it achieved a return on investment of 19% and a 5-year payback period. Whereas fast pyrolysis yielded 24 wt.% of biochar with 77 % elemental carbon content, while the return on investment is 17% with a 6-year payback period. Both modes of pyrolysis using date pits proved to be feasible in the MENA region for the production of high-quality biochar.

Keywords: Date pits, Pyrolysis, Biochar, Biocrude, MENA.

1. Introduction

The date palm (*Phoenix dactylifera*) is a tropical tree that has been cultivated in the MENA region for over 7000 years (Sarraf et al., 2021). Nearly 8 million tonnes of date fruits were produced in the MENA region in 2019 (Dates, Crops and livestock products., 2021). The date fruit is very rich in carbohydrates, salts, minerals, dietary fibers, vitamins, fatty acids, amino acids, and proteins. The fruit contains a large seed known as date pit which is generally used as animal feeds and soil conditioners. It is also used as a functional food and an adsorbent for removing pollutants from water. However, the usage of date pits for the aforementioned applications is little considering its massive availability in the MENA region, which is approximately 800,000 tonnes. It is mostly treated as a waste material. As both landfilling and burning of date pits, cause serious environmental problem, an environmentally benign route of valorising the date pits need to be explored. The valorisation of date pits for the production of biochar and bio-oil through pyrolysis has been found as a potential and sustainable solution for the date pits

accumulation issue by many researchers. Hence, in this study, the pyrolysis method of valorising date pits is considered.

Pyrolysis prediction models are commonly used to forecast pyrolysis product distribution. The models are based on the reactions that take place during the pyrolysis process. Many pyrolysis regulating parameters, such as heating rate and pyrolysis temperature, are considered during the simulation. In this way, several models have been developed by researchers to predict the pyrolysis process (Neves et al., 2011; Sharma et al., 2006). The pyrolytic kinetics model proposed by Song (Song, 2016) has been adapted in this study to predict the yields of bio-oil, biochar, and syngas yields. The model uses empirical equations to satisfy the mass and energy balances, and empirical relationships to represent the overall patterns of product distribution as a function of temperature.

The objective of the current prediction model is to predict product composition of date pits by two pyrolysis modes: slow and fast pyrolysis. Besides, the study investigates the economic feasibility of the two process' modes to produce a cost-effective biochar from date pits.

2. Materials and methods

2.1. Date pits preparation

The date pits are obtained from a local syrup producer and rinsed twice in water to eliminate dirt and impurities. The moist stones are then dried for around 24 hours at 323.15 K. The date pits are then crushed into powder in a blender and stored in an airtight container for subsequent analyses. An SDT-650 Thermogravimetry analyser is used for the proximate analysis, while a Euro-vector Euro EA 3000 CHN elemental analyser is used for the ultimate analysis. The proximate and ultimate analyses are carried out using ASTM D7582-12 and ASTM D 3176-8 standards, respectively.

2.2. Model development

A simplified process flow diagram of date pits pyrolysis is illustrated in Figure 1. The following assumptions guide the development of the prediction model. Date pit samples are pyrolysed in an isothermal reactor with a nitrogen inert environment. The drying of samples occurs at the start of the reaction, and the samples are pyrolysed and converted into water, bio-oil, ash, biochar, and volatile gases such as CO₂, CO, CH₄, and H₂ in the first stage of pyrolysis, as specified in equations 1, 2, and 3 (Swagathnath et al., 2019). The bio-oil is further cracked into volatiles again in the second step, which occurs as the temperature rises. Depending on the biomass composition, the generated gas products split into different gaseous substances during this phase.

$$\text{Total liquid yield} = Y_{\text{bio-oil},F} + Y_{\text{H}_2\text{O},F} + \text{Moisture content of samples} \quad (1)$$

$$\text{Total volatile gas yield} = Y_{\text{H}_2,F} + Y_{\text{CO},F} + Y_{\text{CH}_4,F} + Y_{\text{CO}_2,F} \quad (2)$$

$$\text{Total char yield} = 0.106 + 2.43 * \exp(-0.66 * T * 10^{-2}) \quad (3)$$

The composition of carbon, hydrogen, and oxygen in biochar is calculated using the following equations.

$$\text{Carbon content} = 0.93 - 0.92 * \exp(-0.42 * T * 10^{-2}) \quad (4)$$

$$\text{Hydrogen content} = (-0.41 * 10^{-2}) + (0.10 * \exp(-0.24 * T * 10^{-2})) \quad (5)$$

$$\text{Oxygen content} = 0.07 + 0.85 * \exp(-0.48 * T * 10^{-2}) \quad (6)$$

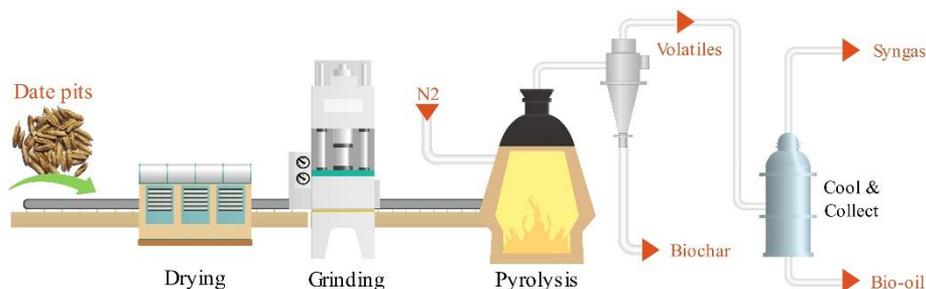


Figure 1: A simplified process flow diagram of date pits' pyrolysis.

2.3. Techno-economic assessment

The economic feasibility of date pits pyrolysis is conducted for a plant feed capacity of 20 tonne/hr and the assumptions highlighted in Table 1. The key pyrolysis equipment costs and labour requirement are adapted from earlier technical reports (Humbird et al., 2011; Wright et al., 2010). All prices are scaled up and inflated to the year 2019 using the Chemical Engineering Plant Cost Index (CEPCI).

Table 1: Assumptions for the techno-economic analysis.

Parameters	Values
Location of the proposed plant	Qatar
Analysis year	2019
Plant lifespan (years)	25
Discount rate (%)	20
Plant capacity (t/h)	20
Annual operating hours (h/y)	8000
Biomass price (\$/t)	100

3. Results and discussion

3.1. Date pits attributes

The proximate and ultimate attributes of the date pit samples are presented in Table 2.

Table 2: Proximate and ultimate analyses result of the date pit samples (this study).

Analysis	Composition (wt. %)
Proximate analysis (adb.%):	
Moisture	7.73
Volatile matter	68.11
Fixed carbon ^d	20.41
Ash	3.75
Ultimate analysis (daf.%):	
Carbon	46.30
Hydrogen	6.70
Nitrogen	0.83
Oxygen ^d	46.17
Sulphur	0.00
Chlorine	0.00

*adb.- air dried basis; ^d-calculated by difference; *daf- dry and ash-free basis

The date pit samples have a relatively high volatile content (68%) and a very low ash content (4%), indicating that the date pit could be an ideal feedstock for the pyrolysis process. Because of the low levels of nitrogen (~1%), and sulphur (~0%), NO_x and SO_x emissions from date pit pyrolysis are unlikely to exert any significant environmental burden.

3.2. Effect of pyrolysis type on pyrolysis products distribution

The effect of pyrolysis type on the pyrolysis products distribution of date pit samples is illustrated in Figure 2. Whereby, the slow pyrolysis produced a bio-oil yield of 19.6% while the fast pyrolysis generated a bio-oil yield of 37.7%. With respect to biochar yield, the slow and fast pyrolysis produced 50.2% and 24.6% yield, respectively. Both processes produced low syngas yields of 1.0% and 6.7%, respectively. The increased production of bio-oil and syngas in fast pyrolysis is due to its high temperature and heating rate as high temperatures and heating rates promotes secondary cracking of products leading towards more bio-oil and syngas production. A higher yield of biochar in slow pyrolysis is due to the process' low temperature, during which a significant amount of the biomass' volatile matter remains in solid state. Encinar *et al.* (Encinar et al., 2000) and Karaosmanoğlu *et al.* (Karaosmanoglu et al., 1999) also noted a similar trend of products distribution, whereby, an increased biochar production at low temperatures and increased bio-oil and syngas generation at higher temperatures have been reported.

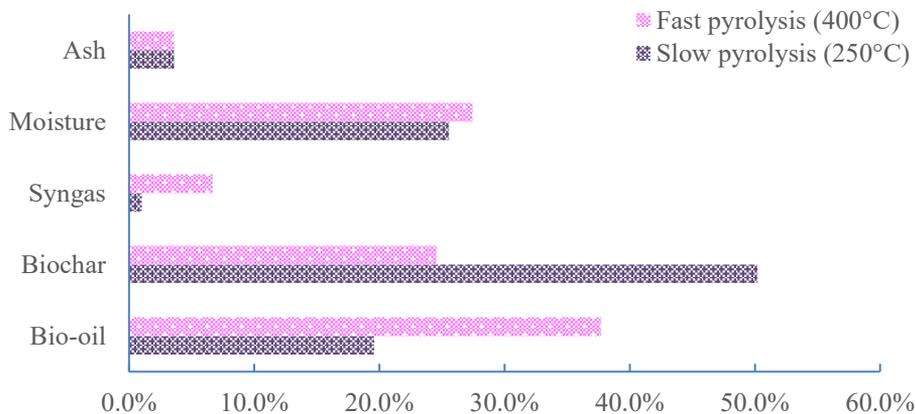


Figure 2: Slow and fast pyrolysis' product distribution of date pit samples

3.3. Effect of pyrolysis type on the biochar composition

The effect of pyrolysis type on the biochar composition of date pit samples is presented in Figure 3.

With the increase in temperature, an increase in the carbon content and at the same time a decrease in the hydrogen and oxygen contents can be observed. The increase in the carbon content is due to carbonisation or enrichment of char which is favoured at high temperatures. The decrease in the hydrogen and oxygen contents of char is attributed to the devolatilisation reactions which are dominant at temperatures between 350 and 600°C. The effect of temperature on the composition of biochar was also studied by Duman et al. (2011) and Bonelli et al. (2003). The aforementioned studies also indicated a similar increase in the carbon content and a decrease in the hydrogen and oxygen contents of char with the increase of operating temperature.

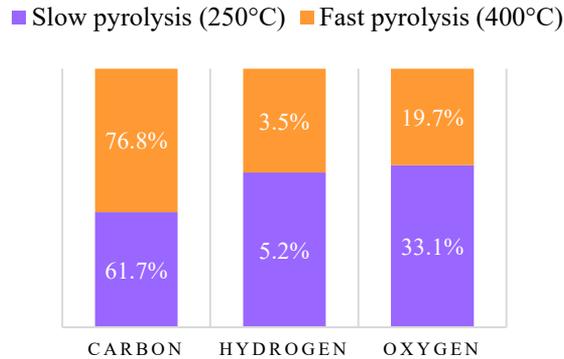


Figure 3: The effect of pyrolysis type on the biochar composition of date pit samples.

3.4. Techno-economic assessment

A summary of the technoeconomic assessment of both pyrolysis processes is presented in Table 3. Capital expenses (CAPEX) of fast pyrolysis is higher due to the need for bio-oil and syngas handling equipment (i.e. condensor, liquid and gas separator, liquid and gas filters, etc.), as fast pyrolysis produces higher quantities of bio-oil and syngas. Similarly, the operating expenses (OPEX) of fast pyrolysis is also higher, which is due to higher utility, labour, and maintenance cost associated with the handling of bio-oil and syngas. With respect to biochar production, slow pyrolysis offers a return on investment (ROI) of 17% as compared to an ROI of 16% for fast pyrolysis. In addition, the slow pyrolysis is offering a payback period of 5.3 years, whereas fast pyrolysis is offering 5.8 years to payback the capital investment.

Table 3: Summary of technoeconomic assessment of slow and fast pyrolysis.

Economic parameter	Slow pyrolysis (250°C)	Fast pyrolysis (400°C)
CAPEX (M US \$)	41.66	65.06
OPEX (M US \$/year)	21.09	21.67
Annual sales (M US \$/year)	29.00	32.86
Annual profit (M US \$/year)	7.91	11.19
Return on investment (%)	19%	17%
Payback period (Years)	5.3	5.8
Bio-oil LCOF (US \$/kg)	0.416	0.432
Biochar LCOF (US \$/kg)	0.206	0.249
Syngas LCOF (US \$/kg)	0.365	0.238

Conclusion

The availability of date pits in the Middle East and North Africa region (MENA) is enormous. The current study employed a prediction model to study the feasibility of pyrolysis of date pits considering the MENA scenario. The study considered two pyrolysis modes; slow (at 250°C) and fast (at 400°C) and investigated the effect of temperature on the product distribution of both pyrolysis processes. Besides, this study presented a techno-economic assessment of the two processes. The outcome of the study

is encouraging, indicating that pyrolysis of date pits is quite feasible in the MENA region. The techno-economic analysis indicates that slow pyrolysis is more ideal as far as biochar production is concerned, where it offers 19% return on investment and a 5-year payback period while the fast pyrolysis is offering 17% return on investment with a 6-year payback period. The outcome of this study provides an insight for the planning, design, and commercialisation of date pits-based pyrolysis plants in the MENA region.

Acknowledgments

This publication was made possible by GSRA grants no. GSRA7-1-0407-20016 from Qatar National Research Fund (a member of Qatar Foundation). The contents herein are solely the responsibility of the author[s].

References

- A.K. Sharma, M.R. Ravi, S. Kohli, 2006. Modelling product composition in slow pyrolysis of wood. *SESI J.* 16, 1–11.
- B. Song, 2016. Biomass pyrolysis for biochar production: kinetics, energetics and economics. *Biochar* 227–238.
- Dates, Crops and livestock products., 2021. , FAOSTAT.
- D. Humbird, R. Davis, L. Tao, C. Kinchin, D. Hsu, A. Aden, P. Schoen, J. Lukas, B. Olthof, M. Worley, D. Sexton, D. Dudgeon, 2011. Process Design and Economics for Conversion of Lignocellulosic Biomass to Ethanol: : Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover. United States. <https://doi.org/10.2172/1013269>
- D. Neves, H. Thunman, A. Matos, L. Tarelho, A. Gómez-Barea, 2011. Characterization and prediction of biomass pyrolysis products. *Prog. Energy Combust. Sci.* 37, 611–630. <https://doi.org/10.1016/J.PECS.2011.01.001>
- F. Karaosmanoglu, E. Tetik, E. Göllü, 1999. Biofuel production using slow pyrolysis of the straw and stalk of the rapeseed plant. *Fuel Process. Technol.* 59, 1–12. [https://doi.org/10.1016/S0378-3820\(99\)00004-1](https://doi.org/10.1016/S0378-3820(99)00004-1)
- G. Duman, C. Okutucu, S. Ucar, R. Stahl, J. Yanik, 2011. The slow and fast pyrolysis of cherry seed. *Bioresour. Technol.* 102, 1869–1878. <https://doi.org/10.1016/j.biortech.2010.07.051>
- G. Swagathnath, S. Rangabhashiyam, K. Parthasarathi, S. Murugan, P. Balasubramanian, 2019. Modeling Biochar Yield and Syngas Production During the Pyrolysis of Agro-Residues, in: *Green Buildings and Sustainable Engineering*. Springer, Singapore, Singapore, pp. 325–336. https://doi.org/10.1007/978-981-13-1202-1_28
- J.M. Encinar, J.F. González, J. González, 2000. Fixed-bed pyrolysis of *Cynara cardunculus* L. Product yields and compositions. *Fuel Process. Technol.* 68, 209–222. [https://doi.org/10.1016/S0378-3820\(00\)00125-9](https://doi.org/10.1016/S0378-3820(00)00125-9)
- M.M. Wright, DE, D.E. Daugaard, J.A. Satrio, R.C. Brown, 2010. Techno-economic analysis of biomass fast pyrolysis to transportation fuels. *Fuel* 89, S2–10.
- M. Sarraf, M. Jemni, I. Kahramanoğlu, F. Artés, S. Shahkoomahally, A. Namsi, M. Ihtisham, M. Brestic, M. Mohammadi, A. Rastogi, 2021. Commercial techniques for preserving date palm (*Phoenix dactylifera*) fruit quality and safety: A review. *Saudi J. Biol. Sci.* 28, 4408–4420. <https://doi.org/10.1016/J.SJBS.2021.04.035>
- P.R. Bonelli, E.G. Cerrella, A.L. Cukierman, 2003. Slow pyrolysis of nutshells: Characterization of derived chars and of process kinetics. *Energy Sources* 25, 767–778. <https://doi.org/10.1080/00908310390207819>

Design of sustainable processes for CO₂ capture

Ana Gabriela Romero-García,^{a,c} Nelly Ramírez-Corona,^b Eduardo Sánchez-Ramírez,^a Heriberto Alcocer-García,^a Cataldo De Blasio,^c Juan Gabriel Segovia-Hernández^{a*}

^a*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato, Gto., 36050, México.*

^b*Departamento de Ingeniería Química, Alimentos y Ambiental, Universidad de las Américas Puebla. ExHda. Santa Catarina Mártir s/n, San Andrés Cholula, Puebla, 72820, México*

^c*Åbo Akademi University, Faculty of Science and Engineering, Laboratory of Energy Technology, Rantakatu 2, Vaasa, 65100, Finland.*

**Corresponding author. Email: gsegovia@ugto.mx*

Abstract

Global warming is one of the main environmental problems affecting the world. The continuous emissions of Carbon Dioxide (CO₂) released into the atmosphere have generated a critical environmental situation worldwide. It is reported that 70% of CO₂ emissions are related to electricity production. For this reason, if there is an increase in energy demand therefore it is expected to increase CO₂ emissions. Globally, different alternatives have been sought to reduce the environmental footprint related to CO₂ emissions, highlighting the implementation of capture and storage plants as a promising alternative. In this work it is presented the multiobjective optimization for the simulation of a CO₂ capture plant coupled to a power plant considering a stochastic optimization having as objective function the minimization of the environmental implications (Ecoindicator 99), the control properties (Condition Number), as well maximize the economic indicator (Return on investment). The analysis considered the most used fuels in the power plant. From the results, the design with the best overall performance is when natural gas is burned. Presenting the lower environmental impact with 22549.43kEcopoints and a return on investment of 73.24%. This results in a sustainable process with less environmental impact.

Keywords: CO₂ Capture Plant, biogas, coal, natural gas, associated gas.

1. Introduction

Electricity production from the burning of fossil fuels, is one of the main sources of CO₂ emissions. Moreover, it is projected the growth of electricity demand. With this scenario, it is necessary to develop sustainable alternatives for power generation, as well as alternatives to mitigate CO₂ emissions. Due to global environmental problems, it is imperative to seek industrial processes in accordance with the United Nations Organization sustainable development goals for green chemistry and sustainability. An alternative to convert electricity production into a sustainable process is the implementation of CO₂ capture and storage plants (CCS). Studies on post-combustion capture technologies focus mainly on chemical absorption using aqueous mono-ethanolamine (MEA) solution. MEA is considered as a good choice of solvent due to its great capacity to capture CO₂ (Nagy, T. et al., 2013) and also because its properties are well

known which facilitates modeling studies (Wagner et al., 2013). Despite its high efficiency, MEA is considered highly toxic so that its implementation entails a high environmental impact. Moreover, no studies report a full design considering environmental impact and the process economies as a sustainable indicator.

In this work aims to present a novel proposal for the simulation of a CO₂ capture plant coupled to a power plant. In which it is presented the optimum design from a sustainable point of view, considering a stochastic optimization having as objective function the minimization of the environmental implications of capture plants using MEA as solvent (Ecoindicator 99), the minimization of Condition Number (CN) as an indicator of the control properties of the system, as well maximize the return on investment (ROI) as an economic indicator. The analysis considered the most used fuels in the power plant: biogas, coal, natural gas, and associated gas. All the cases were standardized to recover at least 95% of the CO₂ produced. The results indicate that for a fuel feed flow of 1000 kmol/h it was necessary to vary the air ratio to achieve the CO₂ recovery objective.

2. Study Case and Methodology

The study case presents a power plant coupled to a post-combustion CO₂ capture process (see Figure 1) considering the most used fuels for electricity production. The simulation was carried out in the software ASPEN PLUS. For the power plant it was considered a single-stage air compressor system as a simplified model for electricity production. According to Luyben (2013) it is possible to represent the electricity production in this simplified model considering the operating parameters as compression ratio, etc. As well, it is considered a fuel feed flow of 1000 kmol/h and a molar ratio of fuel and air of 1:30 to achieve the total combustion for all the fuels used in this study (Luyben, 2013). To describe the thermodynamic properties involved in the combustion, the Peng-Robinson method was used (Hasan, M. et al., 2012). For the CO₂ capture plant, a chemical absorption using an aqueous solution of Monoethanolamine (MEA) at 30% by weight as solvent was considered (Nagy, T. et al., 2013). The proportion of 30 wt% it is important to take in consider because MEA is strongly corrosive in high concentrations. As well, the viscosity also increases affecting the efficient of capture (Kohl, 1985).

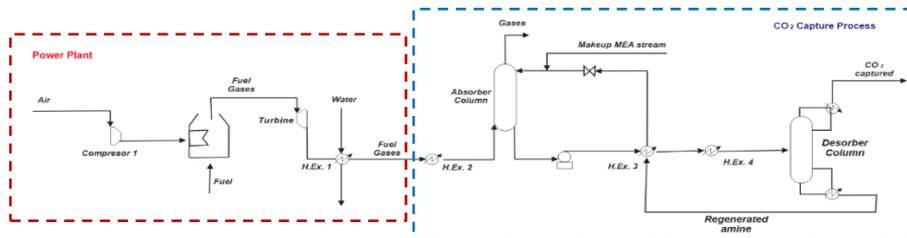


Figure 1. Scheme of a power plant coupled to a post-combustion CO₂ capture process using MEA 30wt% as solvent.

The capture process includes two columns, the absorption and desorption units. In the absorption the flue gases enter at the bottom of the column and the MEA at the top. Therefore, there is an interaction between the liquid and gas phases, driving the CO₂ to the liquid phase at the bottom of the absorber column. In the desorber column it is important to regenerate the solvent and make the desorption of the CO₂, which strongly depends on the reboiler duty. The capture process, is complex because the chemical absorption includes several dissociation reactions (Wagner et al., 2013). In order to model

this electrolyte behavior, it is important to consider all the kinetics as well as the binary interaction parameters. Likewise, the thermodynamic model Electrolyte Non-Random Two Liquids was implemented to consider the dissociation of ionic species. For the combustion process, the main design aspects are related to finding the amount of air to be fed and the optimum combustor pressure because of the trade-off between compressor work and turbine power (Luyben, 2013). The amount of air used to complete the combustion is very important. This factor directly influences the capture, as it is stated that the capture efficiency depends on the composition of flue gases. To ensure high capture efficiency, flue gases must be high in CO₂. For the capture process, some design aspects related to the process configuration and operational conditions should be taken into consideration to reduce the energy consumption and use of MEA. For the desorber column, in order to capture the maximum amount of CO₂ from the flue gas obtained from the power plant, it is necessary to manipulate the distillate flow and the reflux ratio. All the cases were standardized to a purity of 99 mol% CO₂ and at least 95% recovery of the CO₂ produced during the combustion. This directly affects the energy requirements of the process and therefore directly influences environmental and cost indicators.

The proposed model, has a significant number of variables to be optimized to have an optimal design and operation conditions, by aiming a maximum capture of CO₂, the best control criteria, maximum return on investment and lowest environmental impact. In this context, due to the electrolytic behaviour present in the liquid phase, the CO₂ capture process represents a highly non-linear model. With a certain number of discrete and continuous variables, the CO₂ capture model is suitable for optimizing. The design variables and the operation range, where chosen through a sensibility analysis of the process. The design variables for the study case are shown in Table 1.

Table 1. Design variables for the multi-objective optimization

Process	Variables	Type of Variable	Symbol	Range	Units
Combustion	Air Flow	Continue	Air	10 000-35 000	kmol/h
	Pressure Ratio	Continue	P _{com1}	1 – 8	---
	Compressor Combustion	Continue	P _{react}	4.5 – 9	atm
	Reactor Pressure	Continue	P _{turbine}	1 - 8	atm
Absorber	Turbine Pressure	Continue	P _{turbine}	1 - 8	atm
	Number of stages	Discrete	N _{abs}	4-99	---
	Fed Stage Diameter	Discrete	NF _{abs}	4-99	---
	Solvent Flow	Continue	D _{abs}	0.5 – 3.5	m
Desorber	Solvent Flow	Continue	MEA	43000 – 44 000	kmol/h
	Number of stages	Discrete	N _{des}	4-99	---
	Fed Stage Diameter	Discrete	NF _{des}	4-99	---
	Reflux Ratio	Continue	D _{des}	0.5 – 3.5	m
			RR	0.1 - 75	---

Reboiler Duty	Continue	Q	0.03 – 151.38	GJ/h
------------------	----------	---	---------------	------

2.1. Optimization Methodology

For the multi-objective optimization, it was employed the method of Differential Evolution with Tabu List (DETL) proposed by Srinivas and Rangaiah (2003). This technique works as a combined system between the biological evolution from Differential Evolution technique and the random search method from the Tabu search technique. The objective function involves three important axes to analyze: Return on Investment (ROI) as an indicator of the economy of the process by calculating return on investment due to the coupling of a capture plant to an existing power plant. Condition Number (γ^*) as an indicator of the dynamic behavior of the process considering low condition number for systems with good controllability and Eco-indicator 99 (EI99) to quantify the environmental impact derived from the coupling of the capture process, use of MEA as a solvent. According to Jiménez-González et al. (2011), these axes are needed in the optimization of a process to evaluate green chemistry. It is hoped that the analysis framework developed in this paper can contribute to the use of indices that can assess more than one aspect of green chemistry, in order to be easily integrated into a green process-based optimization. Similar works have been reported by Sánchez-Ramírez et al. (2017) and Contreras-Zarazúa et al. (2019) applying these indexes for distillation columns. The objective function is expressed in function of the variables described in Table 1 as is shown in Eq. (1).

$$\text{Min}[-ROI, \gamma^*, EI99] = f(\text{Air}, P_{\text{com1}}, P_{\text{react}}, P_{\text{turbine}}, N_{i,j}, NF_{i,j}, D_{i,j}, \text{MEA}, RR, Q) \quad (1)$$

Subject to:

$$\begin{aligned} y_{i,f} &\geq x_{i,f} \\ w_{i,f} &\geq u_{i,f} \end{aligned} \quad (2)$$

The objective function is restricted to satisfy the recuperation of at least 95% of the CO₂ produced during the combustion and also to achieve a purity of 99% mol of CO₂. Where $y_{i,f}$ represents the CO₂ recovered in the desorber column and $x_{i,f}$ represents the 95% of the CO₂ produced during the combustion. As well $w_{i,f}$ represents the purity achieved at the desorber column and $u_{i,f}$ represents the purity expected of at least 99% mol of CO₂. The optimization process was done through a visual basic hybrid platform in a Microsoft Excel macro. Where initially the values of the input vector are chosen randomly and sent to ASPEN PLUS. Where it is performed the simulation of the process and register that data as initial value in Excel. Then, the output vector data is sent to Matlab where, according to the methodology presented by Skogestad (2007), the value of the condition number is obtained. After that, the data is sent to Excel to obtain the ROI and EI99. Finally, the results are stored and it is generated a new vector. The values of parameters required by DETL method are: Population size: 120 individuals, Generation number: 1000, Tabu list size: 60, Tabu radius: 0.0001, Crossover fractions: 0.9, Mutation fractions: 0.3.

3. Results

In this section are presented the result obtained from the optimization. In order to analyze the operating conditions for each design it is important to highlight the composition of the flue gases obtained for each fuel used for the simulation. With a CO₂ molar fraction for Biogas, Coal, Natural Gas and Associated Gas of: 0.054, 0.124, 0.050 and 0.042

respectively. It is possible to observe that the flue gasses with the highest concentration of CO₂, is the one that is obtained from burning Coal, Biogas and Natural Gas. This concentration directly affects the parameters related to operational conditions as: solvent requirement and reboiler duty, which impact on economic and environmental indexes. Table 2 shows, the impact of some design parameters on the process performance. For the absorption, when the CO₂ concentration is low, the efficiency in the capture is affected. In order to capture a bigger amount of CO₂, the solvent requirement increases. Moreover, the reflux ratio is directly related to the energy requirements. As higher is the reflux ratio the energy requirements for the capture will increase.

Table 2. Design parameters for the best scenario of CO₂ capture for each fuel

Process	Variables	Biogas	Coal	Natural Gas	Associated Gas
Combustion	Air [kmol/h]	17133.95	20519.78	19602.76	24221.83
	P _{com1} [atm]	3.72	3.9	3.64	4.60
	P _{react} [atm]	5.13	6.46	7.46	6.79
	P _{turbine} [atm]	4.8	2.73	6.89	5.53
Absorber	N _{abs}	39	34	29	33
	NF _{abs}	3	3	3	3
	D _{abs}	3.2	1.18	2.59	2.76
	MEA [kmol/h]	43681	43258.06	43504.63	43509.26
Desorber	N _{des}	33	24	24	22
	NF _{des}	3	3	3	3
	D _{des} [m]	2.36	1.33	2.04	1.45
	RR	3.53	2.24	2.32	3.67
	Q[GJ/h]	470.08	210.50	364.45	491.86
	CO ₂ recovered [ton/h]	67.57	32.72	65.84	70.40
	Q/CO ₂ [GJ/ton]	6.9	6.4	5.5	6.9
CO ₂ Purity [%mol]	0.99	0.99	0.99	0.99	
Objective Function	ROI [%]	64.25	40.86	73.24	77.83
	γ*	20.17	2.39	6.7	34.82
	EI99 [kEcopoints]	28920.05	22615.53	22549.43	30369.51

For the analysis of the sustainability indexes, it can be observed that the better return on investment when implementing a capture process is when associated gas and natural gas are used as fuels, this trend is related to the topology of the process. For these two fuels, the process configuration is smaller impacting directly on construction costs. Also, as the flue gasses for those fuel have a bigger concentration in CO₂, the solvent requirements will be less in comparison with the other two fuels. It is possible to see that the least environmental impact occurs when natural gas is used as a fuel. This can be explained due to several factors, first of all it is the process the highest recovery of CO₂ decreasing the environmental food print related to CO₂ emissions. Likewise, the reduction in environmental impact can be explained because it is the process that has the lowest energy requirements and uses of solvents; these factors have a significant impact on the environmental weighting. On the other hand, it can be observed that the process where coal is used has an environmental score approximate to that obtained with natural gas. However, it is the process with the highest use of solvent, due to the low concentration of CO₂ in its flue gases. As well as the process in which there is a lower recovery of CO₂. About the control properties, it can be seen that the processes where coal and natural gas are used, are those that present best control properties by presenting lower condition

numbers. However, through the optimization process it is possible to obtain the best combination of design variables generating the design that meets the lowest environmental impact, highest return on investment and the best control properties for each fuel studied. The results obtained show that the process in which natural gas is used globally presents a balance in the three objectives of costs, environmental and control properties.

Conclusions

Implementing the CO₂ capture process in power plants has been considered the most mature technology to reduce flue gas emissions associated with electricity production. Although capture plants help to reduce the environmental footprint of CO₂. It is important to consider the type of fuel that is used for electricity production. Due to the CO₂ concentration contained in the flue gasses there will be variations on the design specifications of the capture process. However, before implementation, it is necessary to analyze the environmental implications that arise from the implementation, use of solvents and energy consumption. From the results obtained it can be concluded that the type of fuel is a high weight variable during the capture process, impacting directly on solvent and energy requirements. As the CO₂ concentration decreases; CO₂ recovery and capture efficiency decreases, but solvent and energy requirements are increased. There is also a huge area of opportunity in the study of new solvents; such as ionic liquids and deep eutectic solvents, to improve the efficiency of CO₂ capture, to replace the use of MEA as the main solvent for CO₂ capture and minimize the environmental impact that entails the use of such a toxic solvent.

References

- Contreras-Zarazúa G., Villicaña-García E., Cansino-Loeza B., Vázquez-Castillo J.A., Ponce-Ortega J. M., Segovia-Hernández J.G. 2019. Environmental impact and occupational hazard evaluation on intensified processes to produce diphenyl carbonate, *Computers & Chemical Engineering*, 122, 19-30.
- Hasan, M. F., Baliban, R. C., Elia, J. A., & Floudas, C. A. 2012. Modeling, simulation, and optimization of postcombustion CO₂ capture for variable feed concentration and flow rate. 1. Chemical absorption and membrane processes. *Industrial & Engineering Chemistry Research*, 51(48), 15642-15664.
- Jiménez-González C., Constable D. J. C. (2011). *Green Chemistry and Engineering: A practical Design Approach*. Wiley.
- Kohl, F.R.A., 1985. *Gas Purification*. Gulf Publishing Company
- Nagy, T., & Mizsey, P. 2013. Effect of fossil fuels on the parameters of CO₂ capture. *Environmental science & technology*, 47(15), 8948-8954.
- Luyben William L, L. William L. 2013. *Chemical Process Engineering Principles of Combustion Turbines*. *Energy & Fuels*, 27, 6316-6321. doi: 10.1021/ef401329s
- Sánchez-ramírez, E., Quiroz-ramírez, J. J., Hernández, S., Segovia-hernández, J. G., & Kiss, A. A. (2017). Optimal hybrid separations for intensified downstream processing of biobutanol, 185, 149–159.
- Skogestad, S. y I. Postlethwaite. (2007). *Multivariable feedback control. Analysis and Design*. John Wiley and Sons.
- Srinivas M., Rangaiah G. P. (2007). Differential evolution with TL for solving nonlinear and mixed-integer nonlinear programming problems, *Industrial & Engineering Chemistry Research*, 46, 7126-7135.
- Wagner, E.A.M., 2013. Solubility of carbon dioxide in aqueous solutions of monoethanolamine in the low and high gas loading regions. *J. Chem. Eng. Data* 58, 370–371

Water network optimisation in chemical complexes: a refinery case study

Francisco J.G. Patrocínio^{a,b}, Hugo M.D. Carabineiro^c, Henrique A. Matos^b and Nuno M.C. Oliveira^{a*}

^a*CIEPQPF, Dep. Chemical Engineering, Universidade de Coimbra, 3030-790 Coimbra, Portugal*

^b*CERENA, Dep. Chemical Engineering, IST, Universidade de Lisboa, 1049-001 Lisbon, Portugal*

^c*Petrogal, S.A., 7520-952 Sines, Portugal*

nuno@eq.uc.pt

Abstract

A two-level MINLP model is used for water network optimisation of large sites, composed of several independent plants. At the lowest level, the water consumption within a given plant is first minimised, considering feasible connections between the units of this plant, using a predefined set of interplant connections and capacities. During the second step, the water needs and wastewater emissions of each plant are coordinated, to minimise the network costs for the entire complex.

This strategy was applied to the water network of the Petrogal Sines refinery, with significant reductions in the network operating cost and freshwater consumption, using a mixture of local and global numerical solvers available in the GAMS system.

Keywords: Water networks, Process synthesis, Process design, Total site integration

1. Introduction

Water is of utmost importance in several sectors of the society, and industry consumption accounts for a substantial fraction of this resource's usage. Substantial work has therefore been conducted towards rationalising the design and operation of industrial water networks. Takama et al. (1980) introduced non-linear programming to achieve the optimal design of an oil refinery's total water network. More recent examples are Karuppiah and Grossmann (2006); Ahmetović and Grossmann (2011), introducing nonlinear and mixed-integer nonlinear programming formulations, to account for several particularities of these problems. Interplant designs were also proposed by Chew et al. (2008); Lv et al. (2018); Rubio-Castro et al. (2012); Azmi et al. (2020).

2. Problem statement

The problem of a multi-site industrial water network optimisation can be described by a set of freshwater sources ($a \in AS$), contaminants ($c \in CS$), plants ($f \in FS$), operations ($p \in PS$) and treatment units ($t \in TS$), and a set of discharge effluents ($e \in ES$). A typical design problem consists of devising a network which minimises the total or operational costs involved, considering all interconnections, flow rates and contaminant concentrations. Typical industrial networks include a large number of such units and specifications, and may result in computationally challenging problems. Using a common site arrangement as a set of multiple plants, the existing operations can be further classified as belonging to individual plants or common to the industrial complex. The proposed strategy divides the problem in two instances: first the optimal water network of each

plant is achieved separately (P1), and then the total site network is later optimised, to minimise the network costs for the entire complex (P2).

3. Mathematical formulation

3.1. Model P1 — individual plant network

Model P1 is the optimisation block for each individual plant, described by Eqs.(1–6) and illustrated in Figure 1. The objective function Eq. (1) in P1 is comprised by the expenses related with the utilisation of water from the different sources ($\text{cost}_a^{\text{FW}}$ — cost of freshwater from source a , and F_a^{FW} — total freshwater flow from source a). Eqs. (2–4) correspond to the plant's units total flow (units inlet f_p^{in} and outlet f_p^{out}), from source a ($F_{a,p}^{\text{W}}$), reused from unit p' ($F_{p',p}^{\text{RP}}$), unit discharge (F_p^{D}) and contaminant mass balances (inlet and outlet concentrations: $C_{p,c}^{\text{in}}$, $C_{p,c}^{\text{out}}$), with the possibility of water use from all sources and from the same plant's units outlet, although recycle to-and-from the same unit is not allowed ($p' \neq p$). Nonlinearities will arise in these equations, as flows (continuous variables) are multiplied by concentrations (continuous variables). In Eq. (5) operations are modelled as having a contaminant mass discharge, the possibility of an indissociable treatment section (process units' contaminant removal ($\text{rr}_{p,c}$), and an upper limit in the inlet concentration — Eq. (6). In the present formulation, f_p^{in} and f_p^{out} are fixed parameters; however, they can also be treated as variables, as there may be water consumption or formation during the operations. The solution of P1 for each individual plant provides updated inlet flow water needs for the total site problem considered by model P2.

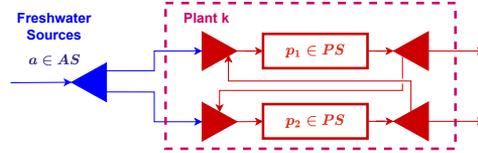


Figure 1: Superstructure of model P1.

$$\min \text{cost}_a^{\text{FW}} F_a^{\text{FW}} \quad (1)$$

$$\text{s.t.} \quad \sum_a F_{a,p}^{\text{W}} + \sum_{p' | p' \neq p} F_{p',p}^{\text{RP}} = f_p^{\text{in}}, \quad \forall p \in \text{PS} \quad (2)$$

$$f_p^{\text{out}} = \sum_{p' | p' \neq p} F_{p,p'}^{\text{RP}} + F_p^{\text{D}}, \quad \forall p \in \text{PS} \quad (3)$$

$$\sum_a F_{a,p}^{\text{W}} C_{a,c}^{\text{W}} + \sum_{p' | p' \neq p} F_{p',p}^{\text{RP}} C_{p',c}^{\text{out}} = f_p^{\text{in}} C_{p,c}^{\text{in}}, \quad \forall p \in \text{PS}, \quad \forall c \in \text{CS} \quad (4)$$

$$f_p^{\text{in}} C_{p,c}^{\text{in}} (1 - \text{rr}_{p,c}) + I_p = f_p^{\text{out}} C_{p,c}^{\text{out}}, \quad \forall p \in \text{PS}, \quad \forall c \in \text{CS} \quad (5)$$

$$C_{p,c}^{\text{in}} \leq C_{p,c}^{\text{in,max}}, \quad \forall p \in \text{PS}, \quad \forall c \in \text{CS} \quad (6)$$

3.2. Model P2 — total site network

Model P2 considers the entire complex, with plants, individual operations not belonging to plants, treatment units and multiple discharge effluents. The plants residual needs and wastewater specifications result from the solution of P1. Figure 2 represents the problem superstructure, and all the possible connections. This second model also reuses some equations from model P1.

Eqs. (7–9) correspond to the processes' total flow (from and to treatment units, $F_{t,p}^{\text{TP}}$ and $F_{p,t}^{\text{PT}}$), and flow to the final discharge ($F_{p,e}^{\text{PD}}$) and contaminant mass balances (here $C_{t,c}^{\text{out}}$ is the outlet concentration from treatment t). Note that reutilising water from the same unit, plant, as well as water from

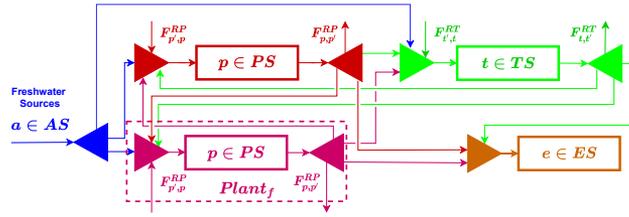


Figure 2: Superstructure of model P2.

the discharges, is not permitted. A new set MP contains the allowed process matches, accordingly to the plant arrangement. As in P1, nonlinearities are present in the contaminant mass balances, as flows are multiplied by concentrations.

$$\sum_a F_{a,p}^W + \sum_{p'|(p,p') \in \text{MP}} F_{p',p}^{\text{RP}} + \sum_t F_{t,p}^{\text{TP}} = f_p^{\text{in}}, \quad \forall p \in \text{PS} \quad (7)$$

$$\sum_a F_{a,p}^W C_{a,c}^W + \sum_{p'|(p,p') \in \text{MP}} F_{p',p}^{\text{RP}} C_{p',c}^{\text{out}} + \sum_t F_{t,p}^{\text{TP}} C_{t,c}^{\text{out}} = f_p^{\text{in}} C_{p,c}^{\text{in}}, \quad \forall p \in \text{PS}, \quad \forall c \in \text{CS} \quad (8)$$

$$f_p^{\text{out}} = \sum_{p'|(p,p') \in \text{MP}} F_{p,p'}^{\text{RP}} + \sum_t F_{p,t}^{\text{PT}} + \sum_e F_{p,e}^{\text{PD}}, \quad \forall p \in \text{PS} \quad (9)$$

Operations are modelled as having a contaminant mass discharge and an embedded treatment section, as in model P1, in Eqs. (5–6).

Similar to process units, Eqs. (10–12) formulate the flows of the treatment units (flow from treatment to discharge, $F_{t,e}^{\text{TD}}$) and mass balances (inlet concentration in treatment, $C_{t,c}^{\text{in}}$). Reutilisation from and to the same treatment is restricted ($t \neq t'$). Contaminant removal ($\text{rr}_{t,c}$), in the treatment units, is achieved through the units' inlet concentration and the units' removal rate — Eq. (13). Effluent flow (F_e^D) and mass balances are inserted for each discharge e in Eqs. (14–15), and the resulting concentration $C_{e,c}^D$ is limited to an upper bound $C_{e,c}^{D,\text{max}}$.

$$\sum_a F_{a,t}^W + \sum_{t'} F_{t',t}^{\text{RT}} + \sum_p F_{p,t}^{\text{PT}} = F_t, \quad \forall t \in \text{TS} \quad (10)$$

$$F_t = \sum_{t'} F_{t,t'}^{\text{RT}} + \sum_t F_{t,p}^{\text{TP}} + \sum_e F_{t,e}^{\text{TD}}, \quad \forall t \in \text{TS} \quad (11)$$

$$\sum_a F_{a,t}^W C_{a,c}^W + \sum_{t'} F_{t',t}^{\text{RT}} C_{t',c}^{\text{out}} + \sum_p F_{p,t}^{\text{PT}} C_{p,c}^{\text{out}} = F_t C_{t,c}^{\text{in}}, \quad \forall t \in \text{TS}, \quad \forall c \in \text{CS} \quad (12)$$

$$C_{t,c}^{\text{in}} (1 - \text{rr}_{t,c}) = C_{t,c}^{\text{out}}, \quad \forall t \in \text{TS}, \quad \forall c \in \text{CS} \quad (13)$$

$$\sum_p F_{p,e}^{\text{PD}} C_{p,c}^{\text{out}} + \sum_t F_{t,e}^{\text{TD}} C_{t,c}^{\text{out}} = F_e^D, \quad \forall e \in \text{ES} \quad (14)$$

$$\sum_p F_{p,e}^{\text{PD}} C_{p,c}^{\text{out}} + \sum_t F_{t,e}^{\text{TD}} C_{t,c}^{\text{out}} = F_e^D C_{e,c}^D, \quad \forall e \in \text{ES}, \quad \forall c \in \text{CS} \quad (15)$$

$$C_{e,c}^D \leq C_{e,c}^{D,\text{max}}, \quad \forall e \in \text{ES}, \quad \forall c \in \text{CS} \quad (16)$$

Effluent costs can be defined accordingly to the discharge concentrations, partitioned into several class intervals ($\delta_{e,c,k}^{\text{cls}}$), with $k \in \text{KS}$, and tarified accordingly to the highest activated concentration. That instance is modelled by defining a subset, corresponding to the special effluent $\text{SES} \in \text{ES}$ — Eq. (17). The highest activated class is multiplied by the discharged flow, resulting in that effluent cost cost_e , in Eq. (18). To avoid nonlinearities, Eq. (18) is further reformulated and linearised.

Eq. (19) computes the network costs, to be minimised similarly to Eq. (1).

$$C_{e,c}^D - (C_{e,c}^{D,\max} + \varepsilon) \delta_{e,c,k}^{\text{cls}} \leq C_{e,c,k}^{\max}, \quad \forall e \in \text{SES}, \quad \forall c \in \text{CS}, \quad \forall k \in \text{KS} \quad (17)$$

$$\text{cost}_e \geq \text{cost}_k \delta_{e,c,k}^{\text{cls}} F_e^D, \quad \forall e \in \text{SES}, \quad \forall c \in \text{CS}, \quad \forall k \in \text{KS} \quad (18)$$

$$\text{cost}^{\text{net}} = \sum_a \text{cost}_a^{\text{FW}} F_a^{\text{FW}} + \sum_t \text{cost}_t F_t + \sum_{e|e \notin \text{SES}} \text{cost}_w F_e^D + \sum_{e|e \in \text{SES}} \text{cost}_e \quad (19)$$

4. Initialisation

A good initial guess is frequently of major importance for local optimisation of nonlinear problems (NLP) and mixed integer non-linear problems (MINLP). To initialise the full network problem P2, two distinct procedures were tested: solving a simpler NLP for the entire network or a linear programming (LP) sequence strategy, where each section of the problem is analysed separately (Figure 3).



Figure 3: LP sequence initialisation strategy.

4.1. NLP initialisation strategy

This technique considers a simpler NLP, by removing the constraint inequalities from the original formulation and replacing the objective with a new function — Eq. (20), minimising the network infeasibilities resulting from a crude network initialisation. Eq. (20) can then be reformulated using suitable smoothing approximations (e.g., based on complementarity analysis (Gopal and Biegler, 1999)). The solution reached is then used as an initial solution for the original MINLP.

$$\min \sum_{c,p} \max\{0, C_{p,c}^{\text{in}} - C_{p,c}^{\text{in},\max}\} + \sum_{e,c} \max\{0, C_{e,c}^D - C_{e,c}^{D,\max}\} \quad (20)$$

$$\text{s.t.} \quad \text{eqs. (5,7–15)} \quad (21)$$

Note that during the application of this strategy, local NLP solvers may not reach an objective value of zero, even when the original problem is feasible. In either case, the solution obtained is considered as the initialisation for the subsequent MINLP solvers.

4.2. LP sequence initialisation strategy

The proposed strategy implies solving each of the following subproblems with LP models.

4.2.1. Water Usage Network (WUN) initialisation

Teles et al. (2008) suggest a combinatorial LP approach to initialise WUN problems. Each operation is dealt individually within an LP block, considering information from the previous units. Zhao et al. (2016) use this methodology, but instead of a combinatorial arrangement they consider the concentration potentials concept (CPDs) (Liu et al., 2009), to select the LP blocks sequence (Figure 4).

4.2.2. Discharge allocation

In order to initialise the water treatment network (WTN) section of P2, the strategy of Castro et al. (2009) was employed.

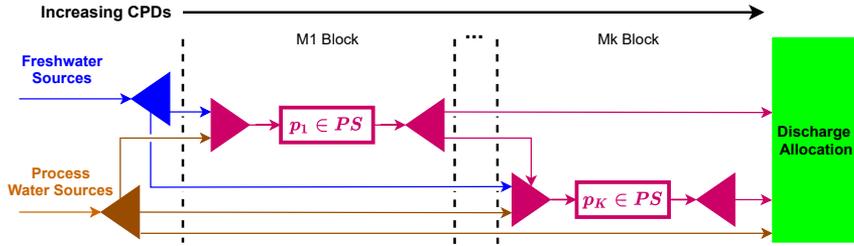


Figure 4: WUN initialisation strategy, based on Zhao et al. (2016).

This method has a limited capacity of dealing with multiple discharges and respective constraints. A LP was designed to allocate wastewater sources ($s \in \text{WWS}$) resulting from the WUN initialisation to the effluent e , accordingly to the discharge constraints and costs, in Eq. (22). The routine has the possibility of diluting the effluents with freshwater ($F_{a,e}^{\text{WD}}$) to eliminate infeasibilities — Eq. (23). Eq. (24) ensures that the wastewater sources allocated to the effluents ($F_{s,e}^{\text{WW}}$). Eq. (25) ensures that the discharge is within the allowed constraints. This equation is particularly relevant as $T \in \text{ES}$, requiring a previous match of the treatments ($t \in \text{TS}$) with each effluent ($e \in \text{ES}$).

$$\min \sum_e \text{cost}_e F_e^D + \sum_{(a,e)} F_{a,e}^{\text{WD}} \quad (22)$$

$$\text{s.t. } f_s^{\text{WW}} = \sum_e F_{s,e}^{\text{WW}}, \quad \forall s \in \text{WWS} \quad (23)$$

$$\sum_s F_{s,e}^{\text{WW}} + \sum_a F_{a,e}^{\text{WD}} = F_e^D, \quad \forall e \in \text{ES} \quad (24)$$

$$\sum_a F_{a,e}^{\text{WD}} C_{a,c}^W + \sum_s F_{s,e}^{\text{WW}} C_{s,c}^{\text{WW}} \prod_t (1 - \text{rr}_{t,c}) \leq F_e^D C_{e,c}^{\text{env}}, \quad \forall e \in \text{ES}, \quad \forall c \in \text{CS} \quad (25)$$

4.2.3. Water Treatment Network (WTN) initialisation

Castro et al. (2009) employ a sequence of LP blocks, each one handling a different treatment unit, and utilising information relative to the previous LPs and the subsequent ones; that strategy can be used to initialise WTN problems. All operations sequences are tested (using a combinatorial procedure), and the arrangement with the minimum objective value is selected. This method becomes prohibitive for medium-large problems. Instead, as recommended by Fan et al. (2019), for WTN problems, units were sorted by decreasing treatment cost, and the resulting sequence was inserted into the method of Castro et al. (2009). This strategy was applied independently for each effluent e , considering the treatment units allocated to that effluent line.

5. Application example

The Galp Sines refinery is an industrial complex located in the south of Portugal, with a production cap of 220 kbpd, accounting for approximately 70% of the country's refining capacity. The existing water network was optimised, employing several initialisation methods: using a reference refinery network configuration, and the previous sections techniques (Table 1). The version of the problem considered used 5 classes of contaminants, 3 individual plants involving 21 operations, one source of raw fresh water, and 2 discharge lines, with dynamic tariffs. The local MINLP solver SBB, and the global MINLP solver ANTIGONE were used with GAMS 33.2, in a AMD Ryzen 7 4800H computer with 16 GB of RAM, considering a maximum limit of 3600 s of CPU time.

Method	Fixed initialisation solution (€h ⁻¹)	MINLP local solver solution (€h ⁻¹)	MINLP global solver solution (€h ⁻¹)	Optimality gap (%)
No initialisation	—	Infeasible		11.4
Reference network	950.9	812.8	800.2	4.9
NLP initialisation	1591.1	1151.3		11.7
LP sequence initialisation	916.7	800.2		4.7

Table 1: Application results to the Sines refinery network.

The procedure resulted in reductions of 15.8% and 7% in cost and freshwater consumption respectively, increasing the environmental sustainability of the complex. Due to the extra effort incurred during the initialisation phase, the LP sequence initialisation displayed one of the lower gaps in the final solutions reached among the techniques utilised.

6. Conclusion

The application of methodology developed provided an efficient method for the optimisation of the industrial example considered. Within this problem, the impact of the initial solution considered is noticeable in the final results, both from the local MINLP solver and in the final global optimality gap. These results stress the importance of the availability of good initialisation techniques for the efficient solution of similar problems of comparable complexity.

Acknowledgements: The authors gratefully acknowledge the financial support provided by Galp Energia and Fundação para a Ciência e Tecnologia, through Ph.D. grant PD/BDE/142833/2018.

References

- E. Ahmetović, I. E. Grossmann, 2011. Global superstructure optimization for the design of integrated process water networks. *AIChE Journal* 57 (2), 434–457.
- M. Azmi, M. Arifin, Z. Handani, 2020. Optimization of inter-plant water network design involving multiple contaminants. *IOP Conference Series: Materials Science and Engineering* 736, 022078.
- P. Castro, J. Teles, A. Novais, 2009. Linear program-based algorithm for the optimal design of wastewater treatment systems. *Clean Technologies and Environmental Policy* 11, 83–93.
- I. M. L. Chew, R. Tan, D. K. S. Ng, D. C. Y. Foo, T. Majoji, J. Gouws, 2008. Synthesis of direct and indirect interplant water network. *Industrial & Engineering Chemistry Research* 47 (23), 9485–9496.
- X.-Y. Fan, J. J. Klemeš, X. Jia, Z.-Y. Liu, 2019. An iterative method for design of total water networks with multiple contaminants. *Journal of Cleaner Production* 240, 118098.
- V. Gopal, L. T. Biegler, 1999. Smoothing methods for complementarity problems in process engineering. *AIChE Journal* 45 (7), 1535–1547.
- R. Karuppiah, I. Grossmann, 2006. Global optimization for the synthesis of integrated water systems in chemical processes. *Computers & Chemical Engineering* 30 (4), 650–673.
- Z.-Y. Liu, Y. Yang, L.-Z. Wan, X. Wang, K.-H. Hou, 2009. A heuristic design procedure for water-using networks with multiple contaminants. *AIChE Journal* 55 (2), 374–382.
- Z. Lv, Y. Song, C. Chen, B. Jiang, H. Sun, Z. Lyu, 2018. A novel step-by-step optimization method for interplant water networks. *Journal of Environmental Management* 213, 255–270.
- E. Rubio-Castro, J. Ponce-Ortega, M. Serna-González, M. El-Halwagi, 2012. Optimal reconfiguration of multi-plant water networks into an eco-industrial park. *Computers & Chemical Engineering* 44, 58–83.
- N. Takama, T. Kuriyama, K. Shiroko, T. Umeda, 1980. Optimal water allocation in a petroleum refinery. *Computers & Chemical Engineering* 4 (4), 251–258.
- J. Teles, P. Castro, A. Novais, 2008. Lp-based solution strategies for the optimal design of industrial water networks with multiple contaminants. *Chemical Engineering Science* 63 (2), 376–394.
- H.-P. Zhao, Z.-F. Wang, T.-C. Chan, Z.-Y. Liu, 2016. Design of regeneration recycling water networks by means of concentration potentials and a linear programming method. *Journal of Cleaner Production* 112 (5), 4667–4673.

Optimal design of solar-aided hydrogen production process using molten salt with CO₂ utilization for polypropylene carbonate production

Wanrong Wang, Nan Zhang, and Jie Li*

Centre for Process Integration, Department of Chemical Engineering and Analytical Science, The University of Manchester, Manchester M13 9PL, UK

Abstract

In this work, optimal design for solar steam methane reforming using molten salt (SSMR-MS) with CO₂ capture and utilisation is investigated by employing a machine-learning based optimisation framework. The results obtained show that a 65.77 % reduction can be obtained compared with the existing SSMR-MS. The optimal Levelised Cost of Hydrogen Production (LCHP) is 1.26 \$ kg⁻¹ which represents about half of the LCHP in the existing SSMR-MS. The captured CO₂ can produce around 485.73 kt polypropylene carbonate annually.

Keywords: Hydrogen production; solar energy; CO₂ utilization; Machine learning; Optimization

1. Introduction

Global warming has been considered as the biggest environmental problem in this fast developing world. Anthropogenic CO₂ emission by using fossil fuel plays a key role in the global temperature rise as seen today. Therefore, it is urgent to shift to low-carbon technologies. To achieve this target, using renewable energies, clean fuels such as hydrogen are considered as longer-term solutions and carbon capture, utilization and storage is considered as short- to mid-term solution. Among all these solutions, hydrogen shows off excellent potentials as a commercial clean fuel. Furthermore, it is also an essential feedstock in many chemicals such as ammonia (Ohs et al., 2019). In 2020, the European Commission have stated that hydrogen economy is an important part for the economy recovery after COVID-19. However, conventional hydrogen production primarily utilizes natural gas and oil-based feedstock for steam reforming which contributes to large amount of CO₂ emissions. To reduce CO₂ emission, using renewable energy and carbon capture and utilisation in hydrogen production process should be investigated significantly (Alper et al., 2017).

Among the renewables, solar hydrogen production has received great attention in recent years due to its high capability to fulfill global energy demands (Koumi Ngoh et al., 2012). Wang et al. has studied the optimisation of solar steam methane reforming using molten salt (SSMR-MS) to reduce TAC and CO₂ emissions (Wang et al., 2021). In their work, it is shown that the optimal LCHP is still higher than the conventional methane steam reforming. What's more, the CO₂ removal model is designed at a constant separation efficiency, which could lead to inaccurate cost and utility calculation for this unit. Considering the coke formation in process, the lower bound of steam to methane ratio should also be adjusted. Therefore, an integrated rate-based CO₂ removal model in

SSMR-MS along with CO₂ utilization for polypropylene carbonate production is investigated in this work. This is the main novelty of this work.

In this work, the employment of machine learning in process optimisation from (Wang et al., 2021) is extended for optimal design of SSMR-MS with integration of CO₂ capture and utilization. The artificial neural network (ANN) is used to address this problem by obtaining functions that approximates the deterministic functions of TAC, hydrogen production rate, molten salt duty and gas flowrate from CO₂ capture unit with process variables in SSMR-MS. Then the problem is solved using a hybrid optimisation technique. The solution is then validated in Aspen Plus V8.8 and SAM. The most striking result is that there is a significant reduction in TAC by 65.77 % and CO₂ emissions by 68.97 % compared to the existing SSMR-MS process. With 485.73 kt polypropylene carbonate annually, the profit is considerable. The Levelised Cost of Hydrogen Production (LCHP) reduces from 1.26 \$ kg⁻¹ to 2.40 \$ kg⁻¹ compared to the optimal case in Wang et al. (2021).

2. Problem description

As shown in Figure 1, the SSMR-MS integrated with rate-based CO₂ capture model and polypropylene carbonate production process is developed in this work. There is a detailed description of SSMR-MS can be found in Wang et al. (2021). The pre-reforming process is non-adiabatic with a co-current flow mode. Solar energy is delivered using molten salt as media to the hydrogen production process. Methyl diethanolamine (MDEA) is used for CO₂ absorption and solution regeneration has the same equations as those in Moiola et al. (2016). Polypropylene carbonate is produced using the captured CO₂.

The objective is to minimize total annualized cost (TAC) of the integrated system.

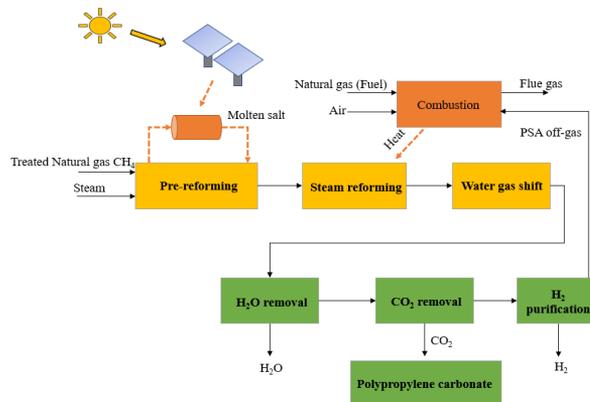


Figure 1: Block diagram of the integrated system (revision from Wang et al., 2021)

3. Mathematical formulation

The integrated system using rigorous models is modelled in Aspen Plus V8.8. However, the evaluation of the objective in this model as well as constraints is computationally expensive. To reduce complexity, surrogate model is applied to obtain a function $\hat{f}(X)$ that approximates the function f . ANN is employed as surrogate model for optimisation. ANN is particularly effective for modelling high-dimensional and highly non-linear

problems, because of its ability to learn and its usefulness in nonlinear processing compared with other techniques (Ibrahim et al., 2018).

In this work, a rigorous rate-based CO₂ removal model for CO₂ capture is firstly developed in Aspen Plus V8.8. To integrate the CO₂ removal model to SSMR-MS process, a surrogate model using ANN is constructed because of the difficulty in convergence in CO₂ removal model. To construct the ANN surrogate model, input variables and output variables should be selected to describe the rigorous CO₂ removal process precisely. The input variables include stream component flowrate of for CH₄, H₂O, CO, CO₂, H₂ which are denoted as $F_{in,CH_4,MDEA}$, $F_{in,H_2O,MDEA}$, $F_{in,CO,MDEA}$, $F_{in,CO_2,MDEA}$ and $F_{in,H_2,MDEA}$, respectively and stream temperature ($T_{in,MDEA}$). The input variables are organised as below,

$$\mathbf{z} = [F_{in,CH_4,MDEA}, F_{in,H_2O,MDEA}, F_{in,CO,MDEA}, F_{in,CO_2,MDEA}, F_{in,H_2,MDEA}, T_{in,MDEA}]^T.$$

The output variables include the outlet stream flowrates of CH₄, H₂O, CO, CO₂ and H₂ in the CO₂ removal process which are predicted using individual ANN surrogate models, as shown in Eqs. 1-5.

$$F_{out,CH_4,MDEA} = ANN_1(\mathbf{z}) \quad (1)$$

$$F_{out,H_2O,MDEA} = ANN_2(\mathbf{z}) \quad (2)$$

$$F_{out,CO,MDEA} = ANN_3(\mathbf{z}) \quad (3)$$

$$F_{out,CO_2,MDEA} = ANN_4(\mathbf{z}) \quad (4)$$

$$F_{out,H_2,MDEA} = ANN_5(\mathbf{z}) \quad (5)$$

These constructed ANN surrogate models are implemented in the rigorous SSMR-MS process by using the user model in Aspen Plus through Excel Link (Fontalvo, 2014). Then, this new SSMR-MS process is used to construct a new surrogate model for optimisation through extending the optimisation framework of Wang et al. (2021). Latin hypercube sampling method is used for sample generation.

In the integrated system, the independent input variables including molar flowrate of natural gas into pre-reformer F_{NG} , steam to methane ratio $\gamma_{S/C}$, operating temperature of reformer T_R , high-temperature water gas shift (HWGS) reactor T_{HWGS} , low-temperature water gas shift (LWGS) reactor T_{LWGS} , tube length of pre-reformer L_{PR} , reformer L_R , HWGS reactor L_{HWGS} and LWGS reactor L_{LWGS} , tube number in pre-reformer N_{PR} , reformer N_R , HWGS reactor N_{HWGS} and LWGS reactor N_{LWGS} vary between lower and upper bounds. The input variables are organised as a vector \mathbf{x} as below,

$$\mathbf{x} = [F_{NG}, \gamma_{S/C}, T_R, T_{HWGS}, T_{LWGS}, L_{PR}, L_R, L_{HWGS}, L_{LWGS}, N_{PR}, N_R, N_{HWGS}, N_{LWGS}]^T.$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U \quad (6)$$

The objective function TAC can be calculated as follows,

$$TAC = C_{capital} \cdot ACCR + C_{production} \quad (7)$$

where $C_{capital}$ is total capital investment. $ACCR$ is annual capital charge ratio. $C_{production}$ is the total production cost per year.

The optimization problem using the surrogate models is stated as follows,

$$(PS) \quad Min \quad TAC = TAC_1 + TAC_{solar}$$

$$\begin{aligned}
 \text{s.t.} \quad & TAC_1 = ANN_6(\mathbf{x}) + ANN_7(\mathbf{x}) \\
 & TAC_{solar} = f(Q_{MS}) \\
 & Q_{MS} = ANN_8(\mathbf{x}) \\
 & F_{H_2} = ANN_9(\mathbf{x}) \geq F_{H_2}^{TA} \\
 & \text{Eq. (6)}
 \end{aligned}$$

where TAC_1 is non-solar related cost, TAC_{solar} is the solar related cost. $ANN_6(\mathbf{x})$ is CO_2 removal process related cost. ANN_7 is the non-solar related cost excluding MDEA unit. \mathbf{x} is the set of independent variables in hydrogen production process, Q_{MS} is molten salt duty. The relationship of solar-related equipment cost, and molten salt duty is described using an algebraic linear function $f(Q_{MS})$. The surrogate model comprises 4 artificial neural networks as indicated above in the optimization problem and a linear regression model $f(Q_{MS})$. F_{H_2} denotes the predicated flowrate of hydrogen.

4. Solution algorithm

As illustrated in Figure 2, a hybrid optimisation algorithm is employed to solve the optimisation problem PS (Wang et al., 2021). Three platforms are employed in this work to reduce the computational complexity. During sample generation process, Matlab is used to generate the sample points and import the sample points to Aspen Plus. The output variables are collected from Aspen Plus simulation results. For each sample generation, the new SSMR-MS process in Aspen Plus would call Visual Basic Application (VBA) in Excel (Fontalvo, 2014) to transfer data from the ANN surrogate model constructed for CO_2 removal model. The hybrid solution algorithm is implemented in MatLab R2019a.

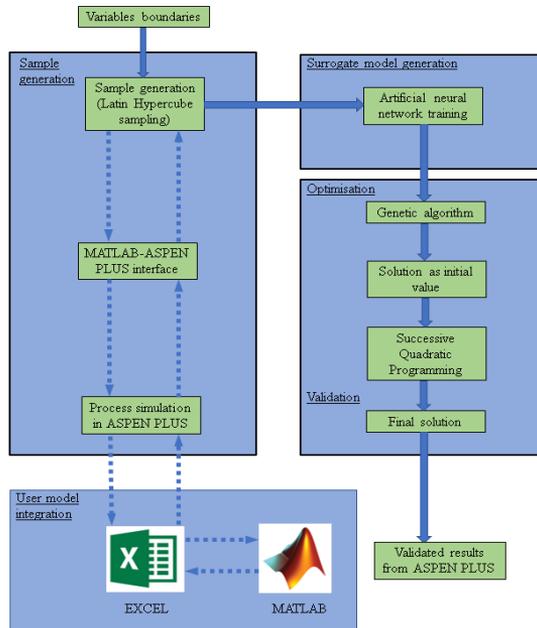


Figure 2: Flowchart of the extended design methodology

5. Computational studies

The results obtained using the extended optimisation framework are shown in Table 1. The required hydrogen production rate is 2, 577 kmol h⁻¹ with 99.9 vol% hydrogen purity. Other data can be referred to Wang et al. (2021). It can be observed that the optimal TAC is 166.52 M\$ y⁻¹. The optimal steam to methane ratio is 2.7.

Table 1: Optimization results from the surrogate models

Item	Optimal value
$Y_{S/C}$	2.7
T_R (°C)	963.0
T_{HWGS} (°C)	379.1
T_{LWGS} (°C)	199.1
L_{PR} (m)	11.1
L_R (m)	11.9
L_{HWGS} (m)	4.6
L_{LWGS} (m)	4.5
N_{PR}	4,021
N_R	50
N_{HWGS}	2159
N_{LWGS}	2036
F_{NG} (kmol h ⁻¹)	786.6
Q_{MS} (MW)	14.61
F_{H_2} (kmol h ⁻¹)	2,577
TAC (M\$ y ⁻¹)	166.52

The optimal results of the independent variables from surrogate models are used as inputs in Aspen Plus V8.8 to calculate all dependent variables for validation. The validated results for Q_{MS} , F_{H_2} and TAC are 14.37 MW, 2577.6 kmol h⁻¹, 165.46 M\$ y⁻¹ respectively. After comparison, it shows that the largest difference between validation results and predicted results from surrogate models is within 1 %. That means the ANN surrogate model has high accuracy for prediction.

Heat integration is conducted for heat recovery in hydrogen production process. The final results are provided in Table 2. After heat integration, a 6.7 % reduction on TAC (155.33 M\$ y⁻¹) is obtained. With integration of poly production, TAC is further significantly reduced to 41.86 M\$ y⁻¹ because of the high profit of poly production.

As shown in Table 2, a comparison is also made between optimal results in this work and the best results from Wang et al. (2021) and the conventional steam methane reforming (SMR). Without poly production, TAC in this work is 33.03 M\$ y⁻¹ higher than that from Wang et al. (2021). The reason is that with the integration of rate-based CO₂ removal model, the cost related to the CO₂ capture unit is increasing. However, with poly production, the TAC decreases 65.77 % which is significantly lower than that in Wang et

al. (2021). As a consequence, LCHP reduces 47.50 % along with the reduction on TAC. CO₂ emission drops from 423.90 kt y⁻¹ to 131.53 kt y⁻¹ and the captured CO₂ can produce 485.73 kt polypropylene carbonate annually.

What is the most striking is that the LCHP in this work is 37.00 % lower than the conventional SMR and CO₂ emissions decreases from 502.90 kt y⁻¹ to 131.53 kt y⁻¹. Through integrating solar energy and CO₂ utilisation into the hydrogen production process, it indicates that the hydrogen can be used in a renewable, sustainable and economical manner.

Table 2: Comparative optimization results

Item	Optimal Case	Wang et al. (2021)	Conventional SMR
Q_{MS} (MW)	14.37	10.20	20.00
F_{H_2} (kmol h ⁻¹)	2,577.6	2577.30	2,577.0
TAC without poly production (M\$ y ⁻¹)	155.33	122.30	90.90
TAC with poly production (M\$ y ⁻¹)	41.86	-	-
LCHP (\$ kg ⁻¹)	1.26	2.40	2.00
CO ₂ emission (kt y ⁻¹)	131.53	423.90	502.90
Polypropylene carbonate (kt y ⁻¹)	485.73	0.00	0.00

6. Conclusion

In this paper, the optimisation-based framework using machine learning techniques is extended for optimal design of SSMR-MS integrated with CO₂ capture and utilization for large-scale hydrogen production. The computational results demonstrate a significant reduction in TAC by around 65.77 % can be achieved compared with the existing SSMR-MS (Wang et al., 2021). Captured CO₂ can produce around 485.73 kt yr⁻¹ polypropylene carbonate. The LCHP is 37.00% lower than the conventional SMR. In the future, more process options for different pre-reformer operating conditions are expected to evaluate.

References

- Alper, E. and Yuksel Orhan, O., 2017, CO₂ utilization: Developments in conversion processes. *Petroleum*, 3, 109-126.
- Fontalvo Alzate, J., 2014. Using user models in Matlab® within the Aspen Plus® interface with an Excel® link. *Ingeniería e Investigación*, 34, 39-43.
- Ibrahim, D., Jobson, M., Li, J. & Guillén-Gosálbez, G., 2018, Optimization-based design of crude oil distillation units using surrogate column models and a support vector machine, *Chem. Eng. Res. Des.*, 134, 212-225.
- Koumi Ngoh, S. & Njomo, D., 2012, An overview of hydrogen gas production from solar energy, *Renewable and Sustainable Energy Reviews*, 16, 6782-6792.
- Moioli, S., Pellegrini, L., Picutti, B. and Vergani, P., 2013, Improved Rate-Based Modeling of H₂S and CO₂ Removal by Methyl-diethanolamine Scrubbing. *Industrial & Engineering Chemistry Research*, 52, 2056-2065.
- Wang, W., Ma, Y., Marouf-mashat, A., Zhang, N., Li, J. and Xiao, X., 2021, Optimal design of large-scale solar-aided hydrogen production process via machine learning based optimisation framework. *Applied Energy*, 305, 117751.
- Oh, B., Falkenberg, M. and Wessling, M., 2019, Optimizing hybrid membrane-pressure swing adsorption processes for biogenic hydrogen recovery. *Chemical Engineering Journal*, 364, 452-461.

Hierarchical Approach for Solvent Selection in Circular Economy

Pranav Majgaonkar,^{a,b*} Kai Kruber,^b Venkat Aryan,^c Ronny Hanich,^a Davide Pico,^a Mirko Skiborowski^b

^a*Environmental Engineering, Fraunhofer Institute for Chemical Technology (ICT), 76327 Pfinztal, Germany*

^b*Institute of Process Systems Engineering, Hamburg University of Technology, 21073 Hamburg, Germany*

^c*Products, Fraunhofer Institute for Environmental, Safety, and Energy Technology (UMSICHT), 46047 Oberhausen, Germany*
pranav.majgaonkar@ict.fraunhofer.de

Abstract

Chemical recycling allows upcycling of polymers into their monomers and/or other valuable chemicals and contributes to a resource-efficient future. This demands selective depolymerization of a particular fraction from a mixed waste stream and thereby, necessitates the use of appropriate solvents. The solvent not only governs the effective depolymerization of the polymer but also influences the sustainability of the process. Conventionally, the tendency of the solvent to dissolve the polymer is considered as the only selection criterion. However, careful consideration should also be given to its interaction with other process reagents and its sustainability. This research proposes a systematic approach to identify potential solvents suitable for a particular recycling strategy. The hierarchical approach begins with identifying potential solvents using computer-aided molecular design based on the Hansen Solubility Parameters (HSPs) of the target polymer. Further, their interactions with other process reagents are also considered. Subsequently, they are screened for their sustainability through a comprehensive life-cycle assessment and some short-listed systems are further evaluated using process simulation. The application of this methodology for developing a recycling strategy for PLA is illustrated.

Keywords: chemical recycling, solvent selection, sustainability

1. Introduction

Fossil-based plastics have gained immense importance because of their durable performance and versatile properties. However, based on their production and subsequent waste management, it is estimated that 12000 Mt of plastic waste will be disposed in landfills and/or oceans by 2050 (Meys et al., 2020). About 50% of this waste originates from the packaging sector (Meys et al., 2020). According to the EU Directive 2019/852 on Packaging and Packaging Waste, the targeted recycling rate is ca. 55% by 2030. While both mechanical and chemical recycling contributes towards a circular economy, the latter represents an infinite strategy as it can effectively depolymerize different grades of polymers from varied value chains. Possible routes for chemical recycling include gasification, pyrolysis and solvolysis. Gasification involves combustion of plastic waste in the presence of oxygen to produce syngas. During pyrolysis, plastic waste is combusted

in the absence of oxygen to produce basic platform chemicals such as benzene, toluene and xylene. These strategies can handle heavily contaminated, mixed waste streams. Solvolysis results in the production of constituent monomers and/or other valuable chemicals. It is suitable for different addition and condensation polymers such as polycarbonate (PC), polyethylene terephthalate (PET), poly-lactic acid (PLA), etc. and has a potential to further reduce the greenhouse gas (GHG) emissions in comparison to other waste-management strategies (Meys et al., 2020).

Though circularity is commonly implemented in process engineering through recovery of excess reagents and/or heat integration, the use of process systems engineering in transforming a linear economy into a sustainable circular economy is quite rare (Avraamidou et al., 2020). Further, research directed towards chemical recycling is primarily focused on optimizing its economy and lack a systematic guideline to incorporate sustainability (Glavič et al., 2021). For instance, previous efforts directed towards solvolysis considered the ability of the solvent to dissolve the polymer as the only criterion for solvent selection (Sherwood, 2020). However, careful consideration should also be given to its sustainability and the energy required for its recovery (Vollmer et al., 2020). Different computer-aided molecular design (CAMD) methods based on thermodynamic predictions are proposed for appropriate solvent selection in several applications (Austin et al., 2016). Recent efforts directed towards integrating CAMD in conventional process development incorporate different scales, ranging from molecular design to process optimization (Gertig, et al., 2020). In addition, integration of predictive life-cycle assessment (LCA) tools in a single multi-scale system has also been strongly advocated (Fleitmann et al., 2021; Gertig et al., 2020; Glavič et al., 2021). With a similar objective, the current work proposes a hierarchical approach for solvent selection in efforts directed towards chemical recycling (solvolysis), involving successive refinement through incorporation of multiple criteria.

2. Methodology

The proposed hierarchical approach involves three steps. Firstly, an appropriate CAMD method is used to identify potential solvents, which are further screened for their interaction with other process reagents. The sustainability of promising candidates is assessed by a comprehensive LCA in accordance to the principles of green chemistry. Further, the performance of some potential candidates is evaluated by process simulation.

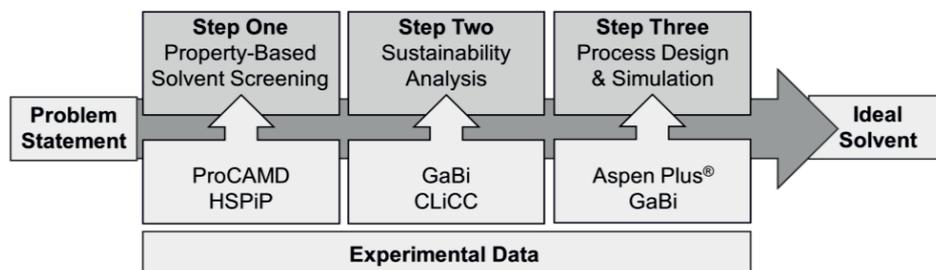


Figure 1: Proposed Hierarchical Approach for Solvent Selection in Circular Economy

2.1. Solvent Screening based on Hansen Solubility Parameters (Step One)

CAMD is an analytical tool aimed at guiding solvent selection for target applications such as liquid-liquid extraction and/or refrigerant and polymer design (Austin et al., 2016). In this approach, the preliminary solvent screening was performed in ProCAMD v1.41 from

PSE for SPEED¹ with the help of a well-established method ("generate-and-test") based on group-contribution methods (Harper & Gani, 2000). During this screening, different alcohols, aldehydes, amines, esters, ethers, ketones and molecules containing chlorine with a boiling point higher than and a melting point lower than 303 K were identified. The generated molecules have a minimum number of two and a maximum number of six groups with two functional groups. Subsequently, these candidates were further evaluated based on their HSPs (δ_D^S , δ_P^S , δ_H^S) relative to that of the polymer ($\delta_D^{PLA} = 18.9$, $\delta_P^{PLA} = 7.6$, $\delta_H^{PLA} = 4.6$) at 298 K (Esmacili et al., 2019). These predictions were further validated with the experimental data obtained from HSPiP v5.3.07. The most promising solvents were evaluated on the basis of the distance between the polymer and the solvent (D_{s-PLA}) relative to the interaction radius of PLA ($R_{o,PLA} = 9.74$) (Agrawal et al., 2004).

$$D_{s-PLA} = \sqrt{4 \times (\delta_D^{PLA} - \delta_D^S)^2 + (\delta_P^{PLA} - \delta_P^S)^2 + (\delta_H^{PLA} - \delta_H^S)^2} \dots\dots (1)$$

Analysis of the interaction of these solvents with other process reagents is a subsequent pre-requisite for further process development. Thus, potential candidates screened on the basis of their relative energy difference ($RED = D_{s-PLA}/R_{o,PLA}$) were analyzed for the possibility of an azeotrope. This was based on vapor-liquid equilibrium predictions performed in ProCAMD v1.41 using the UNIFAC approach and was validated with the help of available experimental data.

2.2 Life-Cycle Assessment of Short-Listed Solvents (Step Two)

LCA is the most common method to assess environmental impacts of products and processes (Monsiváis-Alonso et al., 2020a). Prior to process analysis, the sustainability of the short-listed candidates was evaluated by a comprehensive LCA, in accordance to the ISO 14040 and ISO 14044 standards. Datasets for life-cycle inventory were obtained from commercial databases (ecoinvent v3.7) for some solvents while others were analyzed through predictive life-cycle impact assessment (LCIA) using the Chemical Life Cycle Collaborative (CLiCC) tool. This tool employs artificial neural networks to predict the impact of different organic molecules based on their chemical structure (Song et al., 2017). Environmental impact of these solvents was primarily considered in four categories, namely, acidification potential, climate change, cumulative energy demand (resource use) and human toxicity (human health).

2.3 Evaluation of Process Performance and Sustainability (Step Three)

Based on the afore-mentioned screening, one or more solvents with preferential properties were selected for further experimental investigation. In order to estimate the energy required for solvent recovery, an appropriate down-streaming concept was analyzed through process simulation in Aspen Plus[®]. In addition to this, the sustainability of the down-streaming strategy was also assessed by a preliminary LCA (primarily considering energy-related emissions).

3. Case Study – Chemical Recycling of PLA

Amongst other bio-plastics, PLA is the most promising polyester and is often regarded as a sustainable alternative to petro-chemical plastics (Majgaonkar et al., 2021). However, its production from first-generation biogenic feedstock requires repurposing of land and thus, presents a direct competition to the rising food demand (Atiweh et al., 2021). Further, converting natural landscapes to cultivate feedstock for bio-plastics can lead to higher net GHG emissions than the current GHG savings achieved by replacing fossil-based plastics (Piemonte & Gironi, 2011). Therefore, bio-plastics will only be sustainable

¹ <https://www.pseforspeed.com/> (accessed on 06.02.2022; 13:00)

if non-food crops or waste residues are used for their production and if they are duly recycled at their end-of-life. In the past decade, several efforts have been directed towards chemical recycling of PLA (McKeown & Jones, 2020). As opposed to other attempts, alcoholysis of PLA to produce lactate esters in the presence of primary alcohols such as ethanol or methanol represents an economically and environmentally sustainable recycling strategy (Majgaonkar et al., 2021). However, the insolubility of PLA in alcohols necessitates an appropriate solvent, which was assessed with the help of the proposed hierarchical approach. In an initial screening, about 4200 different primary molecules (excluding isomers) were identified. From these candidates, a compact yet comprehensive list of solvents capable of dissolving PLA was obtained (listed in the Table 1). In addition to conventional solvents, acetyl acetone and ethyl lactate are two new potential candidates identified by the CAMD approach. These candidates were further evaluated to identify the possibility of forming azeotropes with ethanol and methanol. It was observed that, MEK, DCM and chloroform form an azeotrope with both ethanol as well as methanol.

Table 1: Results of the Solvent Screening by ProCAMD v1.41 and HSPiP v5.3.07

Source	Solvent	Solubility Parameters			D_{S-PLA}	D_{S-PLA}/R_o
		δ_D^S	δ_P^S	δ_H^S		
CAMD Screening	acetyl acetone	17	11	6.8	5.52	0.586
	methyl ethyl ketone (MEK)	15.6	7.8	5.5	6.54	0.671
	acetone	15.7	9.1	6.5	6.79	0.697
	ethyl lactate	16.7	7.7	13.1	9.49	0.974
HSPiP Database	cyclohexanone	17.8	8.4	5.1	2.39	0.246
	butyl benzoate	20	5.1	5.2	3.38	0.347
	N-methyl pyrrolidone (NMP)	18	12.3	7.2	5.66	0.580
	tetrahydrofuran (THF)	16.8	5.7	8	5.87	0.603
	N,N-dimethylformamide (DMF)	17.4	13.7	11.3	9.54	0.980
Literature	dichloromethane (DCM)	18.2	6.3	6.1	2.67	0.275
	chloroform	17.8	3.1	5.5	5.09	0.522

Comprehensive LCA of potential solvents, summarized in Table 2 and Table 3, serve as an indicator to assess their sustainability. It can be observed that, acetone, acetyl acetone and ethyl lactate have the most attractive environmental performance. However, acetone outperforms acetyl acetone and ethyl lactate in all impact categories. Further, it has been experimentally demonstrated that, acetone is a better solvent as compared to ethyl lactate with regards to the dissolution of PLA (Gironi et al., 2016). Thus, acetone was chosen as the preferred solvent for further experimental investigation and subsequent process evaluation.

Majgaonkar et al. demonstrated the use of acetone in solvent-assisted depolymerization of PLA for the production of ethyl lactate. The ternary mixture thus obtained contains acetone, ethanol and ethyl lactate, which can be easily separated in a series of two distillation columns. The system was represented by the NRTL equation and the resulting

process sequence was simulated in Aspen Plus[®]. The energy required for this separation amounts to 4.89 MJ/kg_{product}.

Table 2: Heat Map of the Results of LCA from GaBi using EF 3.0

Impact Category	Unit	Solvents				
		acetone	cyclohexanone	DMF	NMP	THF
Acidification	[mole H ⁺ eq./kg]	0.0026	0.0047	0.003	0.006	0.046
Resource Use	[MJ/kg]	55.4	88.1	37.5	93.3	115
Climate Change	[kg-CO ₂ -eq./kg]	1.58	3.65	1.98	4.52	6.86

Table 3: Heat Map of the Results of LCA from the CLiCC tool using predictive LCIA (CED = Cumulative Energy Demand; GWP = Global Warming Potential)

Impact Category	Unit	Solvents		
		acetyl acetone	butyl benzoate	ethyl lactate
Acidification	[mole H ⁺ eq./kg]	1.24	1.23	0.0515
CED	[MJ/kg]	143	121	83.3
GWP	[kg-CO ₂ -eq./kg]	5.32	8.27	3.61

As opposed to the afore-mentioned process, Zeus Industrial Products Inc. developed a chemical recycling strategy for PLA employing chloroform as a solvent and methanol as the depolymerizing agent. The resulting ternary system required a series of three distillation columns, including a pressure-swing distillation sequence, which escalated the energy demand by ca. 20% (5.9 MJ/kg_{product}). The sustainability of these distinct approaches for chemical recycling of PLA was further analyzed in a comparative LCA, wherein the former process outperformed the latter in almost all impact categories (Aryan et al., 2021).

4. Conclusion and Outlook

This study proposes a hierarchical approach for solvent selection to compliment efforts directed towards a circular economy. About 4200 potential solvents were screened based on their ability to dissolve PLA to identify the most promising candidates. Further, integration of LCA and subsequent process simulation allows for the development of a competitive yet sustainable concept with minimal experimental efforts, as illustrated with regards to chemical recycling of PLA. The approach is efficient and flexible and can be extended to other polymers with further introduction of additional constraints.

Acknowledgements

This research stems from an on-going project titled Fraunhofer Cluster Circular Plastics Economy CCPE, funded by the Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.

References

- Agrawal, A., Saran, A. D., Rath, S. S., & Khanna, A. (2004). Constrained Non-Linear Optimization for Solubility Parameters of Poly(Lactic Acid) and Poly(Glycolic Acid)—Validation and Comparison. *Polymer*, 45(25), 8603–8612.
- Aryan, V., Maga, D., Majgaonkar, P., & Hanich, R. (2021). Valorisation of Poly(Lactic Acid) (PLA) Waste: A Comparative Life-Cycle Assessment of Various Solvent-Based Chemical Recycling Technologies. *Resources, Conservation and Recycling*, 172(5), 105670.
- Atiweh, G., Mikhael, A., Parrish, C. C., Banoub, J., & Le, T.-A. T. (2021). Environmental Impact of Bioplastic Use: A Review. *Heliyon*, 7(9), e07918.
- Austin, N. D., Sahinidis, N. V., & Trahan, D. W. (2016). Computer-Aided Molecular Design: An Introduction and Review of Tools, Applications, and Solution Techniques. *Chemical Engineering Research and Design*, 116, 2–26.
- Avraamidou, S., Baratsas, S. G., Tian, Y., & Pistikopoulos, E. N. (2020). Circular Economy - A Challenge and An Opportunity for Process Systems Engineering. *Computers & Chemical Engineering*, 133(7), 106629.
- Esmaceli, M., Pircheraghi, G., Bagheri, R., & Altstädt, V. (2019). Poly(Lactic Acid)/Co-Plasticized Thermoplastic Starch Blend: Effect of Plasticizer Migration on Rheological and Mechanical Properties. *Polymers for Advanced Technologies*, 30(4), 839–851.
- Fleitmann, L., Kleinekorte, J., Leonhard, K., & Bardow, A. (2021). COSMO-susCAMPD: Sustainable Solvents from combining Computer-Aided Molecular and Process Design with Predictive Life-Cycle Assessment. *Chemical Engineering Science*, 245(2), 116863.
- Gertig, C., Leonhard, K., & Bardow, A. (2020). Computer-Aided Molecular and Processes Design based on Quantum Chemistry: Current Status and Future Prospects. *Current Opinion in Chemical Engineering*, 27, 89–97.
- Gironi, F., Frattari, S., & Piemonte, V. (2016). PLA Chemical Recycling Process Optimization: PLA Solubilization in Organic Solvents. *Journal of Polymers and the Environment*, 24(4), 328–333.
- Glavič, P., Pintarič, Z. N., & Bogataj, M. (2021). Process Design and Sustainable Development—A European Perspective. *Processes*, 9(1), 148.
- Harper, P., & Rafiqul Gani. (2000). A Multi-Step and Multi-Level Approach for Computer-Aided Molecular Design. *Computers & Chemical Engineering*, 24(677-683).
- Majgaonkar, P., Hanich, R., Malz, F., & Brüll, R. (2021). Chemical Recycling of Post-Consumer PLA Waste for Sustainable Production of Ethyl Lactate. *Chemical Engineering Journal*, 423(12), 129952.
- McKeown, P., & Jones, M. D. (2020). The Chemical Recycling of PLA: A Review. *Sustainable Chemistry*, 1(1), 1–22.
- Meys, R., Frick, F., Westhues, S., Sternberg, A., Klankermayer, J., & Bardow, A. (2020). Towards a Circular Economy for Plastic Packaging Wastes – The Environmental Potential of Chemical Recycling. *Resources, Conservation and Recycling*, 162(6), 105010.
- Monsiváis-Alonso, R., Mansouri, S. S., & Román-Martínez, A. (2020a). Life-Cycle Assessment of Intensified Processes towards Circular Economy: Omega-3 Production from Waste Fish Oil. *Chemical Engineering and Processing - Process Intensification*, 158(7), 108171.
- Piemonte, V., & Gironi, F. (2011). Land-Use Change Emissions: How Green are the Bioplastics? *Environmental Progress & Sustainable Energy*, 30(4), 685–691. doi:10.1002/ep.10518
- Sherwood, J. (2020). Closed-Loop Recycling of Polymers using Solvents: Remaking Plastics for a Circular Economy. *Johnson Matthey Technology Review*, 64(1), 4–15.
- Song, R., Keller, A. A., & Suh, S. (2017). Rapid Life-Cycle Impact Screening Using Artificial Neural Networks. *Environmental Science & Technology*, 51(18), 10777–10785.
- Vollmer, I., Jenks, M. J. F., Roelands, M. C. P., White, R. J., van Harmelen, T., Wild, P. de, Weckhuysen, B. M. (2020). Beyond Mechanical Recycling: Giving New Life to Plastic Waste. *Angewandte Chemie (International ed. in English)*, 59(36), 15402–15423.

A Framework to Facilitate Decision Making for Infrastructure Options Analysis of Distribution and Utilities Systems in Chemical Production Plants

Marcello Di Martino^{a,b}, Iosif Pappas^{a,b}, Anh Tran^{a,b}, R. Cory Allen^{a,b}, Russell R. Husfeld^c, Sam Eleff^c, Scott G. Moffatt^c, Styliani Avraamidou^d, Burcu Beykal^{e,f} and Efstratios N. Pistikopoulos^{a,b,*}

^a*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA*

^b*Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843, USA*

^c*Ascend Performance Materials, Houston, TX 77002, USA*

^d*Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA*

^e*Department of Chemical and Biomolecular Engineering, University of Connecticut, Storrs, CT 06269, USA*

^f*Center for Clean Energy Engineering, University of Connecticut, Storrs, CT 06269, USA*
stratos@tamu.edu

Abstract

The consequences of investment decisions regarding a chemical plant's distribution and utilities system can be crucial for the operability, and thus the success of a chemical plant. This work presents a novel framework for investment decision making regarding the distribution and utilities systems of a chemical plant, based on multi-stage infrastructure planning and production scheduling, incorporating all relevant process and material timescales. The literature has classically investigated industrial distribution systems only for a single material or energy system and utilities systems only in terms of heat or work exchange networks and their respective integration. However, the evaluation of several relevant material streams holistically within the distribution system of one chemical plant together with several interconnected utilities, and the incorporation of different timescales has not yet been presented in a decision making framework. To do so, a mixed-integer linear programming (MILP) problem is formulated to investigate decision making regarding future investment decisions based on multi-stage infrastructure and maintenance planning for chemical plants. The applicability and capabilities of the developed framework are demonstrated through an application to the distribution and utilities system of a chemical plant in South-East Texas.

Keywords: infrastructure planning, production scheduling, multi-scale modeling.

1. Introduction

Investigating the feasibility and economic implications of a chemical plant's distribution and utilities system is of utmost importance for generating revenue, improving energy use and plant resilience, reducing material losses, and consequently staying competitive within a global economy. Especially, the consequences of investment decisions regarding these systems can be crucial for

the operability and thus the success of a chemical plant.

The optimization of distribution and utilities systems of chemical plants have been examined in the literature for over 30 years (Sahinidis et al. (1989), Papoulias and Grossmann (1983)). However, distribution systems have traditionally only been investigated for a single distribution process material or energy system (Pena et al. (2019)). Regarding the utilities system, it can be observed that optimization has mostly been done in terms of heat or work exchange networks and their respective integration (Pavão et al. (2020)). Recently, integrated planning and scheduling solution strategies have been applied to chemical process systems (Beykal et al. (2022)). Further, the optimization of process integration within a plant and potentially for industrial symbiosis based on different scales has been investigated by Kantor et al. (2020). In addition, multi-stage infrastructure planning and production scheduling, incorporating all relevant process and material timescales, can be utilized to analyze the implications of possible process modifications and future industrial investment decisions (Bi et al. (2020)).

Thus, this work introduces a generic decision making framework which can be applied to a chemical plant's utilities and distribution system. One of the main novelties of the proposed approach is the generic nature of the tool, enabling fast and efficient analyses of competing process and investment alternatives for chemical production facilities. Further, it allows the investigation of various process objectives depending on the facility's goals, i.e. minimizing investment costs or maximizing produced energy. Therefore, various plant operational setpoints can be simulated and automatically visualized.

In Section 2, a general overview of the developed framework and possible applications are given, while in Section 3 the mathematical model is described in detail. In Section 4, the framework is applied to a representative steam utilities system case study to illustrate the capabilities. Lastly, the conclusion and a brief future outlook is stated.

2. Framework Overview

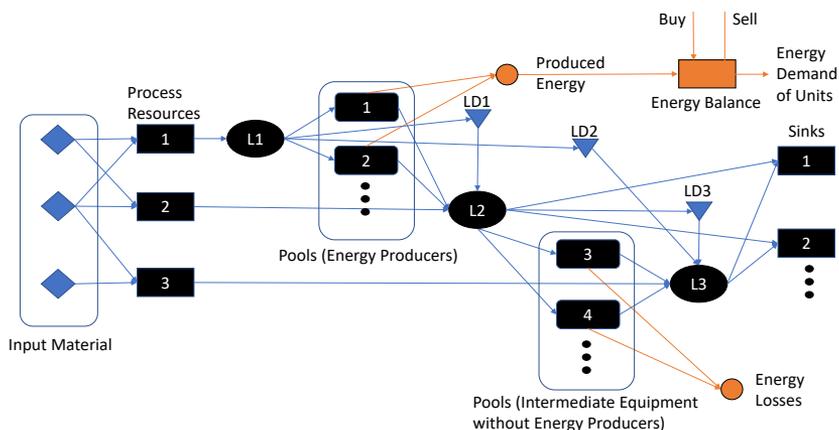


Figure 1: Superstructure of a representative utilities process.

To illustrate to what kind of processes the framework can be applied, a representative utilities superstructure is shown in Figure 1. In this case, input materials (blue diamonds) are used to generate process resources of different quality (black rectangles). These resources are of different quality levels (L1 to L3) and can consequently be used for various purposes. Higher quality resources can be letdown (LD - blue triangles) for lower quality usage if needed or used in intermediate

equipment, e.g. for energy generation purposes (rectangles with rounded corners). All generated resources are used to satisfy material demands at process sinks (black rectangles). Comparing the overall produced energy with the energy demand of all units results in an overall energy balance which shows whether there is excess energy to be sold or additional energy to be bought. If the process involves equipment which requires energy in the form of a material flow, i.e. turbine driven pumps, the energy used to power this type of equipment is summarized as an energy loss, since this material flow cannot be used for energy production. In case of distribution systems, the given superstructure can be further simplified since material transformation or energy production does not occur throughout these systems. In this case, our main goal is to fulfill the material sink demands, while considering an applicable storage system and scheduling of the incoming material from various suppliers through different means of transportation, e.g. by pipeline or truck. Thus, for distribution systems it is essential to consider a time dependency of variables during the mathematical modeling, whereas for utilities steady-state assumptions are applicable. The developed framework can handle time dependent systems of equations. However, for simplicity and since the developed case study is based on a steam utilities system the time index has been neglected in the mathematical modeling equations. By adding an additional time index to all material and energy flows the given system of equations can be modified to a time dependent system.

To derive a generic decision making framework which is capable of optimizing the distribution and utilities system of an industrial chemical plant, a mathematical model is developed in the Python Pyomo environment. All process specific information, such as the involved units and materials, are defined in an Excel data exchange file, which is loaded into the optimization program. Thus, historic plant data is processed to define relevant process data in terms of identifying the correct equipment classification as well as equipment specific constraints and the relevant material and energy conversions of the respective units. This data is automatically processed and assigned in the optimization model. After solving the model, the feasibility of the solution is checked. Further analyses in terms of relevant material and energy streams and their respective implications for the obtained solution can also be conducted. It is important to note that the tool is capable of automatically visualizing the generated solution in a superstructure summarizing all material and energy streams.

3. Mathematical Modeling

The basis of the decision making tool are mass and energy balances based on a set notation, which results in a more compact system of equations. To this end, various sets for all available units and material flows are introduced as follows. All available units of a process (set called *Units*) are separated into material producers or suppliers (*P*), intermediate equipment (*EQ*) and material sinks (*S*). Additionally, all energy producing units are summarized in a set called energy producers (*EP*), whereas units with an energy demand are classified in the set called energy sinks (*ES*). For further analysis of solutions and trade-offs additional sets are introduced which consist of all material producing and intermediate equipment (EQ_2), as well as one consisting of EQ_2 excluding all energy producing equipment (EQ^*). Material flows are generally classified as the set of resources (*Res*), which can be divided into the subsets key resources (*Key*) and other resources (*NK*). In this case, key resources describe the focus of the process, e.g. for steam at various steam pressure levels, whereas other resources summarize all non steam resources like natural gas, air or boiler feed water and include the subset of inputs (*In*), $In \subseteq NK$.

Apart from the set notation, energy and material conversion factors are introduced. These conversion factors can be deduced from actual plant data or from ASPEN simulations. These conversion factors are equipment and material specific and summarized in the previously mentioned Excel file, which therefore enables fast and easy modifications, if necessary. The definition of the material and energy conversion factors are given in Eq. 1 and 2.

Conversion Factors

$$f_{EQ,i,j} = \frac{\text{Output flow } j \text{ of } EQ}{\text{Input flow } i \text{ of } EQ} \forall EQ, i \in Res, j \in Res, i \neq j \quad (1)$$

$$ef_{EQ,i} = \frac{\text{Energy produced or consumed in } EQ}{\text{Input flow } i \text{ of } EQ} \forall EQ, i \in Res \quad (2)$$

The material conversion factors are derived for all equipment and respective input and output flows. The energy conversion factors are calculated for all equipment and respective input flows. The key idea of the generic tool is to introduce artificial resource pools to which generated material throughout the process can be booked and from which consumed material throughout the process can be taken. Therefore, book-keeping variables for all resources are introduced as x_{Res} . Generally, in the case of key resources this variable has to be zero for generated solutions. For other resources the amount produced or needed can be deducted from the variable value. Further, material flows from equipment to the resource pools can be introduced as $x_{EQ,Res}$ and flows from resource pools to equipment as $x_{Res,EQ}$ in $\frac{klb}{h}$. Based on these definitions the following mass balance equations can be introduced (Eq. (3) overall balances of *Key*, Eq. (4) fulfill all demands at *S*, Eq. (5) supply of *Key*, Eq. (6) conversion of *Key* in *EQ*, Eq. (7) overall balances of *NK*):

Mass Balancing

$$\sum_P x_{P,Key} + \sum_{EQ} x_{EQ,Key} = x_{Key} + \sum_S x_{Key,S} + \sum_{EQ} x_{Key,EQ} \forall Key \quad (3)$$

$$x_{Key,S} = x_s^{Demand,Key} \forall S, Key \quad (4)$$

$$x_{P,Key} = f_{P,In,Key} \cdot x_{In,P} \forall P, In, Key \quad (5)$$

$$f_{EQ,i,j} \cdot x_{i,EQ} = x_{EQ,j} \forall EQ, i, j \in Res, i \neq j \quad (6)$$

$$x_{NK} = \sum_{EQ_2} x_{EQ_2,NK} - x_{NK,EQ_2} \forall NK \quad (7)$$

For the energy balance, additional variables are introduced. The energy produced or consumed in a unit is given by n_{Units} . Further, n_{Sell} and $n_{Purchase}$ can be introduced as variables for selling a surplus of energy generated in the plant or buying energy in the case of an energy deficiency throughout the system. Thus, the energy balances can be summarized as follows (Eq. (8) overall energy balance, Eq. (9) overall energy demand, Eq. (10) energy needed for production of *K*, Eq. (11) energy produced):

Energy Balancing

$$\sum_{EP} n_{EP} - n_{ES} + n_{Purchase} - n_{Sell} = 0 \quad (8)$$

$$n_{ES} = \sum_P n_P + \sum_{EQ} n_{EQ}^{Demand} \quad (9)$$

$$n_P = ef_{P,In} \cdot x_{In,P} \forall P \quad (10)$$

$$n_{EP} = ef_{EP,Key} \cdot x_{Key,EP} \forall EP \quad (11)$$

$$n_{EQ^* \setminus P} = ef_{EQ^* \setminus P,Key} \cdot x_{Key,EQ^* \setminus P} \forall EQ^* \setminus P \quad (12)$$

Further, Eq. (12) derives the energy needs for all units summarized in EQ^* without key resource producing units. This special set includes units which require energy in the form of material flows, e.g. turbine driven pumps. Therefore, $\sum_{EQ^* \setminus P} n_{EQ^* \setminus P}$ specifies energy lost for energy generation due to equipment specific choices and can be understood as an opportunity to increase energy efficiency throughout the system by trying to reduce this value.

Originally, binary decision variables have been introduced for equipment and flow decision making. However, since competing equipment or investment decisions are taking resources from the same artificial resource pool, and flows from a resource pool to an equipment can be zero, the introduced binaries were discovered to be redundant. Apart from that, neglecting binary decision variables resulted in better computational performance, while no process information was lost.

4. Case Study

To highlight the capabilities of the developed tool, we apply our framework to the steam utility system of a chemical production plant in South-East Texas. The production plant generates four different steam pressure levels ($P_1 > P_2 > P_3 > P_4$, summarized in *Key*) as byproducts from various processes, such as boiler and reactor systems (*P*). Further, the steam is used to satisfy demands of various material (*S*) and energy sinks (*ES*) using two different sized turbines using different pressure levels of steam as input (*EP*), as well as a set of intermediate turbine driven pumps for various applications in other plant processes ($EQ^* \setminus P$). When necessary, steam from higher pressure levels can be let down to lower pressure levels. Additionally, the steam process requires feed water, air and natural gas (set of *NK*). A representative process operation point is summarized in the top part of Figure 2.

The goal of the case study is to minimize the overall cost by investigating a process alternative denoted by neglecting one steam producer, the turbine operating at a higher pressure level, as well as modifying all turbine driven pumps to motor driven pumps. The energy and material demands are assumed to be constant. For this particular case study, the objective function is introduced based on the investment and maintenance costs of the selected equipment. Linear correlations for the total installed cost (TOC) of equipment dependent on the flow rate have been derived with ASPEN Capital Cost Estimator (AspenTech, 2012). Further, the annual maintenance cost of a unit are estimated based on its TOC. The process visualization for the modified case is given in the bottom part of Figure 2. The selected modification reduced the overall cost by approximately \$7.9 MM. Subsequently, by only investigating this scenario, steps towards a process cost reduction of approximately 70% can be deducted.

Overall, the tool can not only be used to match streams for heat or work exchange, comparable to more conventional methods such as PINCH analysis, but also to evaluate present and future operational set-points as well as influences of investment decisions and process alternatives underlining the multi-stage nature and capabilities. This further enables the transition planning and scheduling for an operational process transition from the current setup to a desired future one. However, one potential drawback to conventional methods is a less rigorous implementation of thermodynamic restrictions.

5. Conclusion

We have presented a comprehensive generic framework with the capabilities of optimizing the distribution and utilities systems of an industrial chemical production plant. Our proposed approach is based on developing a mathematical superstructure model which encompasses the incoming and outgoing material and energy flows, as well as their respective conversions for each unit of the plant. By defining different investment alternatives for the facility based on the demand and costs, we are able to identify and recommend optimal investment opportunities to reduce the overall cost. The effectiveness of our framework was demonstrated through its application to the steam utility system of an industrial chemical plant. The results indicate a significant reduction in overall cost, as well as the plant's energy needs. Our next steps include the possible interconnection of various utility processes, e.g. steam, water and electricity networks and take into account network

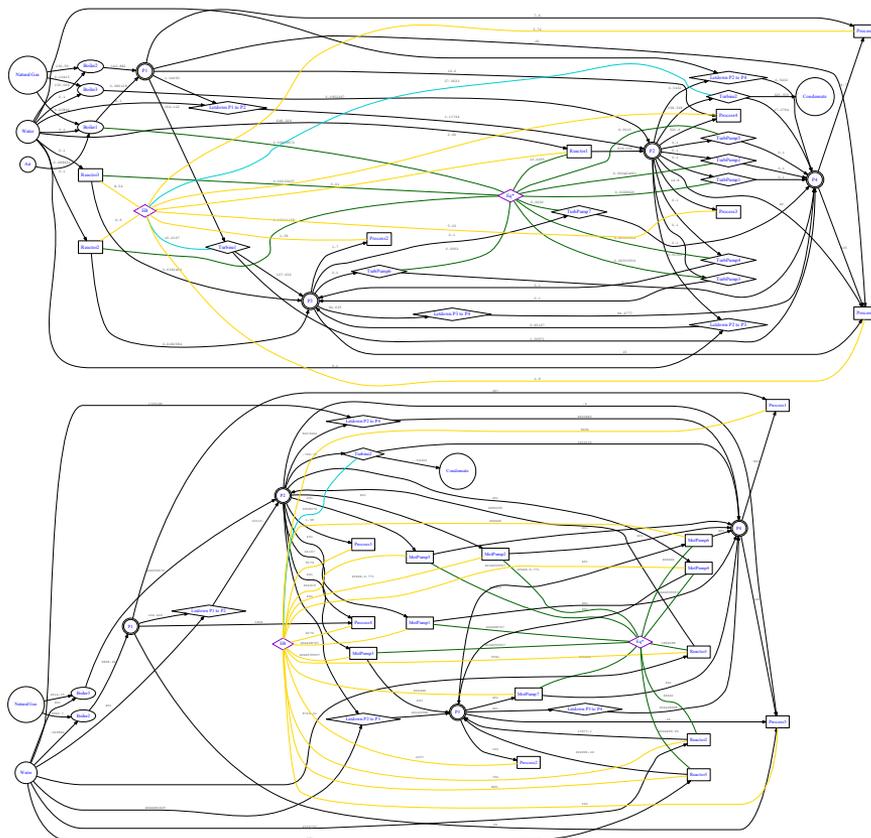


Figure 2: Automatic visualization of process simulation. Top: Standard case; Bottom: Modified case.

disruption scenarios for optimizing plant resilience and respective cost and energy trade-offs.

References

- AspenTech, 2012. Aspen Capital Cost Estimator. Aspen Technology, Inc Burlington, MA.
 URL <https://www.aspentech.com/en/products/engineering/aspen-capital-cost-estimator>
- B. Beykal, S. Avraamidou, E. N. Pistikopoulos, 2022. Data-driven optimization of mixed-integer bi-level multi-follower integrated planning and scheduling problems under demand uncertainty. *Computers & Chemical Engineering* 156, 107551.
- K. Bi, B. Beykal, S. Avraamidou, I. Pappas, E. N. Pistikopoulos, T. Qiu, 2020. Integrated modeling of transfer learning and intelligent heuristic optimization for a steam cracking process. *Industrial & Engineering Chemistry Research* 59 (37), 16357–16367.
- I. Kantor, J.-L. Robineau, H. Büttün, F. Maréchal, 2020. A mixed-integer linear programming formulation for optimizing multi-scale material and energy integration. *Frontiers in Energy Research* 8, 49.
- S. A. Papoulias, I. E. Grossmann, 1983. A structural optimization approach in process synthesis—i: Utility systems. *Computers & Chemical Engineering* 7 (6), 695–706.
- L. V. Pavão, J. A. Caballero, M. A. Ravagnani, C. B. Costa, 2020. An extended method for work and heat integration considering practical operating constraints. *Energy Conversion and Management* 206, 112469.
- J. G. C. Pena, V. B. de Oliveira, J. L. F. Salles, 2019. Optimal scheduling of a by-product gas supply system in the iron- and steel-making process under uncertainties. *Computers & Chemical Engineering* 125, 351–364.
- N. Sahinidis, I. Grossmann, R. Fornari, M. Chathrathi, 1989. Optimization model for long range planning in the chemical industry. *Computers & Chemical Engineering* 13 (9), 1049–1063.

An Optimization Model for Expansion Planning of Reliable Power Generation Systems

Seolhee Cho^a and Ignacio E. Grossmann^{a*}

^a *Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

* *Corresponding Author: grossmann@cmu.edu*

Abstract

This paper aims to develop an optimization model for the expansion planning of reliable power generation systems. To achieve this goal, we propose an optimization model that minimizes the total cost using Generalized Disjunctive Programming (GDP). Specifically, the model determines both investment decisions (number, size, location, and time of generators to install, retire, and decommission) and operation decisions (number of operating/backup generators, operating capacity, and expected power output) by imposing penalties when the demand is not satisfied, and a system has low reliability. The model is verified through an illustrative example of two regions with four power stations over five operating periods.

Keywords: Reliability, Expansion Planning, Power systems, Optimization

1. Introduction

Expansion planning for power generation systems aims to determine the optimal number, size, location, and time to install and/or retire the power generators, while minimizing the total cost over a given time horizon (Conejo et al., 2016). Expansion planning works have been reported in the literature. For instance, Lara et al. (2018) develop a multi-region and multi-period MILP optimization model by considering short-term operation problems (e.g., unit commitment) and investment decisions. Given the large size of the model, it is solved with a Nested Benders decomposition method. Li et al. (2022) extend the expansion model developed by Lara et al. (2018) by integrating generation and transmission. Although studies on integrating the reliability and expansion planning have been reported (Moreira et al., 2017), there is still modest research that comprehensively accounts for the impact of reliability in expansion planning.

Reliability is the probability that a system or component can perform its required function without failure for a given time (Sherwin & Bossche, 1993). Securing high reliability, or more precisely high availability, in the design and operation of power systems is essential since power generation systems aim to provide uninterrupted electric power to customers. One method to improve reliability at the design phase is to add redundant or backup units, which allows the systems to operate even if one or multiple generators fail (Kim & Kim, 2017). This approach is known as ‘reliability-based design optimization,’ and various studies on this topic have been reported (Ye et al., 2018; Ortiz-Espinoza et al., 2021; Chen et al., 2022). However, since power systems operate dynamically due to time-varying power demand, reliability is also influenced by the operational strategies that power systems use to satisfy the load demand. Specifically, backup units have a dual role in power generation systems. They can remain as backup units in case of low power demand or change to operating units when the power demand is high. This dual purpose of backup

units depending on the load demand must be considered in designing reliable power generation systems.

One of the conventional methods used to evaluate the reliability of power systems is “ $N-1$ reliability”. The $N-1$ reliability assumes that a power system can withstand an unexpected failure of a single component (Ballireddy & Modi, 2019). This implies that power systems may not function properly if multiple units fail simultaneously. The failures of multiple generators may reduce the power output but not necessarily fail the entire system. Hence, a rigorous method anticipating every possible failure state and selecting the proper number and size of the backup generators should be developed to design and plan reliable power generation systems. The specific goal of this work is to develop an optimization model for the expansion planning of reliable power generation systems. The model is formulated using Generalized Disjunctive Programming (GDP) that determines both investment (number, size, location, and time of generators to install, retire, and decommission) and operation decisions (number of operating/backup generators, operating capacity, and expected power production level) to minimize the total cost including unmet demand and low reliability penalties.

2. Problem statement

Given are regions $r \in R$, with power stations $k \in K_r$, parallel generators $j \in J_k$, and discretized capacities of generator $p \in P_k$. There are two types of power stations: existing ($k \in K_r^{EX}$) and potential ($k \in K_r^{NW}$) power stations. A set of power station designs $h \in H_k$ and corresponding operation modes $m \in M_{k,h}$, time periods (year) $t \in T$, and sub-periods (season) in each year $n \in N$ are also given. Specifically, $h = 1$ indicates that one generator is available in power stations k and $h = H$ means all generators that can be installed in power stations k are available. Likewise, $m = 1$ represents one generator is operated, $m = M$ refers to the mode in which all generators are operated. Each power station k has different failure states $s \in S_{m,k,h}$ depending on the design h and operation mode m . The failure states can also be classified into successful operation states ($S_{m,k,h}^F$) and partial operation states ($S_{m,k,h}^P$). ‘Successful operation states’ indicate the operation modes in which the power generation capacity is sufficient to satisfy the load demand, whereas ‘Partial operation states’ refer to the operation modes in which the power generation capacity is insufficient to meet the load demand as it can only produce electric power at a limited level. The major assumptions in this model are: (i) Each power station has a maximum number of available power generators, (ii) Storage systems are not included, (iii) Operational problems such as unit commitment are not included.

3. Model formulation

The model is developed by Generalized Disjunctive Programming (GDP) (Trespalcios and Grossmann, 2014), which can be expressed in terms of Boolean and continuous variables, algebraic constraints, disjunctions, and logic propositions. We also introduce binary variables since several investment decisions are reformulated as algebraic constraints.

3.1. Investment constraints

The binary variable $y_{j,p,k,r}^{EX}$ indicates the existence of unit j with discrete capacity p in power station k of region r . The binary variables $y_{j,p,k,r,t}^{IN}$ and $y_{j,p,k,r,t}^{AV}$ state the installation and availability of unit j in year t , respectively. The binary variables are also represented with corresponding Boolean variables ($Y_{j,p,k,r,t}^{IN}, Y_{j,p,k,r,t}^{AV}$). $IC_{j,k,r,t}, AC_{j,k,r,t}, TAC_{k,r,t}$ are

installed capacity, available capacity, and total available capacity, respectively. There are three different investment decisions for generators such as early decommission, retirement due to lifetime expiration, and lifetime extension. However, equations for the three decisions are not provided here due to space limitation. Eqns. (4) – (6) correspond to the capacity constraints, and Eqn. (7) states that only one capacity p can be selected for each generator j .

$$\begin{aligned}
 & \sum_{p \in P_k} y_{j,p,k,r}^{EX} + \sum_{p \in P_k} \sum_{t \in T} y_{j,p,k,r,t}^{IN} \leq 1 \quad \forall j \in J_k, k \in K_r, r \in R & \text{(1) Generator installation} \\
 & \left. \begin{aligned}
 y_{j,p,k,r,t}^{IN} & \Rightarrow \bigwedge_{q \in [t, t+LT_{j,k,r}]} y_{j,p,k,r,q}^{AV} \quad \forall j \in J_k, p \in P_k, k \in K_r^{NW}, r \in R, t, q \in T \\
 y_{j,p,k,r,t}^{AV} & \Rightarrow \bigvee_{q \in [t-LT_{j,k,r,t}]} y_{j,p,k,r,q}^{IN} \quad \forall j \in J_k, p \in P_k, k \in K_r^{NW}, r \in R, t, q \in T
 \end{aligned} \right\} & \begin{aligned}
 & \text{(2) Generator in potential} \\
 & \text{power stations} \\
 & \text{(3) Generator in existing} \\
 & \text{power stations}
 \end{aligned} \\
 & y_{j,p,k,r,t}^{AV} = 1 \quad \forall (j, p, k, r) \in Pre, t = 1 \\
 & IC_{j,k,r,t} = \sum_{p \in P_k} \delta_{k,p} y_{j,p,k,r,t}^{IN} \quad \forall j \in J_k, k \in K_r, r \in R, t \in T \quad \text{(4)} \quad AC_{j,k,r,t} = \sum_{p \in P_k} \delta_{k,p} y_{j,p,k,r,t}^{AV} \quad \forall j \in J_k, k \in K_r, r \in R, t \in T \quad \text{(5)} \\
 & TAC_{k,r,t} = \sum_{j \in J_k} AC_{j,k,r,t} \quad \forall k \in K_r, r \in R, t \in T \quad \text{(6)} \quad \sum_{p \in P_k} y_{j,p,k,r,t}^{AV} \leq 1 \quad \forall j \in J_k, k \in K_r, r \in R, t \in T \quad \text{(7)}
 \end{aligned}$$

3.2. Operation and system reliability constraints

There are two Boolean variables related to investment and operation decisions. $Z_{k,r,h,t}$ is true if design h is selected for power station k of region r in time t (Eqn. (a)). $W_{k,r,m,h,n,t}$ is true if power station k of region r is in operation mode m during sub-period n in time t for design h (Eqn. (b)). The binary variable $x_{j,p,k,r,n,t}$ indicates the operation of unit j with discrete capacity p in power station k of region r during sub-period n of time t . $OC_{j,k,r,n,t}$ represents an operating capacity, $P_{s,k,r,n,t}$ is the probability of station k of region r being in failure state s during sub-period n of time t , and $EP_{s,k,r,n,t}$ corresponds to the expected power output. $P_{k,r,n,t}^F$ and $P_{k,r,n,t}^P$ are the probabilities of station k being in successful and partial operations.

$$\left[\begin{array}{l} \bigvee_{h \in H_k} \\ \bigvee_{m \in M_{k,h}} \end{array} \left[\begin{array}{l} \sum_{j \in J_k} \sum_{p \in P_k} y_{j,p,k,r,t}^{AV} Z_{k,r,h,t} = i \quad i = 0, 1, \dots, h \\ \sum_{j \in J_k} \sum_{p \in P_k} x_{j,p,k,r,n,t} = i \quad i = 1, \dots, m \\ \sum_{p \in P_k} \varphi_j \delta_{k,p} x_{j,p,k,r,n,t} \leq OC_{j,k,r,n,t} \quad \forall j \in J_k \\ OC_{j,k,r,n,t} \leq \sum_{p \in P_k} x_{j,p,k,r,n,t} \delta_{k,p} x_{j,p,k,r,n,t} \quad \forall j \in J_k \\ P_{k,r,n,t}^F = \sum_{s \in S_{m,h,k}^F} \left\{ \prod_{j \in J_{s,m,h}^O} \lambda_{j,k} \prod_{j \in J_{s,m,h}^N} \gamma_{j,k} \right\} \\ P_{k,r,n,t}^P = \sum_{s \in S_{m,h,k}^P} \left\{ \prod_{j \in J_{s,m,h}^O} \lambda_{j,k} \prod_{j \in J_{s,m,h}^N} \gamma_{j,k} \right\} \\ EP_{s,k,r,n,t} = \sum_{j \in J_{s,m,h}^O} \rho_n OC_{j,k,r,n,t} P_{s,k,r,n,t} \quad s = 1, s \in S_{m,k,h} \\ EP_{s,k,r,n,t} = \sum_{j \in J_{s,m,h}^N} \rho_n x_{j,k,r,n,t} AC_{j,k,r,n,t} P_{s,k,r,n,t} \quad s \neq 1, s \in S_{m,k,h} \end{array} \right] \quad \forall n \in N \quad \left. \begin{array}{l} \text{(a) Investment decision} \\ \text{(8) Number of available unit} \\ \text{(b) Operation \& reliability decision} \\ \text{(9) Number of operating unit} \\ \text{(10) Operating capacity of} \\ \text{available unit} \\ \forall k \in K_r^{TH}, r \in R, t \in T \\ \text{(11) Successful and partial} \\ \text{operation reliability} \\ \text{(12) Expected production by} \\ \text{successful and partial} \\ \text{operation} \end{array} \right\}$$

The objective function is to minimize the total cost (Eqn. (13)), which includes the investment costs $(\theta_{j,p} y_{j,p,k,r,t}^{IN})$, the operating costs $(\varphi_{j,p} x_{j,p,k,r,n,t})$, the feedstock costs

($\kappa_k FS_{k,r,n,t}$), the downtime penalty ($\varepsilon DT_{k,r,n,t}$), and the unmet demand penalty ($PN_{n,t}$). The unmet demand penalty is determined with the disjunction shown in Eqn. (14).

$$\min COST = \sum_{j \in J} \sum_{p \in P_k} \sum_{k \in K_r} \sum_{r \in R} \sum_{n \in N} \sum_{t \in T} (\theta_{j,p})_{j,p,k,r,t}^{IN} + \varphi_{j,p} x_{j,p,k,r,n,t} + \kappa_k FS_{k,r,n,t} + \varepsilon DT_{k,r,n,t} + PN_{n,t} \quad (13)$$

$$\left[\begin{array}{l} \sum_{k \in K_r} \sum_{r \in R} TEP_{k,r,n,t} < \sum_{r \in R} \rho_n D_{r,n,t} (1 + R_t^{min}) \\ PN_{n,t} = (\sum_{r \in R} D_{r,n,t} - \sum_{r \in R} TEP_{r,n,t}) \alpha \end{array} \right] \vee \left[\begin{array}{l} \sum_{r \in R} TEP_{r,n,t} \geq \sum_{r \in R} \rho_n D_{r,n,t} (1 + R_t^{min}) \\ PN_{n,t} = 0 \end{array} \right] \quad \forall n \in N, t \in T \quad (14)$$

where ρ_n is a duration (hours) of sub-period n , $D_{r,n,t}$ is a power demand, and R_t^{min} is a reserve margin related to the peak demand of time t . The GDP given by (1)–(14) can be reformulated into a Mixed-Integer Nonlinear Programming (MINLP) using Big-M (BM) and/or Hull Reformulation (HR) (Trespalcios & Grossmann, 2014). In this paper we use the Big-M method. Also, the MINLP model is reformulated as MILP model by using exact linearization of Eqn. (12) (Garcia-Herreros et al., 2015).

4. Illustrative example

To illustrate the application of the proposed model, we consider a simple case consisting of two regions, each with two power stations (i.e., coal power plants (1 existing, 1 potential), and natural gas power plants (1 existing, 1 potential)) as shown in Figure 1(a). Each power station is assumed to have a maximum of three parallel generators, and there are three capacities for the generators. Note that renewable power generators are not considered in this example. Figure 1(b) displays power demands during five time periods, each with four sub-periods. Table 1 shows the technical and economic parameters.

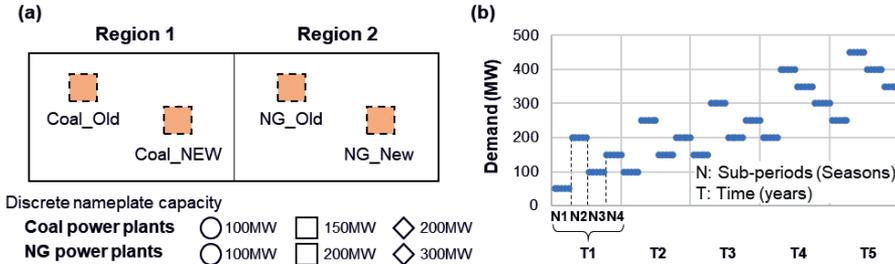


Figure 1. (a) Geographical representation of illustrative example, (b) Power demand

Table 1. Parameters of the illustrative example

Parameters	Symbols	Values	Parameters	Symbols	Values
Nameplate capacity (MW)			Installation cost (k\$/unit)		
Coal plants	$\delta_{k,p}$	100, 150, 200	Coal plants	$\theta_{j,p}$	50, 64, 76
Natural gas plants		100, 200, 300	Natural gas plants		60, 91, 116
Unit reliability			Start-up cost (\$/unit)		
Coal plants	$\lambda_{j,k}$	0.93	Coal plants	$\varphi_{j,p}$	50, 80, 100
Natural gas plants		0.95	Natural gas plants		60, 100, 150

Additionally, the deterministic expansion planning without reliability is obtained, and its expected cost is calculated under the risk of generator failures. Table 2 shows the numerical results of the cases, and the results are obtained with CPLEX in GAMS 32.1.0 on an Intel Core i7-10510U CPU, 1.80GHz. It should be noted that the disjunctive constraints proposed in this paper make the model require longer computational times (almost 2000 sec) than the deterministic model without disjunctions (less 1 sec).

Table 2. Numerical results of the example

	# Binary variables	# Cont. variables	# Constraints	CPU (sec)
Deterministic case (No reliability)	1,363	15,762	62,888	0.203
Reliability-constrained case	1,646	15,773	63,659	1.983

A detailed comparison of the deterministic and reliability-constrained cases can be found in Figure 2. As shown in Figure 2(a), when reliability is not considered, the model newly installs one large coal generator (200MW) in region 1 and one large natural gas generator (300MW) in region 2. On the other hand, the proposed model decides to install three large natural gas generators in region 2 (See Figure 2(c)). Note that both cases are assumed to have two existing coal generators and one natural gas generator. Figures 2(b) and 2(d) show the operation results of the deterministic and the reliability-constrained cases during the last period (T5), respectively. The deterministic case has one backup generator in region 1 during spring (N1) and winter (N4), but all generators participate in power generation during summer (N2) and fall (N3) due to relatively high power demand. On the other hand, the reliability-constrained case has two backup generators (one in region 1 and one in region 2) during all periods. This result is because the system availability can be improved by having more backup generators.

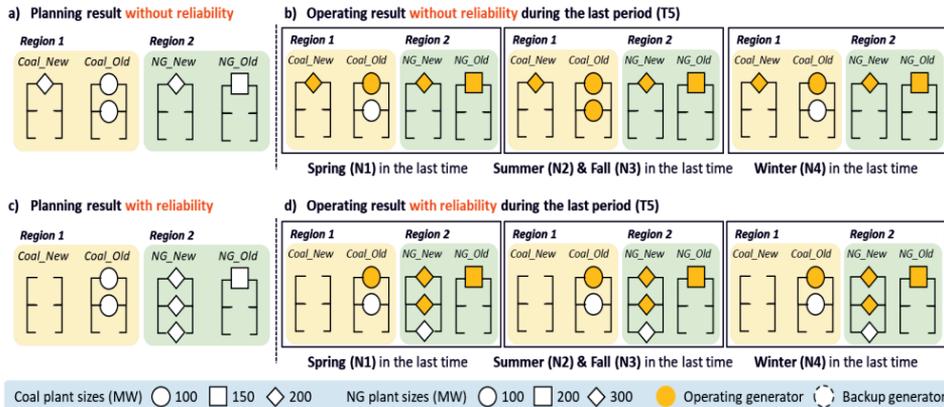


Figure 2. (a) and (b): investment and operation results of deterministic case; (c) and (d): investment and operation results of reliability-constrained cases during period 5 (T5)

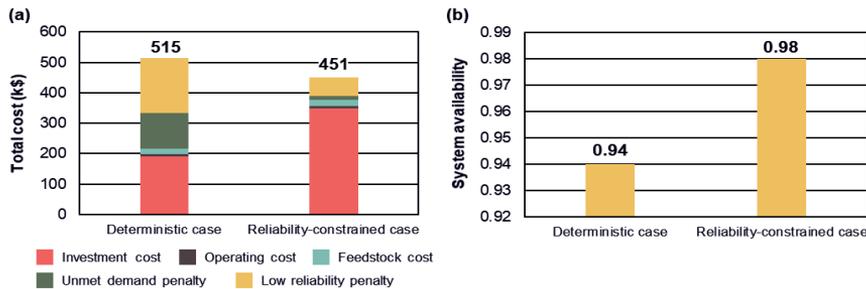


Figure 3. (a) Total cost and (b) System availability of the cases

Figure 3(a) shows the cost contribution of the two cases. Due to backup generators, the reliability-constrained case requires a higher investment cost (\$192,000 vs. \$348,000). The investment cost is compensated by lower penalties (unmet demand penalty: \$115,000 vs. 12,703; low reliability penalty: \$182,774 vs. \$61,807). The deterministic case has a poor performance in the case of generator failure due to a lack of flexibility. On the other hand, the reliability-constrained case has enough slack capacity to reallocate power

demands when the generator failure occurs. Such flexible design of the reliability-constrained case enables the system to have higher availability than the deterministic cases (0.94 vs. 0.98), as shown in Figure 3(b). This example proves that the proposed model is more effective for designing reliable power systems than expansion models in which generator failures are not considered.

5. Conclusions

This paper has presented an optimization model for the expansion planning of reliable power generation systems. The proposed GDP model optimizes investment decisions (number, size, location, and time of generators to install, retire, and decommission) and operation decisions (number of operating/backup generators, operating capacity, and expected power production level). It was identified that the proposed model effectively designed a reliable power generation system. Future work will develop decomposition methods for large-scale problems and involve other operation problems such as economic dispatch and unit commitment to evaluate the reliability more precisely. Moreover, renewable technologies such as wind turbines and solar panels will be included, and Markov chain theory will be applied to enumerate all possible failure states explicitly.

Acknowledgement

This work was conducted as part of the Institute for the Design of Advanced Energy Systems (IDAES) with support through the Simulation-Based Engineering, Crosscutting Research Program within the U.S. Department of Energy's Office of Fossil Energy and Carbon Management.

References

- Ballireddy, T. R. R., & Modi, P. K. (2019). Power System Expansion Planning Incorporating Renewable Energy Technologies with Reliability Consideration: A State of Art Literature. *International Journal of Recent Technology and Engineering*
- Chen, Y., Ye, Y., Yuan, Z., Grossmann, I. E., Chen, B. (2022). Integrating stochastic programming and reliability in the optimal synthesis of chemical processes, *Computers and Chemical Engineering*
- Conejo, A. J., Baringo Morales, L., Kazempour, S. J., & Siddiqui, A. S. (2016). Investment in Electricity Generation and Transmission. *Springer International Publishing*
- Garcia-Herreros P, Zhang L, Misra P, Arslan E, Mehta S, Grossmann, I. E. (2015). Mixed-integer bilevel optimization for capacity planning with rational markets, *Computers and Chemical Engineering*
- Kim, H., & Kim, P. (2017). Reliability–redundancy allocation problem considering optimal redundancy strategy using parallel genetic algorithm. *Reliability Engineering & System Safety*
- Lara, C. L., Mallapragada, D. S., Papageorgiou, D. J., Venkatesh, A., & Grossmann, I. E. (2018). Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. *European Journal of Operational Research*
- Li, C., Conejo, A. J., Liu, P., Omell, B. P., Sirola, J. D., & Grossmann, I. E. (2022). Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems. *European Journal of Operational Research*
- Moreira, A., Pozo, D., Street, A., & Sauma, E. (2017). Reliable Renewable Generation and Transmission Expansion Planning: Co-Optimizing System's Resources for Meeting Renewable Targets. *IEEE Transactions on Power Systems*
- Ortiz-Espinoza, A. P., Ye, Y., Grossmann, I. E., & Jiménez-Gutiérrez, A. (2021). Multi-objective optimization for the incorporation of safety and reliability considerations in process design. *Computer Aided Chemical Engineering*
- Sherwin, D. J., & Bossche, A. (1993). The Reliability, Availability and Productiveness of Systems. *Springer Netherlands*
- Trespalacios, F., & Grossmann, I. E. (2014). Review of mixed-integer nonlinear and generalized disjunctive programming methods. *Chemie-Ingenieur-Technik*
- Ye, Y., Grossmann, I. E., & Pinto, J. M. (2018). Mixed-integer nonlinear programming models for optimal design of reliable chemical plants. *Computers and Chemical Engineering*

Application of an ontology-based decision support system for the design of emulsion-based cosmetic products

Juliana Serna^{a,b}, Jose L. Rivera-Gil^{a,b}, Alex Gabriel^a, Javier A. Arrieta-Escobar^{b,c}, Vincent Boly^a, Véronique Falk^c, Paulo C. Narváez-Rincón^b

^a*Université de Lorraine, Equipe de Recherche des Processus Innovatifs, ERPI-ENSGSI, 8 Rue Bastien-Lepage, 54000 Nancy, France*

^b*Universidad Nacional de Colombia, Departamento de Ingeniería Química y Ambiental, Grupo de Procesos Químicos y Bioquímicos, 111321 Bogotá, Colombia*

^c*Université de Lorraine, Laboratoire Réactions et Génie des Procédés, CNRS-LRGP-ENSIC, 1 Rue Grandville, 54000 Nancy, France*
jsernaro@unal.edu.co

Abstract

The decision-making process for the design of formulated products faces different challenges because of its intrinsic complexity. On the one hand, it is not sequential, but iterative due to the fragmented and heterogeneous nature of available information. On the other hand, there is not a unique design workflow because it changes from company to company according to its context and specific requirements. The lack of structure of knowledge for product formulation requires developing a robust knowledge representation to show coherently and explicitly concepts, models, and data. Furthermore, this representation must allow design teams to use it flexibly and to adapt it to specific design contexts. In view of the above, this work proposes an ontology for formulated products with emphasis on cosmetic emulsions. This ontology integrates concepts from emulsion science, cosmetic formulation, expert knowledge, and design heuristics in a systematic and accessible way. It was done based on the recent work of our research group in Chemical Product Design. This document shows an overview of the ontology and one of its possible applications: verification of the formulation of a skin care cream. As a conclusion, it was found that the ontology enables the access to precise information according to design requirements. It is a versatile and useful information tool for the design of emulsion-based products.

Keywords: Product design, Cosmetic emulsions, Ontology

1. Introduction

Formulated chemical products, such as cosmetics, are complex systems whose properties are defined by a synergistic action of ingredients, composition, and production process. Their design is a scientific and organizational challenge, as it requires the management of heterogeneous information from multiple sources and the application of multidisciplinary knowledge (Zhang et al., 2020). Available information for their design is huge and diverse in nature: there are multiple ingredient databases, property models, heuristics, knowledge from colloidal science, regulations, recommendations, guidelines, etc. Additionally, in the field of Chemical Product Design (CPD), there are several methodologies and computer-aided methods. Examples are as a design methodology for cosmetic formulations incorporating heuristics (Arrieta-Escobar et al., 2020), a methodology for formulated products applying expert knowledge (Serna et al., 2021), and

an integrated framework for formulated products including liquids, solids, emulsions (Zhang et al., 2017), among others. Despite its abundance, the application of the existing information for product formulation of is not straight-forward (Zhang et al., 2020). Design methodologies are not adapted to practical design processes, which are iterative and change according to the design context. Additionally, available information is not centralized, has different forms (qualitative and quantitative), and it is frequently fragmented. Thus, it is necessary to create tools capable of managing real design contexts and available information to solve practical problems. In view of the above, this work proposes OntoCosmetic, an ontology to systematically represent knowledge for the formulation of cosmetic products of the type of homogenous mixture or emulsion. It is proposed to support data storage and analysis, information modeling and decision-making for cosmetic formulation. It is based on the knowledge and information gathered and developed by our design team in CPD (Arrieta-Escobar et al., 2020) (Serna et al., 2021) and it has a structure that can be extended to represent more general information in the field of formulated chemical products. The ontology was developed following the ontology creation methodology described by Gabriel (2019) and formalized using the software Protégé¹. In this document, section 2 presents a short introduction to ontologies and an overview of OntoCosmetic. Section 3 presents an application of the ontology for the verification of a formulation in relation to a heuristic database. Finally, Section 4 highlights some conclusions and perspectives.

2. An ontology for cosmetic product design

2.1. Concept of ontology

Ontology is “an explicit specification of a conceptualization” which is an abstract, simplified view of the world that is represented for some purpose (Gruber, 1993). It is holistic since it contains all relevant concepts and relationships between concepts to model a specific subject or phenomenon (Hailemariam & Venkatasubramanian, 2010). It is also as precise as required because it is normally defined within a specific domain of knowledge and with a clear purpose (Gabriel et al., 2019). In software engineering, ontologies are seen as “a means to efficiently build the knowledge-based necessary to effectively support engineering work processes” (Marquardt et al., 2010).

The main elements of an ontology are: *Classes* which are concepts or sets containing instances (e.g. the class ‘Ingredient’ in a cosmetic formulation). *Relations* which show the relation between classes (e.g., the relation ‘has’, connecting ‘Formulation’ with ‘Ingredient’). *Instances* which are the elements of the classes (e.g., ‘Argan oil’, which belongs to the class ‘Ingredient’). *Rules* which are logical statements that can be used to make inference (e.g., heuristics).

2.2. An ontology for cosmetic product formulation

OntoCosmetic was constructed based on the ontology OntoCAPE (Marquardt et al., 2010), a very elaborated ontology for process design in the field of Chemical Engineering. In OntoCAPE, products and processes are technical systems, which are developed in an engineering design process to meet defined requirements (Marquardt et al., 2010). There, technical systems are represented through five viewpoints: requirements, function, realization, behavior, and performance. OntoCosmetic uses part of this description to define the Product System and adapts it to the application of cosmetic formulation.

Figure 1 shows an overview of OntoCosmetic. It is organized in two interconnected blocks: A) Product System, containing all the relevant concepts describing a cosmetic formulation, B) Product Model, containing a knowledge base of models to support practical decision-making during the formulation process.

¹ <http://protege.stanford.edu/>

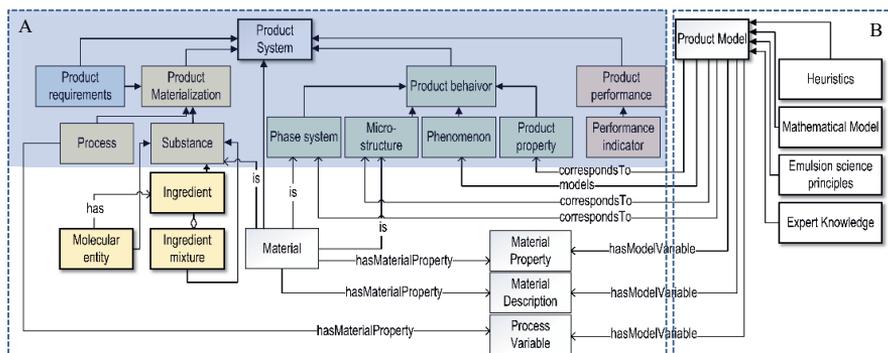


Figure 1. Overview of the OntoCosmetic ontology. A) Product System B) Product Model. The dark area corresponds to the concepts adapted from OntoCAPE

In OntoCosmetic, Product System is divided in four viewpoints:

- Product requirements: It corresponds to the product desired behavior, represented by product specifications and target values. It is a 'to be' relation.
- Product materialization: It corresponds to the decisions that designers can make and implement to achieve product requirements. For example, the selection of ingredients. Product materialization has the modules of process and substances. The latter is a super class of the ingredient class which is connected to a database of cosmetic ingredients.
- Product behavior: It corresponds to the real behavior of the product whether it is expected or not. It can be characterized quantitatively, through product properties, or qualitatively, through a qualitative product description. Product behavior can also be related to a phenomenon, to the micro-structure of the product (in the case of emulsions, microstructure is frequently characterized by particle size and particle size distribution), and to the phase system (which describes the product as a thermodynamic system).
- Product performance: It corresponds to the evaluation of the Product System specifically in relation to product requirements. It indicates whether the product has the desired performance (as described by product requirements) or not. The evaluation is done through target value indicators and sustainability indicators. In this case, in contrast to product requirements, indicators are related to the product system with a 'as is' relation.

The ontology has been developed with focus on cosmetic applications. Thus, it contains concepts and databases specifically related to this subject. In the ontology, ingredients are classified in subclasses that enable their search based on cosmetic related categories. Figure 2 shows some sub-classes of the class Ingredient with their definitions. Figure 3 shows some attributes of the sub-class Emollient with their definitions.

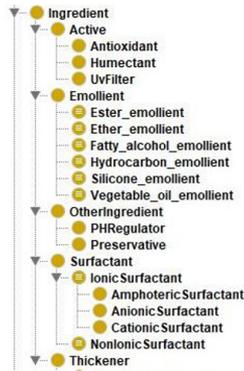


Figure 2. Extract of Ingredient sub-classes in OntoCosmetic

Table 1. Extract of Ingredient sub-classes in OntoCosmetic

Function	Description
Ingredient	Chemical compounds that are part of a formulation
Emollient	Oils used in cosmetic emulsions for several purposes: to make skin and hair softener, to serve as solvents for liposoluble actives. Some of them can act as occlusive ingredients.
Surfactant	Ingredient used in emulsions for three purposes: to decrease the energy required during emulsification, to adsorb rapidly at the interface of recently formed droplets, and to generate a barrier for long-term emulsion stability.

By its part, Product Model contains the models and tools that enable designers to make decisions. Product models are classified in the following classes:

- Heuristic: It is a guideline or rule of thumb expressed in natural language that gives a practical instruction to achieve a product requirement. It can be transformed into an algebraical mathematical model.
- Expert knowledge: It is a qualitative relation between the classes of product behavior and product materialization. They were established with the help of experts in emulsion science. They are used in combination with emulsion science principles.
- Mathematical model: It is an equation that enables a quantitative definition of ingredients in relation to product requirements.
- Emulsion science principles: They are scientific descriptions of emulsion phenomena, which explain some product system behaviors and can be used for decision-making in combination with expert knowledge.

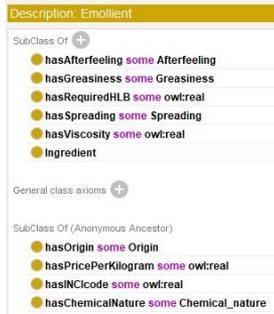


Figure 3. Some attributes of the class emollient

Table2. Some attributes of the class emollient

Function	Type	Description
requiredHLB	Real	It determines the compatibility of oils with the surfactant system
greasiness	String	It that reflects the oiliness of the skin after a lotion is applied. Allowed values: {Very high, High, Medium, Low}
viscosity	Real	It is the is a measure of the resistance of a fluid to deformation.
origin	String	Ingredient source. Allowed values: {natural, from nature raw materials, from natural and synthetic raw materials, synthetic}

As well, Product models are classified into sub-classes according to their applicability (general emulsions, cosmetic emulsions, specific cosmetic applications). Additionally, their implementation is related to the product materialization system (a model can be related to one or more ingredient class, sub-class, or attribute), and their effect is related to the classes of product behavior and product performance systems. Examples of product models and their relations are presented in Table 3.

Table 3. Some product models and their relations

Product model	Description	Application	Relations	System	Range
Heuristic	For a good sensorial, at least one emollient of each spreading type should be used: high, medium, and low.	Cosmetic emulsion	isImplementedBy	Product Materialization	Emollient; Spreading
			hasEffectOn	Product Performance	Performance indicator; Sensorial
Emulsion Science	Emulsion stability is affected by the phenomenon gravitational separation.	General	hasEffectOn	Product Performance	Performance indicator; Stability
		General	isRelatedtoPhenomenon	Product Behavior	PhenomenonGravitationalSeparation
Emulsion Science	Gravitational separation can be controlled with the addition of a thickener.	General	isImplementedBy	Product Materialization	Thickener; Concentration
		General	isRelatedtoPhenomenon	Product Behavior	PhenomenonGravitationalSeparation

3. Application: Checking if a given formulation follows the heuristics

The Ontology was used to verify if a given cosmetic formulation follows or not a group of heuristics, and thus if it is recommended based on them. The verification was done using the Semantic Web Rule Language (SWLR) in Protégé, where heuristics can be written as logic rules involving instances of classes. The formulation checked against the heuristics is shown in Table 4 and the considered heuristics are presented in Table 5. As result, the reasoner classified the formulation as recommended according to selected heuristics (ConformFormulation).

Table 4. Formulation checked against some heuristics

Class	Ingredient	%
Emollient	Ethylhexyl Palmitate	4
Emollient	Olive oil	3
Emollient	Caprylic/Capric Triglyceride	9
Surfactant	Glyceryl Stearate Citrate	2.5
Surfactant	Oleth-3	0.5
Surfactant	Sorbitan Monostearate	1
Preservative	Cosgard	0,7
Antioxidant	Tocopherol	0,5
Humectant	Glycerin	3
Thickener	Xanthan gum	0.1
	Water	To 100

Table 5. Heuristics considered to test the formulation using the ontology

Product model	Description	System	Range
Heuristic	at least one emollient of each spreading type should be used: high, medium, and low	Product Materialization	Emollient; Spreading
Heuristic	Use a minimum level of 2% and up to 5% of surfactants	Product Materialization	Surfactant; SurfactantComposition
Heuristic	At least one ingredient should be of the natural origin	Product Materialization	Ingredient; Origin
Heuristic	surfactant/emollient ratio should be between 1:4 to 1:6	Product Materialization	SurfactantComposition; EmollientComposition
Heuristic	HLB should match the RHLB required	Product Materialization	Surfactant; Emollient; SurfactantComposition; EmollientComposition

As shown in Figure 4, the formulation is introduced to the system as an instance with relations to ingredients (composition) and rules (heuristics that the formulation should

follow). Based on this information, the reasoner activates the required rules and classifies the formulation as *ConformFormulation*, when it follows the rules or as *NonConformFormulation*, when it does not. The formulation must be experimentally validated to verify that it meets the requirements.

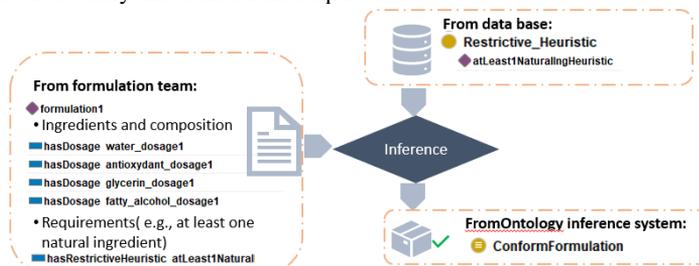


Figure 4. Results of the verification in Protégé

4. Conclusions and perspectives

The article presents a brief overview of *OntoCosmetic*, an ontology for cosmetic product design. Currently, the ontology formalizes concepts of product formulation. It can be used to create information systems with common definitions, to call ingredient and formulation information through a SPARQL² request, to apply heuristics through the activation of rules. The perspective of this ontology is to create an application programmer interface (API) to allow the integration of knowledge into a software. This software would assist the decision-making process during design as well as visualize data and optimize the formulation of cosmetic products considering multiple criteria.

References

- Arrieta-Escobar, J. A., Camargo, M., Morel, L., Bernardo, F., Orjuela, A., & Wendling, L. (2020). Methodology for the Design of Formulated Products Incorporating Consumer Preferences and Heuristic Knowledge. <https://doi.org/10.22541/au.158879155.54817449>
- Gabriel, A., Chavez, B. P., & Davy, M. (2019). Methodology to design ontologies from organizational models: Application to creativity workshops. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 33(2), 148–159. <https://doi.org/10.1017/S0890060419000088>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5, 199–220. <https://doi.org/10.1006/knac.1993.1008>
- Hailemariam, L., & Venkatasubramanian, V. (2010). Purdue ontology for pharmaceutical engineering: Part I. Conceptual framework. *Journal of Pharmaceutical Innovation*, 5(3), 88–99. <https://doi.org/10.1007/s12247-010-9081-3>
- Marquardt, W., Morbach, J., Wiesner, A., & Yang, A. (2010). *OntoCAPE A Re-usable Ontology for Chemical Process Engineering*. In RWTHedition. Springer.
- Serna, J., Narváez Rincón, P. C., Falk, V., Boly, V., & Camargo, M. (2021). A Methodology for Emulsion Design Based on Emulsion Science and Expert Knowledge. Part 1: Conceptual Approach. *Industrial & Engineering Chemistry Research*, 60(7), 3210–3227. <https://doi.org/10.1021/acs.iecr.0c04942>
- Zhang, L., Fun, K. Y., Zhang, X., Fun, H. K., & Ng, K. M. (2017). An integrated framework for designing formulated products. *Computers and Chemical Engineering*, 107, 61–76. <https://doi.org/10.1016/j.compchemeng.2017.05.014>
- Zhang, L., Mao, H., Liu, Q., & Gani, R. (2020). Chemical product design – recent advances and perspectives. *Current Opinion in Chemical Engineering*, 27(November 2019), 22–34. <https://doi.org/10.1016/j.coche.2019.10.005>

² SPARQL Protocol and Resource Description Framework (RDF) Query Language

Synthesis of extractive distillation structures for the purification of ethanol

Tomoki Nakamura,^a J. Rafael Alcántara-Ávila,^{a*} Julián Cabrera-Ruiz^{b*}

^a*Department of Chemical Engineering, Kyoto University, Kyoto, 615-8510, Japan*

^b*Departamento de Ingeniería Química, Universidad de Guanajuato. Noria Alta S/N, Guanajuato, 36050, Mexico*

Corresponding authors email: jrafael@cheme.kyoto-u.ac.jp; j.cabrerar Ruiz@ugto.mx

Abstract

This work presents a synthesis problem for the separation and purification of Ethanol using extractive distillation. Glycerol was the chosen solvent to separate the azeotropic mixture. The synthesis problem was based on the Infinite-DimensionAl State-space (IDEAS) approach to discretize the liquid composition space represented by a ternary diagram. Then, a distillation module was assigned to each subspace, and a superstructure representation was used in which all distillation modules connect to generate an optimal distillation structure. The synthesis problem was formulated as a linear programming (LP) problem in which the utility cost was minimized subject to material and energy balances. Furthermore, the optimal structure was derived without the need to set any preestablish distillation structure. After the optimization problem is solved, the best connection between distillation modules is known. Therefore, a realistic and feasible distillation structure was interpreted at a post-optimization step.

Keywords: Process Synthesis, Extractive Distillation, LP, Process Simulation, Ethanol production

1. Main Text

Ethanol obtained from lignocellulosic biomass can be used in the chemical industry as an intermediate material for the synthesis of esters and ethers and as a solvent in the production of paint, cosmetics, sprays, perfumery, medicine, and food, among others. Furthermore, mixtures of anhydrous ethanol and gasoline can be used as fuels, reducing environmental contamination and improving gasoline's octane index (Gil et al., 2014). In order to obtain high-purity ethanol over 99.5 mol%, the water-ethanol azeotrope (89 mol%) must be surpassed.

Conventional distillation fails to obtain high-purity ethanol because as the water-ethanol mixture approaches the azeotrope, their vapor and liquid compositions become the same. Thus, further separation is not possible. However, by adding a third component acting as a mass separating agent (MSA), the vapor liquid equilibrium relationships between water and ethanol change so as to make possible ethanol purification. Therefore, this work deals with the extractive distillation of ethanol using glycerol as solvent.

The water-ethanol separation has been extensively researched in the past decades where the separation alternatives have been mainly related with extractive distillation (Dai et al., 2014), azeotropic distillation (Luyben, 2006), pressure-swing distillation (Mulia-Soto and Flores-Tlacuahuac, 2011), and membrane-assisted distillation (Kunnakorn et al., 2013).

In this work, the synthesis of extractive distillation structures for the purification of ethanol is studied through a low aggregation module-based superstructure framework consisting of the following steps: 1) superstructure representation, 2) modules generation, 3) linear programming (LP) optimization, 4) solution interpretation, and 5) solution validation.

2. Problem statement

Given the feed conditions and product specification for the separation of a water-ethanol mixture, the used solvent, and the operating pressure, a synthesis problem based on distillation modules can be proposed to derive optimal distillation structures.

Synthesis problems can be divided according to their aggregation level. Superstructures of high aggregation level comprise a set of equipment or tasks subject to a preestablished connectivity while superstructures of low aggregation level are based on a phenomenological approach, which combines cooling, heating, mixing, splitting, reaction, etc. (Alcantara-Avila et al., 2021). The use of low aggregation superstructures has been already proposed for the purification of ethanol. Kuhlmann et al., (2018) adopted the Phenomena Building Blocks (PBBs) framework for the synthesis of a membrane-assisted distillation process while Tian and Pistikopoulos, (2019) adopted a Generalized Modular Representation Framework (GMF) for the synthesis of an extractive distillation process. This work is based on the Infinite Dimensional State-space (IDEAS) framework proposed by Drake and Manousiouthakis, (2002) in which a process can be represented by a network of units operations.

2.1. Superstructure representation

The proposed superstructure in this work is composed by five types of modules: a heating module, a cooling module, two feed modules, two product modules, and a set of distillation modules. Figure 1 shows the representation of a distillation module.

Since the liquid composition space of the ternary mixture is discretized in subspaces, by definition, a distillation module is a subspace in which the liquid and vapor compositions and enthalpies are uniquely define given the pressure is known.

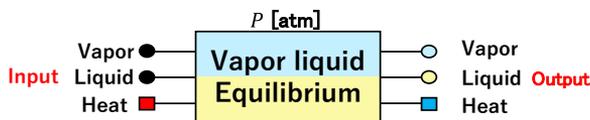


Figure 1. Conceptual representation of a distillation module

For each module i , the liquid molar fraction of component k (x_{ik}), vapor molar fraction of component k (y_{ik}), liquid enthalpy (h_i^L), and vapor enthalpy (h_i^V) are uniquely defined and entered as parameter in the optimization problem. The combination of modules will result in the optimal process.

Figure 2 shows the proposed superstructure representation in this work. The modules in the superstructure contain heat and mass inputs and-or outputs. For example, the heating, feed, and solvent modules have only outputs while cooling and product modules have only inputs, and the distillation module have both. The superstructure considers all possible flow connections between distillation modules and other modules.

3. Mathematical formulation

The minimize utility cost (UC) minimization was taken as objective function because the distillation cost is largely dominated by UC. Equation 1 shows the objective function.

$$UC = \sum_{i \in S_M} (c_H Q_i^H + c_C Q_i^C) \quad 1$$

where S_M the set of distillation modules, c_H and c_C are the cost of heating and that of cooling per unit amount of energy. Q_i^H and Q_i^C are the heating amount entering a module i and that of cooling leaving a module i .

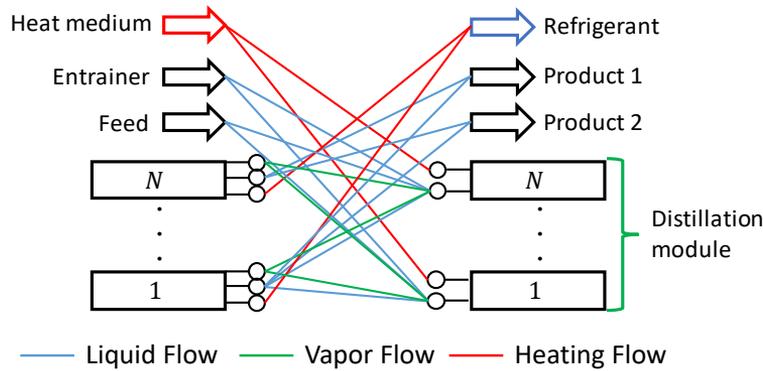


Figure 2. Conceptual representation of the module-based superstructure

Equation 2 shows the overall mass balance while Eq. 3 shows the component mass balance for the two products in each distillation module

$$\sum_{i \in S_M} (L_{ji} + V_{ji}) + L_{Fi} + L_{Si} - \sum_{j \in S_M} (L_{ij} + V_{ij}) - \sum_{p \in S_p} L_{ip}^P = 0 \quad 2$$

$$\sum_{i \in S_M} (L_{ji} x_{jk} + V_{ji} y_{jk}) + L_{Fi} x_k^F + L_{Si} x_k^S - \sum_{j \in S_M} (L_{ij} x_{ik} + V_{ij} y_{ik}) - x_{ik} \sum_{p \in S_p} L_{ip}^P = 0 \quad 3$$

where L_{ij} and V_{ij} are the liquid and vapor molar flows from module i to module j , L_{Fi} is the feed molar flow to module i , L_{Si} is the liquid solvent molar flow to module i , L_{ip}^P is the product molar flow amount from module i . x_k^F is the liquid molar fraction of component k in the feed, and x_k^S is the liquid molar fraction of component k in the solvent, respectively.

Equation 4 shows the overall heat balance in each distillation module

$$\sum_{j \in S_M} (L_{ji} h_j^L + V_{ji} h_j^V) + L_{Fi} h^F + L_{Si} h^S + Q_i^H - \sum_{j \in S_M} (L_{ij} h_i^L + V_{ij} h_i^V) - h_i^L \sum_{p \in S_p} L_{ip}^P - Q_i^C = 0 \quad 4$$

where h^F and h^S are enthalpy of feed, and the solvent, respectively.

Equations 5 and 6 are additional equations for the connectivity between the feed module and distillation modules, and between distillation modules and product modules, respectively.

$$\sum_{i \in S_M} (L_{Fi}) - L^F = 0 \quad 5$$

$$x_k^P \sum_{j \in S_M} (L_{jk}^P) - \sum_{j \in S_M} (L_{jk}^P x_{jk}) \geq 0 \quad \{k = 1, 2\} \quad 6$$

where L_{Fi} is the liquid feed flow to module i , and x_k^P is the molar purity target of component k . Moreover, since all the proposed equations in this section are linear, the optimization problem can be solved as a linear programming (LP) problem.

3.1. Case study

The feed flow rate of the water-ethanol mixture is 100 kmol/h, and its liquid molar fraction of ethanol is 0.894 and that of water in 0.106. The molar fraction specifications for ethanol, water and glycerol must be more than 0.95. The pressure in all modules is 1 atm. The feed and the product flows are saturated liquid. The heating cost is 10 \$/GJ and the cooling cost is 0.5 \$/GJ.

As for the discretization of the liquid compositions for the distillation modules, each of ethanol, water, and glycerol molar composition was discretized with the width of 0.02.

The vapor liquid equilibrium and each phase enthalpy were calculated by using NRTL model. The calculations done in Aspen Plus V10®. The optimization was done by IBM ILOG CPLEX Optimization Studio 12.8.0.

4. Results and Discussion

Figures 3 and 4 show the plot of the connectivity of liquid and vapor flows between distillation modules in a ternary diagram.

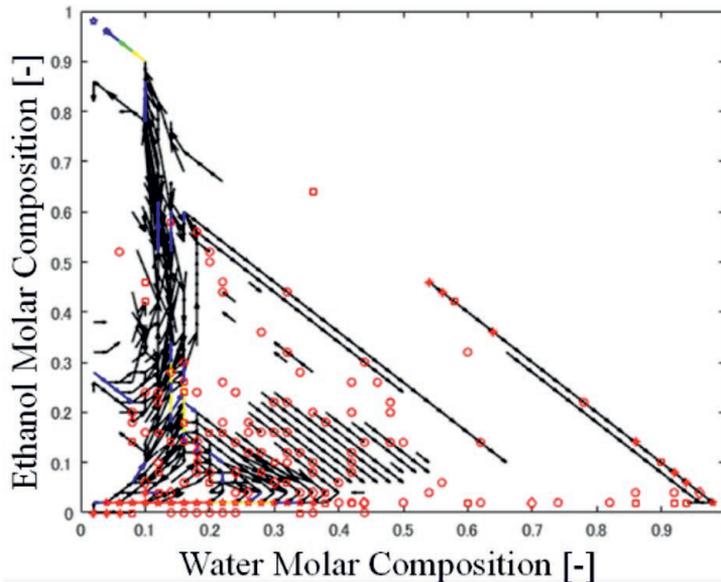


Figure 3. Liquid flow path of the optimal solution

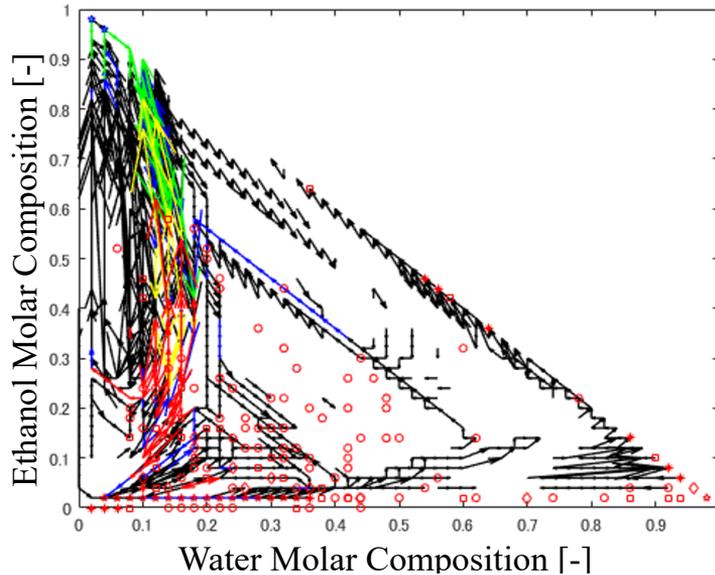


Figure 4. Vapor flow path of the optimal solution

The utility cost of the optimization result was 9.36 \$/h. Table 1 shows the description of colors and shapes for the ternary diagrams in Figures 2.

Table 1. The colors and shapes of modules and flows in Figs. 3 and 4

Shapes and colors	explanation	amount
→	Flow direction	More than 1 kmol/h
→	Flow direction	Less than 17 kmol/h
→	Flow direction	Less than 25 kmol/h
→	Flow direction	Less than 33 kmol/h
→	Flow direction	More than 33 kmol/h
○, ○	Amount of heating or cooling	More than 0 kJ/h
□, □	Amount of heating or cooling	Less than 5000 kJ/h
◇, ◇	Amount of heating or cooling	Less than 10000 kJ/h
*, *	Amount of heating or cooling	Less than 30000 kJ/h
★, ★	Amount of heating or cooling	More than 30000 kJ/h

It can be seen from Figs. 3 and 4 that cooling is necessary to separate Ethanol while heating is necessary to separate Water and Glycerol. Also, most of liquid flows are less than 1 kmol/h. However, modules with low molar fraction of water have liquid flows less than 17 kmol/h. On the other hand, there are large flows of vapor in the region with low molar fraction of water.

Figure 5 shows the interpreted solution of the optimization results in Figs. 3 and 4. The interpreted solution exhibits several features: 1) there is no need to use a condenser in the first column, 2) water is separated in second column, 3) thermal coupling between stages makes possible to increase the internal flow of vapor streams, and 4) there is no need to cool the recycled solvent streams since the process will use this energy directly when mixed with the feed steam. The utility cost of the interpreted solution is 60.8 \$/h.

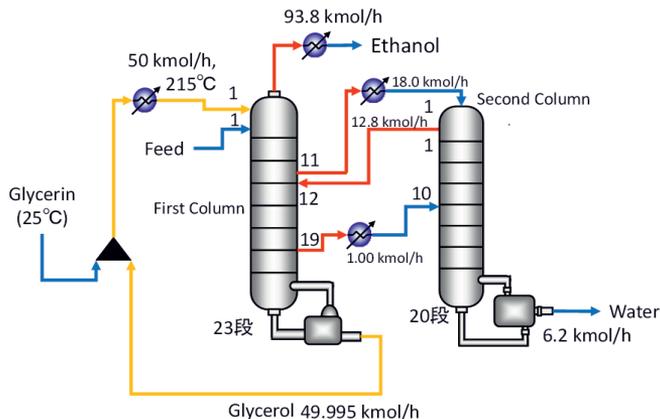


Figure 5. Interpreted distillation structure

5. Conclusions

This work presented a module-based framework for the synthesis of extractive distillation column. The modules are collected and represented in a ternary diagram. The interpreted solution showed an intensified distillation column with a cost much higher than the cost obtained after solving the LP problem. The derived solution was done without any preestablished connectivity. Further improvements at the interpretation step must be done.

References

- J.R. Alcantara-Avila, Z.Y. Kong, H.-Y. Lee, J. Sunarso, 2021, Advancements in Optimization and Control Techniques for Intensifying Processes. *Processes* 9, 12, 2150.
- C. Dai, Z. Lei, X. Xi, J. Zhu, B. Chen, 2014, Extractive distillation with a mixture of organic solvent and ionic liquid as entrainer, *Ind. Eng. Chem. Res.*, 53, 15786–15791.
- J.E. Drake, V. Manousiouthakis, 2002. IDEAS approach to process network synthesis: minimum utility cost for complex distillation networks, *Chem. Eng. Sci.*, 57, 3095–3106.
- I.D. Gil, L.C. García, G. Rodríguez, 2014, Simulation of ethanol extractive distillation with mixed glycols as separating agent, *Brazilian J. Chem. Eng.*, 31, 259–270.
- H. Kuhlmann, H. Veith, M. Möller, K.P. Nguyen, A. Górak, M. Skiborowski, M., 2018, Optimization-Based Approach to Process Synthesis for Process Intensification: Synthesis of Reaction-Separation Processes. *Ind. Eng. Chem. Res.* 57, 3639–3655.
- D. Kunnakorn, T. Rirksomboon, K. Siemanond, K. Aungkavattana, P. Aungkavattana, N. Kuanchertchoo, N. Chuntanaler, P. Hemra, K. Kulprathipanja, S. James, R.B., Wongkasemjit, S., 2013. Techno-economic comparison of energy usage between azeotropic distillation and hybrid system for water-ethanol separation. *Renew. Energy* 51, 310–316.
- W.L. Luyben, 2006, Control of a multiunit heterogeneous azeotropic distillation process, *AIChE J.* 52, 623–637.
- J.F. Mulia-Soto, A. Flores-Tlacuahuac, 2011,. Modeling, simulation and control of an internally heat integrated pressure-swing distillation process for bioethanol separation. *Comput. Chem. Eng.* 35, 1532–1546.
- Y., Tian, E.N. Pistikopoulos, 2019, Generalized Modular Representation Framework for the Synthesis of Extractive Separation Systems. *Comput. Aided Chem. Eng.* 47, 475–480.

Optimal design for flexible operation with multiple fluctuating input parameters

Bastian Bruns^{a,*}, Marcus Grünewald^a and Julia Riese^a

^a*Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany
bastian.bruns@fluidvt.rub.de*

Abstract

Feasibility is a crucial aspect when dealing with flexible operation of processes, as it is, e.g., the case when incorporating renewable energies. Process design and operation should thus account for the operational fluctuations. When flexible operation with multiple fluctuating input parameters is required, a design that increases the feasible region spanned by these parameters can be advantageous. In this paper, we study the optimal design of a fixed bed reactor for the CO₂ hydrogenation to methanol that is operated under variable capacities. We introduce a design approach to optimize the multidimensional operating window of the process. Dynamic optimization is used as an overarching framework to adapt the design of the fixed bed reactor, increasing the flexibility of the process in terms of H₂ and CO₂ feed capacity.

Keywords: Design optimization, flexibility, power-to-x, dynamic operability, methanol production

1. Introduction

Chemical process plants are constantly subject to disturbances and need to respond appropriately in order to meet safety, environmental or operational constraints. Early works found that these disturbances should be considered as uncertainties during the design phase of processes. Mathematical tools for steady-state (Swaney and Grossmann (1985)) and dynamic analysis (Dimitriadis and Pistikopoulos (1995)) were developed to provide processes able to overcome operational flexibility limitations. These tools are able to determine flexibility of a process system by means of a hyperrectangle that is centered at the nominal point in the space of uncertain parameters. The length of the sides of the hyperrectangle are defined by the nearest boundary. This measure can then be used to adapt design parameters to provide feasibility in uncertain conditions. Most of the studies analyzing flexibility under uncertainty are, however, based on open loop processes.

Therefore, there have been efforts towards the consideration of design and control actions of processes subject to disturbances. Early research showed that an integrated approach to design and control processes is superior to a sequential one when dealing with these disturbances during operation (Swartz and Kawajiri (2019)). However, these integrated design and control strategies were usually developed to maintain an optimal operation around a nominal steady-state operating point.

Due to recent developments in chemical engineering such as the increasing incorporation of renewable energies, flexible operation of chemical plants has received considerable interest (e.g. Caspari et al. (2019); Otashu and Baldea (2019); Fischer and Freund (2020)). Flexible operation implies significant deviations from steady-state operating conditions and requires that process design, and operation account for these fluctuations in advance. This includes that control actions are found that allow a feasible and fast transient behavior. When these fluctuations are the main

concern, a system that can compensate for larger deviations in its operating conditions is advantageous. However, there are few design approaches (e.g. Cao et al. (2015)) that consider the size of the operating window of a process.

In this work, we introduce a design approach that considers the multidimensional operating window of a process. The aim of the design optimization is to provide a process design that is highly flexible towards multiple fluctuating input parameters. We demonstrate the approach on the production process of green methanol, as this is a realistic process scenario for flexible operation (Mbatha et al. (2021)). We consider that CO₂ is produced by absorption in a fossil power plant, operated for frequency regulations. Hydrogen required by the process is produced by a water electrolysis that operates with renewable energies. Although there is an optimum in the H₂ and CO₂ ratio, we assume that each flow rate varies independently, neglecting possible storage capacities.

2. Design Problem

During flexible operation, certain intended variables undergo significant deviations from their nominal operating conditions. These fluctuating variables are limited by process boundaries and thus span the operating window of the process. Boundaries can be, e.g., process performance, or environmental, and safety issues. In the space of uncertain parameters the operating window corresponds to the feasible region of the process. The size of the operating window determines the feasible combinations of the intended fluctuating variables. Thus, increasing the size of the operating window, increases the feasible deviations of the corresponding variables. It becomes evident that flexible operation benefits from an increased size of the operating window, as it allows for more feasible operating points.

We consider the load ξ of a process or its equipment as an auxiliary variable that determines the operating point in accordance with the intended changing variables (flexibility dictating parameter β) during flexible operation. The definition of the operating window considered here follows the approach proposed recently in Bruns et al. (2021). The load is then a function of the vector of design d and operational variables x , and the flexibility dictating parameters:

$$\xi = f(d, x, \beta) \quad (1)$$

Consequently, the operating window is the space of all feasible loads bounded by all considered process limitations. To optimize the operating window of the process, we determine the maximum ranges over which each flexibility dictating parameter may vary, while remaining in the operating window. The size of the operating window ξ^{ow} is then characterized by the product of the difference between the maximum and minimum deviation of each parameter:

$$\xi^{\text{ow}} = \prod_{i=1}^n (\xi_n^{\text{max}} - \xi_n^{\text{min}}) \quad (2)$$

where n is the number of flexibility dictating parameters, and ξ_n^{max} and ξ_n^{min} are the farthest loads on the process boundaries. This results in a multidimensional rectangle, where the length of its sides correspond to the difference between maximum and minimum deviation of parameters β .

This representation of the operating window can be incorporated into design optimization by means of dynamic optimization. By considering a time horizon τ , in which the flexibility dictating parameters represent the degrees of freedom that are available during operation, the farthest loads on the process boundaries can be determined.

3. Methanol Synthesis Modeling

The investigated green methanol synthesis process is shown in Figure 1. Fresh feed of H₂ and CO₂ enters the process and is then mixed with the recycle stream. Before entering the methanol reactor,

the gas mixture is preheated to 498 K. Methanol is then formed in the tube bundle reactor over a commercial Cu/Zn/Al₂O₃ catalyst at an inlet pressure of 55 bar. In the tube bundle reactor 150 tubes with a length of 8 m and a tube diameter of 0.058 m are assembled in parallel. The catalysts particle diameter is 0.006 m with a density of 1175 kg m⁻³. The void fraction in the reactor is 0.4. The tubes are surrounded by a shell where an isothermal heat exchange fluid regulates the temperature in the reactor. The produced gas mixture is then condensed after the reactor in a separator in order to remove and recycle unreacted gases from methanol and water. 10 % of the recycled gases are purged from the process.

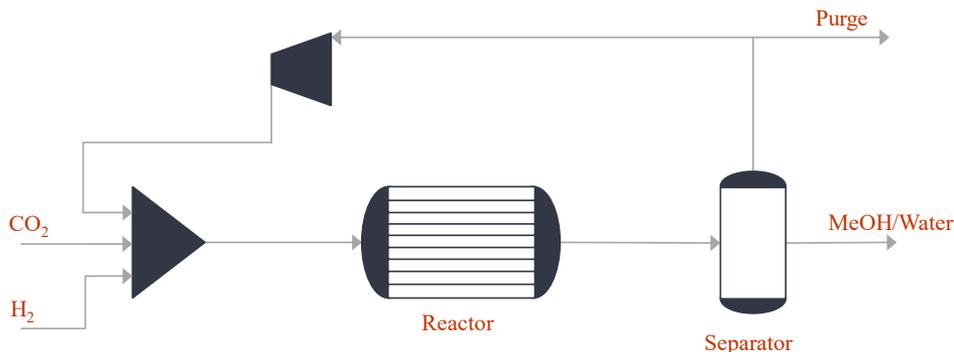


Figure 1: Methanol synthesis process.

The tube bundle reactor is represented by a one dimensional plug-flow model, considering heat capacities and weight of the catalyst bed, and the reactor material. The model consists of component material balances and an energy balance in the form of partial-differential equations. Reaction rates, pressure drop correlation and heat transfer are considered in the form of algebraic equations. We employed reaction kinetics by Bussche and Froment (1996) to calculate the reaction rates of each component over the length of the reactor. The overall heat transfer coefficient considered in the model between reactor wall and heat exchange fluid is 1000 J s⁻¹ m⁻².

The separator is modeled by a set of differential-algebraic equations, calculating the vapor-liquid equilibrium, while considering liquid and gas hold-up in the vessel. Valves and the compressor in the process are modeled to represent the typical behavior of the process equipment by efficiency relationships and pressure drops. The entire process is solved pressure-driven in Aspen Custom Modeler v10.

4. Optimal Process Design

The design decision variables for the methanol reactor are the length of the reactor ($l \in [7, 9]$ m) and the diameter of the tubes ($d_t \in [0.02, 0.065]$ m). Operational decision variables are the heat exchange fluid temperature ($T_{cw} \in [473.15, 543.15]$ K) and the reactor inlet temperature ($T_{in} \in [473.15, 543.15]$ K). Process constraints are: 1) minimum productivity of 1000 kg h⁻¹, 2) minimum CO₂ conversion of 0.1, 3) maximum split fraction of the purge of 0.15, 4) maximum pressure drop of 1.5 bar, and 5) maximum hot spot temperature of 543.15 K. Due to the pressure-driven simulation of the methanol synthesis process, the split fraction of the purge is not a fixed value, but rather calculated according to the pressure in the system. The split fraction is thus a value that changes during operation and should not exceed a certain value to prevent excessive loss of reactant.

We want to design the methanol reactor such that its operating window in the space of the considered flexibility dictating parameters is optimized. The flexibility dictating parameters for this case

are the H₂ feed flow rate $\beta_1 = \dot{n}_{H_2}$ and the CO₂ feed flow rate $\beta_2 = \dot{n}_{CO_2}$. The objective function is thus defined as follows:

$$\begin{aligned} \xi^{\text{ow,MeOH}} &= \prod_{i=1}^2 \left(\xi_n^{\text{upper}} - \xi_n^{\text{lower}} \right) \\ &= \prod_{i=1}^2 \left(\mathbf{f}(\mathbf{d}\mathbf{v}, \mathbf{x}, \beta_n^{\text{upper}}, \tau + \Delta t) - \mathbf{f}(\mathbf{d}\mathbf{v}, \mathbf{x}, \beta_n^{\text{lower}}, \tau) \right) \end{aligned} \quad (3)$$

where \mathbf{f} represents all model equations considered in the methanol synthesis process and ξ_n^{upper} and ξ_n^{lower} are the upper and lower load depending on the respective flexibility dictating parameter value. At the process boundaries ξ_n^{upper} and ξ_n^{lower} equal ξ_n^{max} and ξ_n^{min} . $\tau + \Delta t$ is the time horizon, where the time step Δt considers the dynamic behavior of the process and needs to be sufficiently large so that steady-state is reached again. For the methanol synthesis process we chose $\Delta t = 500$ s. β_1 and β_2 are degrees of freedom in addition to the design and operational decision variables that are optimized at $t = \tau$. We thus maximize the rectangle resulting from the farthest loads in the feasible region in the space of β_1 and β_2 . The dynamic optimization is performed by means of scripts integrated in the simulation software.

We investigate two scenarios: first, solely the design decision variables are optimized and second, both design and operational variables of the reactor are optimized. Dynamic feasibility is assessed for both scenarios by means of path constraints. We also use the dynamic optimization framework to test the investigated operational variables to be adapted as advanced control strategy during operation to satisfy the path constraints. Both scenarios are compared to the base case, which was described in Section 3.

5. Results and Discussion

The optimization results for both scenarios and for the base case are shown in Table 1. The process model is highly nonlinear so that various initial guesses were required to find the optimal solution for each scenario. In the base case no decision variables were considered, and by that we just evaluate the size of the operating window. The maximum and minimum deviation for both H₂ (318.39 kmol h⁻¹ and 96.22 kmol h⁻¹) and CO₂ (106.19 kmol h⁻¹ and 30.44 kmol h⁻¹) feed flow rates are determined. The resulting scaled operating window size is 16.83. We performed the same optimization with reactor length and tube diameter as decision variables in scenario 1. The optimal design has a reactor length at the lower design boundary and a tube diameter at the upper design boundary. The size of the operating window increases by 54.03 %. By additionally considering the inlet temperature and the thermal fluid temperature during the design optimization, as in scenario 2, the determined operating window can be further enhanced compared to the base case (60.56 %). The optimal inlet temperature changed only insignificantly, while the thermal fluid temperature has increased by 7.99 K.

For both considered scenarios, we can observe an increase of the determined operating window size. However, because the size of the operating window is approximated by the rectangle drawn by the farthest feasible operating points, we must evaluate the actual operating window, since the shape of the operating window may differ significantly from a rectangle. By sampling the operating points in the space of β_1 and β_2 and drawing all feasibility boundaries, we can graphically assess the actual operating window of the process (see Figure 2). As can be noticed by plotting the optimized operating window sizes for the base case and scenario 2 (black dashed line), the rectangle is not able to sufficiently approximate the actual operating window size. However, the rectangle size gradient allows to push the boundaries of the operating window. In case of the methanol process, the productivity and the split fraction boundary are moved, resulting in an increase of the actual size of the operating window. This is due to the fact that productivity and split

Table 1: Design optimization results

Results	Unit	Base case	Scenario 1	Scenario 2
Design variables				
Reactor length	m	8	7	7
Tube diameter	m	0.058	0.065	0.065
Operational variables				
Inlet temperature	K	498	498	498.33
Thermal fluid temperature	K	523	523	530.99
Flexibility dictating parameter				
$\dot{n}_{H_2}(\xi^{\max})$	kmol h ⁻¹	318.39	339.76	360.01
$\dot{n}_{H_2}(\xi^{\min})$	kmol h ⁻¹	96.22	94.76	95.95
$\dot{n}_{CO_2}(\xi^{\max})$	kmol h ⁻¹	106.19	136.29	132.01
$\dot{n}_{CO_2}(\xi^{\min})$	kmol h ⁻¹	30.44	30.48	29.68
$\xi_{\text{ow,MeOH}}$		16.83	25.92	27.02
Objective function improvement	%		54.03	60.56

fraction are highly dependent on temperatures, concentrations and pressures in the recycle stream. A shorter reactor length, larger tube diameter and higher coolant temperature have a beneficial impact on productivity of the reactor and equilibrium of the reactions.

In a last step, we tested dynamic feasibility of the system for all scenarios, assuming dynamic transition between the farthest operating points as worst-case test. For the methanol synthesis process none of the relevant boundaries are violated. The temperature in the reactor reacts sluggish, never reaching critical temperatures for the catalyst. It should be noted that dynamics of the process are highly dependent on reaction kinetics and modelling approaches. Depending on the assumptions made, such large steps between operating points might lead to critical conditions in the reactor. At this point, the design approach proposed here can incorporate advanced control strategies such as adapting the coolant temperature according to the operating point at which operation takes place to prevent operational failure.

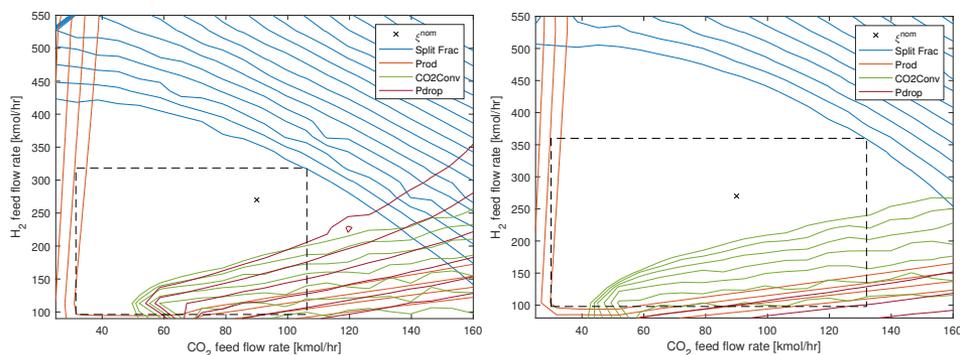


Figure 2: Operating window (white space) and size of the rectangle for the base case (left) and scenario 2 (right).

Finally, it is worth comparing the results achieved by the proposed approach with the outcome of typical mathematical tools from the literature such as the flexibility index, since the objective is rather similar. Comparing the two, we can observe a different design outcome with the proposed

approach. This is due to the fact that the flexibility index is measured by a hyperrectangle which is centered at the nominal operating conditions. By adapting design and operational decision variables, the flexibility index can be influenced, which means changing the size of the hyperrectangle. However, since the length of the sides of the hyperrectangle are bound by the closest process boundary, an increase in the size of the hyperrectangle would not directly lead to an increase in the size of the operating window. We can observe this when considering the examined methanol synthesis process. At the nominal operating condition ξ^{nom} with a H_2 feed flow rate of 270 kmol h^{-1} and a CO_2 feed flow rate of 90 kmol h^{-1} , changing the considered design and operational variables pushes the split fraction boundary, however, shifts the CO_2 conversion boundary towards the nominal operating condition. While the operating window increases, the flexibility index would not adequately represent this aspect, as it would only indicate the maximum deviation to the closest process boundary, which then could be the CO_2 conversion.

6. Conclusions

In this work, we present a design approach to optimize the multidimensional operating window of processes and its equipment in the space of intended fluctuating parameters. We use a dynamic optimization framework to incorporate the operating window size into the design decision and test the approach on the process of green methanol synthesis. The aim is to design a methanol reactor that is highly flexible towards the feed flow rate of H_2 and CO_2 . We track the movement of the operating window size and its boundaries by the size of a multidimensional rectangle spanned from the farthest operating points. Results show that the rectangle is not able to approximate the actual operating window size sufficiently. However, the rectangle size gradient considered in combination with graphical assessment leads to an increase of the actual operating window size of the methanol synthesis process. While the proposed approach does not give a direct sizing of the operating window, it enables to determine parameters that push the process boundaries in order to increase the size of the operating window with relatively low computational effort. Future work should incorporate the actual size of the operating window in order to prevent false movement in the shape of the operating window.

References

- B. Bruns, F. Herrmann, M. Grünwald, J. Riese, 2021. Dynamic design optimization for flexible process equipment. *Industrial & Engineering Chemistry Research* 60 (20), 7678–7688.
- K. Bussche, G. Froment, 1996. A steady-state kinetic model for methanol synthesis and the water gas shift reaction on a commercial $\text{Cu/ZnO/Al}_2\text{O}_3$ catalyst. *Journal of Catalysis* 161 (1), 1–10.
- Y. Cao, C. L. Swartz, M. Baldea, S. Blouin, 2015. Optimization-based assessment of design limitations to air separation plant agility in demand response scenarios. *Journal of Process Control* 33, 37–48.
- A. Caspari, C. Offermanns, P. Schäfer, A. Mhamdi, A. Mitsos, 2019. A flexible air separation process: 1. design and steady-state optimizations. *AIChE Journal* 65 (11), e16705.
- V. D. Dimitriadis, E. N. Pistikopoulos, 1995. Flexibility analysis of dynamic systems. *Industrial & Engineering Chemistry Research* 34 (12), 4451–4462.
- K. L. Fischer, H. Freund, 2020. On the optimal design of load flexible fixed bed reactors: Integration of dynamics into the design problem. *Chemical Engineering Journal* 393, 124722.
- S. Mbatha, R. C. Everson, N. M. Musyoka, H. W. Langmi, A. Lanzini, W. Brilman, 2021. Power-to-methanol process: a review of electrolysis, methanol catalysts, kinetics, reactor designs and modelling, process integration, optimisation, and techno-economics. *Sustainable Energy Fuels* 5, 3490–3569.
- J. I. Otashu, M. Baldea, 2019. Demand response-oriented dynamic modeling and operational optimization of membrane-based chlor-alkali plants. *Computers & Chemical Engineering* 121, 396–408.
- R. E. Swaney, I. E. Grossmann, 1985. An index for operational flexibility in chemical process design. part i: Formulation and theory. *AIChE Journal* 31 (4), 621–630.
- C. L. Swartz, Y. Kawajiri, 2019. Design for dynamic operation - a review and new perspectives for an increasingly dynamic plant operating environment. *Computers & Chemical Engineering* 128, 329–339.

Design and Optimization of Membrane System for Gas Separation

Shivom Sharma* and François Maréchal

École Polytechnique Fédérale de Lausanne, IPESÉ, Sion, 1950, Switzerland
*shivom.sharma@epfl.ch

Abstract

Membrane separation may substitute conventional energy intensive technologies, and it could have cost benefits and lower environmental footprints. Several membrane processes have been developed to achieve higher purity and recovery of products. This study uses a generic membrane superstructure (or system) that facilitates all possible inter- and intra-connections among different membrane stages (or units) which are arranged in series-parallel configurations. The developed mathematical model for the membrane system is a mixed integer non-linear programming (MINLP) problem. The mathematical model is implemented in AMPL, and BARON solver is used to solve it. The MINLP model of the membrane system can choose membrane from a Membrane Database, which has a number of polymeric and inorganic (graphene, carbon molecular sieve, zeolite and metal-organic frameworks) membranes. In this work, two industrial case studies of gas separation are considered: post-combustion CO₂ capture and biogas upgradation by CO₂ removal. The selection of CO₂ removal technology depends on plant location, production capacity and product quality specifications. The chosen applications have challenges in terms of energy consumptions, economics and environmental burden. The separation performance of the membrane system is evaluated and compared for same membrane in all membrane stages. Two optimization problems were solved for each membrane: minimization of total area of membranes and minimization of total mechanical power. For both applications, best performing membranes were identified to target the minimum separation cost.

Keywords: Membrane System, Post-combustion CO₂ Capture, Biogas Upgradation.

1. Introduction

Membrane separation is one of the emerging technologies that has the potential to replace traditional energy intensive separation technologies. In process industry, solvent absorption (amine absorption), solid adsorption (pressure swing adsorption) and cryogenic distillations are used to separate gas mixtures (Tock, 2013; Leung et al., 2014). Recently, gas separation using membranes has received considerable attention for industrial applications, namely air separation, syngas ratio adjustment, hydrogen recovery in refinery, post-combustion CO₂ capture and biogas upgradation by CO₂ removal (Ismail, 2015). Membrane separation has several advantages over conventional gas separation technologies, e.g., no use of chemicals, mild operating conditions, simple installation and easier operation, and flexibility to integrate with other separation technologies.

Several studies have explored post-combustion CO₂ capture from coal and natural gas power plants, using membranes (Kárászov et al., 2020). Zhang et al. (2014) studied post

combustion carbon capture, and amine-based capture system had higher energy consumptions and environmental impact compared to the membrane process. Arias et al. (2016) optimized performance of multi-stage membrane superstructure for capturing CO₂ from flue gases. Lee et al. (2018) optimized membrane superstructure for CO₂ capture from coal power plants, and showed the benefits of using different membranes in different membrane stages. Scholz et al. (2013) performed detailed analysis of biogas upgradation into biomethane using several types of membranes. Finally, Sun et al. (2015) reviewed several biogas upgrading technologies, including cryogenic separation, physical and chemical absorptions, pressure swing adsorption, membrane separation, hydrate formation and biological methods.

Post-combustion CO₂ capture and biogas upgradation have several challenges such as energy consumptions, capital and operating costs. The selection of CO₂ removal technologies depends on plant location, production capacity, product quality specifications, availability of financial resources, environmental regulations and energy integration with CO₂ emitting plant or industrial site. In order to achieve the required purity and capture rate, membranes are arranged in complex series-parallel configurations. This arrangement gives numerous degrees of freedom for membrane system design. In this study, a generic superstructure of membrane modules/units, with all possible inter- and intra-connections, is used (see Figure 1). A mixed integer non-linear programming problem of membrane superstructure has been developed in AMPL (A Mathematical Programming Language). A database of several membranes has been used, and optimization method can choose any membrane from the database.

Two important industrial case studies of gas separation are considered: post-combustion CO₂ capture and biogas upgradation by removing CO₂. There are eight (M1-M8) and seven (m1-m7) membranes respectively in the membrane databases for post-combustion CO₂ capture and biogas upgradation. The separation performance of the membrane system was evaluated and compared for all membranes. For each membrane, two optimization problems were solved: minimization of total area of membrane (TAM), and minimization of mechanical power (TP). The optimization results allow to identify best performing membranes for both applications. In the mathematical model of the membrane system, different membranes can also be used in different membrane stages to improve the separation performance. For both applications, the final solution always contains same membranes in both membrane stages, as Membrane Database has limited number of membranes. In case of large number of membranes in the Membrane Database, the proposed approach can identify best performing membrane clusters, based on the membrane permeability and selectivity. These findings could be useful to the membrane researchers for further improving the performances of their membranes.

2. Membrane Superstructure Model

Figure 1 presents generic membrane superstructure or system. Fresh feed is compressed [$C(F)$] and cooled-down [$HE(F)$] before it enters the membrane system. The fresh feed can go to any membrane stage or unit in the membrane system. Figure 1 shows i^{th} membrane stage of the membrane system, inside the dotted line. Each membrane stage has a membrane module [$MEM(i)$], a mixer [$M(f,i)$], two splitters [$S(r,i)$, $S(p,i)$] on retentate and permeate sides, a compressor [$C(i)$] and cooler [$HE(i)$] for permeate stream. The membrane stage mixer [$M(f,i)$] is used to mix the fresh feed and retentate and/or permeate recycled from the same or different membrane stages.

Finally, there are two mixers $[M(r), M(p)]$ for both product streams. A turbine $[T(P)]$ and a heater $[HE(P)]$ are used to recover the mechanical power, from the retentate side product.

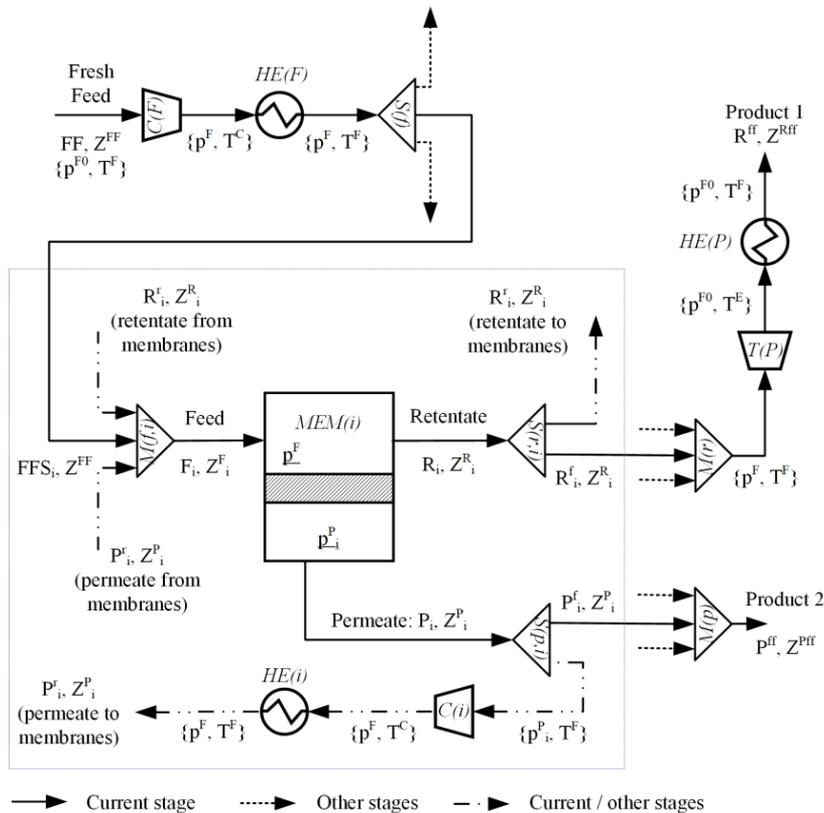


Figure 1: Membrane superstructure for gas separation

The mathematical model for the membrane superstructure is a mixed integer non-linear programming (MINLP) problem. Table 1 summarizes balances and equations for different units in membrane system.

Table 1: Summary of MINLP model for membrane system

Units	Balances or Equations
Membrane system	Mass and component balances
Splitters $[S(F), S(r,i), S(p,i)]$	Mass balance
Mixers $[M(f,i), M(p), M(r)]$	Mass and component balances
Membranes $[MEM(i)]$	Mass and component balances, membrane transport
Compressors, turbines $[C(F), C(i), T(P)]$	Equations for calculating outlet temperatures and powers ($\eta_C = 0.8, \eta_T = 0.85$)
Heaters, coolers $[HE(F), HE(i), HE(P)]$	Heat balance equation for calculating heat duties
Constraints	Limits on the product purities

3. Post-combustion CO₂ Capture

The flue gases contain mainly N₂, O₂, CO₂ and H₂O. This work considers separation of water prior to the use of membrane separation for CO₂ capture. It is assumed that the feed contains 14% CO₂ and remaining 86% N₂. The CO₂ and N₂ mixture have a flow rate of 10 mol/s at 1 bar pressure and 40 °C temperature. The membrane database has eight membranes (M1-M8), and Figure 2(c) shows CO₂ permeance and CO₂/N₂ selectivity for different membranes. M1-M3 are polymeric membranes whereas M4-M8 are inorganic (graphene, carbon molecular sieve, zeolite, metal-organic frameworks) membranes. The feed side pressure for membrane units can vary between 5 and 13 bar (Minh et al., 2008). The CO₂ and N₂ streams from the membrane system have 95% purities.

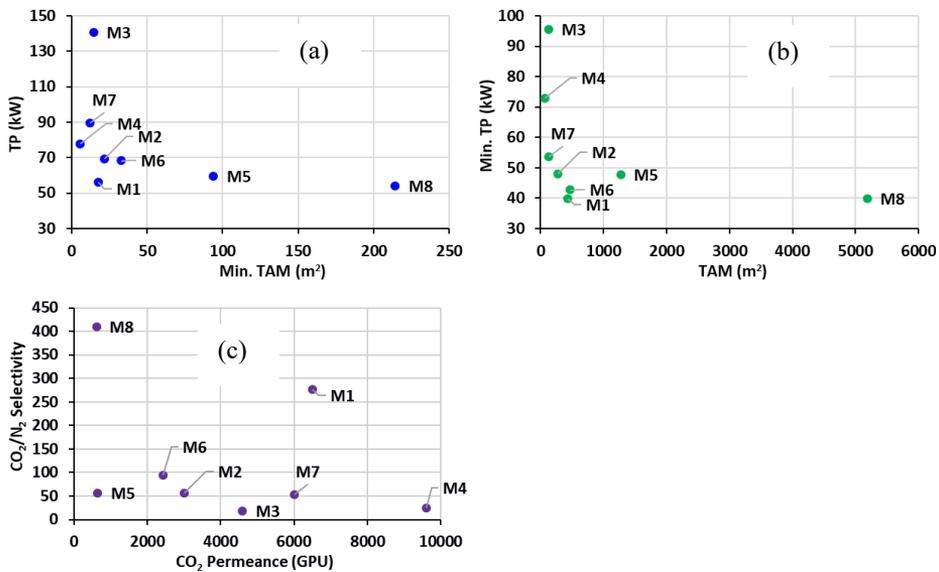


Figure 2: Post-combustion CO₂ capture using different membranes: (a) minimization of total area of membranes, (b) minimization of total mechanical power, (c) CO₂ permeance versus CO₂/N₂ selectivity for all membranes.

For this study, the membrane system has two membrane stages. For each membrane in the membrane database, two optimization problems were solved: minimization of total area of membranes, and minimization of total mechanical power (feed compression + stage compression – product expansion). The optimization problems have 252 variables and 266 constraints, and they were solved using BARON (v21.1.13) solver in AMPL (v20210220), with a maximum solution time of 30 minutes. Figure 2(a) presents optimization results for minimum total area of membranes. This figure also presents related values of total mechanical power obtained for different solutions. Similarly, Figure 2(b) presents results for minimum total mechanical power (2nd optimization problem) for all membranes. To minimize the total cost of separation, both total area of membranes and total mechanical power are equally important. It can be seen from Figures 2(a) and 2(b) that membrane M1 has the best performance for separating a mixture of CO₂ and N₂. Further, membrane M1 has best compromise between permeance and selectivity. Membrane M4 has very high permeance but low selectivity, whereas membrane M8 has very high selectivity but low permeance.

4. Biogas Upgradation by CO₂ Removal

The biogas contains 38% CO₂ and remaining 62% CH₄. The biogas has a flow rate of 10 mol/s at 1 bar pressure and 40 °C temperature. The membrane database for biogas upgradation has seven membranes (m1-m7), and Figure 3(c) shows CO₂ permeance and CO₂/CH₄ selectivity for different membranes. m1-m3 are polymeric membranes whereas m4-m7 are inorganic (graphene, carbon molecular sieve, zeolite, metal-organic frameworks) membranes. Two membrane stages were considered for separating CO₂ and CH₄ mixture. The feed side pressure for membrane units can vary between 5 and 13 bar. The CO₂ and CH₄ streams from the membrane system have 95% purities.

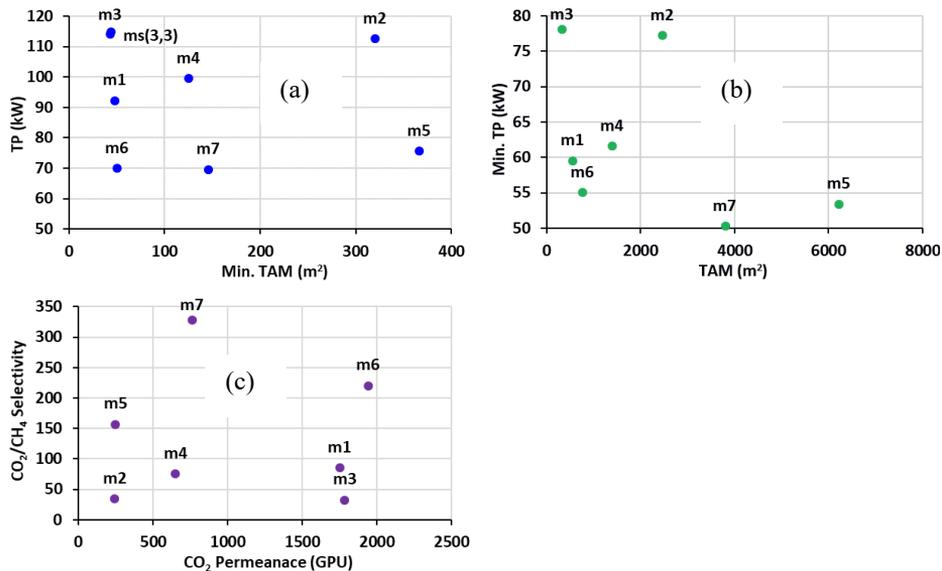


Figure 3: Biogas upgradation by CO₂ removal using different membranes: (a) minimization of total area of membranes, (b) minimization of total mechanical power, (c) CO₂ permeance versus CO₂/CH₄ selectivity for all membranes.

Similar to previous case study, two optimization problems were solved for all the membranes: minimization of total area of membranes, and minimization of total mechanical power. Both optimization problems have 252 variables and 266 constraints, and they were solved using BARON solver in AMPL, with a maximum solution time of 30 minutes. Figure 3(a) presents optimization results for minimum total area of membranes. This figure also present values of total mechanical power obtained for different solutions. Figure 3(b) presents results for minimum total mechanical power for all membranes, along with related total area of membranes. For minimum total cost of separation, solutions obtained for membrane m6 are better than other membranes, as these solutions are nearer to the corner. Membrane m6 has best compromise between permeance and selectivity, as shown in Figure 3(c).

The mathematical model of membrane system can use different membranes in different stages to improve the separation performance. For biogas upgradation, we minimized total area of membranes, and this optimization problem has 250 variables and 268 constraints. The optimal solution uses membrane 3 in both stages (see solution ms(3,3) in Figure 3(a)). Figure 4 provides detail of solution ms(3,3). If we have many membranes

in the Membrane Database, it is possible to identify best performing membrane clusters, for separating a gas mixture.

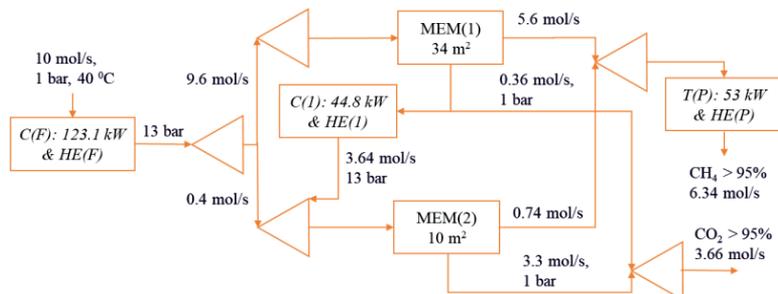


Figure 4: Details of solution ms(3,3); membrane 3 was used in both membrane stages

5. Conclusions

This study develops a mathematical model for multi-stage membrane superstructure. The model is a mixed integer non-linear programming problem that has been implemented in AMPL, and solved using BARON solver. Two case studies, namely post-combustion carbon capture and biogas upgradation were solved using the developed mathematical optimization problem. The optimization results present optimal flows and pressure levels inside the membrane system, and also allow to identify best performing membranes. The proposed approach can identify best performing membrane clusters, based on the membrane permeability and selectivity. This knowledge could be useful to the membrane developers for further improving the performances (selectivity versus permeance) of membranes. The future studies will focus on separating a gas mixture with three components, using different number of membrane stages.

References

- A. F. Ismail, Khulbe K.C., Matsuura T., 2015, Application of gas separation membranes. In: Gas Separation Membranes: Polymeric and Inorganic. Springer
- A.M. Arias, Mussati M.C., Mores P.L., Scenna N.J., Caballero J.A. and Mussati S.F., 2016, Optimization of multi-stage membrane systems for CO₂ capture from flue gas, International Journal of Greenhouse Gas Control, 53, pp. 371-390.
- D.Y.C. Leung, Caramanna G. and Marto-Valer M.M., 2014, An overview of current status of carbon dioxide capture and storage technologies, Renewable and Sustainable Energy Reviews, 39, pp. 426-443.
- L. Tock, 2013, Thermo-environomic optimisation of fuel decarbonisation alternative processes for hydrogen and power production, PhD Thesis, EPFL Switzerland.
- M. Kárászov, Zach B., Petrusov Z., Červenková V., Bobák M., Šyc M., Izák P., 2020, Post-combustion carbon capture by membrane separation, Review, Separation and Purification Technology, 238.
- M. Scholz, Melin T. and Wessling M., 2013, Transforming biogas into biomethane using membrane technology, Renewable and Sustainable Energy Reviews, 17, pp. 199-212.
- Q. Sun, 2015, Selection of appropriate biogas upgrading technology-a review of biogas cleaning, upgrading and utilization, Renewable and Sustainable Energy Reviews, 51, 521-532.
- S. Lee, Binns M. and Kim J.K., 2018, Automated process design and optimization of membrane-based CO₂ capture for a coal-based power plant, 563, pp. 820-834.
- T. H. Minh, Allinson G.W., and Wiley D.E., 2008, Reducing the cost of CO₂ capture from flue gases using membrane technology, Ind. Eng. Chem. Res. 47(5), pp. 1562-1568.
- X. Zhang, Singh B., He X., Gundersen T., Deng L. and Zhang S., 2014, Post-combustion carbon capture technologies: Energetic analysis and life cycle assessment, International Journal of Greenhouse Gas Control, 27, pp. 289-298.

Assessment of Carbon Capture Technologies for Waste-to-Energy System

Shivom Sharma^{a,*}, Rafael Castro-Amoedo^a, Jaroslav Hemrle^b, and François Maréchal^a

^a*École Polytechnique Fédérale de Lausanne, Industrial Process Energy Systems Engineering (IPESE), Sion, 1950, Switzerland*

^b*Hitachi Zosen Inova AG, Zürich, 8005, Switzerland*

**shivom.sharma@epfl.ch*

Abstract

Municipal solid waste is a mixture of urban and industrial waste, consisting of biodegradable fractions, such as food waste, waste wood or paper, but also fossil-based fractions, among which plastics, textiles, metals, glass and aluminum. Depending on the type of waste and recovery technique, biogas, biofuels, heat, electricity and metals, are possible value-added products. As both biogenic and fossil carbon are present among waste fractions, the reduction and capture of carbon is crucial in the deployment of sound waste management technologies. There are several physicochemical CO₂ capture technologies, and they have their own benefits, challenges and limitations. Some techniques are in the development phase, and they need to be evaluated for their possible integration within waste-to-energy system. We have developed a waste-to-energy superstructure, including digestion, gasification and incineration as the main waste treatment technologies. The latter is the main contributor of CO₂ emissions. The developed superstructure includes three options for CO₂ capture from flue-gases: amine absorption, temperature swing adsorption and membranes. Amine absorption and membranes are considered for biogas upgradation, whereas pressure swing adsorption and membranes are evaluated for syngas upgradation. This study systematically generates and compares a number of decarbonization options for waste-to-energy system. The formulated optimization problem is a mixed integer linear programming problem, and total annual cost is considered as the performance criterion for generating decarbonizing options. For carbon capture from flue-gases, amine absorption and temperature swing adsorption found to be better options compared to membrane separation

Keywords: Waste-to-Energy; Anaerobic digestion; Gasification; Incineration; CO₂ Capture.

1. Introduction

Waste generation has increased significantly in the past few years. Close to 2 billion tons of municipal solid waste (MSW) are annually produced (worldbank.org). Despite a hierarchy framework to process and treat waste (reduce, reuse, recycle, dispose), about one third is still mismanaged. In that regard, several technologies have been recently developed and are available with different costs and environmental impacts (Brunner and Rechberger, 2015; Lombardi et al., 2015). For instance, frequently overlooked steps, such as waste collection and sorting are now subject of research and optimization, with advanced collection and separation systems established in numerous countries.

Figure 1 shows a simplified flowchart for different waste management pathways, focusing on three major waste categories: MSW, biomass and plastics. MSW has some unsorted biomass and plastics. The biomass contains 12% whey, 55.1% green waste, 18.4% food waste and 14.5% others. Plastics have 74% PET and 26% others (LDPE and HDPE). Anaerobic digestion is used to convert biomass into biogas; the organic material is broken down by bacteria into a methane rich gas in the absence of oxygen. Syngas is produced by gasification of plastics in the presence of air or steam at high temperature, while incineration is a thermal recovery process. If required, both the biomass and plastic fractions of MSW can be thermally valorized. Note that plastics are a major contributor in lower heating value of MSW.

In order to reduce the environmental impact of the waste management, CO₂ can be captured, sequestered and mineralized. Process integration and optimization techniques are useful in identifying sustainable and cost-effective strategies for waste management (Rizwan et al., 2020). Münster and Meibom (2011) have explored the integration of waste into energy system. Puchongkawarin and Mattaraj (2020) developed a decision-making tool for the design of the optimal MSW facilities. Recently, Castro-Amoedo et al. (2021) have studied biowaste valorization along with CO₂ removal from biogas. There are several physicochemical CO₂ capture technologies, such as physical adsorption (pressure swing adsorption, temperature swing adsorption), chemical absorption (amine absorption, ammonia scrubbing, selexol process), membrane separation and cryogenic technology (Tock, 2013). Each technology has its own benefits, challenges and limitations. Some of these technologies (e.g., temperature swing adsorption, membranes) are in the development phase, and they need to be evaluated for their possible integration with waste-to-energy system (Sharma and Maréchal, 2019). Each technology has some key requirements, namely materials, heat and electricity. Further, CO₂ source (biogas, syngas, flue-gases) has specific characteristics, such as composition, temperature, pressure and impurities. Hence, there is a strong need to evaluate and compare the integration of different CO₂ capture technologies within waste-to-energy system.

In this study, we have developed a waste-to-energy superstructure, where digestion (AD), gasification (GAS) and incineration (INC) are the main waste treatment options (i.e. conversion units). The superstructure includes several options for CO₂ capture: amine absorption, temperature swing adsorption and membranes for flue-gases; amine absorption and membranes for biogas upgradation; pressure swing adsorption and membranes for syngas. A mixed integer linear programming (MILP) formulation was implemented in AMPL/GLPK. Eight decarbonizing options/scenarios were systematically generated and compared, using total annual cost as objective function.

2. Waste-to-Energy Superstructure

2.1 Description

Municipal waste generated can be divided into three broad categories: biomass, plastics and MSW. The biomass can be converted into biogas and compost, using an anaerobic digester. The plastics can be recycled as monomers or can be used to produce syngas and char inside a gasifier. Finally, MSW can be combusted inside an incinerator to produce heat (or high-pressure steam). The biogas can be used in a furnace and/or upgraded to natural gas using amine absorption (MEA) and membranes (MEM). Syngas can be upgraded to natural gas using pressure swing adsorption (PSA) and membranes, whereas char can be burned in a furnace for heat recovery. The high-pressure steam from the incinerator is used to generate electricity *via* steam cycle. Amine absorption, temperature

swing adsorption (TSA) and membranes are used for CO₂ capture from flue-gases. With proper infrastructure, heat from waste-to-energy system can be supplied to the district heating networks. The captured CO₂ can be converted into methane using surplus electricity or cheap electricity from the grid. Anaerobic digester is present in Cluster 1 (C1), whereas gasifier and incinerator are considered in Cluster 2 (C2). There is no exchange of mass between clusters, but resources (such as electricity, natural gas and water) can be inter-changed between clusters.

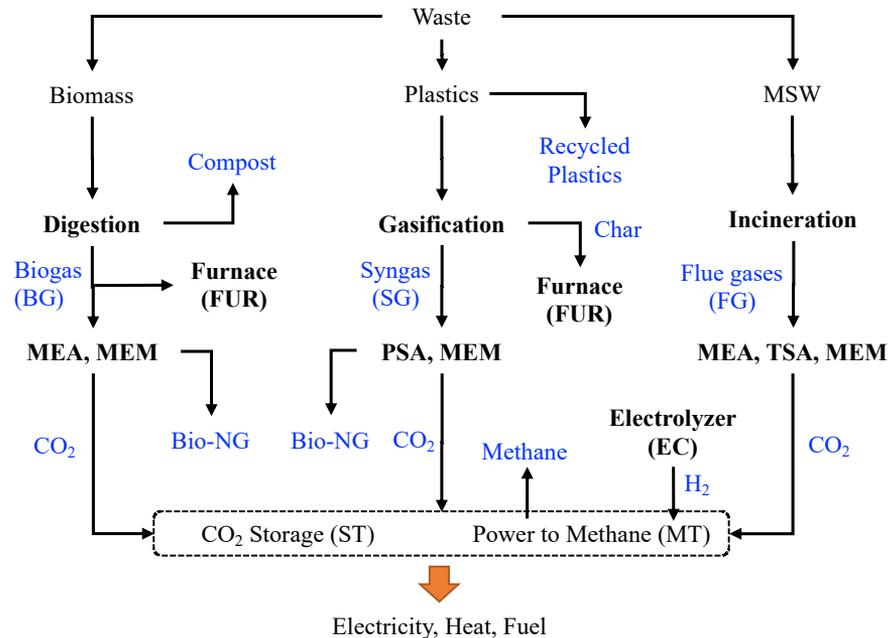


Figure 1: Waste-to-energy superstructure, with carbon capture, utilization and storage.

2.2 Mathematical Formulation

For each technology in the superstructure, linear models were developed that include mass flows, temperature-enthalpy profiles, as well as the use or generation of electricity. Each technology/model also include fixed and variable capital (c_i^{inv1}, c_i^{inv2}) and operating ($c_{i,t}^{op1}, c_{i,t}^{op2}$) costs. Both binary ($\Psi_{i,t}, \Psi_i$) and continuous ($f_{i,t}, f_i$) variables are used. The former accounts for the existence of a unit and the latter for its size. Based on a given objective function, the optimizer defines connections among technologies, levels of use of technologies (material flows) and heat flows. Interested readers are referred to Castro-Amoedo et al. (2021) for more details on the problem formulation.

The optimization problem can be solved for capital cost, operating cost or total annual cost (TAC) (see Eq. 1), in which the first element on the right-hand side concerns operating costs, while the second element pertains to annualized capital expenditures, with i interest rate (0.08) and n assumed lifetime (20 years). Constraints were added to ensure: (i) minimum (f_i^{min}) and maximum (f_i^{max}) units capacities (Eq. 2); (ii) resource consumption (Eq. 3 and 4), with $\dot{m}_{r,i,t}^+$ the reference quantity of resource $\forall r \in R$ needed in each unit; (iii) mass balances (Eq. 5) for each unit; (iv) heat cascade formulation (Eq. 6 and 7) based on Maréchal and Kalitventzeff (1998), where residual heat ($\dot{R}_{t,k}$) is

transferred, according to the 2nd law of thermodynamics, from higher ($k + 1$) to lower temperature level (k). The mathematical formulation is written in AMPL (v20210220) and solved by CPLEX (v12.7.0.0), on a Microsoft Windows v10.0.18363 machine equipped with a 2.4 GHz Intel(R) Xeon (R) 8 core processor and 16 GB RAM.

$$\min TAC (\$/y) = \sum_{t \in T} \sum_{i \in L} (c_{i,t}^{op1} \cdot \psi_{i,t} + c_{i,t}^{op2} \cdot f_{i,t}) \cdot t_t^{op} + \frac{i(1+i)^n}{(1+i)^n - 1} \cdot \sum_{i \in Lu} (c_i^{inv1} \cdot \psi_i + c_i^{inv2} \cdot f_i) \quad \text{Eq. 1}$$

$$f_i^{min} \cdot \psi_{i,t} \leq f_{i,t} \leq f_i^{max} \cdot \psi_{i,t}, \quad \forall i \in L, \quad \forall t \in T \quad \text{Eq. 2}$$

$$\dot{M}_{r,i,t}^+ = f_{i,t} \cdot \dot{m}_{r,i,t}^+, \quad \forall r \in R, \quad \forall i \in L, \quad \forall t \in T \quad \text{Eq. 3}$$

$$\dot{M}_{r,i,t}^- = f_{i,t} \cdot \dot{m}_{r,i,t}^-, \quad \forall r \in R, \quad \forall i \in L, \quad \forall t \in T \quad \text{Eq. 4}$$

$$\sum_{i \in L} \dot{M}_{r,i,t}^- = \sum_{i \in L} \dot{M}_{r,i,t}^+, \quad \forall r \in R, \quad \forall t \in T \quad \text{Eq. 5}$$

$$\sum_{i \in L} \dot{Q}_{i,t,k} \cdot f_{i,t} + \dot{R}_{t,k+1} - \dot{R}_{t,k} = 0, \quad \forall t \in T, \quad \forall k \in K, \text{ with } T_{k+1} \geq T_k, \quad \text{Eq. 6}$$

$$\dot{R}_{t,k} \geq 0, \quad \dot{R}_{t,k_{max}+1} = \dot{R}_{t,1} = 0, \quad \forall t \in T \quad \text{Eq. 7}$$

3. Waste-to-Energy: A Case Study

For this case study, the availabilities of waste are: biomass (BM) = 1.89 t/h, plastics (PL) = 0.25 t/h and MSW = 15 t/h. Maximum heat demands of districts heating networks (DHN) are 500 kW and 15,000 kW respectively for clusters C1 and C2. For DHN, water supply and return temperatures are 60 and 55 °C, respectively. The selling price of biogas/syngas is 15.5 \$/MWh, and selling price of bio-SNG is 30.9 \$/MWh. There is a carbon tax (99 \$/t-CO₂) on CO₂ release into environment. If captured CO₂ is converted into methane, there is carbon credit of 99 \$/t-CO₂. The storage of CO₂ costs 20.6 \$/t-CO₂, without any carbon tax or credit. We have generated several scenarios as shown in Table 1, to analyze different decarbonizing strategies. Scenarios S1, S3, S5 and S7 consider CO₂ storage, whereas scenarios S2, S4, S6 and S8 convert captured CO₂ into methane.

Table 1: Decarbonizing options/scenarios for waste-to-energy system

	Biogas CO ₂ capture	Syngas CO ₂ capture	Flue-gases CO ₂ capture			CO ₂ Storage	CO ₂ to Methane
			MEA	MEM	TSA		
S1	√	√				√	
S2	√	√					√
S3	√	√	√			√	
S4	√	√	√				√
S5	√	√		√		√	
S6	√	√		√			√
S7	√	√			√	√	
S8	√	√			√		√

Table 2 presents results for all decarbonizing scenarios. The capital cost (ACC) is calculated yearly, whereas operating cost (MOC) is evaluated on monthly basis. These results can virtually be divided into several parts (*via* horizontal lines): economics and

waste flows, heat and electricity generation, anaerobic digestion and biogas upgradation, plastics gasification and syngas upgradation, CO₂ capture from flue-gases using amine, membrane and TSA processes, and methanation (CO₂ to methane). The flue-gases from incinerator contain 9,461 kg/h CO₂. The amounts of captured CO₂ by MEA (8,431 kg/h) and TSA (8,516 kg/h) processes are comparable, whereas only 6,445 kg/h CO₂ is captured by MEM process.

Table 2: Optimization results for different scenarios (SN: steam network, PV: photovoltaics, Env: environment, ELE: electricity)

	S1	S2	S3	S4	S5	S6	S7	S8
MOC, \$/m	-794000	-225000	-984000	4977000	-614000	4085000	-932000	5110000
ACC, \$/y	114000	1130000	357000	10730000	539000	8708000	697000	11164000
MSW-INC, t/h	15	15	15	15	15	15	15	15
BM-AD, t/h	2	2	2	2	2	2	2	2
PL-GAS, t/h	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
C1 DHN, kW	143	386	143	386	143	386	143	386
C1 SN, kW	113	176	113	176	113	176	113	176
C1 PV, kW	42	42	42	42	42	42	42	42
C2 DHN, kW	15000	15000	15000	15000	15000	15000	15000	15000
C2 SN, kW	17284	17572	19898	24216	18534	21835	16966	21097
C2 PV, kW	145	145	145	145	145	145	145	145
C1 BG-MEA, Nm ³ /h	365	365	365	365	365	365	365	365
C1 MEA-Bio NG, kW	1920	1920	1920	1920	1920	1920	1920	1920
C1 MEA-CO ₂ -ST, kg/h	245		245		245		245	
C1 MEA-CO ₂ -MT, kg/h		245		245		245		245
C2 Char-FUR, kg/h	49	49	49	49	49	49	49	49
C2 FUR-CO ₂ , kg/h	77	77	77	77	77	77	77	77
C2 SG-PSA, kW	1505	1505	1505	1505	1505	1505	1505	1505
C2 PSA-Bio NG, kg/h	3677	3677	3677	3677	3677	3677	3677	3677
C2 PSA-CO ₂ -ST, kg/h	643		643		643		643	
C2 PSA-CO ₂ -MT, kg/h		643		643		643		643
C2 FG-CO ₂ -Env, kg/h	9461	9461						
C2 FG-MEA, kg/h			51318	51318				
C2 MEA-CO ₂ -Env, kg/h			1030	1030				
C2 MEA-CO ₂ -ST, kg/h			8431					
C2 MEA-CO ₂ -MT, kg/h				8431				
C2 FG-MEM, kg/h					51318	51318		
C2 MEM-CO ₂ -Env, kg/h					3020	3020		
C2 MEM-CO ₂ -ST, kg/h					6445			
C2 MEM-CO ₂ -MT, kg/h						6445		
C2 FG-TSA, kg/h							51318	51318
C2 TSA-CO ₂ -Env, kg/h							946	946
C2 TSA-CO ₂ -ST, kg/h							8516	
C2 TSA-CO ₂ -MT, kg/h								8516
C1,2 Total CO ₂ -MT, kg/h		888		9319		7333		9404
C1,2 MT-CH ₄ , kW		4482		47065		37034		47495
C1,2 EC-H ₂ , kg/h		161		1695		1333		1710
C1,2 EC-O ₂ , kg/h		645		6777		5333		6839
C1,2 ELE-EC, kW		6722		70581		55537		71225
C1,2 ELE-Market, kW	14860	8348	14828	-52891	12634	-40734	14363	-54208

For scenarios S4, S6 and S8, the electricity produced by the waste-to-energy system is not enough for the methanation, due to large quantity of captured CO₂. Hence, electricity is imported from the grid, and these scenarios are not profitable (negative MOC means profit). The flue-gases from incinerator are the main contributor of CO₂ emissions. We did not compare the decarbonizing options based on capital cost due to large variabilities in the cost estimation. MEA and TSA are heat driven processes, whereas MEM process requires electricity for compression. Further, MEA and TSA processes have comparable CO₂ product quality, which is better than MEM process. Scenarios S3 (MEA, AOC = -984,000 \$/m) and S7 (TSA, AOC = -932,000 \$/m) are more profitable than scenario S5 (MEM, AOC = -614,000 \$/m). More CO₂ has been captured in scenarios S3 and S7, compared to scenario S5. Further, scenario S5 (12,634 kW) exports less electricity to the grid compared to scenarios S3 (14,828 kW) and S7 (14,363 kW). Hence, MEA and TSA are better choices for decarbonizing the waste-to-energy sector.

4. Conclusions

This study presents a waste-to-energy superstructure, comprising digestion, gasification and incineration as the main treatment options. Amine absorption, membrane separation and pressure and temperature swing adsorptions were considered for carbon capture along with methanation for converting captured CO₂ into methane. A mixed integer linear programming problem was formulated for waste-to-energy system. In total, eight decarbonizing options/scenarios were compared, and appropriate cost, credit and tax were used for CO₂ storage, CO₂ conversion to CH₄, and CO₂ emissions to the environment, respectively. For carbon capture from flue-gases, amine absorption and temperature swing adsorption outperformed membrane process, based on the operating cost, amount of captured CO₂ and quality of the CO₂ product.

References

- C. Puchongkawarin, Mattaraj, S., 2020, Development of a superstructure optimization framework for the design of municipal solid waste facilities, *Sustain Environ Res*, 30, 27.
- F. Maréchal and Kalitventzeff, B., 1998, Energy Integration of Industrial Sites: Tools, Methodology and Application. *Appl. Therm. Eng.*, 18, pp 921-933.
- L. Lombardi, Carnevale E. and Corti A., 2015, A review of technologies and performances of thermal treatment systems for energy recovery from waste, *Waste Management*, 37, pp. 26-44.
- L. Tock L., 2013, Thermo-environmental optimisation of fuel decarbonisation alternative processes for hydrogen and power production, PhD Thesis, EPFL Switzerland.
- M. Münster and Meibom P., 2011, Optimization of use of waste in the future energy system, *Energy*, 36(3), pp. 1612-1622.
- M. Rizwan, Saif Y., Almansoori A. and Elkamel A., 2020, A multiobjective optimization framework for sustainable design of municipal solid waste processing pathways to energy and materials, *Int J Energy Research*, 44(2), pp. 771-783.
- P.H. Brunner and Rechberger H., 2015, Waste to energy - key element for sustainable waste management, *Waste Management*, 37, pp. 3-12.
- R. Castro-Amoedo, Morisod N., Granacher J. and Maréchal F., 2021, The role of biowaste: a multi-objective optimization platform for combined heat, power and fuel, *Front. Energy Res.* 9:718310.
- S. Sharma and Maréchal F., 2019, Carbon dioxide capture from internal combustion engine exhaust using temperature swing adsorption, *Front. Energy Res.* 7:143.

Process Integration and Techno-Economic Assessment of a Green Biorefinery Demonstration Scale Platform for Leaf Protein Production

Thalles A. Andrade, Morten Ambye-Jensen

Aarhus University, Department of Biological and Chemical Engineering, 8200 Aarhus N, Denmark

Abstract

Green biorefinery corresponds to sustainable processing of green biomass, such as grasses and legumes with a high content of leaf protein and soluble sugars to produce multiple products. The extraction of proteins from leaves represents an attractive alternative solution to the high European dependency on soy imports. A demonstration facility for R&D in Aarhus University Foulum, Denmark, optimizes the separation of protein-rich concentrates from green biomass. Besides the leaf protein concentrate, the facility also produces a fibrous pulp that can be used for ruminant feed, biomaterials, or bioenergy, and a nutrient-rich residual juice, that can be used for biogas production and fertilizer.

The green biorefinery was simulated using Aspen Plus V12 for the processing of grass-clover. The process included the maceration of the harvested biomass; mechanical fractionation into a fiber-rich press cake and a green juice; heat treatment of the green juice for precipitation of soluble leaf proteins; centrifugation and drying of the precipitated protein; and anaerobic digestion of the residual brown juice for biogas production. The simulation was validated based on data obtained from the demonstration plant, resulting in an overall yield of 40 % crude protein in the protein concentrate. The feasibility of the implementation of the biorefinery on a commercial scale was calculated using Aspen Process Economic Analyzer.

Keywords: green biorefinery; leaf protein concentrate; process modeling; techno-economic assessment.

1. Introduction

The increasing world population and increased prosperity lead to increased demands for nutritious food. Simultaneously, the climate changes caused by greenhouse gas emissions and the ecological problems of nutrient leaching, pesticide use, and soil carbon depletion caused by intensive agricultural practices call for the development of alternative and more sustainable ways to produce food, feed, biobased materials, fuels, and energy. This can be done through integrated biorefineries. Perennial green biomass also referred to as forage crops, stands out as a promising feedstock with low environmental impact and is suitable for producing both protein-rich food and feed, biochemicals, and bioenergy (Manevski et al., 2018, Njakou Djomo et al., 2020).

Green biorefining stimulates and increases the synergy of the local agricultural sector (Corona et al., 2018a). Else, the production of leaf protein from forages decreases the dependency on protein-rich feed imports, especially soy, used for monogastric animals. According to FAO, soybean meal imports in Europe increased from 19.1 million tonnes in 1990 to around 26.2 million tonnes in 2019. The majority of the imported protein feeds

are used in pig and poultry meat production. Denmark is the fourth biggest pig producer in the EU, after Spain, Germany, and France, and imports around one million tonnes of protein for feed per year (EUROSTAT, 2021). This makes the biorefining of green forages an attractive local protein production in Denmark.

Besides being rich in protein, forage crops (that includes grass, clover, and lucerne) are a source of fibers, soluble sugars, amino acids, lactic acid, biogas, among other products. This multiple-product green biorefinery system allows developing greener alternatives to conventional products, employing different routes, and integrating different technologies chosen according to the target products. In this scenario, the green biorefinery is presented in a simple form producing a feed protein concentrate, a fibrous pulp, and biogas from the residual liquid. The process was simulated in Aspen Plus V12, based on the results obtained in the green biorefinery demonstration platform located at the Aarhus University Foulum research center, Denmark (Corona et al., 2018; Feng et al., 2021). The anaerobic digester was simulated based on studies previously developed by Rajendran et al. (2014) and Llano et al. (2021). Techno-economic assessment of the biorefinery was carried out at Aspen Process Economic Analyzer V12.

2. Materials and Methods

2.1. Process description

Grass-clover was chosen as the feedstock used in the green biorefinery. The crop is named ForageMax 45 and is a mixture of tall fescue, ryegrass, and red- and white clovers. The grass-clover mixture was harvested and processed at AU Foulum research center. The biomass is harvested whole and directly collected using a GrassTech GT140 harvester and processing generally starts within one hour after the forage is harvested. At the demonstration plant, the grass-clover is first shredded in a stationary cutter to a theoretical length of 4-5 cm. Then, crushed leaves are sent to a CirTech P25 Twin screw press where the solid and liquid streams are mechanically separated. The pressed fiber pulp was set to be sold as ruminants feeding. The pressed green juice is filtered in a 50 μm -filter bowl screen and pumped through a two-step heat exchanger, where the first step recovers heat from the residual liquid (brown juice) and heats the green juice to around 65°C and the second adds additional heat to reach 85 °C. The protein precipitates by heat and solids are separated in a decanter centrifuge producing a protein paste and the residual brown juice. The protein paste, that contains approximately 47 % of dry matter content, is sent to a vacuum drier to reach a dry matter content greater than 95 %. Dried LPC was set to be commercialized as monogastric feeding. Meanwhile, the brown juice was established to be fermented by anaerobic digestion in a biogas plant, due to its high content of easily digested organic compounds. The biogas and liquid digestate can increase the process's profitability. Table 1 presents the composition of the fractions that were used in the simulation.

2.2. Process modeling and simulation

The integrated process was simulated in Aspen Plus V12. The thermodynamic model NRTL was chosen because it correlates and calculates the mole fractions and activity coefficients of different compounds and facilitates the liquid and the gas phase in the biogas production (Rajendran et al., 2014). The feed flow rate was 40 tonnes per hour. Due to its complexity and variable composition, the biomass was defined as a nonconventional solid. In this case, the density and enthalpy of the grass-clover were estimated based on the proximate, ultimate, and sulfate analysis. Gates–Gaudin–Schuhman model was used for the particle size distribution, assuming a max grass length of 30 cm. The maximum particle length was reduced to 5 cm in the cutter. The screw

press was modeled using two stages: an RYield reactor to convert the shredded nonconventional biomass into the biomass composition shown in Table 1, and a separator to split the fiber and green juice streams.

Table 1. Chemical composition (%DM) of the fractions in the grass-clover biorefinery.

Component	Biomass	Fiber	GJ	BJ	LPC
Protein	21.6	17.8	29.2	6.6	47.0
Carbohydrates	15.0	6.7	31.7	50.6	16.9
Hemicellulose	16.7	21.2	7.6	8.9	6.7
Cellulose	19.6	28.7	1.3	0.1	2.2
Lignin	9.0	13.3	0.3	0.3	0.4
Ash	13.1	8.2	23.0	31.3	16.4
Others	5.0	4.1	6.8	2.2	10.4
DM content (%)	18.0	32.8	9.4	4.7	>95.0

The kinetic model in the anaerobic digester was evaluated at thermophilic conditions at 55 °C since the operation in the range of 50 to 65 °C is the best regime for maximizing biogas production (Llano et al., 2021). The anaerobic digester was divided into two steps in the simulation. First, a stoichiometric reactor was used to simulate the hydrolysis reactions based on the extent of the reaction. Hydrolysis of cellulose, hemicellulose, sugars, and lipids are modeled in this reactor. Then, a CSTR was used where the amino acids degradation and the acidogenic, acetogenic, and methanogenic reactions were modeled on a kinetic basis. A residence time of 15 days was defined. The chemical reactions, as well as the fractional conversions and kinetic constants, were found in the literature (Llano et al., 2021; Rajendran et al., 2014). The stream from the digester was cooled down in a flash column to remove water, improving the biogas composition.

2.3. Techno-economic assessment

The techno-economic evaluation was performed in Aspen Process Economic Analyzer V12 based on the process simulated in Aspen Plus. Capital expenses (CAPEX) are determined as a factor of the total purchase of equipment. It includes direct (equipment, installation, instrumentation) and indirect (engineering, construction, contingency) costs, and working capital. Equipment costs were either estimated from the designed process equipment or obtained from the market quotation. Fresh grass-clover is seasonal and only available to harvest between May and October in Denmark. Therefore, the biorefinery should operate during these months, in a total of 4200 hours per year of operation. All operating expenses (OPEX) are based on Danish prices. This includes the raw materials, utilities, and labor costs, as well as the taxes rate and product sales. Electricity, cooling water, and steam are the process utilities required. The LPC selling price was based on the price of soybean meal. Three different prices were compared: conventional (330 USD/tonne), non-GMO (535 USD/tonne), and organic (840 USD/tonne).

3. Results and Discussion

The integrated green biorefinery flowsheet simulated in Aspen Plus is shown in Figure 1. Table 2 presents an overview of the main streams and compares the simulated and expected production. The simulated production per tonne of biomass is validated by the mass streams from Corona et al. (2018b). The volume of biogas estimated in the simulation is higher than the volume estimated by the expected results. This is because the latter estimates the volume of the upgraded biogas, that is the methane potential

produced. Conversely, the simulated results only consider the biogas production before the upgrade, i.e., the biogas stream contains other components, essentially carbon dioxide.

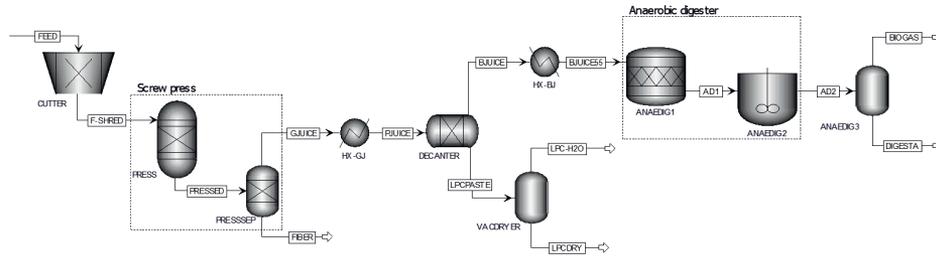


Figure 1: Integrated green biorefinery process diagram to produce feed protein and biogas.

The composition profile of the brown juice, biogas, and liquid digestate fractions, in mole fraction and excluding the water content is shown in Table 3. The macromolecules content was consumed during the anaerobic digestion, leading to the production of components such as ammonia, CO₂, and methane. Methane and CO₂ contents obtained in the biogas phase agree with the results expected in the literature (Penteado et al., 2019).

Table 2. Overview of the main process streams in the integrated green biorefinery.

	Biomass	Fiber	GJ	BJ	LPC	Biogas
Flow rate (tonne/h)	40.0	14.7	25.3	22.5	1.4	0.29
DM content (%)	18.0	32.8	9.4	4.7	98.1	0.0
Simulated input and product (in kg) per tonne (DM) of biomass	1000	668.1	331.9	145.9	184.6	38.43*
Expected input and product (in kg) per tonne (DM) of biomass	1000	668.0	332.0	146.0	186.0	9.15*

*Unit: m³

Table 3. Mole fractions of brown juice, biogas, and digestate excluding the water content.

Component	BJ	Biogas	Digestate
Carbohydrates, lipids	0.798	0.000	0.024
Cellulose, hemicellulose, lignin	0.202	0.000	0.023
Ammonia	0.000	0.001	0.048
Carbon dioxide	0.000	0.405	0.461
Methane	0.000	0.593	0.159
Others	0.000	0.001	0.285

The capital and operating expenses summary, as well as the total product sales when LPC is commercialized using the conventional soybean meal price, are reported in Table 4. The cost of the raw material corresponded to around three-quarters of the total operating cost. In this scenario, the total product sales are smaller than the total operating cost, which means that the expenses are greater than selling profits. Therefore, under those conditions, the implementation of the green biorefinery is unprofitable.

Sensitivity analysis investigated the feasibility of implementing the integrated grass-clover biorefinery to produce LPC, biogas, and fiber. Besides the selling price of the protein, the analysis evaluated the rate of return in case the biomass was estimated cheaper, the CAPEX was reduced and the biorefinery operated the whole year. Figures 2 and 3 presents the variation of the profitability index as a function of the LPC price and

fluctuation on the CAPEX. Figure 3 reports the index when the biomass cost was estimated 20 % cheaper than the cost reported.

Table 4. Techno-economic assessment of the integrated green biorefinery.

CAPEX	Mio. USD	OPEX	Mio. USD/year
Total purchase of equipment	1.96	Raw material	4.72
Direct costs	4.15	Utilities	0.11
Indirect costs	3.09	Labor and maintenance	0.45
Fixed capital investment	7.24	Total operating cost	6.05
Total project capital cost	9.33	Total product sales	5.25

The profitability index shows the relative profitability of the project, showing the present value of the benefits relative to the costs. An index greater than one suggests that the project is profitable. Index lower than one was observed in all the cases when LPC was sold using the conventional price. In the case of selling the LPC at non-GMO prices, the process would be profitable only if the biorefinery operates throughout the whole year, with a biomass price at least 20 % cheaper than the estimated (Figure 3b). Conversely, selling the LPC at the organic price for soybean meals resulted in a profitability index greater than one in different scenarios (Figures 2b, 3a, and 3b).

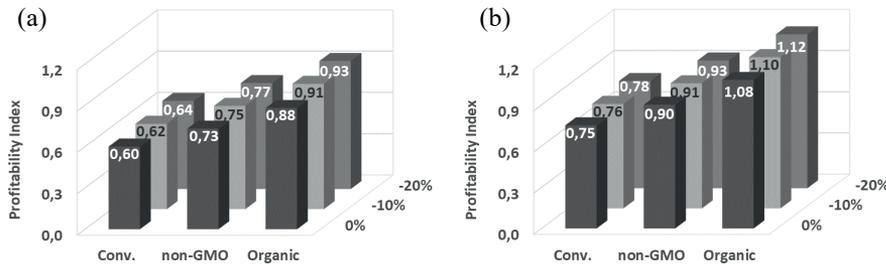


Figure 2: Profitability index of the biorefinery operating during (a) 6 and (b) 12 months per year.

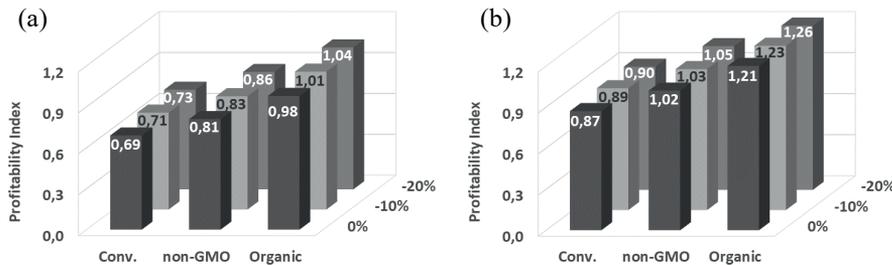


Figure 3: Profitability index of the biorefinery operating during (a) 6 and (b) 12 months per year, at a 20 % cheaper biomass cost.

In the case of annual production, the biorefinery payout return varied between 5 and 6 years, while this return could be reduced to around 4 years if the biomass was cheaper than the estimated value. However, those scenarios are not viable in a Danish environment, where the annual operation is not realistic. To obtain yearly production, the biorefinery needs to integrate the processing of biomasses that are available during winter. This could be grass clover silage, but then the LPC product would be replaced by amino acids, as the soluble protein degrades during ensiling, and the processing should be

adapted to recover amino acids, using membrane filtration instead of precipitation and centrifugation. This scenario was not carried out within this study. Integration of other routes in the biorefinery to obtain higher-value products is also strongly recommended. Upgraded biogas, for example, could increase the total product sales, since biomethane is more expensive than non-upgraded biogas. The same could be obtained separating the white protein, used for human consumption, from the green LPC. The use of fiber and brown juice in other applications also needs further investigation. A complete techno-economic assessment could indicate the best scenario to be implemented.

4. Conclusions

An integrated biorefinery was simulated using data from a demonstration-scale platform. The techno-economic assessment suggested that the process could be profitable for 11 out of 36 scenarios evaluated. Payout return of around 10 years is expected for an organic LPC. Moreover, the cost of biomass has a major effect on profitability. Whole year production would be a good way of increasing the profit, reducing the return time to 5-6 years. However, the process needs to be adapted to other feedstock with different product opportunities. It could also be placed in countries where a continuous annual operation on grasses is viable. The integration of different routes is further recommended.

References

- AspenTech, 2021, Aspen Plus V12 and Aspen Process Economic Analyzer V12, available at: <https://www.aspentech.com/en/products/full-product-listing>.
- A. Corona, M. Ambye-Jensen, G.C. Vega, M.Z. Hauschild, M. Birkved, 2018a, Techno-environmental assessment of the green biorefinery concept: Combining process simulation and life cycle assessment at an early design stage, *Science of the Total Environment*, 635,100-111.
- A. Corona, R. Parajuli, M. Ambye-Jensen, M.Z. Hauschild, M. Birkved, 2018b, Environmental screening of potential biomass for green biorefinery conversion, *Journal of Cleaner Production*, 189, 344-357.
- EUROSTAT, 2021, <http://ec.europa.eu/eurostat/data/database>, accessed 8th November 2021.
- FAOSTAT, 2021, <http://www.fao.org/faostat/en/#data>, accessed 8th November 2021.
- L. Feng, A.J. Ward, M. Ambye-Jensen, H.B. Møller, 2021, Pilot-scale anaerobic digestion of by-product liquid (brown juice) from grass protein extraction using an un-heated anaerobic filter, *Process Safety and Environmental Protection*, 146, 886-892.
- S. Höltinger, J. Schmidt, M. Schönhart, E. Schmid, 2013, A spatially explicit techno-economic assessment of green biorefinery concepts, *Biofuels, Bioproducts and Biorefining*, 8, 325-341.
- T. Llano, C. Arce, D.C. Finger, 2021, Optimization of biogas production through anaerobic digestion of municipal solid waste: a case study in the capital area of Reykjavik, Iceland, *Journal of Chemical Technology and Biotechnology*, 96(5), 1333-1344.
- K. Manevski, P.E. Lærke, J.E. Olesen, U. Jørgensen, 2018, Nitrogen balances of innovative cropping systems for feedstock production to future biorefineries, *Science of the Total Environment*, 633, 372-390.
- S. Njakou Djomo, M.T. Knudsen, L. Martinsen, M.S. Andersen, M. Ambye-Jensen, H.B. Møller, J.E. Hermansen, 2020, Green proteins: An energy-efficient solution for increased self-sufficiency in protein in Europe, *Biofuels, Bioproducts and Biorefining*, 14(3), 605-619.
- A.T. Penteado, J.C. Schöneberger, E. Esche, H.R. Godini, G. Wozny, J.U. Repke, 2018, Sequential Flowsheet Optimization: Maximizing the Exergy Efficiency of a High-Pressure Water Scrubbing Process for Biogas Upgrade, *Computer-Aided Chemical Engineering*, 43,1329-1334.
- K. Rajendran, H.R. Kankanala, M. Lundin, M.J. Taherzadeh, 2014, A novel process simulation model (PSM) for anaerobic digestion using Aspen Plus, *Bioresource Technology*, 168, 7-13.
- M. Santamaría-Fernández, M. Lübeck, 2020, Production of leaf protein concentrates in green biorefineries as alternative feed for monogastric animals, *Animal Feed Science and Technology*, 268, 1-20.

The effect of alternative fuels on calcium-looping

Ana Amorim,^{a,b} Ismail Mohamed,^a Luíza Marques,^a Rui M. Filipe,^{b,c} Henrique A. Matos,^b

^a *c⁵Lab – Sustainable Construction Materials, 2795-242, Linda-a-Velha, Portugal*

^b *CERENA, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal*

^c *Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, 1959-007 Lisbon, Portugal*

aamorim@c5lab.pt

Abstract

Calcium-looping (CaL) is a promising post-combustion CO₂ capturing technology, showing high compatibility with the cement industry. Nevertheless, this technology is still under development, and its implementation still leads to higher operating costs. Aiming cost reduction, the influence of using alternative fuels, commonly used in the kiln of cement plants, in the CaL calciner was studied in terms of key performance indicators, such as the amount of fuel needed, the carbonator's heat duty, the air separation unit (ASU) energy consumptions, and the amount of CO₂ captured in the process. Thus, a CaL process was modeled using Aspen Plus, incorporating a carbonator model developed using Python. Different random compositions of petroleum coke, tire chip, waste-derived fuels, olive pomace, dangerous liquid residues, and cork residues were studied. It was concluded that the use of petroleum coke leads to lower fuel consumption and ASU consumption due to its higher heating value. However, the alternative fuels studied have advantages in terms of CO₂ capture efficiencies and energy produced in the carbonator due to lower sulfur contents. Therefore, the sorbent flowrates would have to increase to obtain the same capture efficiency using petroleum coke, which would result in a higher process cost. Alternative fuels are promising in CaL and should be a topic for further studies.

Keywords: Calcium-looping, cement, CO₂ capture, alternative fuels, process modeling

1. Introduction

The cement industry is one of the primary industrial sources of CO₂ emissions, responsible for 7 % of the anthropogenic emissions (IEA, 2018). Carbon capture utilization and storage (CCUS) technologies emerge as a strategy to achieve carbon neutrality by 2050, according to the objectives defined in the 13th United Nations Sustainable Goals. A promising CCUS technology is the calcium-looping (CaL) process. This process presents a high synergy with the cement industry since a Ca-based sorbent, limestone, is also a raw material for cement production. This process comprehends two circulating fluidized bed (CFB) reactors, a calciner, and a carbonator. The kiln flue gas is fed to the carbonator, in which the CO₂ is captured, reacting with CaO to form CaCO₃. The carbonation reaction is exothermic and the heat released is used to produce electrical energy through a steam cycle. The CaCO₃ formed in the carbonator enters the calciner to be regenerated to CaO. As the calcination is endothermic, energy must be supplied to the calciner through oxyfuel combustion. The most common fuel used in CaL pilot-plants is still coal (Hornberger et al., 2017; Arias et al., 2018). However, there is an effort to replace it with greener fuels, such as biomass (Alonso et al., 2014). The use of alternative fuels (AFs) leads to the maintenance of fossil fuel reserves and is also reported to enable cost

reduction in cement plants (Chatziaras et al., 2014). Thus, they are partially replacing fossil fuels to fulfill the kiln's energy requirements. The calcium looping process must follow as the cement industry evolves towards using higher percentages of alternative fuels. This work aims to assess the influence of these fuels on CaL performance and evaluate them as an alternative to fossil fuels. The alternative fuels studied were the ones usually consumed in Portuguese cement plants.

2. Model development

2.1. The experimental procedure

An experimental laboratory-scale fluidized bed reactor (FBR) system with an electric oven was used to obtain the multicyclic sorbent deactivation curve of a CaO-based sorbent. The sorbent used was a Portuguese limestone provided by an industrial partner with particle sizes of 250-355 μm . The carbonation and calcination were conducted at atmospheric pressure and 700 °C and 970 °C, respectively. The FBR is fed with synthetic mixtures (at 20 °C, 1 bar, 5 mL/min) of 25 mol% CO₂ and 75 mol% air during carbonation and 80 mol% CO₂ and 20 mol% air during calcination. A limestone sample with approximately 5 g was loaded into the reactor for each experiment. Four deactivation curves were obtained, and the average sorbent conversion was considered.

2.2. The process model

The CaL process model was developed using the commercial software Aspen Plus. A diagram of the process is presented in Figure 1. The calciner ($T = 970$ °C, $p = 1$ atm) was modeled by minimizing the free energy of Gibbs. 100 combinations of 6 different fuels (petroleum coke, waste-derived fuels, tire chip, olive pomace, dangerous liquid residues (DLR), and cork residues) with different random compositions were provided to the calciner to fulfill its energy requirements. The proximate and ultimate analysis of each of these fuels is shown in Table 1. These analyses were provided by an industrial partner and are presented on a dry basis. The sulfur present in all the fuel samples was considered organic. The fuels were burnt in oxyfuel conditions, and the amount of fuel needed was estimated to maintain the calciner's temperature at 970 °C. A 4 mol% O₂ excess was considered in the calciner. A fresh sorbent stream was added to the calciner, and a sorbent purge was withdrawn to reach a steady-state with an average sorbent conversion. High-efficiency cyclones were placed at the exit of both the calciner and the carbonator to separate the CO₂ concentrated stream and the CO₂ free flue gas from the sorbent stream, respectively. They had a separation efficiency higher than 99 wt% and a pressure drop of approximately 0.01 atm. The composition of the flue gas being treated in the carbonator was based on the annual emissions of a Portuguese cement plant (Cimpor, 2019). Currently, the model is customized to describe a pilot-plant unit; therefore, the process was designed to treat a hundredth part of the actual emissions. A heat exchanger between the CO₂ concentrated stream leaving the calciner and the sorbent stream leaving the carbonator was considered in the model, as suggested by Martínez et al. (2013). This heat integration leads to a reduction in the calciner's energy requirements, decreasing fuel consumption. The CO₂ concentrated stream was compressed with moisture removal.

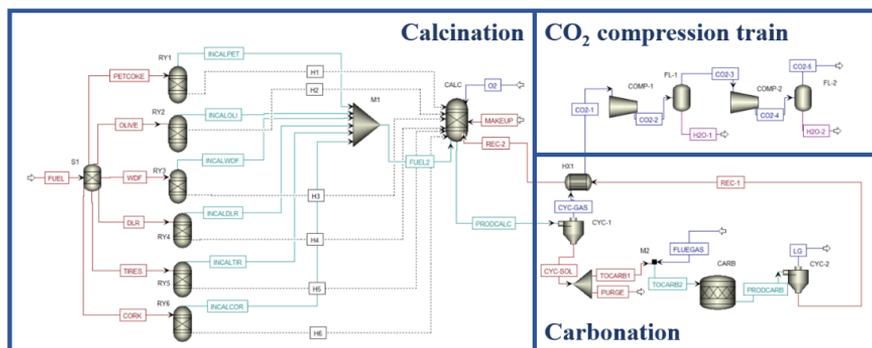


Figure 1. Diagram of the CaL process model developed using Aspen Plus. Solid streams are presented in red, gas streams in blue, liquid ones in pink and gas-solid streams in green.

Table 1. Proximate and ultimate analysis of fuel samples.

Fuel analyses		Petroleum coke	Waste derived fuels	Tire chip	Olive pomace	Dangerous liquid residues	Cork residues
Proximate	Moisture	9.0	26.9	1.4	12.1	22.0	12.1
	Fixed Carbon	87.2	66.2	21.3	90.9	100.0	94.0
	Volatiles	12.0	0.0	63.1	0.0	0.0	0.0
	Ash	0.8	33.8	15.6	9.1	0.0	6.0
Ultimate	Ash	0.80	33.80	15.58	9.10	0.0	6.00
	C	85.80	46.90	71.41	49.80	63.20	51.40
	H	3.92	5.56	4.55	5.83	9.47	5.91
	N	0.00	0.85	0.58	1.59	2.35	0.00
	O	0.00	0.54	0.01	0.00	0.00	0.00
	S	6.16	0.58	1.59	0.16	0.57	0.00
O	3.34	11.77	6.28	33.52	24.41	36.69	

2.3. The carbonator model

The carbonator model followed the work developed by Romano (2012). The model was implemented in Python and returned the CO₂ capture efficiency. This value was inserted on the process model, yielding ash and sulfur flowrates, which entered the carbonator model as an input. This process was repeated until convergence was reached. The deactivation over multiple carbonation/calcination cycles was considered based on the semi-empirical law proposed by Grasa and Abanades (2006). Based on the empirical curves, the deactivation constant, k , and residual sorbent conversion, X_r , were estimated at 0.086 and 4.78, respectively. The experimental fitting of the deactivation is represented in Figure 2. Since there is no combustion occurring in the laboratory calciner, the experiments were performed in the absence of SO₂. The deactivation caused by the presence of sulfur was assumed to follow the correlation proposed by Romano (2012). In the carbonator, the sorbent recycling flowrate to CO₂ flue gas flowrate ratio and the fresh sorbent flowrate to CO₂ flue gas flowrate ratio were set to 6 and 0.5, respectively. The CFB hydrodynamics was considered by applying the model by Kunii and Levenspiel (2000), as proposed by Romano (2012). A reactor height of 40 m and a mean superficial velocity of 5 m/s were considered.

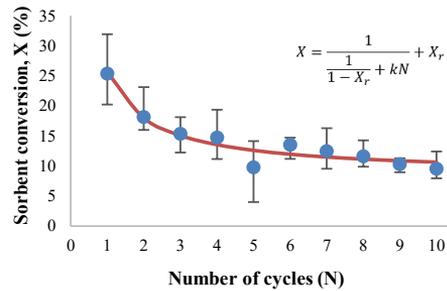


Figure 2. Comparison between model and experimental multicyclic sorbent deactivation. The model deactivation constant, k , and residual sorbent conversion, X_r , were estimated at 0.086 and 4.78, respectively. The experimental data was obtained for a Portuguese limestone (250-355 μm).

3. Results

3.1. Amount of fuel needed

The influence of the fuel mixture composition on the amount of fuel needed in the calciner is presented in Figure 3. The increase of the petroleum coke fraction in the fuel mix corresponds to a decrease in the amount of fuel needed in the calciner. On the other hand, an increase in the waste-derived fuel (WDF) and olive pomace contents results, in general, in an increase in the fuel requirements in the calciner. This increase is explained by the carbon content of the fuels. A higher carbon content results in a higher calorific value, thus in a lower amount of fuel needed in the calciner. Petroleum coke is the fuel with higher C content, and both WDF and olive pomace have low carbon contents.

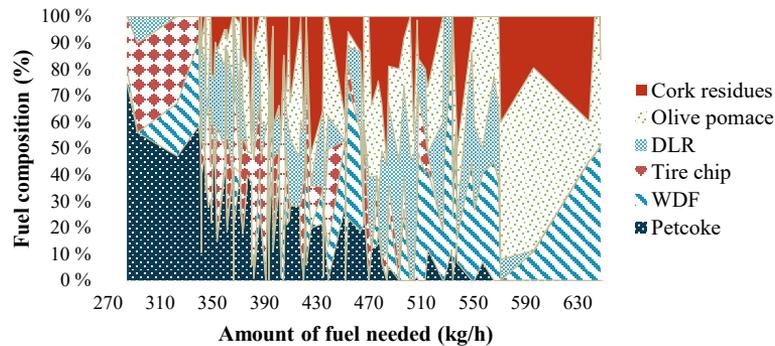


Figure 3. Variation in the mass flowrate of total fuel needed in the calciner as a function of fuel composition variations.

3.2. Air separation unit energy consumptions

The oxygen needs in the calciner and thus the air separation unit electric consumptions decrease with the increased use of petroleum coke, as represented in Figure 4. This decrease was expected, since a higher content of petroleum coke results in lower fuel consumption, thus a lower amount of oxygen required for the combustion reactions. It is possible to see the reverse trend when using higher amounts of olive pomace.

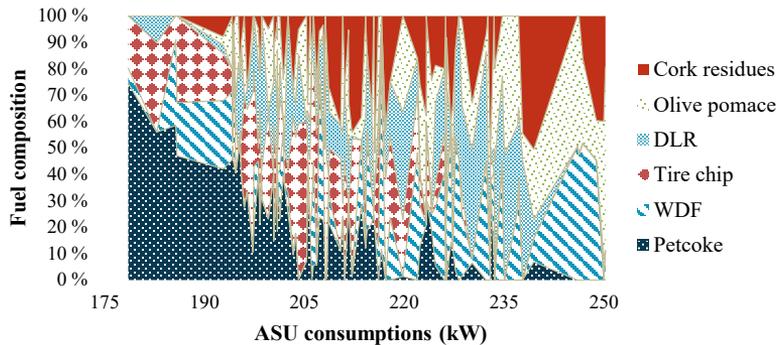


Figure 4. Variation of the air separation unit (ASU) consumptions as a function of fuel composition variations.

3.3. CO₂ capture in the carbonator

As depicted in Figure 5, the increase of the petroleum coke content results in a reduction of the CO₂ capture efficiency. This reduction is related to the high sulfur content of this fuel, as can be seen in Figure 2. Therefore, for presenting the same capture efficiency as the one obtained using the AFs studied, the recycling sorbent stream flowrate would have to increase, thus increasing process costs.

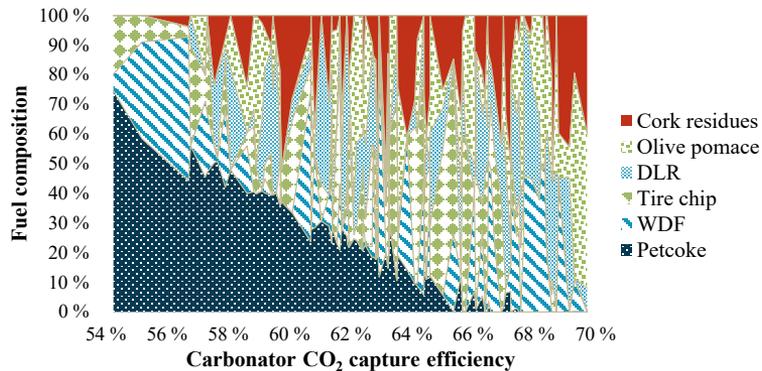


Figure 5. Variation of the carbonator CO₂ capture efficiency as a function of fuel composition variations.

3.4. Carbonator's heat duty

The carbonation reaction is exothermic, thus releasing energy. The influence of each of the fuels in the energy released in the carbonator is presented in Figure 6. It is possible to conclude that the petroleum coke content also significantly impacts the energy released in the carbonator. With the increase of this content, the energy released decreases because of the lower CO₂ capture efficiencies obtained using this fuel. The lower captures mean the extent of the reaction is lower, thus producing less energy.

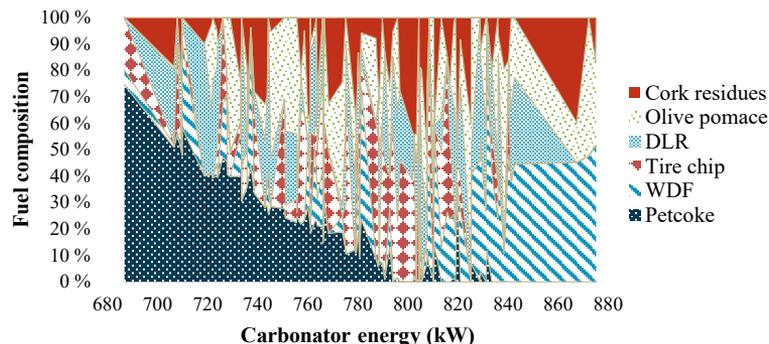


Figure 6. Variation of the energy released in the carbonator as a function of fuel composition variations.

4. Conclusions

The use of fuels commonly consumed in the cement plant, including AFs, was evaluated in terms of total fuel required, ASU electric consumptions, CO₂ capture in the carbonator, and heat released during carbonation. It was concluded that petroleum coke had the highest heating value, thus requiring lower flowrates to fulfill the calciner's energy requirements. Despite the higher AFs' flowrates required, they may reduce fuel costs since their market values are lower than fossil fuels. However, the higher amounts of fuel also lead to higher O₂ usage and, therefore, ASU's electric energy consumption, which corresponds to higher process costs. On the other hand, petroleum coke has a high sulfur content, resulting in lower carbonator CO₂ capture for equal sorbent flowrates; therefore, higher sorbent flowrate and, thus, higher costs are required to keep CO₂ capture efficiency. Consequently, it can be concluded that these AFs are a promising alternative for the CaL process and should be a topic for further studies, including a global comparative economic analysis, where the carbon tax savings due to the negative CO₂ emissions of AFs is considered.

References

- M. Alonso, M. E. Diego, C. Pérez, J. R. Chamberlain, and J. C. Abanades, 2014, Biomass combustion with in situ CO₂ capture by CaO in a 300 kW_{th} circulating fluidized bed facility, *International Journal of Greenhouse Gas Control*, 29, 142–152.
- B. Arias, M. Diego, A. Méndez, M. Alonso, and J. Abanades, 2018, Calcium looping performance under extreme oxy-fuel combustion conditions in the calciner, *Fuel*, 222, 711–717.
- N. Chatziaras, C. Psomopoulos, and N. Themelis, 2014, Use of alternative fuels in cement industry, 12th International Conference on Protection and Restoration of the Environment, 521–529.
- Cimpor, 2019, Declaração Ambiental Atualizada - Alhandra.
- G. S. Grasa and J. C. Abanades, 2006, CO₂ capture capacity of CaO in long series of carbonation/calcination cycles, *Industrial and Engineering Chemistry Research*, 45, 26, 8846–8851.
- M. Hornberger, R. Spörl, and G. Scheffknecht, 2017, Calcium Looping for CO₂ Capture in Cement Plants - Pilot Scale Test, *Energy Procedia*, 6171–6174.
- IEA, 2018, Technology Roadmap - Low-Carbon Transition in the Cement Industry.
- D. Kunii and O. Levenspiel, 2000, The K-L reactor model for circulating fluidized beds, *Chemical Engineering Science*, 55, 20, 4563–4570.
- A. Martínez, Y. Lara, P. Lisbona, and L. M. Romeo, 2013, Operation of a Cyclonic Preheater in the Ca-Looping for CO₂ Capture, *Environmental Science and Technology*, 47, 19, 11335–11341.
- M. C. Romano, 2012, Modeling the carbonator of a Ca-looping process for CO₂ capture from power plant flue gas, *Chemical Engineering Science*, 69, 1, 257–269.

Simultaneous Synthesis of Metabolic and Process Engineering for the Production of Muconic Acid

Dimitriou Konstantinos^{a,b}, Kokossis Antonis^a

^a*School of Chemical Engineering, National Technical University of Athens, 9, Iroon Polytechniou St, Athens 15780, Greece*

^b*Faculty of Technology, Policy and Management, Delft University of Technology, 4, Mekelweg, Delft 2628 CD, Netherlands*

Abstract

The paper presents a systems integration approach for the computational strain design workflow for the identification of reaction eliminations that reshape network connectivity in a way that both biomass production and revenue are simultaneously maximized by utilizing a bilevel optimization framework. The research expands the outer problem of the bilevel approach that is established in the literature with the use of a superstructure scheme that addresses several design options simultaneously with the several options to select the pathways. The superstructure scheme is laid out to reduce model and optimization complexities. We assess the effect of the new optimization goal, for varying number of metabolic interventions to the downstream separation network and the bioprocess revenue. We then compare the enriched model to the previous analysis, which only aims to maximize the production of a target metabolite. To showcase the functionality and effectiveness of the developed model we applied the workflow to a muconic acid producing strain of *S. cerevisiae* (iMM904 GEM) that includes the necessary heterologous pathways. Overall, this computational framework could be an important step to bridge the gap between strain design and process engineering.

Keywords: Systems metabolic engineering, superstructure optimization, bilevel programming, OptKnock, DBTL

1. Background and Motivation

Metabolic engineering has played a central role in the production of chemicals and materials but it can be time consuming and costly to produce strains suitable for industrial production (Choi et al., 2019). Still, recent advances in metabolic pathway reconstruction and in silico modelling have enabled the development of microbial strains suitable for efficient cell-factories. Directed evolution, for instance, has assisted in the expansion of the possible spectrum of products that can be derived from host strains. It involves a cyclic process that aims to reveal functional gene variants through an alternation between gene diversification and screening (Arnold, 2018). It has also made the manufacturing of enzymes and biofuels more environmentally friendly, by shifting the input from petroleum-based to renewable sources such as lignocellulosic.

Emerging infrastructures of Design-Build-Test-Learn (DBTL) paradigms hold a tremendous promise to innovate with new products and novel integrated processes that can be intensified for the highest efficiency. DBTL is a cycle that starts with the Design of the strains based on the industrial specifications through a computational workflow. The next steps are Build, where the strains are manufactured with protein engineering techniques, and then Test with omics analysis and high throughput measurements.

Finally, the Learn step derives insightful solutions, by using data management and life cycle analysis, that are fed back into the first step.

The current DBTL cycle approach in biofoundries and biorefineries, however, remains a disconnected sequence of tasks with an undisclosed potential to integrate. Specifically, the genetic modifications of strains are usually driven by the overproduction of target metabolites without considering downstream processing which can account for up to 80% of the overall bio-production cost (Kiss et al., 2015). The explosion of data masks poor interactions between upstream science and engineering. The Design step typically produces hundreds, thousands, or even millions of candidates, but Build stages proceed with a few choices, often made by heuristics and intuition. The selected choices hold no guarantees for quality and all inheritance and knowledge generated by the in-silico work is lost and unexplored. For instance, by utilizing convex analysis models on Genome Scale Models (GSM) the behavior of cells can be explained and predicted. But, still these models fail to incorporate the downstream processing steps that are ahead. Additionally, the connection between the Design and Learn accounts for the weakest of all interactions in the DBTL cycle. Generally, the Design deploys over-simplified formulations for mass and balances and neglects important aspects of process engineering. Lastly, yields in pilot or industrial scale installations dramatically inferior to initial promises. As such, there is an instrumental potential for the synthesis of in-silico methods meant to optimize the metabolism of organisms and the design of downstream processes.

2. Methodology

To address the issues raised we propose to expand the mathematical formulation of a metabolic reconstruction tool (OptKnock) by including the downstream processing step in the form of a superstructure. In this part, we will present the extension of the model, then explain the mathematical formulation utilized and present the enlisted superstructure.

2.1. Model Extension

The problem faced in this work is the following. Given the target product, the microorganism with constrained based model available and reconstruction options to select, as well as the sequences of process engineering trains with options, optimize the process efficiency, select the metabolic reconstruction and select the process engineering options most suitable to the problem.

The objective function is the maximization of the profit, which is calculated by the final product flow and value, as well as the downstream costs. The problem is constrained by the outer problem, which is the process performance, and the inner problem, which is the cell performance, and together they form a bilevel optimization problem. To simplify the final optimization problem, it was decided to solve it using linear programming, which led to the utilization of simplified engineering models for unit operations and shortcut cost models. Additionally, the engineering processes focused only on the extracellular products as a part of this study.

2.2. Mathematical Formulation

Optknock is a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization (Burgard et al., 2003). The initial OptKnock optimization problem consists of an inner problem and an outer problem. The inner problem aims to maximize the biomass and the outer problem aims to maximize the target product. These two targets are competing for the carbon intake and are referred as a bilevel optimization problem.

The novel approach in this research is that the outer problem is expanded to include new process engineering objectives and the maximization goal is shifted from product maximization to profit optimization. Since profit is derived directly from the desired final product flow, the product is still maximized, but now process engineering is also accounted for. Specifically, the downstream processing is translated in the form of a superstructure. Each processing stage (S) includes the selected possible separation processes (T). Separation steps include pretreatment, cell removal, product isolation, concentration, purification and refinement. The existence of each process is denoted by binary variables. The superstructure's input stream is a variable determined by the optimization problem and is separated into five components: product, liquid by-product, water, cells, and solid by-product. Each process redistributes the stream's components through linearized model equations. The problem can be constrained based on the specific bio-process characteristics and objectives.

Our method entails the identification and the categorization of these technologies as well as the insertion of the model equations and their economic parameters into the optimization problem. The model has been developed in GAMS environment and solved using the BARON global optimization solver. On the tables below, the mathematical formulation is presented.

Table 1 (left): The mathematical representation of the model with the new maximization goal

Table 2 (right): Representation of the additional sets of constraints

maximize Profit	Target Maximization	Outer Problem	Input Data	$F(i) = P(i) + Lbp(i) + W(i)$
subject to				$P_{in} = TargetMetabolite$
maximize $v_{biomass}$	Biomass Maximization	Inner Problem		$Lbp_{in} = \sum Liquid\ Byproducts$
subject to				$W_{in} = WaterContent$
				$C_{in} = Biomass$
				$Sbp_{in} = \sum Solid\ Byproducts$
$\sum_{j \in I} S_{ij} v_j = 0, \quad \forall i \in I$	Mass Balance			$\forall j, \sum_{i \in I} y_i \leq 1, \quad y_i \in \{0,1\}$
$V_{EX_glc}(\epsilon) \geq -v_{EX_glc}^{uptake}(\epsilon)$	Nutrient intake requirement			$y_{u_2} - y_{u_1} \leq 0$
$V_{EX_O_2}(\epsilon) \geq -v_{EX_O_2}^{uptake}(\epsilon)$	Oxygen intake requirement			
$V_{ATPM} = v_{ATPM}^{min}$	ATP requirement			
$V_{biomass} \geq f v_{biomass}^{max,WT}$	Biomass production requirement			$\sum C_i - U \cdot y_i \leq 0$
$LB_j y_j \leq v_j \leq UB_j y_j, \quad \forall j \in J$	Limits of reaction flows			$\sum Lbp_i - U \cdot y_i \leq 0$
$\sum_{j \in I} (1 - y_j) \leq K$	Allowed knockouts		$\sum P_i - U \cdot y_i \leq 0$	
$y_j \in \{0,1\}, \quad v_j \in R, \quad \forall j \in J$			$\sum Sbp_i - U \cdot y_i \leq 0$	
New constraints			$\sum W_i - U \cdot y_i \leq 0$	
			Process Constraints	$P_{t2B} \geq X_f$
				$W_{t2B} \leq W_f$
				$Sbp_{t2B} \leq Sbp_f$
				$Lbp_{t2B} \leq Lbp_f$
			Mass Balance Equations	$\forall Z, \sum_{i \in I} X_i \cdot Z = \sum_{i \in I+1} Z$

2.3. Enlisted Downstream Processing

The downstream processing of each bioproduction might be different but the general steps for the treatment of the fermentation broth are similar. The difference lies on the specific process that is chosen. For our case, we selected a broad spectrum of separation processes for each step but these can be adjusted to each specific case. In the case that a step should be omitted we added the option BYPASS where the process flows remain the same for the following step. Below we present the superstructure illustration.

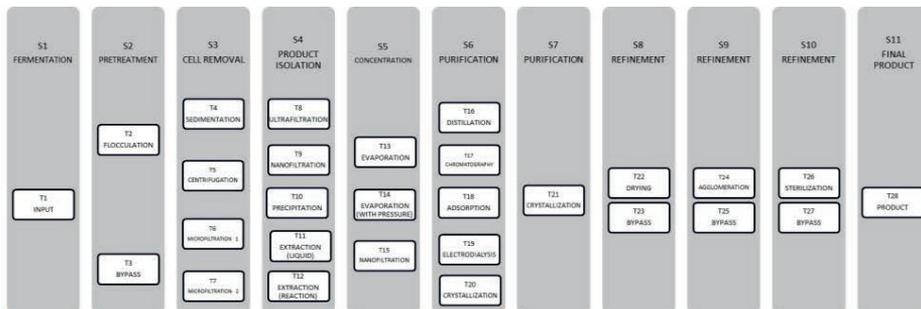


Figure 1: Representation of the enlisted downstream processing superstructure

3. Illustration Cases

Here we will showcase the optimization problem and present a case study for the utilization of the proposed methodology and tool. The case is the production of muconic acid from the strain of *S. cerevisiae* (iMM904 GEM) that includes the necessary heterologous pathways.

3.1. Optimization Problem

The degrees of freedom of the problem include the biological system, meaning the reactions that could be knocked out and the processes engineering objectives that are the separation trains and equipment for a 5 year operation. Firstly, a parametric analysis will be conducted by using only the metabolic reconstruction tool where glucose intake (substrate) and minimum allowed biomass will be set to derive the deleted reactions and output distributions. It is assumed that up to 2 reactions can be knocked out, the separation task can be repeated but with a single and different piece of equipment, no separation stages are repeated in a sequence and the separation mixtures are ideal (e.g. no azeotropes and recycles are allowed). Then, after calculating the initial product flows, they will be inserted to the disconnected superstructure for the downstream processing and the results will be compared to the proposed approach, which is solving the two problems simultaneously. It is expected that designing the biological system and the separation together will yield a more profitable production.

3.2. Case Study

Muconic acid is considered a “bio-privileged molecule” as it enables the production of a variety of useful products such as bio-plastics and energy (Bentell et al., 2001). It can be produced by renewable resources such as lignocellulosic biomass through bio-production (Wu & Maravelias, 2019). Therefore it crucial to build strains, by introducing heterologous enzymes and genes, which can produce muconic acid economically. *S. cerevisiae* is a microorganism that has been used for the production of muconic acid and has been chosen for this study (Wang et al., 2020).

3.2.1. Parametric Analysis (Cases A, B, C)

As it was discussed in this part we conducted a parametric analysis on the biological system. Specifically, we assessed the effect of substrate intake (glucose) and minimum biomass on the final concentration of the product, the liquid by-product and the water. It was revealed that increasing the minimum biomass, with the same glucose intake, led to significantly increased product formation, with a slight decline in the liquid-by product and the same water content. By returning the glucose intake to its initial value and decreasing the minimum biomass it, the results showcased that the product was again

higher than the initial value, but the liquid by-product and the water contents got reduced. The results can be seen on the Table 3.

3.2.2. Conventional Approach (Case C)

Here we set the glucose intake to 0.5 mmol/(dDW·h) and the minimum biomass to 0.01 h⁻¹ and solved the biological system to find the metabolites distribution which were then used to feed the downstream processing superstructure. By solving the problem we found the deleted reactions, the chosen processes, as well as the total profit, which was 1,7 mil. \$. Again, the overall results can be seen on the Table 3.

3.2.3. Simultaneous Approach (Cases D, E, F, G)

By solving the optimization problem simultaneously we were able to reveal different metabolic reconstruction options (deleted reactions), which led to a different selected downstream processing scheme and a more profitable production that reached 16.1 mil. \$ (Case D). This significant increase in profits could not be revealed from the disconnected approach. For Case E, the problem is further constrained by demanding the minimum product flow at the final stage to be 40 kg/h. This forces the algorithm to comply with this restriction and thereby choosing the suitable parameters that decrease the overall profits to 14.9 mil. \$. Similarly, in the following run (Case F) the restriction of maximum water content at the final stage decreases the overall profit to 13.9 mil. \$. Finally, in Case G the two previous restrictions are relaxed and the effect of increased glucose intake is examined. Specifically, at 0.1 mmol/(dDW·h) substrate intake the profits increase to 37.9 mil. \$, a significant increase from Case D, where glucose intake was just the half amount.

3.3. Summary of Results

On the following table, a summary of the results is presented.

Table 3: Summary of all the cases inputs and outputs

Scenario	Conventional			Simultaneous			
	A	B	C	D	E	F	G
Input							
Glucose Intake [mmol/dDW·h]	1	1	0.5	0.5	0.5	0.5	1
Minimum Biomass [h ⁻¹]	0.01	0.05	0.01	0.01	0.01	0.01	0.01
Min Product Flow [kg/h]				0	40		
Maximum Water Flow [kg/h]				0		0.02	
Output							
Deleted Reaction 1	r_3529	r_0892	r_0253	r_0073	r_4025	r_2220	r_3028
Deleted Reaction 2	r_3530	r_1836	r_1994	r_0806	r_4608	r_4608	r_3908
Product [mmol/dDW·h]	0.16	0.47	0.31	0.41	0.41	0.41	0.85
Liquid By-product [mmol/dDW·h]	3.02	2.79	0.71	0.54	0.54	0.54	0.91
Water [mmol/dDW·h]	4.35	4.36	1.82	1.77	1.77	1.77	3.45
Product S11 [Kg/h]				36.6	40.1		
Water S11 [kg/h]				0.027		0.005	
Superstructure							
S1			T1	T1	T1	T1	T1
S2			T3	T3	T3	T3	T3
S3			T6	T6	T6	T6	T6
S4			T12	T10	T8	T8	T12
S5			T13	T13	T13	T13	T13
S6			T18	T17	T20	T20	T20
S7			T21	T21	T21	T21	T21
S8			T23	T23	T23	T23	T23
S9			T25	T25	T25	T25	T25
S10			T27	T27	T27	T27	T27
S11			T28	T28	T28	T28	T28
Profit [mil. \$]			1.7	16.1	14.9	16.9	37.9

From the results above it is evident that there are significant deviations between the simultaneous and the staged approach, including differences in the knockout reactions. At the same time different types of separation trains for each case are observed, where the number and the types of DSP both change as different engineering constraints are used. Even in the simultaneous approach, each case has resulted into different sets of knockouts. Engineering constraints hold a lesser impact on profitability, holding a similar order of magnitude, so the approach can be good for targeting provided that the design stage retains DOF for reaction knockouts and that the optimization is allowed to switch different types of separation

4. Future Work

The present work contributes to the connection between the Design and Learn steps of the DBTL cycle, but still follows a simplified approach on the superstructure construction. Therefore, the development of a complete superstructure with all the possible downstream processes and their respective more realistic equations that describe them, could yield more accurate results. Moreover, this work does not contribute to the Design to Build steps, which consists of the scaling-up, that is affected by kinetic parameters and dynamics. An approach that simultaneously solves the GSMs and kinetics is still needed. Additionally, systematic methods that connect GSMs with the real world are important and therefore complete superstructures have to ‘wait’ before this part of systematization is addressed. The proposed methodology described in the paper has value mostly as a targeting and scoping tool and not as an all-encompassing solution.

References

- Arnold, F. H. (2018). Directed Evolution: Bringing New Chemistry to Life. *Angewandte Chemie - International Edition*, 57(16), 4143–4148. <https://doi.org/10.1002/anie.201708408>
- Bentell, L., Sandberg, H., Lagneborg, R., & Lindblad, B. (2001). *CO 2 emissions of the Swedish steel industry*. 420–425.
- Burgard, A. P., Pharkya, P., & Maranas, C. D. (2003). *OptKnock : A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization*. <https://doi.org/10.1002/bit.10803>
- Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., & Lee, S. Y. (2019). Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends in Biotechnology*, 37(8), 817–837. <https://doi.org/10.1016/j.tibtech.2019.01.003>
- Kiss, A. A., Grievink, J., & Rito-Palomares, M. (2015). A systems engineering perspective on process integration in industrial biotechnology. *Journal of Chemical Technology and Biotechnology*, 90(3), 349–355. <https://doi.org/10.1002/jctb.4584>
- Wang, G., Özmerih, S., Guerreiro, R., Meireles, A. C., Carolas, A., Milne, N., Jensen, M. K., Ferreira, B. S., & Borodina, I. (2020). Improvement of cis, cis-Muconic Acid Production in *Saccharomyces cerevisiae* through Biosensor-Aided Genome Engineering. *ACS Synthetic Biology*, 9(3), 634–646. <https://doi.org/10.1021/acssynbio.9b00477>
- Wu, W., & Maravelias, C. T. (2019). Identifying the Characteristics of Promising Renewable Replacement Chemicals. *IScience*, 15, 136–146. <https://doi.org/10.1016/j.isci.2019.04.012>

Renewable hydrogen supply chain for transport application in Corsica island

T. Moustapha Mai^a, C. Cristofari^a, C. Azzaro-Pantel^b, E. Carrera^b

^a*UMR CNRS 6134 Renewable Energy Laboratory, Scientific Centre Georges Peri, University of Corsica, F20000 Ajaccio, France.*

^b*Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, 31029 Toulouse, France*

Moustapha-mai_m@univ-corse.fr

Abstract

This paper aims to present an optimization approach for the design and planning of a hydrogen supply chain in Corsica Island. The formulation developed is based on a multi-period approach using a Mixed Integer Linear Programming (MILP) methodology. Three configuration scenarios of the Hydrogen Supply Chain (HSC) considering three single objective case studies are investigated related to the minimization of the total daily cost, of greenhouse gas emissions and of a risk-based index, respectively. The production of hydrogen is expected to cover the future demand of the fuel cell electrical vehicles (FCEVs) fleet planned for the 2030 period. The key innovation proposed here is to reconcile design and strategic planning over a short period using monthly fluctuations of hydrogen demand according to the predicted fuel consumption of the territory. The single objective optimizations for the case study show a decentralized distribution of production, storage and distribution units across the territory, with a lowest hydrogen cost of 8 €/kg during the period of high demand. This approach can be used to explore appropriate incentive mechanisms to boost the hydrogen economy in an isolated territory.

Keywords: Hydrogen, optimization, low-carbon fuel, Mixed-Integer Linear Programming, isolated area.

1. Introduction

The urgency to fight climate change has stimulated more and more interest in shifting investments from fossil-based energy to renewable sources. In that context, islands could play a key role in global development by becoming perfect places for demonstration of new clean technologies and pathways for sustainable development (Krajačić et al., 2008). In these isolated territories, the impact of the fossil fuel price is even higher due to its transport and their grid network is less stable. These reasons make the grid management more difficult and increase the cost of electricity production (Lamas, 2016). The massive integration of renewable energies (RE) on the small island grids could further increase the disturbances between production and demand because of their intermittency. In that context, hydrogen can be viewed as an “energy vector” and play a major role in decarbonization if generated by electrolysis using renewable energy surplus (Krajačić et al., 2008). Due to actual hydrogen demand scale, today investments remain yet too risky for wide-scale green hydrogen production that could compress costs, creating the so-called chicken-and-egg problem.

According to several studies, the economic viability of a hydrogen production system associated with renewable energy sources depends strongly on the identification of an

optimal configuration of the hydrogen supply chain (HSC) (Seo et al., 2020; Won et al., 2017). The HSC design has been mostly studied at the strategic/tactical level compared to the operational levels. The inputs of such models are constituted by a set of options for the production, storage and transportation, while the outputs are relative to the type, number, location and capacity of the production, storage and transportation units. Most works devoted to hydrogen supply chain modelling are based on mathematical programming approaches, mainly Mixed Integer Linear Programming (MILP) and generally limited to single objective (cost minimization) (Won et al., 2017) or bi-criteria assessment, i.e., cost-environment or cost-risk (Kim and Moon, 2008). These contributions generally involve a rolling horizon approach in which the overall horizon is divided into strategic periods during which capital investments are made at the very beginning and which can last up to 10 years. The decision variables made at a given period are then coupled over the large time-span, since the optimization strategy is viewed globally. As reported in (Agnolucci et al., 2013) only a few infrastructure optimization studies explored the spatial and temporal dynamics of demand which was identified as a key sensitive parameter with sources of uncertainty that can be addressed by a stochastic approach (Kim et al., 2008).

The objective of this work is to demonstrate deployment of hydrogen technologies in an integrated manner in a remote territory, here Corsica Island. The formulation is based on the previously developed long-term MILP approach for HSC infrastructure design (De Leon Almaraz, 2014). The key innovation we propose here is to extend this framework to the combined design and planning over a shorter period using monthly fluctuations of hydrogen demand according to the fuel consumption of the territory.

2. Presentation of the methodological framework

2.1. Principles

The methodological framework for the design and operation of a future HSC is treated regarding a geographical and multi-period approach in order to supply hydrogen demand profile of a given geographic area. The optimization problem was formulated as a Mixed Integer Linear Programming (MILP) model using GAMS[®] modeling system with CPLEX 12 solver (De Leon Almaraz, 2014). Three optimization objectives are considered in this study: an economic criterion that consists of the minimization of the total network cost, both in terms of capital and operating expenditures, an environmental criterion related to the global Greenhouse Gas (GHG) emissions of the supply chain and a risk index (Kim and Moon, 2008) subject to: supply, demand, mass conservation and technical performance. A set of techno-economic parameters is considered: production, storage, and transport options, possible locations, available energy sources, capital and operation cost, technical features (efficiency, capacity, lifetime, load factors, storage capacities) and GHG emissions of the various technologies. Several constraints are involved considering continuous, integer and binary variables. The integer decision variables represent the location of facilities, sizing decisions and the selection of suitable production technologies and of transportation modes between facilities. The binary (respectively continuous) variables indicate the hydrogen transport direction (respectively flows). The model then provides as outputs the HSC optimal configuration.

2.2. Model assumptions

The assumptions of the model can be summarized by:

- The territory is divided into grids;
- Production units can only be installed on specified grids;
- Only photovoltaic (PV) and wind power are considered for energy sources;

- Hydrogen is produced only by electrolyzers and supplied in gaseous form;
- Only transport by tanker-trucks may exist between grids;
- Produced hydrogen is dedicated to transport application and distributed via refueling stations.

2.3. Mathematical formulation

The novelty of this work is based on the use of finer grains for time interval definition (i.e. monthly intervals) that require the definition of additional parameters such as the number of operating hours and the power of the electrolyzers.

2.3.1 Constraints

The total availability of primary energy sources in a grid g during time period t ($Atot_{gt}$) is given as a sum of three terms, that correspond to the initial average availability of renewable energy sources ($A0_{egt}$), the import of primary energy sources from the grid network ($IPES_{egt}$) and the rate of consumption of the primary energy sources (ESU_{gt}):

$$Atot_{gt} = \sum_e (A0_{egt} + IPES_{egt}) - ESU_{gt} \quad \forall e, t, g; g \neq g'$$

ESU_{gt} is equal to the product of gamma (γ_{pj}) the rate of utilization of primary energy source by the daily hydrogen production rate for each production plant type p and size j (PR_{pjigt}).
 $ESU_{gt} = \sum_{pji} (\gamma_{pj} \times PR_{pjigt}) \quad \forall g, t; g \neq g'$

The Energy Availability EA_{gte} is the product of the energy power source capacity in each grid during period t and for each energy source (ESP_{gte}) multiplied by the number of hours per month (Mh_t) and the capacity factor for each energy source (CF_{et}) divided by the total number of days per month (dM_t).

$$EA_{gte} = \frac{\sum (ESP_{gte} \times Mh_t \times CF_{et})}{dM_t} \quad \forall e, g, t; g \neq g'$$

The installation of production (IP_{pjigt}), storage (IS_{sjigt}), and distribution (IFS_{fsjigt}) units only occurs at the first period t .

$$IP_{pjigt} = 0, \forall t \neq 1; \quad IS_{sjigt} = 0, \forall t \neq 1; \quad IFS_{fsjigt} = 0, \forall t \neq 1;$$

The minimum and maximum daily production capacity ($PCAP^{min}_{p,i,j}$) and ($PCAP^{max}_{p,i,j}$) is respectively calculated by considering the lower and upper daily capacity factor of electrolyzer type p and size j given in hours per month (respectively $ELCF^{min}_{p,j,i}$ and $ELCF^{max}_{p,j,i}$) and the assigned power of each given electrolyser in kW ($PElec_{p,i,j}$). Nd is the number of days per month and gamma (γ_{pj}) is the rate of utilization of primary energy source (in kWh/kg).

$$PCap_{pij}^{min} = \frac{\left(\frac{ELCF_{pij}^{min}}{Nd}\right) \times PElec_{pij}}{\gamma_{pj}}; \quad PCap_{pij}^{max} = \frac{\left(\frac{ELCF_{pij}^{max}}{Nd}\right) \times PElec_{pij}}{\gamma_{pj}}; \quad \forall p, i, j$$

2.3.2 Objective functions

The monthly average cost of hydrogen ($Mcost_t$) is determined by summing the capital costs of all units over the lifetime of the units (lf) with the maintenance and operating costs ($Opex_t$) of the units over the sum of the total demand for each grid g of product type i in time period t (DT_{igt}) multiplied by the number of days per month (dM_t).

$$Mcost_t = ((Capex_t \div lf) + Opex_t) / \sum_{ig} (dM_t \times DT_{igt}); \quad \forall e, g, t; g \neq g'$$

The total global warming potential is determined from the total daily greenhouse gas (GHG) emissions of the production, storage and transport units. Similarly, the relative

risk of hydrogen activities is determined by risk ratings calculated based on a risk index method including the risk of production, storage and transport (Kim and Moon, 2008).

2.4. Study case

2.4.1 Demand evaluation

Corsica is a French region that is greatly impacted by tourist activities. According to the annual evolution of the fuel consumption in the island, 2.5% of the equivalent energy could be substituted by hydrogen fuel in 2030 (“Le Schéma Régional Climat, Air, Energie (SRCAE) de Corse,” n.d.). This corresponds to a daily hydrogen demand ranging from 4.3 t/day in January to 8.1 t/day in August. The fluctuating hydrogen demand expressed in energetic values is presented in figure 1.

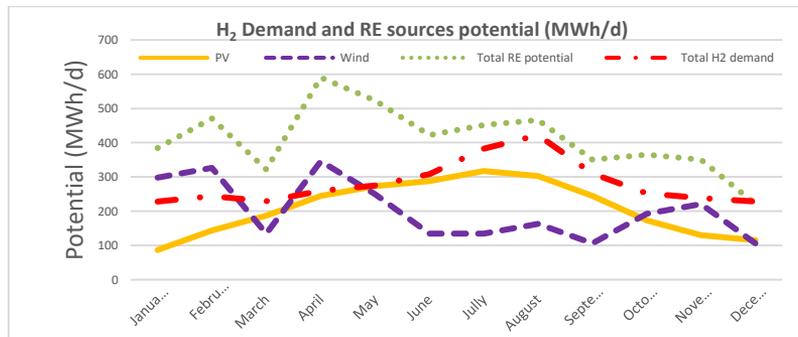


Figure 1: H₂ Demand and RE source potential

2.4.2 Techno-economic parameters

Table 1 presents the minimum and maximum values of the most significant parameters for production, storage, transport and distribution units.

Min-Max	Production unit	Compressor unit	Storage unit	Transport unit	Distri. unit
Capacity	300-5000 (kW)	126 (kg/h)	50-30000 (kg)	670 (kg)	20-1300(kg/d)
CAPEX	1038-3500 (€/kW)	635 (€/kW)	500 (€/kg)	746 (€/kg)	410-1480 (k€)
O&M	0,11-0,20 (€/kg)	0,007 (€/kg)	0,006-0,02 (€/kg)	trip dependent	0,15-0,39 (€/kg)
Efficiency	37,80-52 (kWh/kg)	2,66 (kWh/kg)	-	-	-

Tableau 1: Techno economic data of the production, storage, transport and distribution units

Two types of electrolysis technologies are used, i.e., Alkaline (AE) and Proton Exchange Membrane (PEM) electrolyzers, seen as the most suitable ones for the supply chain in Corsica as they are mature technologies and do not require large heat sources to operate.

2.4.3 Energy sources (cost and potential)

The cost projections used in this work assume that renewable energies will cover 87% of the French electricity demand in 2050 and that 70% of the French nuclear park will be kept. This hypothesis projects the cost of electricity produced by PV and wind power at respectively 16 €/MWh and 35 €/MWh in 2030.

The power generation capacity of the renewable energy sources (PV and wind respectively 60 MW and 40 MW) allocated to the production of hydrogen has been determined according to the available resources on the island and the capacity factor of each technology. Figure 1 shows the forecasted production of these sources for year 2030.

3. Results and discussions

The territory is divided into 9 grids for this study. The results of the simulations show that the model has treated a total of 134308 single equations, 53861 single variables and 24192 discrete variables. The single optimization run led to a maximum CPU time 22s for an instance with CPLEX solver provided with Gams 28.2 (using 7 threads of an Intel Core i5 computer). Figure 2 shows the distribution of the different types of electrolyzers selected and the value of each objective function for each optimization case. The number, size and technology chosen for each configuration is different depending on the objective function used. Large production, storage and distribution units are the most used to reduce the operating cost by mutualizing the maintenance and operation costs. The average production cost obtained is the lowest 12.19 €/kg against 28.03 €/kg and 26 €/kg for the GHG and risk criteria respectively (i.e. expressed as a monthly average cost of production that differs from the levelized cost of hydrogen (LCOH) which takes into account the life cycle of the plant and the discount rate). This cost evolves according to the monthly production rate of hydrogen and the availability of energy sources, reaching its lowest value in August, i.e. 8 €/kg.

A total of 6 transport units are found with cost optimization and one less for the other criteria. The transport from zone 7 to 1 has been replaced by a production unit directly located in zone 1.

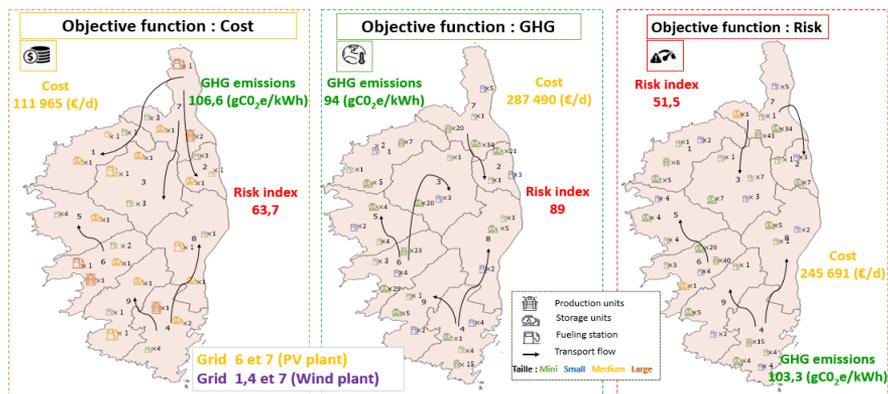


Figure 2: Hydrogen supply chain configuration

The objective based on minimizing GHG emissions leads to a more decentralized configuration with much smaller equipment and more efficient alkaline-type electrolyzers. The amount of GHG emissions is 94 gCO₂e/kWh and although logically most favorable compared to the other criteria, the gap is not significantly high compared to the result obtained with the cost criterion since the technologies considered for hydrogen production involve a green energy mix. Finally, when the optimization criterion chosen is to reduce the total risk index, almost all the equipment chosen are of minimal size because they present less risk. The risk index at this level is 51.5 against 67.3 and 89 for the cost and GHG emission criteria respectively.

According to the season, the distribution of hydrogen can vary. In winter, the distribution rate is at a minimum (670 kg/truck) while in summer it reaches almost 1t/day with no additional transport units. Although the demand in zone 7 is the highest, it provides 56% of the hydrogen supply due to its high production potential. Figure 3 shows the distribution rate for the cost criteria in January and August.

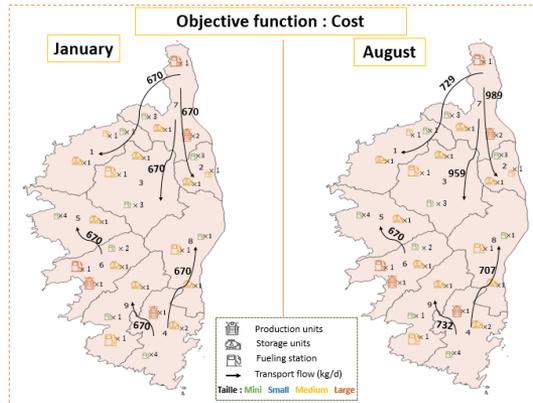


Figure 3: hydrogen distribution during winter and summer

Since photovoltaic energy is the cheapest, it is used up to 74% for the production of hydrogen in order to obtain the most affordable costs. However, the use of photovoltaic energy is still higher than wind energy even for the two other criteria (56% and 65% for the GHG and risk criteria respectively) since the availability of solar energy is higher during periods of high consumption. Photovoltaic power plants are only available in grids 6 and 7 and wind power plants only in grids 1, 4 and 7.

4. Conclusion and perspectives

This paper presented a methodological framework for the design and operation of HSC in Corsica island considering monthly variations in demand. The single objective optimizations for the case study show a decentralized distribution of production, storage and distribution units across the territory. The production cost is still quite high compared to large-scale production chain, but this approach can be used as a framework to support authorities to invest in this field in isolated areas. The next step of this work is to look for the optimal configuration for the three optimization criteria taken simultaneously (multi-objective optimization). This approach can be used to explore appropriate incentive mechanisms to kick-start the hydrogen economy in an isolated territory since the hydrogen cost still remains high even when cost minimization is considered. This optimization approach can also be extended by integrating the geographical and topographical constraints of Corsica.

References

- Agnolucci, P., Akgul, O., McDowall, W., Papageorgiou, L.G., 2013. The importance of economies of scale, transport costs and demand patterns in optimising hydrogen fuelling infrastructure: An exploration with SHIPMod (Spatial hydrogen infrastructure planning model). *International Journal of Hydrogen Energy* 38, 11189–11201. <https://doi.org/10.1016/j.ijhydene.2013.06.071>
- Kim, J., Lee, Y., Moon, I., 2008. Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy* 33, 4715–4729. <https://doi.org/10.1016/j.ijhydene.2008.06.007>
- Kim, J., Moon, I., 2008. Strategic design of hydrogen infrastructure considering cost and safety using multiobjective optimization. *International Journal of Hydrogen Energy* 33, 5887–5896. <https://doi.org/10.1016/j.ijhydene.2008.07.028>
- Krajačić, G., Martins, R., Busuttill, A., Duić, N., da Graça Carvalho, M., 2008. Hydrogen as an energy vector in the islands' energy supply. *International Journal of Hydrogen Energy* 33, 1091–1103. <https://doi.org/10.1016/j.ijhydene.2007.12.025>
- Lamas, J., 2016. Energy supply in isolated areas: an outline of the current situation and the potential of hydrogen technologies for distributed power generation. *Global Resource Management* 2, 31–53.
- Le Schéma Régional Climat, Air, Energie (SRCAE) de Corse [WWW Document], n.d. . Agence d'aménagement durable, d'urbanisme et d'énergie de la Corse. URL https://www.aue.corsica/Le-Schema-Regional-Climat-Air-Energie-SRCAE-de-Corse_a31.html (accessed 10.26.21).
- Seo, S.-K., Yun, D.-Y., Lee, C.-J., 2020. Design and optimization of a hydrogen supply chain using a centralized storage model. *Applied Energy* 262, 114452. <https://doi.org/10.1016/j.apenergy.2019.114452>
- Sofia De Leon Almaraz, 2014. Multi-Objective Optimisation of A Hydrogen Supply Chain. Institut National Polytechnique de Toulouse (INP Toulouse), INP Toulouse.
- Won, W., Kwon, H., Han, J.-H., Kim, J., 2017. Design and operation of renewable energy sources based hydrogen supply system: Technology integration and optimization. *Renewable Energy* 103, 226–238. <https://doi.org/10.1016/j.renene.2016.11.038>

Optimal configuration of a biodiesel production network using oil from black soldier fly larvae

Dulce María Aguilar-Murguía,^a Sergio Iván Martínez-Guido,^b Juan Fernando García-Trejo,^b Salvador Hernández,^c Claudia Gutiérrez-Antonio^b

^a*Universidad Michoacana de San Nicolás de Hidalgo, C. de Santiago Tapia 403 Centro, Morelia 58030, Mexico*

^b*Universidad Autónoma de Querétaro, Cerro de las Campanas s/n Las Campanas, Querétaro 76010, Mexico*

^c*Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, Mexico*

Abstract

Before the pandemics, society was facing challenges regarding the climate change problem as well as the decline in oil well's production. Now, the economic recovery, which also must be sustainable, represents a bigger challenge. In this context, the development of sustainable biofuels is a key area to accelerate economic recovery. In particular, biodiesel is one of the most produced biofuels worldwide. Nonetheless, all the existent plants are based on vegetable or animal oils (edible, non-edible, and waste); however, these bioenergy sources have limited availability, especially considering that oils derived from cultivated crops compete with lands with other edible crops, affecting food safety. Therefore, the current efforts are focused on the search for other oil sources. Particularly, biodiesel production using oil derived from black soldier fly larvae has been proposed (Jung et al., 2022), since the cultivation of this insect requires reduced spaces of non-fertile lands, and its life cycle is 24 days; thus, this oil has advantages as bioenergy source. Additionally, these insects can be fed with waste (as animal manure, food and agricultural waste, among others); these residues are transformed into biomass with high (35-40%) content of lipids (Feng et al., 2018). Therefore, in this work the optimal configuration for a biodiesel production network is proposed, considering as raw material the oil derived from black soldier fly larvae. Hence, a mathematical model was proposed where all the involved unit operations required for the biodiesel production supply chain were included; for this, a B10 mixture was considered to satisfy the Mexican's demand. As objective functions were considered the CO₂ minimization, and the profit maximization. Results show that is possible to satisfy the current total diesel demand in Mexico using a B10 mixture, generating a national benefit of 3.529E9 USD/y, which represent 0.27% of increase of national Growth Domestic Product; in addition, 90% fewer emissions are released, in comparison with those derived from conventional diesel use.

Keywords: Biofuels, optimization, supply chain, waste management, circular economy.

1. Introduction

Globally energy security and climate change mitigation are the two main drivers of the transformation of the current energy system; a transition from an economy based on fossil fuels to one based on renewable sources, which allows to sustain current economic activities and reduced net carbon emissions. Particularly, transport sector consumes 30% of the total primary energy, being 98% of this consumption obtained from fossil fuels, which is traduced into 8.5 Gt of CO₂ emissions. According to Net Zero by 2050 report

(EIA, 2021), to achieve a de-carbonization in the transport sector is possible, if new policies focused in promote modal shifts and energetic efficiency operations are implemented; in addition to an energy source transition, including electric vehicles and biofuels implementation. Thus, the substitution of fossil fuels, in transport sector, by alternative sources such as solar, wind, biomass has received the greatest attention, given the growing demand and the problems associated with this need (Im-orb et al., 2021).

Particularly, biodiesel is a biofuel alternative to conventional diesel consumption; this is obtained from renewable sources, such as vegetable oils and animal fats through a transesterification reaction with methanol in presence of an acid or basic catalyst. However, nowadays the biodiesel production is mainly obtained from first and second-generation biomass, which results in disadvantages in economic, competitiveness, and use land terms. Specifically, in first-generation biodiesel production, approximately 70% of the total cost is generated by raw materials production, given low profitability. Therefore, new raw materials to reduce costs and avoid competition with food supply are needed; some authors have proposed the use of oils obtained from insects, particularly the one extracted from the black soldier fly larvae (*Hermetia illucens*). In this way, Leong et al. (2015) and Nguyen et al. (2018) proposed that there are many advantages linked to the use of black soldier fly larval oil (BSFO), due to its high-fat content, high reproductive speed, and short life cycles; coupled with the ability to consume organic waste or residues without any application of recovery, being this the main sustainable advantage over other raw materials.

Hence, in this work the supply chain optimization of biodiesel production using BSFO as raw material is proposed, through a mathematical model approach. Current conventional diesel demand in Mexico was used as a case study. In addition, organic residues generated in each State in Mexico are considered as food sources for the black soldier fly larvae production; in addition, this biological treatment contributes to a sustainable management residues strategy.

2. Problem statement

According to the National Biomass Atlas (ANBIO, 2020), Mexico generates 278 million tonnes of organic waste per year, with an estimated energy potential of 2,980 PJ. These kind of residues in developing countries are burned, or dumped into a final disposal site, contributing with negative environmental consequences (Pradhan et al., 2019). On the other hand, the Secretariat of Energy (SENER, 2018) expected a diesel national demand of 405 Mbd in 2021, with annual growth of 1.57%; 30% of the total diesel demand is produced in Mexico, while the 70% is imported. In this way, Mexico is the 13th emitting country of CO_{2e}, with almost 737 million tonnes; 64% of these emissions are given by the energy sector, while 18.5% is released by transport sector, 7% generated by the residues management and 3.9% by the final disposition sites (INECC, 2019). Consequently, in the recent COP26, Mexico committed to reduce 25% of its greenhouse gas emissions, in addition to decrease its emissions from the industry sector through the generation of 35% of clean energy by 2024 and 43% by 2030; these actions enrolled to achieve the objective of keeping the temperature increase below 1.5 °C. Thus, the use of new renewable energy sources that do not compromise food security, ensure economic development, and help achieve the commitments agreed is required; thus, the use of oil from this insect is a promissory raw material for biodiesel production.

3. Methodology

The biodiesel production network in Mexico described in Figure 1 is proposed. As first step, the places (being the 32 states of Mexico) where organic wastes are collected (harvested) are considered. Particularly, only the food loss waste (approximately 1,840,853 tonne/y) are taken into consideration; therefore, if the residues flux is representative in one site, then a black soldier fly biotransformation plant (BP) is installed; this plant will serve as waste management biological center. Then, in each installed BP all the oil from the cultivated larvae is extracted; with the collected oil in each site is produced pure biodiesel through the transesterification process. The final consumer of biodiesel is the Mexican transport sector; in order to supply this biofuel, a B10 mixture is considered to avoid additional costs for engine modifications. In Figure 1, sub-indexes i , k , j , and l have the possibility to take values from 1-32, while s only can take values from 1-6 (considering current conventional refineries).

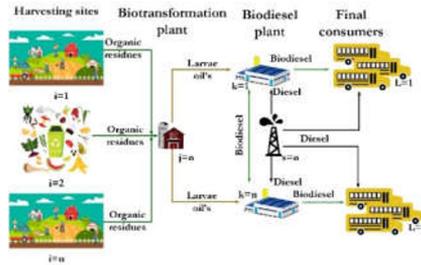


Figure 1. Biodiesel supply chain

3.1 Mathematical model

The proposed approach considers the current situation of organic residues in Mexico and its diesel demand. The organic residues flux in each site ($FR_{i,t}^O$) has the possibility to be dumped in a final disposal site ($RF_{i,j,t}^{s-dump}$) or to be used as feed in the black soldier fly larvae production plants ($RFS_{i,j,t}^{s-oil}$); the sum of both fluxes are equal to the total flux of organic waste produced in each site (i) per week (t), as equation 1 describes. In equation 2, the flux of oil ($OF_{j,t}^{Ext}$), extracted from each tonne of larvae produced (β^{oil}), using the organic residues received as feed ($RFR_{j,t}^O$) in each installed black soldier fly production plant, is calculated.

$$FR_{i,t}^O = \sum_j RFS_{i,j,t}^{s-oil} + \sum_j RF_{i,j,t}^{s-dump}, \forall i \in I, t \in T \quad \text{Eq. 1}$$

$$OF_{j,t}^{Ext} = RFR_{j,t}^O \cdot \beta^{oil}, \forall j \in J, t \in T \quad \text{Eq. 2}$$

Afterward, the collected flux of larvae oil is sent ($OF_{j,k,t}^{sent}$) to the transesterification plant (k) as equation 3 shown; while in equation 4, the biodiesel flux produced per week ($FP_{k,t}^{Bio}$) in each transesterification plant is calculated through the multiplication of the received larvae oil flux ($LO_{k,t}^{rece}$) per a respective conversion factor (μ), tonne of pure biodiesel produced per tonne of larvae oil.

$$LO_{k,t}^{rece} = \sum_j OF_{j,k,t}^{sent}, \forall k \in K, t \in T \quad \text{Eq. 3}$$

$$FP_{k,t}^{Bio} = LO_{k,t}^{rece} \cdot \mu, \forall k \in K, t \in T \quad \text{Eq. 4}$$

Thus, with the pure biodiesel produced in each plant is possible to obtain the B10 mixture (10% of pure biodiesel [$FP_{k,t}^{Bio}$] and 90% of conventional diesel [$FD_{k,t}^{Conv}$]). The expected flux of B10 ($Bio_{k,t}^{Mix}$), in tonnes in each installed plant, is calculated with Equation 5. Then, the B10 flux transported ($Bio_{k,l,t}^{Mix-sent}$) to the consumption sites (l) is computed with Equation 6. Current fuel demand in each site ($FD_{l,t}^{Fuel}$) is lower to sum of available B10 flux ($Bio_{l,t}^{Mix-avai}$) plus conventional diesel flux ($CD_{l,t}^{use}$), as equation 7 describes. Conventional diesel consumption is considered when is not possible to satisfy the fuel demand with B10.

$$Bio_{k,t}^{Mix} = FP_{k,t}^{Bio} + FD_{k,t}^{Conv}, \forall k \in K, t \in T \quad \text{Eq. 5}$$

$$Bio_{l,t}^{Mix-avai} = \sum_k Bio_{k,l,t}^{Mix-sent}, \forall l \in L, t \in T \quad \text{Eq. 6}$$

$$FD_{l,t}^{Fuel} \leq Bio_{l,t}^{Mix-avai} + CD_{l,t}^{use} \quad \text{Eq. 7}$$

Total cost (TC) is calculated through equation 8, in which are summed all the generated transportation cost (C^{Tra}), the cost by the conventional diesel consumption (C^{usoDC}), the cost by the waste management (C^{wm}), and the production of B10 (C^{PB10}). Waste management and B10 production cost include fixed and variables cost for each installed plant. Similarly, the CO₂ emissions (TE) include those due to transportation (E^{Tra}), the use of conventional diesel (E^{usoDC}), the use of B10 (E^{UsoB10}), B10 production (E^{PB10}), and those released by the waste management strategies (E^{wm}).

$$TC = C^{Tra} + C^{usoDC} + C^{wm} + C^{PB10} \quad \text{Eq. 8}$$

$$TE = E^{Tra} + E^{usoDC} + E^{UsoB10} + E^{wm} + E^{PB10} \quad \text{Eq. 9}$$

Equation 10 shows the profit balance (*Profit*), where total costs is subtracted to generated sales by B10 (*Sales*), plus the economic gain by food losses waste management ($Pric^{wast}$). This last earning is considered if the residues are sent to biotransformation plant; hence, the current cost of the final disposal site is removed. Lastly, a multiobjective function was included (*O.F.*) for the profit maximization (*Max profit*) and emissions minimization (*Min CO₂*), Equation 11.

$$Profit = Sales + Pric^{wast} - TC \quad \text{Eq. 10}$$

$$O.F. = Max\ profit; Min\ CO_2 \quad \text{Eq. 11}$$

4. Results

The mathematical optimization model was codified in GAMS® platform, and it includes 47,558 constrains, 273,734 single, and 192 discrete variables. This model was solved as a Mixed Integer Linear Programming problem using the CPLEX solver. The test was carried out in an Intel® Core-i7, with a CPU 3.60 GHz and 8 GB RAM, with an execution time of 1.093 s/assay. The results show that is possible to produce 561,976 tonnes of pure biodiesel per year, this production represents 5.2% of the total diesel demand in Mexico; however, nowadays is impossible the use of pure biodiesel in the current engines available in Mexican transport sector. Thus, 6,811,732 tonne/y of B10 can be produced and used directly in the transport sector without modify the vehicle's engines. Thus, approximately 58% of current diesel demand in Mexico is satisfied with B10, and the rest with conventional diesel. To achieve this amount of B10 production is required the installation of biotransformation plants and biodiesel plants in each State of Mexico as Figure 2 shows. Note that biotransformation plants in white color represent those that received organic wastes from CDMX; this means the plants of 24 States, from 32, receive residues from CDMX, the biggest waste generator. Particularly, CDMX contributes with almost 70% of the food loss waste produced in Mexico, where 6% of the Mexican population

lives. On the other hand, biodiesel plants in yellow color (located in 11 States) are those with the capacity to share B10 with CDMX, which is the city with the bigger flux of B10, due its bigger population. Additionally, arrows in red and black color indicate the fluxes of residues and B10 transported, respectively.



Figure 2. Mexican biodiesel supply chain distribution.

According to the CO₂ emissions quantification performed in GREET software (2021), the raw material generation used in biodiesel production, crops production and residual materials are 39% and 8% (respectively) of the total CO₂ emissions generated by each tonne biodiesel produced. In this way, each tonne of food loss transformed by BSF releases 0.88 kg CO₂ (Pang et al., 2020); therefore, as is shown in Table 1 each tonne of BSFO produce approximately 12 kg of CO₂. This value is 98.8% lower that the emissions generated by the crops production, and 83.3% lower in comparison with other residual products. On the other hand, if all the Mexican food losses waste are sent to a final disposal site are generated approximately 11,497,200 tonne of CO₂, based on the values of emissions indicated by Scukling et al. (2021). However, if these residues are bio-transformed into biodiesel using the BSF, then emissions are reduced in 86%. Table 1 also shows the CO₂ emissions generated by each tonne of biodiesel produced; particularly, the production of 1 tonne of biodiesel from BSFO is 41% and 78% lower in terms of emissions respect to the use of waste cooking oil and soybean oil, respectively.

Table 1. CO₂ emissions comparison

		Raw material production	Total emissions
	Type of raw material	(CO ₂ kg/raw material tonne)	(CO ₂ kg/tonne biodiesel)
Residual	Lard+chicken fat	72	956 ⁰
	Waste cooking oil	72	956 ⁰
	Tallow	72	956 ⁰
Annual	Castor bean	1,031	2,654 ⁰
Annual Mechanized	Soybean	1,031	2,654 ⁰
Perennials	Palm	1,031	2,654 ⁰
Black soldier fly (BSF)	BSFO	12	562 ^c

Note: ⁰ Value reported by Costa et al., 2013; ^ε calculated value using GREET.

5. Conclusion

Food losses waste can be transformed to biodiesel through its processing with the cultivation of black soldier fly, oil extraction from larvae, and its transesterification. This approach promotes a circular economy based on residues management. Moreover, it is possible to satisfy 58% of the national demand using B10 mixtures; at the same time, the food loss waste is processed, reducing the environmental impact, and new green business can be created. To achieve a total de-carbonization of the transport sector new policies, alternative energy sources, as well as higher investment are required.

References

- A. Costa, L.B. Oliveira, M.P.E. Lins, A.C.M. Silva, A.O. Pereira-Jr, L.P. Rosa, 2013, Sustainability analysis of biodiesel production: A review on different resources in Brazil, *Renewable and Sustainable Energy Reviews*, 27, 407-412
- ANBIO-National Atlas of Biomass, 2018, Atlas Nacional de Biomasa. Available on: <https://dgel.energia.gob.mx/ANBIO/>
- Energy Information Administration “EIA”. 2021. Net zero by 2050: A roadmap for the global energy sector. Available on: <https://www.iea.org/reports/net-zero-by-2050>
- Greenhouse Gases, Regulated Emissions, and Energy Use in Technologies Model by Argonne National Laboratory “GREET” (2021), available on: <https://greet.es.anl.gov/>
- H.C. Nguyen, S.H. Liang, S.S. Chen, C.H. Su, J.H. Lin, C.C. Chien, 2018, Enzymatic production of biodiesel from insect fat using methyl acetate as an acyl acceptor: Optimization by using response surface methodology, *Energy Conversion and Management*, 158, 168-175
- J. Suckling, A. Druckman, R. Small, F. Cecelja, M. Bussemaker, Supply chain optimization and analysis of *Hermetia illucens* (black soldier fly) bioconversion of surplus foodstuffs, *Journal of Cleaner Production*, 321, 128711
- K. Im-orb, A. Arpornwichanop, L. Simasatitkul, 2021, Process intensification approach for design and optimization of biodiesel production from palm fatty acid distillate, *Biotechnology Reports*, 30, e00622
- National Institute of Ecology and Climate Change “INECC”, 2019, First Biennial Update to the United Nations Framework Convention on Climate Change, available on: <https://www.gob.mx/inecc/articulos/comunicaciones-nacionales?idiom=es>
- P. Pradhan, P. Gadkari, A. Arora, S.M. Mahajani, 2019, Economic feasibility of agro waste pelletization as an energy option in rural India, *Energy Procedia*, 158, 3405-3410
- S. Jung, J.M. Jung, Y.F. Tsang, A. Bhatnagar, W.H. Chen, K.Y. A. Lin, E.E. Kwon, 2022, Biodiesel production from black soldier fly larvae derived from food waste by non-catalytic transesterification, *Energy*, 238, 121700
- S.W. Leong, S.R. Mohamed-Kutty, A. Malakahmad, C.K. Tan, 2015. Feasibility study of biodiesel production using lipids of *Hermetia illucens* larva fed with organic waste, *Waste Management*, 47, 84-90
- Secretariat of Energy “SENER”, 2018, Diagnosis of the oil industry in Mexico, available on: https://www.gob.mx/cms/uploads/attachment/file/177622/Prospectiva_de_Energ_as_Renovables_2016-2030.pdf
- W. Feng, L. Qian, W. Wang, T. Wang, Z. Deng, F. Yang, J. Xiong, C. Wang, 2018, Exploring the potential of lipids from black soldier fly: New paradigm for biodiesel production (II) Extraction kinetics and thermodynamic, *Renewable Energy*, 119, 12-18
- W. Pang, D. Hou, J. Chen, E.E. Nowar, Z. Li, R. H., J.K. Tomberlin, Z. Yu, Q. Li, S. Wang, 2020, Reducing greenhouse gas emissions and enhancing carbon and nitrogen conversion in food wastes by the black soldier fly, *Journal of Environmental Management*, 260, 110066

Optimization of Sink Locations in Carbon Integration Networks

Sabla Y. Alnouri ^a, Ilkan Sarigol ^b, Dhabia Al-Mohannadi ^c, Hadi Jaber ^d

^a*Gas Processing Centre, College of Engineering, Qatar University, Doha, Qatar*

^b*Kranz Wolfe Associates, Oud-Strijdersstraat 15, 3020 Herent, Vlaams-Brabant, Belgium*

^c*Chemical Engineering Program, Texas A&M University at Qatar*

^d*College of Engineering, Abu Dhabi University, Abu Dhabi, United Arab Emirates*

Abstract

The detrimental environmental impacts of climate change that are a result of high atmospheric CO₂ concentrations have prompted global efforts to limit the continuous increase in greenhouse gas (GHG) emissions. Many are now relying on the deployment of various carbon capture, utilization and storage (CCUS) methods which have been found reliable for reducing CO₂ levels in the atmosphere. Various industries are shifting towards the decarbonisation of their operations through the use of a combination of various CCUS activities. This could involve the conversion of CO₂ into value-added chemicals, the utilization of CO₂ for enhanced oil recovery (EOR), the injection of CO₂ into geological formations and/or oceans, the biological fixation of CO₂, and other similar activities. Such operations are often referred to as CO₂ “sinks”. It is often crucial for industries to identify which CCUS operations to deploy, especially when faced with many choices, since factors such as the cost of implementation and the sink efficiency play a significant role in the sink selection process. In this work, a mathematical model that helps identify optimal CO₂ sink locations within industrial clusters is developed and utilized. Identifying optimal sink locations is an important factor that needs to be considered as part of a CCUS network planning problem. The proposed mathematical model was found to be very useful for identifying optimal CCUS sink selections, and their respective locations. A total of 22.7% of savings in transportation costs have been realized within the network, when compared to the case of having sink locations prescribed upfront.

Keywords: CCUS, Climate Change, Integration, Optimization, Spatial

1. Introduction

Carbon Capture, Utilization, and Storage (CCUS) technologies aim to remove CO₂ from the atmosphere in a safe and effective manner. When talking about the exact terminology used for each, (1) capture refers to the process of capturing CO₂ from exhaust and/or reformed gases at stationary sources, via standardized technologies that are applied pre- or post- combustion, (2) utilization refers to the use of captured CO₂ for the production of value-added chemicals and other useful industrial products, (3) storage (also known as sequestration) refers to the process of transferring captured CO₂ into geological formations and/or oceans, in which it is stored permanently. The investigation efficient CCUS schemes has been the subject of many research papers.

Liang et al. (2009) studied the economic value, as well as the investment characteristics of carbon capture and reuse (CCR) plants in China. Li et al. (2012) developed a qualitative assessment that investigates early opportunities CO₂ capture in advanced gasification plants. Yousefi-Sahzabi et al. (2011) presented a review on the applicability of geo-information technology for the screening and analysis of CO₂ emission sources, transportation and storage possibilities. Wilberforce et al. (2021) presented a thorough investigation of the various technologies that can be harnessed to capture carbon dioxide. Hasan (2015) developed a hierarchical and multi-scale framework for the design CCUS networks at minimum investment, operating and material costs. Al-Mohannadi and Linke (2016a) presented a mathematical model to systematically design low cost carbon integration networks for industrial parks through an integrated analysis of sources, utilization and storage options, while accounting for capture, separation, compression and transmission. Following this, Al-Mohannadi and Linke (2016b) presented an extension to the MINLP that is capable of CO₂ planning network transitions over a time horizon. Zhang et al. (Zhang et al., 2020) developed a superstructure optimization-based framework in the form of a multi-objective mixed integer linear programming (MILP) program, for the deployment of CCUS supply chains. Al-Mohannadi and Linke (2020a) presented a Mixed Integer Nonlinear (MINLP) optimization approach which simultaneously considers natural gas distribution combined with (CCUS), and energy integration. Al-Mohannadi et al. (2020b) then introduced a two-step optimization approach that is capable of evaluating CO₂ reduction policies using a Mixed Integer Linear program (MILP) using a 2-step process. Hetti et al. (Hetti et al., 2020) presented a systematic literature review to critically analyze the feasibility of community level carbon capturing, and explored the prospects of CCSU integration towards transitioning into zero-emission community energy systems. D'Amore et al. (2021) presents a supply chain Mixed Integer Linear Programming method for the optimization of a Europe-wide carbon capture and storage supply chain. Mualim et al. (2021) proposed an improved pinch-analysis based method for carbon capture and storage target and network design, which mainly involves targeting the maximum carbon exchange in the carbon capture and storage system, followed by designing a CCUS network.

Even though much of the previous work has looked into the design of efficient carbon integration and CCUS systems, there exist no models that account for the possibility of having CO₂ utilization sink (or sequestration site) location selections, simultaneously. In other words, the possibility of locating more than one type of sink in the same location, and choosing the best sink out of all that would need to be incorporated in that location (or site) is an important problem, especially if there exists a number of location choices that could potentially be assigned from a selection of different CCUS operations. The proposed extension allows this aspect to be incorporated into the main CCUS model.

2. Methodology

It is often common to use the term “source” to refer to the capture stage of CO₂, while both utilization and storage are often referred to as CO₂ “sinks”. Hence, multiple aspects need to be combined for the successful removal and capture of CO₂, followed by an appropriate selection of utilization and storage options. Moreover, before utilization and storage, captured CO₂ would need to be appropriately conditioned to meet the CO₂ sink requirements. This is often achieved through a combination of treatment, compression,

and pumping operations. Upon this, CO₂ can then be dispatched and transported from the source location at the capture site, to appropriate sink locations. This work introduces the respective sink locations as additional variables that need to be mathematically represented and identified within the model. Each CO₂ sink be assigned one optimal location, which is to be chosen from a set of available locations that are appropriately identified. Figure 1 below represent the sink location problem within the context of source-to-sink CO₂ mapping.

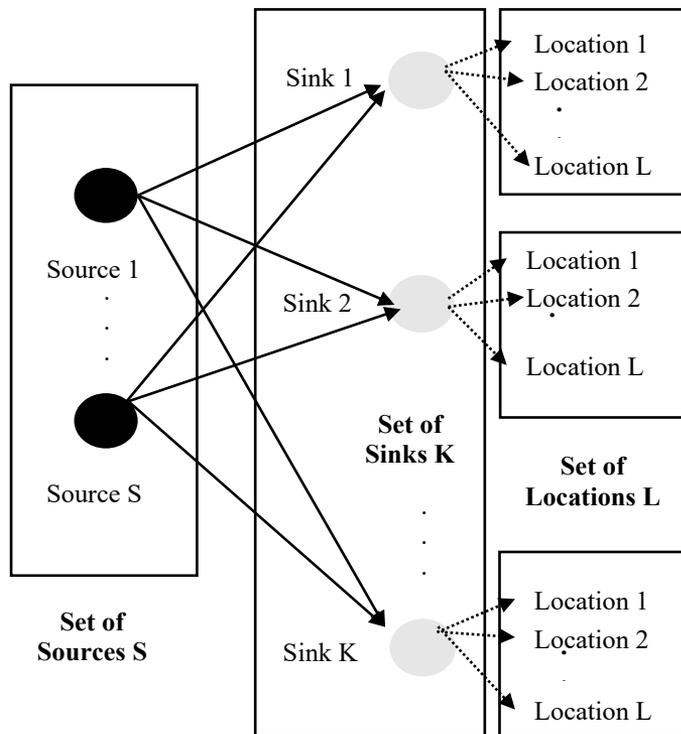


Figure 1. CO₂ Source-Sink mapping with Sink locations

Pipelines are widely used for CO₂ transportation between stationary sources and sinks. However, other modes of transportation may be utilized, depending on the sink type, such as trucks, ships etc. In this work, pipelines were utilized as the primary transportation model. Ultimately, viable CCUS structures must prove to be effective in terms of their ability to reduce atmospheric CO₂ while maintaining its ability to sustain long-term operation. Moreover, identifying effective CCUS schemes based on sink locations must provide economic, as well as environmentally-sustainable solutions. Hence, the cost associated with having a pipeline constructed at each sink location site must be accounted for. Moreover, it is imperative that only one sink site is chosen per source-to-sink allocation, and each sink location in set L must be chosen only once for each sink identified within the CCUS network.

3. Mathematical Formulation

A Mixed Integer Non-Linear Problem (MINLP) has been adopted from Al-Mohannadi and Linke (Al-Mohannadi & Linke, 2016), with the following additional equations that

account for the sink location selection. Even though the overall model is nonlinear, the equations that are needed to account for sink location selections within the problem were all identified to linear, as described below. Superscripts s , k , and l represent the carbon source, carbon sink, and sink locations respectively. S , K and L represent the set of CO₂ sources, CO₂ sinks and sink locations, respectively. T and U represent treated and untreated CO₂ flows, X is a binary variable which takes a value of 1 when source s has a connection with sink k ; and 0 otherwise. Similarly, W is a binary variable which takes the value of 1 when sink k is located at location l , and 0 otherwise:

$$\sum_l W_{kl} = 1 \forall k \in K \quad (1)$$

$$\sum_k W_{kl} = 1 \forall l \in L \quad (2)$$

$$L_{sk} X_{sk} \leq T_{sk} + U_{sk} \forall s \in S, k \in K \quad (3)$$

$$T_{sk} + U_{sk} \leq M_{sk} X_{sk} \forall s \in S, k \in K$$

$$X_{sk}, W_{kl} \in \{0,1\} \forall s \in S, k \in K, l \in L \quad (4)$$

The objective function of this study is set to minimize the total cost of the carbon integration network. The overall cost of the network entails the cost of carbon treatment, compression, transportation and sink processing costs. All carbon-related costs are adapted from Al-Mohannadi and Linke (Al-Mohannadi & Linke, 2016). Moreover, the carbon compression and pumping costs consist of operating and capital costs functions. The central decision variables mainly consist of the carbon source-to-sink flows (treated and untreated) and the respective sink locations. The total cost is obtained by subtracting any obtained revenue, as a result of selling any captured carbon dioxide, from all the previously mentioned costs. Moreover, the model constraints consist of total and component mass balance equations of the different material present throughout the system. Since there are different qualities of sources, an overall mass balance is used to compute actual stream compositions, before they are sent to the sinks. Production and performance constraints were also used, to ensure certain requirements on the amount of material flows, mainly CO₂ capacity constraints, which refer to setting a minimum on the amount of material to be allocated (carbon and ash) or a maximum amount of material that can be accommodated by the available sinks, and composition constraints, which place a restriction on the amount of specific component (carbon) in the flowrate allocated to a specific sink. Moreover, since only one sink site must be chosen per source-to-sink allocation, and no single sink location can be chosen more than once for different sinks. The MINLP combines equations (1-4) above, together with the objective function and mass balance constraints from Al-Mohannadi and Linke (Al-Mohannadi & Linke, 2016). The problem was implemented using What'sBest via MS Excel, and the LINGO solver was used on a laptop with an Intel core i5 processor, 8 GB RAM and a 64-bit Operating System.

4. Case Study

This case study involves the possibility of allocating 4 different types of CO₂ sources, with differing qualities into six different sink possibilities. Table 1 below summarizes the respective flowrate and composition data for each. For each sink, six different location options have been allowed, and the distance data associated with each of those locations are shown in Table 2.

Table 1: Available source and sink data (Al-Mohannadi & Linke, 2016)

Source	Flow MTPD	CO ₂ Composition (%)	Sink	Max.Capacity MTPD	Min.CO ₂ Composition (%)
Ammonia Removal	977	1	Algae	4716.7	6
Steel-Iron	3451	44	Greenhouse	1095.7	94
Power	9385	7	Storage	8847.9	94
Refinery	1092	27	Methanol Sol.	1711.7	99.9
			Urea	1127.1	99.9
			EOR	2913.8	94

Table 2: Sink location data (km)

Source	Location 1	Location 2	Location 3	Location 4	Location 5	Location 6
Ammonia Removal	1.0664	1.7356	0.862	0.9362	0.961	0.9672
Steel-Iron	1.2834	1.9526	1.0432	1.1532	1.178	1.1842
Power	1.7174	1.9446	1.261	1.829	0.5642	0.3162
Refinery	1.5686	1.7958	1.802	1.6802	0.4092	0.5084

In order to demonstrate the proposed approach, a net capture target of 15% (corresponding to a total of 3076 MTPD of captured CO₂) has been imposed on the system. In doing so, only 2 sinks were activated within the network (Methanol and EOR), and the values corresponding to the optimal allocations attained have been provided in Table 3.

Table 3: Summary of optimal source-to-sink allocations and the respective sink locations

	Algae	Greenhouse	Storage	Methanol	Urea	EOR
Ammonia Removal	0	0	0	0	0	977
Steel-Iron	0	0	0	279.1	0	1170.7
Power	0	0	0	0	0	591.3
Refinery	0	0	0	265.4	0	0
Location	-	-	-	Location 5	-	Location 6

While the EOR sink was completely filled to its maximum capacity, the Methanol sink was filled to about 31.8% of its maximum capacity. Moreover, the optimal locations for

each of those 2 sinks has been identified by the solver. Location 5 was assigned for Sink 4 while Location 6 was assigned for Sink 6. The optimal cost of the network was found to be 5,959,719 USD as opposed to a total of 6,013,437 USD when no location optimization is enabled, and each sink is prescribed a location beforehand. Hence, upon enabling the sink location optimization feature, a total of 22.7% of savings in the total transportation costs have been realized within the network, when compared to the case of having the sink locations prescribed upfront. This aspect highlights the importance of sink locations in CCUS networks on the overall transportation costs, and allows for appropriate sink location planning activities, especially within multi-period problems, are studied, and expansion activities are crucial in the system.

5. Conclusions

In this study, a mathematical model that assesses sink location optimization features within carbon integration networks has been introduced. In order to ensure the successful development and implementation of a CCUS network, appropriate planning and location identification of CO₂ sinks is necessary, since significant economic savings in terms of transportation costs can be realised.

References

- A. Yousefi-Sahzabi, K. Sasaki, I. Djamaluddin, H. Yousefi, Y. Sugai, 2011, GIS modeling of CO₂ emission sources and storage possibilities, *Energy Procedia*, 4, 2831-2838.
- d'Amore, F, Romano, M.C, Bezzo, F, 2021. Carbon capture and storage from energy and industrial emission sources: A Europe-wide supply chain optimisation. *Journal of Cleaner Production* 290, 125202.
- D.M, Al-Mohannadi, P, Linke, 2016a, On the systematic carbon integration of industrial parks for climate footprint reduction. *Journal of Cleaner Production*, 112, 4053-4064.
- D.M, Al-Mohannadi, S.Y, Alnouri, S.K, Bishnu, P, Linke, 2016b, Multi-period carbon integration. *Journal of Cleaner Production*, 136, 150-158.
- D.M, Al-Mohannadi, P, Linke, 2020a, Optimal Utilization of Natural Gas in Processing Clusters with Reduced CO₂ Emissions through Material and Energy Integration, *Energy Technology* 8, 1901381.
- D.M, Al-Mohannadi, G, Kwak, P, Linke, 2020b, Identification of optimal transitions towards climate footprint reduction targets using a linear multi-period carbon integration approach, *Computers & Chemical Engineering*, 140, 106907.
- J, Li, X, Liang, J, Gibbins, 2012, Early Opportunity for CO₂ Capture from Gasification Plants in China, *Energy Procedia* 14, 1451-1457.
- M.M.F, Hasan, E.L, First, F, Boukouvala, C.A, Floudas, 2015, A multi-scale framework for CO₂ capture, utilization, and sequestration: CCUS and CCU, *Computers & Chemical Engineering*, 81, 2-21.
- R.K, Hetti, H, Karunathilake, G, Chhipi-Shrestha, R, Sadiq, K, Hewage, 2020, Prospects of integrating carbon capturing into community scale energy systems, *Renewable and Sustainable Energy Reviews* 133, 110193.
- S, Zhang, Y, Zhuang, R, Tao, L, Liu, L, Zhang, J, Du, 2020, Multi-objective optimization for the deployment of carbon capture utilization and storage supply chain considering economic and environmental performance. *Journal of Cleaner Production* 270, 122481.
- T, Wilberforce, A.G, Olabi, E.T, Sayed, K, Elsaid, M.A, Abdelkareem, 2021, Progress in carbon capture technologies, *Science of The Total Environment*, 761, 143203.
- X, Liang, D, Reiner, J, Gibbins, J, Li, 2009, Assessing the value of CO₂ capture ready in new-build pulverised coal-fired power plants in China, *International Journal of Greenhouse Gas Control*, 3, 787-792.

Towards Efficient Bioenergy Systems: Understanding the Role of Soil Sequestration, Supply Chain Design, and Carbon Capture

Caleb H. Geissler^{a,b}, Eric G. O'Neill^{a,b}, Christos T. Maravelias^{a,b,c}

^a*Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08540, USA*

^b*DOE Great Lakes Bioenergy Research Center, USA*

^c*Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ 08540, USA*

maravelias@princeton.edu

Abstract

Bioenergy with carbon capture and storage (BECCS) can produce carbon-negative biofuels from dedicated cellulosic biomass to mitigate global warming; however, the economic and environmental performance of a BECCS system is highly sensitive to carbon incentives and the supply chain (SC) network. Because dedicated bioenergy crops have yet to be planted, there is an opportunity to simultaneously design the SC and choose crop locations that balance the tradeoffs among transportation costs, emissions, biomass yields, and soil carbon sequestration. At the biorefinery, alternative pretreatment and conversion methods lead to different ethanol yields and relative carbon flows among point sources of CO₂ emissions. We present an integrated mixed-integer optimization model for the biofuel SC with a detailed biorefinery model considering CCS that can be used to study the economic and environmental tradeoffs between (i) high resolution SC design and operation, (ii) upstream spatially explicit carbon sequestration in soils by dedicated bioenergy crops, (iii) biorefinery design and operation, and (iv) downstream CCS technology selection and operation. We present a case study with, critically, realistic crop, economic, and environmental data in the US Midwest to explore the effects of key system parameters on the cost and GHG balance of the entire system, and the associated changes in SC and biorefinery design and operation.

Keywords: Sustainability, Supply Chain, Carbon Capture

1. Introduction

An efficiently designed and operated supply chain (SC) from field to product will be required to make cellulosic biofuels competitive with current grain-based fuels and fossil fuels. While a large amount of highly distributed land will need to be converted to biomass to produce an adequate amount of fuel, carefully choosing the location of crop establishment and levels of management (e.g. fertilization) can have a significant impact on crop yields and the greenhouse gas (GHG) balance of the supply system; however, landscape design decisions are typically studied separately from SC and biorefinery design (Field et al., 2018). Crop yields and soil carbon sequestration potentials are field specific and involve complex tradeoffs with downstream SC decisions related to facility location, transportation planning, and inventory management. Optimization modeling at a high spatial resolution can be used to study the above tradeoffs simultaneously to

minimize both the SC costs and GHG emissions. However, the soil carbon sequestration is not the only source of negative emissions. Carbon capture and storage (CCS) technologies installed at the biorefinery have the potential to improve the GHG balance of the system but require additional capital and operational investment.

At the biorefinery, CO₂ is emitted from three different point sources: nearly pure CO₂ from fermentation, ~73 wt.% CO₂ in biogas from anaerobic digestion of wastewater, and ~20 wt.% CO₂ in flue gas from solid residue (primarily lignin) combustion (Humbird et al., 2011). The relative amounts of CO₂ emitted from these point sources varies depending on what pretreatment technology is used. While dilute acid (DA) pretreatment is the most mature, other pretreatments based on ammonia fiber expansion (AFEX) or copper-catalyzed alkaline hydrogen peroxide (AHP) have been developed. In addition to affecting the relative CO₂ emissions from each point source, these pretreatment technologies have different ethanol yields, energy demands, and operating and capital costs. Therefore, a biorefinery model needs to consider different pretreatment technologies and allow for capture from specific point sources independently, which has only recently been considered by Geissler and Maravelias (2021).

Several studies have simultaneously considered SC and CCS optimization for other bioenergy systems (Lainez-Aguirre et al., 2017); however, relaxing assumptions about the availability and location of biomass and modeling a spatially-explicit supply chain is important especially given the opportunity to design the landscape while considering the effect of soil carbon sequestration on the GHG balance. Importantly, modeling the SC and landscape at a high spatial resolution and allowing flexible capture from different point sources of CO₂ at the biorefinery can identify tradeoffs between upstream and downstream designs, operation, costs, and emissions to inform policy decisions. We combine SC and biorefinery design into a flexible integrated model to study the relative economic and environmental impacts under a range of carbon incentives applied to different areas of the system.

In section 2, we present a model for simultaneous optimization of landscape, SC, and biorefinery design. In section 3 we report and discuss the results for a case study in the state of Michigan, USA. Finally, in section 4 we present conclusions and discuss further applications of our research.

2. Supply chain and biorefinery model

The integrated optimization model introduced in this paper consists of three components: a landscape design model, a multi-period logistics network model, and a biorefinery superstructure model. We employ a flexible data workflow to generate realistic high-resolution data for switchgrass grown on marginal lands. First, a geographic raster of marginal land is identified based on a chosen definition (Lark et al., 2020). Next, crop growth is simulated using the SALUS biogeochemical crop model for the chosen lands according to weather, soil, and land use data (Basso and Ritchie, 2015). The resulting biomass yields, soil carbon sequestration potentials, and land areas are processed into a set of fields $f \in \mathbf{F}$ and a set of harvesting sites $j \in \mathbf{J}$ for use in the optimization model. Other model parameters are based on literature values. The model is solved for a representative year broken into time periods $t \in \mathbf{T}$. We use italicized uppercase Latin letters for variables, lower and uppercase Greek letters for parameters, lowercase Latin letters for indices, and uppercase bold Latin letters for sets. The landscape model makes decisions regarding the location of crop establishment, the level of fertilization to apply at each location, and harvest decisions. The associated costs, C^{LAND} , and emissions, GHG^{LAND} , from the landscape model are presented in equations (1) and (2). Emissions

are considered from all on field activities including the annualized impact from crop establishment and soil carbon sequestration.

$$C^{LAND} = \sum_{i,j,t} \lambda_i H_{i,j,t} + \sum_f \rho N_f + \sum_{i,j} \phi_i A_{i,f} \quad (1)$$

$$GHG^{LAND} = \sum_{i,j,t} \Gamma_i^{MG} H_{i,j,t} + \sum_f \Gamma^N F_f + \sum_{i,j} \Gamma_i^{HA} A_{i,f} - \sum_{i,f} GHG_{i,f}^{SOC} \quad (2)$$

Parameter λ_i is the per-Mg cost of harvesting and handling ($H_{i,j,t}$) crop $i \in \mathbf{I}^F$ at harvesting site j . Variable N_f is the amount (kg) of fertilizer applied to field f , and ρ is the cost per-kg of fertilizer. ϕ_i is the annualized per-ha cost of crop establishment and $A_{i,f}$ is the area of field f established with crop i . Emissions are determined similarly with Γ_i^{MG} , Γ^N , and Γ_i^{HA} being the per-Mg crop, per-kg nitrogen fertilizer, and per-ha emissions respectively. Finally, $GHG_{i,f}^{SOC}$ is a decision variable that represents the amount of soil carbon sequestered at field f and is constrained by decision variables related to crop establishment and fertilization. If $GHG_{i,f}^{SOC}$ is positive, the soil at field f sequesters carbon. The logistics model extends work by Ng et al. (2018) to include a more detailed representation of the biorefinery, including CCS (details shown in equations (3)-(6)). The biorefinery superstructure model makes decisions related to the technology selection and interconnection within the facility, the production planning decisions for a representative year, and the capacity of each technology block. A diagram of the biorefinery and the potential interconnections is shown in Figure 1. The SCND model makes decisions for a representative year that include preprocessing depot location, technology, and capacity, biorefinery location, transportation planning, transportation mode, and inventory planning. In the interest of brevity, detailed equations for the SCND model are omitted as a similar logistics model has previously been published (Ng et al., 2018).

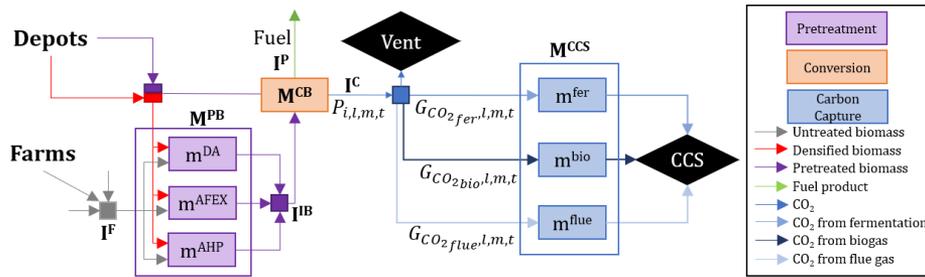


Figure 1. Biorefinery configuration, potential technologies, and potential interconnections.

The pretreatment block is a set of alternatives $\mathbf{M}^{PB} = \{DA, AFEX, AHP\}$. The CCS block $\mathbf{M}^{CCS} = \{FER, FLU, BIO\}$ includes technologies for capturing carbon from each point source (fermentation, flue gas, and biogas respectively). The binary variable $U_{l,m}$ indicates if technology m is installed at biorefinery $l \in \mathbf{L}$, and the capacity $Q_{l,m}$ of each technology is constrained in equation (3) by lower and upper bounds $\epsilon_m^L / \epsilon_m^U$. The variable $P_{i,l,m,t}$ represents the amount of CO_2 , $i \in \mathbf{I}^C$, produced during biomass conversion as different point sources that are available to be consumed by installed CCS technologies. Equations (4)-(6) govern the flow of CO_2 through the biorefinery where $G_{i,l,m,t}^{CCS}$ is the amount of carbon source $i \in \mathbf{I}^C$ consumed by CCS technology $m \in \mathbf{M}^{CCS}$. We introduce two-dimensional set $\mathbf{V} \subset \mathbf{I}^C \times \mathbf{M}^{CCS}$ as the set of valid CCS combinations, so that if element $(i, m) \in \mathbf{V}$, carbon point source i can be captured by CCS technology m .

$$\epsilon_m^L U_{l,m} \leq Q_{l,m} \leq \epsilon_m^U U_{l,m} \quad (3)$$

$$G_{i,l,m,t}^{CCS} \leq \sum_{m' \in \mathbf{M}^{CB}} P_{i,l,m',t} \quad \forall (i, m) \in \mathbf{V}, l, t \quad (4)$$

$$G_{i,l,m,t}^{CCS} \leq Q_{l,m} \quad \forall (i, m) \in \mathbf{V}, l, t \quad (5)$$

$$L_{i',t} = \sum_{l,m \in \mathbf{M}^{PB} \cup \mathbf{M}^{CB}} P_{i',l,m,t}^{IPB-CB} + \sum_{k,m \in \mathbf{M}^{PD}} P_{i',k,m,t}^{ID-PD} + \sum_{l,i \in \mathbf{I}^C, m \in \mathbf{M}^{CCS}} \eta_{i,i',m} G_{i,l,m,t}^{CCS} \quad \forall i' \in \mathbf{I}^B, t \quad (6)$$

The amount of carbon consumed by a CCS technology is bounded by both the amount that is produced in the conversion block (the balance is assumed to be vented if $G_{i,l,m,t}^{CCS} < P_{i,l,m',t}$) and by the installed capacity of appropriate CCS technologies. Finally, $L_{i',t}$, the sale/sequestration of electricity/ CO_2 , $i \in \mathbf{I}^B = \{CSEQ, ELEC\}$, is controlled by equation (6) where $P_{i',l,m,t}^{IPB-CB}$ represents the production of $i' \in \mathbf{I}^B$ from the conversion and pretreatment blocks $m \in \mathbf{M}^{PB} \cup \mathbf{M}^{CB}$, $P_{i',l,m,t}^{ID-PD}$ is the electricity consumption of preprocessing depot technology $m \in \mathbf{M}^{PD}$ at depot $k \in \mathbf{K}$ ($P_{i',l,m,t}^{IPB-CB} = P_{i',l,m,t}^{ID-PD} = 0$ for $i' = CSEQ$) and $\eta_{i,i',m}$ is the conversion coefficient for product i to i' using technology m . We take the convention of variables being ≤ 0 for the consumption of electricity and ≥ 0 for the production. The objective value is given by equation (7) which considers, on an annualized basis, the capital cost, landscape cost, production cost, transportation cost, inventory cost, revenue from byproducts (electricity), and GHG costs respectively.

$$TAC = C^{CAP} + C^{LAND} + C^{PROD} + C^{TRA} + C^{INV} - C^{BY} + C^{GHG} \quad (7)$$

$$C^{GHG} = \sum_t \pi^{SCC} GHG_t^{SC} - \sum_t \pi^{SEQ} L_{CSEQ,t} \quad (8)$$

Equation (8) represents the GHG emissions on a cost basis rather than using an explicit multi-objective approach in order to better compare captured carbon and GHG emissions from the SC (You et al., 2012). To compare the environmental impact of the SC with the impact of CCS at the biorefinery, we introduce two parameters that value CO_2 equivalents (CO_2e) in the objective. The first is the sequestration credit, π^{SEQ} , which applies only to tangible CO_2 sequestered by CCS technologies at the biorefinery. The sequestration credit represents a measurable policy-based tax/credit that can incentivize CCS. The second parameter, π^{SCC} , is the ‘social cost of carbon’ (SCC) which represents the cost impact of all the additional sources of CO_2e (GHG_t^{SC}) that are present from annualized landscape and SC activities, and the impact from replacing grid electricity with excess electricity produced at the biorefinery. Importantly, the SCC also applies to the less readily measurable source of negative emissions, the soil carbon sequestration. By adjusting the two carbon value parameters separately, the integrated model can be used to study the environmental and economic tradeoffs between the landscape, the SC, and the biorefinery. Furthermore, this framework could be easily extended to use multi-objective optimization for further insights into the tradeoff between cost and GHG emissions.

3. Results

We apply the model discussed in section 2 to a case study of Michigan, USA. We use a harvesting site resolution of 8×8 km with 15 potential biorefinery locations, 300 potential depots, and 41,000 potential fields for establishment identified as marginal lands (Lark et al., 2020). The model is implemented in GAMS 36.1 and solved with Gurobi 9.1.2 to a 1% optimality gap on 2.6 GHz Linux cluster machines in under 24h for all instances. The results for a range of values for sequestration credit and SCC are shown in Figure 2. The SC and biorefinery decisions for representative solutions are shown in Figure 3.

and compression, so less electricity is sold when carbon is captured, and therefore a lower total credit is obtained for selling grid electricity. This causes the sequestration credit required to incentivize CCS from a given point source to increase proportionally to the SCC considered. The rate of this increase is determined by the electricity-equivalent energy requirement for capture from that source. It is important to note that the GHG credit for selling grid electricity depends on the electricity production mix, which varies significantly by location and over time. Overall, CCS at the biorefinery has a much greater effect of the GHG balance than SC operation by nearly an order of magnitude.

Applying sequestration credits less than \$25/Mg CO₂ does not incentivize any carbon capture at the biorefinery and therefore has no effect on the total cost. At this level of biofuel production, the SC results in net GHG emissions, so increasing the SCC increases the cost of ethanol production. As the SCC increases, fields with higher soil carbon are chosen, and the rate of increasing total cost diminishes as the SCC increases.

4. Conclusions

We present a model to simultaneously consider high resolution SC design with detailed biorefinery design that includes carbon capture and sequestration. In the case study, we find that CCS at the biorefinery has a much greater impact on the total GHG mitigation of ethanol production than the SC operation. Correspondingly, a sequestration credit has a greater effect on the total cost of ethanol production than a social cost of carbon applied to all other GHG emissions/sequestration. Still, the social cost of carbon impacts what sequestration credit is needed to incentive carbon capture from specific point sources and also effects the optimal SC and landscape configuration. The magnitude of this effect is determined by the energy requirement of capture from that source, and the emissions from local grid electricity, the latter of which varies by location and both of which will change in the future. The proposed model allows for the study of economic and environmental tradeoffs among different stages in the production of biofuels. Its integration of spatially-explicit real-world data makes it a critical tool for future analyses of BECCS systems.

References

- B. Basso, J.T. Ritchie, 2015, Simulating Crop Growth and Biogeochemical Fluxes in Response to Land Management Using the SALUS Model, *The Ecology of Agricultural Landscapes: Long Term Research on the Path to Sustainability*, 252–274.
- J.L. Field, S.G. Evans, E. Marx, M. Easter, P.R. Adler, T. Dinh, B. Willson, K. Paustian, 2018, High-resolution techno-ecological modelling of a bioenergy landscape to identify climate mitigation opportunities in cellulosic ethanol production, *Nat. Energy*, 3, 211–219.
- C.H. Geissler, C.T. Maravelias, 2021, Economic, energetic, and environmental analysis of lignocellulosic biorefineries with carbon capture, *Appl. Energy*, 302, 117539.
- D. Humbird, R. Davis, L. Tao, C. Kinchin, D. Hsu, A. Aden, 2011, Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute acid pretreatment and enzymatic hydrolysis of corn stover, NREL/TP-5100-47764.
- J.M. Lainez-Aguirre, M. Pérez-Fortes, L. Puigjaner, 2017, Economic evaluation of bio-based supply chains with CO₂ capture and utilisation, *Comput. Chem. Eng.*, 102, 213–225.
- T.J. Lark, S.A. Spawn, M. Bougie, H.K. Gibbs, 2020, Cropland expansion in the United States produces marginal yields at high costs to wildlife, *Nat. Commun.*, 11, 1–11.
- R.T.L. Ng, D. Kurniawan, H. Wang, B. Mariska, W. Wu, C.T. Maravelias, 2018, Integrated framework for designing spatially explicit biofuel supply chains, *Appl. Energy*, 216, 116–131.
- F. You, L. Tao, D.J. Graziano, S.W. Snyder, 2012, Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input-output analysis, *AIChE J.*, 58, 1157–1180.

A Multi-Period Planning and Scheduling Strategy for Developing Hydrogen-Based Supply Chains

R. Cory Allen^{a,b}, C. Doga Demirhan^c, Clara F. Heuberger-Austin^d and Efstratios N. Pistikopoulos^{a,b,*}

^a*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, United States*

^b*Texas A&M Energy Institute, Texas A&M University, College Station, TX, United States*

^c*Shell Technology Center, Shell International Exploration and Production Inc., Houston, TX, United States*

^d*Shell Technology Center, Shell Global Solutions International B.V., Amsterdamm, Netherlands
stratos@tamu.edu*

Abstract

Supply chain problems for the production of green electricity and green hydrogen-based fuel for use in utilities and transportation markets are typically represented by large-scale multi-period planning and scheduling models. The solution of such models often presents computational challenges even with state-of-the-art commercial solvers. In this work, we present an effective decomposition based strategy to overcome this challenge featuring a two-stage matheuristic that is embedded in a mixed-integer programming (MIP) solver. The matheuristic first determines a locally optimal solution for the infrastructure planning decisions based on a relaxed solution to the overall problem, and then identifies a locally optimal solution for the operation scheduling decisions. For large-scale supply chain problems, the computational results indicate that the proposed strategy is quite effective.

Keywords: Supply Chain Optimization, Mathematical Programming, Large-Scale Problems

1. Introduction

Due to globalization, liberalization of emerging markets, and population growth, the demands for raw materials and manufactured products are rapidly increasing (Allen et al., 2019). Consequently, supply chains are expanding with alacrity at the global scale and becoming more and more synergistic at the regional scale. At the same time, there has been a societal push towards carbon neutral energy systems that meet their energy requirements via power generated from solar arrays and wind turbines (Shell, 2020; Li et al., 2021).

Unfortunately, the dispatchability of these two types of energy generators are plagued with fluctuations and stochastic intermittencies. To combat these complications, various technologies for storing energy have been put forth, namely, lithium-ion batteries, pumped storage hydro-power, and hydrogen-based dense energy carriers (H-DECs); however, each of these energy storage technologies have their own strengths and weaknesses (Palys et al., 2021). Therefore, there is a need for quantitative frameworks to ensure that the energy systems within supply chains are optimally developed and can operate over a wide range of conditions (Demirhan et al., 2020; He et al., 2021).

In this work, we generalize our previously developed optimization framework for the simultaneous design and operation of energy systems from a single period framework to a multi-period frame-

work (Allen et al., 2021). To address the computational issues that arise in solving multi-period planning and scheduling problems, we have developed a matheuristic that can be integrated into a MIP solver to compute upper bounds at nodes in the branch-and-bound tree – for a detailed review on matheuristics, please see (Archetti and Speranza, 2014). The proposed strategy is agnostic towards the exact MIP formulation of the problem and only assumes: (i) the planning constraints and scheduling constraints can be separated from each other; and (ii) the scheduling constraints can be further divided by their corresponding planning period and representative scenario.

We demonstrate the effectiveness of the framework through the use of a case study wherein we examine the development of a green hydrogen-based supply chain that spans multiple locations over the course of multiple planning periods. We find that the proposed strategy significantly reduces the overall computational time when compared to a state-of-the-art MIP solver.

The outline of this article is as follows: in Section 2 we present the problem statement; in Section 3 we present a MIP formulation for solving multi-period planning and scheduling problems and describe how it can be further decomposed, show the two-stage matheuristic, and outline how the matheuristic can be conjoined with a MIP solver; in Section 4 we present and describe the results to the problem under consideration; and in Section 5 we state some closing remarks.

2. Problem Statement

Consider a central planner who is developing a green hydrogen-based supply chain along the Gulf Coast of Texas and in Southern California over the next thirty years. The central planner would like to produce hydrogen and ammonia via electrolysis and the Haber-Bosch process, respectively, in Houston, Texas and Corpus Christi, Texas. To power these processes the central planner would like to utilize solar and wind farms. To combat the issues in the dispatchability of these generators, the central planner would like to have the ability to store electricity in lithium-ion batteries.

Once the fuels are produced in Texas, they are either transported directly to Long Beach, California or stored to be transported to Long Beach at a later time. After the fuels reach Long Beach, they are: (i) utilized to meet the regions fuel and energy demands as given in Table 1; and/or (ii) stored to be utilized at a later time. It should be noted that electricity can be generated locally via solar farms and/or stored locally in lithium-ion batteries in Long Beach.

Table 1: Energy and fuel demands in Long Beach, California

Material	Planning Period					
	2025	2030	2035	2040	2045	2050
Ammonia [ton/day]	300	350	400	450	500	550
Electricity [GW]	1.0	1.2	1.4	1.6	1.8	2.0
Hydrogen [ton/day]	50	600	900	1100	1300	1500

The cost functions and conversion factors of the processes, storage units, and transportation mechanisms in the problem are taken from a case study in the literature (Allen et al., 2021; Demirhan et al., 2020). The hourly capacity factors for the solar and wind farms are taken from the SIND Toolkit (Hummon et al., 2012) and WIND toolkit (Draxl et al., 2015), respectively.

3. Solution Framework

In this section, we present a generic and compact MIP formulation for solving multi-period planning and scheduling problems. We then describe how the multi-period planning and scheduling problem can be decomposed into a single planning problem and a set of independent scheduling

problems. Finally, we present a two-stage matheuristic that utilizes the decomposed form of the problem to first find a locally optimal integer feasible solution to the planning problem, and then find a locally optimal integer feasible solution to the set of scheduling problems.

3.1. Mathematical Program

The generic and compact MIP formulation of the multi-period planning and scheduling problem is given by Eq. (1). In this formulation there are no underlying assumptions about the structure of planning and scheduling problems except that their corresponding constraint matrices are decomposable. Specifically, it is assumed that planning constraints can be separated from the scheduling constraints and that the scheduling constraints can be further decomposed by their corresponding planning period, $p \in P$, and representative scenario, $s \in S$, where P is the set of planning periods where infrastructure decisions can be made and S is the set of representative scenarios.

$$\min_{x \in X, y \in Y} \{c(x) + d(y) : f(x) = 0; g(x, y) = 0\} \quad (1)$$

The vector, $x \in X$, represents the binary and continuous variables required to describe the planning decisions. The vector, $y \in Y$, represents the binary and continuous variables required to describe the scheduling decisions. The objective functions, $c(\cdot)$ and $d(\cdot)$, capture the capital and operational costs, respectively. The constraints, $f(\cdot) = 0$ and $g(\cdot, \cdot) = 0$, govern the behavior of the planning decisions and scheduling decisions, respectively. It is assumed that a new component of a technology, $t \in T$, can be constructed in each planning period, $p \in P$, at each location, $l \in L$, in the supply chain.

3.2. Decomposition Framework

Figure 1 is an illustration of the block-diagonal decomposable structure of the multi-period planning and scheduling problem. In this illustration, it is assumed that there are three planning periods in which infrastructure decisions can be made – each set of corresponding scheduling problems are represented by a large diagonal block. It is further assumed, that each scheduling problem has been restructured such that it is now composed of three representative scenarios – each of the representative scenarios are illustrated by a small blue diagonal block. From inspection of Figure 1 it is evident that if the planning decisions have been fixed, the problem reduces to nine independent scheduling problems, as given by Eq. (2).

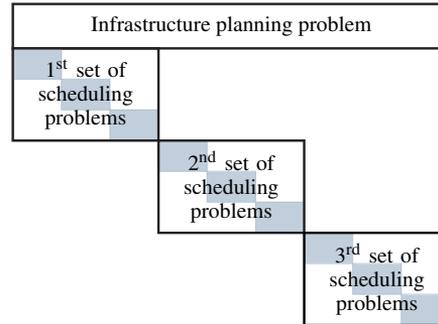


Figure 1: Decomposed formulation of the multi-period planning and scheduling problem

Each of the independent scheduling problems are parameterized by the vector, \bar{x} , which stores a feasible solution for the infrastructure planning decisions. The goal of each scheduling problem is to minimize the operational cost of the system, $d_{p,s}(\cdot)$, during the respective planning period, $p \in P$, and representative scenario, $s \in S$, while simultaneously ensuring that the scheduling constraints, $g_{p,s}(\cdot, \cdot) = 0$, are not violated, where the vector, $y_{s,p} \in Y_{s,p}$, represents the binary and continuous variables required to describe the scheduling decisions as well as any additional slack variables.

$$y_{p,s}^*(\bar{x}) = \arg \min_{y_{p,s} \in Y_{p,s}} \{d_{p,s}(y_{p,s}) : g_{p,s}(\bar{x}, y_{p,s}) = 0\} \quad (2)$$

It should be noted that the size of each scheduling problem is on the order of $|P| \cdot |S|$ times smaller than the original multi-period planning and scheduling problem.

3.3. Two-Stage Matheuristic

In this subsection, we present a two-stage matheuristic, which is given by Algorithm 3.1, for constructing locally optimal integer feasible solutions to multiple-period planning and scheduling problems. In the first stage of the matheuristic, a solution to the infrastructure planning decisions is found by solving a set of independent set covering problems. These set covering problems are parameterized by a relaxed solution to the operational scheduling decisions. In the second stage of the matheuristic, a solution to the operational scheduling decisions is found by solving a set independent scheduling problems. These independent scheduling problems are parameterized by locally optimal solution to the infrastructure planning decisions constructed in the first stage of the matheuristic.

Algorithm 3.1: Two-Stage Matheuristic

input : y – relaxed solution to the operational scheduling decisions
output: locally optimal solution to the planning and scheduling decisions

Function Matheuristic(y):

```

 $\bar{x} \leftarrow \emptyset$ 
foreach  $(t, l) \in T \times L$  do
  |  $\bar{x} \leftarrow \bar{x} \cup x_{t,l}^*(y_{t,l})$ 
end
 $\bar{y} \leftarrow \emptyset$ 
foreach  $(p, s) \in P \times S$  do
  |  $\bar{y} \leftarrow \bar{y} \cup y_{p,s}^*(\bar{x})$ 
end
return  $\bar{x} \cup \bar{y}$ 

```

End Function

The first stage of the matheuristic constructs an integer feasible solution to the infrastructure planning decisions, $\cup_{t \in T} \cup_{l \in L} x_{t,l}^*(\hat{y}_{t,l})$. The parameter, $\hat{y}_{t,l}$, that is passed to the optimization problem, $x_{t,l}^*(\cdot)$, stores a relaxed solution to the scheduling decisions for components of technology, $t \in T$, that are located in location, $l \in L$. In Eq. (3), the goal is to minimize the capital cost, $c_{t,l}(\cdot)$, of constructing the components of a given technology, $t \in T$, over the course of the planning horizon, P , at a specific location, $l \in L$, where the vector, $x_{t,l} \in X_{t,l}$, represents the binary and continuous variables required to describe the infrastructure planning decisions for a specific technology, $t \in T$, at a location, $l \in L$.

$$x_{t,l}^*(\hat{y}_{t,l}) = \arg \min_{x_{t,l} \in X_{t,l}} \{c_{t,l}(x_{t,l}) : f_{t,l}(x_{t,l}) = 0; h_{t,l,p,s}(x_{t,l}, \hat{y}_{t,l}) \geq 0 \forall p \in P, s \in S\} \quad (3)$$

Simultaneously the set covering constraint, $h_{t,l,p,s}(\cdot, \cdot) \geq 0$, ensures that the combined capacities of the components of a given technology are larger than their maximum combined production set points in a relaxed scheduling solution, $\hat{y}_{t,l}$, for each planning period, $p \in P$, and each representative scheduling scenario, $s \in S$. Practically speaking, this constraint ensures that there is always enough production capacity that can operate in parallel for a given technology to cover the production set points given in a relaxed solution to the scheduling problems.

In the second stage of the matheuristic, the locally optimal infrastructure planning decisions, $\bar{x} = \cup_{t \in T} \cup_{l \in L} x_{t,l}^*(\hat{y}_{t,l})$, found in the first stage of the matheuristic are utilized as parameters to generate a locally optimal solution, $\cup_{p \in P} \cup_{s \in S} y_{p,s}^*(\bar{x})$, for the operational scheduling decisions.

3.4. Unification of the Two-Stage Matheuristic and a MIP solver

The matheuristic is integrated into a MIP solver and is called whenever the MIP solver generates a new node in the branch-and-bound tree. Once it is called, a three-stage procedure begins to transpire. In the first stage of the procedure, a relaxed solution to planning and scheduling decisions is accessed from the MIP solver. In the second stage, the solution to these relaxed decisions is passed to the two-stage matheuristic and a locally optimal integer feasible solution is generated. In the final stage, the solution found in the two-stage matheuristic is passed back to the MIP solver as a new upper bound. It should be noted that set covering problems in the first stage of the matheuristic and the scheduling problems in second stage of the matheuristic can be solved in serial or parallel via a MIP solver or a heuristic.

4. Results and Discussion

In this Section, the proposed framework is utilized to solve the aforementioned problem that the central planner is facing. Specifically, the proposed framework is pitted against Gurobi 9.5 utilizing its default parameters through the use of a set of computational experiments that examines how the number of representative scenarios effects the problem’s overall run time (Gurobi Optimization, LLC, 2021). The experiments were performed on a machine with a Intel Xeon W-10885M Processor and 64 GB of RAM.

The MIP formulation of the problem was posed as a mixed-integer quadratic programming (MIQP) problem, please see (Allen et al., 2022) for the constraints and objective function utilized in the model. It should be noted that in the MIQP formulation the bilinear terms arise from the multiplication of the continuous and binary variables in the constraints. Furthermore, we did not linearize these constraints because it was found that Gurobi’s internal linearization toolbox provided a tighter initial linear programming relaxation albeit with a slight increase in the overall computational time.

Table 2 describes the summary statistics for the MIQP formulation for the problem that the central planner is facing. Figure 2 and Figure 3 illustrate the MIP gap of the problem as a function of computational time and the number of representative days for Gurobi when employing its default parameters and for the proposed solution strategy respectively. In both of these bar plot figures, the color of the bar represents the MIP gap at a specific time. It should be highlighted that: (i) when the color of the bar is “black” Gurobi has yet to compute the initial lower bound to the problem; (ii) the y-axis and the color bar scale in both of the figures utilize a \log_{10} scale; and (iii) the problems have been solved a MIP gap of less than 1%.

From inspection of Figure 2 and Figure 3 it is clear that the proposed solution strategy significantly reduces the amount of time that it takes the problem to converge to the desired MIP gap of 1%. Specifically, we find that when the two-stage matheuristic is conjoined with Gurobi, thereby acting as its primal heuristic, the amount of time required to solve the prob-

Table 2: Size of the MIQP formulation for the problem that the central planner is facing

Rep. Days	Binary Variables	Continuous Variables	Constraints
1	97,962	37,482	87,564
2	195,738	74,778	174,756
3	293,514	112,074	261,948
4	391,290	149,370	349,140
5	489,066	186,666	436,332
6	586,842	223,962	523,524
7	684,618	261,258	610,716
8	782,394	298,554	697,908
9	880,170	335,850	785,100
10	977,946	373,146	872,292

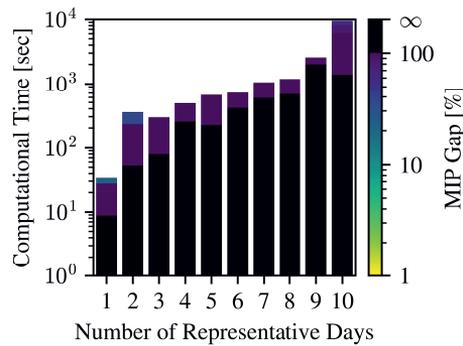


Figure 2: Computational time required to solve the problem utilizing Gurobi employing its default parameters

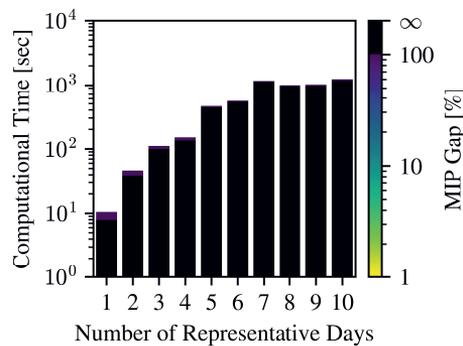


Figure 3: Computational time required to solve the problem utilizing the two-stage matheuristic when it is integrated in Gurobi

lem, once the initial linear programming relaxation was found, was reduced by an average of 93%. This feat is primarily attributed to the strength of the initial linear programming relaxation in which the two-stage matheuristic is utilized to construct a locally optimal solution to the infrastructure planning decisions. This tight relaxation in turn allowed the proposed solution strategy to converge to the desired MIP gap in one iteration for each of the test instances.

5. Conclusion

We have presented a solution strategy for solving large-scale multi-period planning and scheduling problems. We first illustrated that multi-period planning and scheduling problems can be decomposed into a single infrastructure planning and a set of scheduling problems. We then introduced a two-stage matheuristic that allows the problem to be further decomposed into a set of infrastructure planning problems and a set of scheduling problems. We demonstrated the effectiveness of the solution strategy through the use of a case study that examined the development of a national supply chain that produces green electricity and green hydrogen-based fuels for use in utilities and transportation markets. The results show that the proposed strategy significantly reduces the amount of time it takes the problem to converge to the desired MIP gap when compared to Gurobi's default solver. From inspection of the converge plots, Figure 2 and Figure 3, we believe that the framework can be further improved by reducing the amount of time spent computing the initial linear programming relaxation of the problem. This can be accomplished by utilizing a lagrangian decomposition method or by applying a Dantzig-Wolfe reformulation and utilizing a column generation procedure.

6. Acknowledgements

The authors thank Royal Dutch Shell, the National Science Foundation (Grant Numbered 1739977), and the Texas A&M Energy Institute for their financial support.

References

- R. C. Allen, S. G. Baratsas, R. Kakodkar, S. Avraamidou, J. B. Powell, C. F. Heuberger, C. D. Demirhan, E. N. Pistikopoulos, 2021. An optimization framework for solving integrated planning and scheduling problems for dense energy carriers. *IFAC-PapersOnLine* 54 (3), 621–626.
- R. C. Allen, C. D. Demirhan, C. F. Heuberger-Austin, E. N. Pistikopoulos, 2022. A multi-period integrated planning and scheduling approach for developing energy systems. *Optimal Control Applications and Methods*.
- R. C. Allen, Y. Nie, S. Avraamidou, E. N. Pistikopoulos, 2019. Infrastructure planning and operational scheduling for power generating systems: An energy-water nexus approach. In: *Computer Aided Chemical Engineering*. Vol. 47. Elsevier, pp. 233–238.
- C. Archetti, M. G. Speranza, 2014. A survey on matheuristics for routing problems. *EURO Journal on Computational Optimization* 2 (4), 223–246.
- C. D. Demirhan, W. W. Tso, J. B. Powell, C. F. Heuberger, E. N. Pistikopoulos, 2020. A multiscale energy systems engineering approach for renewable power generation and storage optimization. *Industrial & Engineering Chemistry Research* 59 (16), 7706–7721.
- C. Draxl, A. Clifton, B.-M. Hodge, J. McCaa, 2015. The wind integration national dataset (wind) toolkit. *Applied Energy* 151, 355–366.
- G. He, D. S. Mallapragada, A. Bose, C. F. Heuberger, E. Gençer, 2021. Hydrogen supply chain planning with flexible transmission and storage scheduling. *IEEE Transactions on Sustainable Energy*.
- M. Hummon, E. Ibanez, G. Brinkman, D. Lew, 2012. Sub-hour solar data for power system modeling from static spatial variability analysis. Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States).
- C. Li, A. J. Conejo, P. Liu, B. P. Omell, J. D. Sirola, I. E. Grossmann, 2021. Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems. *European Journal of Operational Research*.
- M. J. Palys, I. Mitrai, P. Daoutidis, 2021. Renewable hydrogen and ammonia for combined heat and power systems in remote locations: Optimal design and scheduling. *Optimal Control Applications and Methods*.
- Gurobi Optimization, LLC, 2021. Gurobi optimizer reference manual.
- Shell, 2020. Shell scenarios sketch: A US net-zero CO₂ energy system by 2050.

Holistic Capacity Management and Production Planning in the Pharmaceutical Supply Chain

Simon B. Lindahl,^{a,b} Deenesh K. Babi,^b Krist V. Gernaey,^a Gürkan Sin,^a

^a*Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 228 A, 2800 Kgs. Lyngby, Denmark*

^b*API Manufacturing Development BIO, Novo Nordisk A/S, Novo Allé 1, 2880 Bagsværd, Denmark*

gsi@kt.dtu.dk

Abstract

Pharmaceutical companies are responsible for the consistent delivery of life-saving medicines to patients. Production must be planned to ensure supply-demand balancing and time-dependent capacity constraints must be handled through capacity planning. This work describes an aggregated 4-step framework for mid-term decision support on integrated production and capacity planning including the relevant MILP mathematical models. The framework has been developed by combining theory with observed requirements in industry and it is applied to a case study abstracted from industrial API manufacturing to provide an example of its industrial relevance.

Keywords: MILP, Planning & Scheduling, Pharmaceutical Manufacturing, Capacity Planning

1. Introduction

Pharmaceutical companies must coordinate all parts of their complex supply chain (SC) to deliver life-saving medicines to patients consistently at low cost (Marques et al., 2020). Manufacturing of active pharmaceutical ingredient (API) is a crucial step that is constrained by low manufacturing flexibility, high capital investment and long building & approval timelines. Multi-product manufacturing lines are introduced to increase flexibility but the added system complexity makes it necessary to plan production on the mid-term time horizon spanning months to a few years. On this time horizon, short-term operational effects and long-term capacity changes can be overlooked. Typically, manufacturing at minimum cost is desired which results in a capacity constrained planning problem and the highly regulated pharmaceutical environment means that only certain limited capacity adjustments can be made. Researchers in process systems engineering (PSE) and operations research (OR) have developed models for production planning (Maravelias and Sung, 2009; Hu and Hu, 2016) and capacity planning (Smirnov et al., 2021) but the combined problem of planning production and capacity on the mid-term horizon has received limited attention despite its industrial importance. In this work, we aim to bridge the gap by developing a framework for evaluating capacity in a multi-product, multi-stage, multi-line campaign production environment, which is typical in pharmaceutical manufacturing. The development is based on the planning literature within PSE and OR combined with insights from a pharmaceutical company, and the framework integrates mathematical optimization models for campaign scheduling. The news value consists of creating production plans while accounting for capacity projects

to both provide production plans, expected inventories of all materials and effects of capacity changes as a function of both capacity level and implementation timing.

2. Problem Statement

The capacity and production planning problem can be stated as follows: Given a multi-product, multi-stage, multi-line API manufacturing system, demand profiles for all products and initial inventories of all products and intermediates. Determine optimal system-wide mid-term production plans and any required capacity changes, their amount and timing. This is subject to maintaining all products above their safety levels and specific features of API manufacturing such as changeover and release times.

3. Methodology

In this section, a 4-step framework for combined, model-based production and capacity planning in the pharmaceutical industry is presented. Such an approach is necessary to evaluate mid-term capacity requirements that are variable in both amount and timing along with regular production planning across a complex system. Next, the model is described followed by the framework for evaluating production plans and capacity.

3.1. Campaign scheduling model

The basis of the framework is a campaign scheduling model that is used to assess production capability across the system and estimate mid-term inventory levels for products and intermediates through campaign plan generation. The model is formulated as a discrete time, multi-period, multi-product, multi-stage, multi-line mixed integer linear programming (MILP) model based on the state-task network (STN) (Kondili et al., 1993). It includes new release time parameters Rel_{is} , continuous slack variables for representation of demand backlogging and/or safety inventory violations Δ_{st} , inclusion of time dependency in parameters for both capacity and safety storage level V_{ijt}^{max} , $C_{s,t}^{min}$, respectively, and a new continuous variable that allows the activation and selection of required capacity changes ∇_{ijt} . Overall, the model must be able to determine material balances through the allocation of production to lines while respecting capacity and storage limitations and time requirements for cleaning and quality control of products. Equation 1 determines the continuous material balance variable S_{st} , of product s at time t as a function of continuous variables for production and consumption B_{ijt} , and parameters for task stoichiometry ρ'_{is} and ρ_{is} , and demand D_{st} . Equation 2 enforces a safety inventory. Equation 3 uses binary allocation variables W_{ijt} , to describe allocation of manufacturing to lines through full backwards aggregation and states that only one product can be produced at a time. Equation 4 sets the production amount which is (de)activated by the allocation variable and constrained by the upper system capacity which has a fixed and variable component. Equations 5 is used to enforce sequence-dependent changeover times such that a new task can only start when enough time has passed after the previous task has finished. Equation 6 states that all slack variables must be below a maximum value and Equation 7 enforces all capacity increase variables below their corresponding allocation variable which maintains the link between production and allocation. Equation 8 constrains all capacity variables below a maximum value.

$$S_{st} = S_{s,t-1} + \sum_{i \in T_s} \rho'_{is} \sum_{j \in K_i} B_{i,j,t-P_{is}-Rel_{is}} - \sum_{i \in T_s} \rho_{is} \sum_{j \in K_i} B_{ijt} - D_{st} + \Delta_{st} - \Delta_{st-1}, \forall s, t \quad (1)$$

$$C_{s,t}^{min} \leq S_{s,t}, \quad \forall s, t \quad (2)$$

$$\sum_{i \in I_j} \sum_{t'=t}^{t-p_i+1} W_{ijt'} \leq 1, \quad \forall j, t \tag{3}$$

$$B_{ijt} \leq (W_{ijt} + \nabla_{ijt})V_{ijt}^{max}, \quad \forall i, t, j \in K_i \tag{4}$$

$$\sum_{i \in I_j^{(k)}} W_{ijt} + \sum_{i \in I_j^{(k)}} W_{ij,t-p_i-\theta} \leq 1, \quad \forall t, j, \theta = 0 \dots \tau_{jkk'} - 1 \tag{5}$$

$$\Delta_{st} \leq \Delta^{max}, \quad \forall s \in \text{Products}, t \tag{6}$$

$$\nabla_{ijt} \leq W_{ijt}, \quad \forall i, t, j \in K_i \tag{7}$$

$$\nabla_{ijt} \leq \nabla^{max}, \quad \forall i, t, j \in K_i \tag{8}$$

Different objective functions may be used depending on the questions or problems that are addressed. Objective O1 states that product inventories are maximized on a set of time points. Objective functions O2 and O3 are used in reformulations of the original problem to obtain minimum safety stock violation/backlogging or minimum capacity increase.

$$\text{maximize } \sum_t S_{st}, \forall s \in \text{Products}, t \in \text{Time Points} \tag{O1}$$

$$\text{minimize } \Delta^{max} \tag{O2}$$

$$\text{minimize } \nabla^{max} \tag{O3}$$

3.2. Framework for capacity and production planning

The framework is applicable to production planning problems within pharmaceutical manufacturing. The overall objective of the problem will determine the specific objective function and relevant scenarios to be investigated. It consists of four main steps: (1) system categorization and data collection, (2) baseline supply capability, (3) capacity analysis and (4) supply decision analysis. Baseline performance is tracked through steps 1 & 2, step 3 is added to escalate from a production planning to a capacity planning problem which is demanded by practitioners and step 4 shows the effects of changing the system capacity, operation, etc. Each of the steps will now be described in further detail.

3.2.1. System characterization and data collection

The characterization step is used to establish system boundaries and describe manufacturing of each product and intermediate. System boundaries must be established based on the problem definition to ensure that conclusions are valid and to limit complexity and required assumptions. The system characterization results in the overview in Figure 1 showing production pathways, lines that are set up for specific production and where these lines exist. The data collection step defines model parameters based on internal company systems and people. The data could come from finance systems or it could be extracted from data historians and aggregated into main operations to avoid describing individual pieces of equipment. Future capacities from projects that have already been approved can be added along with any planned downtime. The projects are an outcome of step 3 in the framework and the planned downtime comes out of steps 3 or 4 depending on whether it results in changed capacity or not.

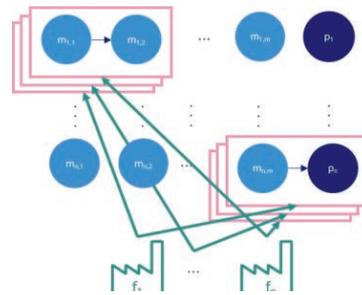


Figure 1: System characterization describing products (dark blue), intermediates (light blue), production lines (pink) and facilities (green).

3.2.2. Baseline supply capability

In step 2, the campaign scheduling model is set up with an objective function(s) that depends on the decision maker (DM) and the type of analysis/type of questions asked.

Next, the model is solved, and the results are reported in terms of key performance indicators (KPIs), for example, minimum inventories, cost or number of changeovers, and illustrations of the plans and inventories. It is our experience that DMs often want to see real plans and inventories rather than just KPIs since these are more relatable and provide a better understanding. If all KPIs are within specifications, step 3 is skipped, otherwise the capacity analysis is carried out.

3.2.3. Capacity analysis

In step 3, the future capacity of the system is explored either through a direct or an indirect approach. In the direct approach, one or more capacity changes are considered and introduced either as scenarios or directly into the model. These changes could originate from modelling activities on the system that result in physical changes or changed operating parameters. In the indirect approach, optimal capacity changes are investigated solely based on the campaign scheduling model to suggest required capacity increases that are given as input to other modelling activities.

3.2.4. Supply decision analysis

In step 4, the effects of any changes to the system (capacity, shutdown, product priority etc.) are determined. The results include scenario-dependent production plans, expected inventories, effects of (capacity) changes and suggested capacity implementation timings and final levels. Relevant KPIs are used to compare baseline supply with the different scenarios to provide the DM with the needed information to decide.

4. Case Study

The case study is abstracted from an industrial manufacturing network that converts raw materials to API on multiple multi-product lines. Plans up to a certain time are fixed due to constraints on manning, material supply, fixed deliveries etc. and used to project initial inventories for all materials along with the state of each equipment (producing, idle, changeover in progress). Demand increases in the pharmaceutical industry typically lead to a situation where capacity becomes constrained. Therefore, in this case study inventories maximization was selected throughout the horizon for all final products. The next sections will describe the solution of the case study through the 4-step framework.

4.1. System characterization and data collection

The baseline system consists of four products (referred to as A-D) and seven production lines in three separate stages (A, B and C) with intermediate storage as shown in Figure 2 (dark blue). Upstream production (USP) lines (A) produce the precursor which is purified in two downstream production (DSP) stages, DSP1 (B) and DSP2 (C), to the final API. The product-line allocations (Table 1) show that not all products visit all stages and final products are therefore defined by the last stage a product visits. As shown in Figure 2 the system used in the current case study (dark blue) was reduced from a larger system by assuming sufficient supply of material from Line A3 to Line B3. Demand data was given by the company as an average over two years with weekly deliveries and all final products have a six-month safety storage level and require four weeks release time for quality control after production. The results based on constant demands are applicable since the demand volatility is less than the final product safety storage levels and no stockouts will thus occur when implementing the generated plans. Table 1 gives yearly

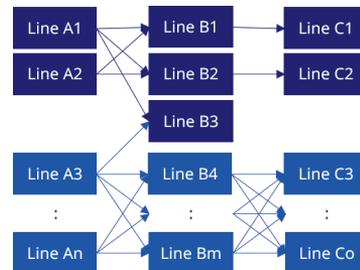


Figure 2: System reduction in a 3-stage environment with USP (A), DSP1 (B) and DSP2 (C)

(46 week) capacities in final product equivalents (scaled by yearly demand) and yields for all product-line combinations that exist. Raw material is unconstrained and therefore the yield in USP is set to 100%. Production is shut down 6 weeks per year (weeks 7, 29-31, 42, 52) and after each shutdown, the first week is spent on ramping-up. Table 1 provides ramp-up coefficients for each line as a fraction of full capacity. Initial levels in final product yearly demand equivalents are: 9 months for final products and 1 month for intermediates. Changeover times are sequence-independent and last 3 weeks in USP and 1 week in DSP.

Table 1: Capacity and yield overview in the format ‘Capacity (Yield)’ and ramp-up coefficients as a fraction of total capacity in the first week after a shutdown. Capacities are reported in units of yearly final product demand equivalent.

	USP		DSP1			DSP2	
	Line A1	Line A2	Line B1	Line B2	Line B3	Line C1	Line C2
Product A	1.5 (1.0)		1.5 (0.5)	1.5 (0.5)	1.5 (0.5)		
Product B	1.5 (1.0)	1.5 (1.0)	1.5 (0.5)			1.5 (0.9)	
Product C		1.5 (1.0)		1.5 (0.5)			1.5 (0.5)
Product D					1.5 (0.5)		
Ramp-up coefficient	0.5	0.5	0.5	0.5	0.5	0.5	0.5

4.2. Baseline supply capability

The campaign scheduling model was set up in Pyomo and solved with CPLEX on a 2-year horizon with a weekly discretization and objective function O1 maximizing final product inventories every half year which weights all products equally and highlights the need for high inventory levels throughout the horizon rather than solely at the end. All slack variables and capacity increase variables were constrained to 0. The results are given in Table 2 (Scenario I) which shows that all KPIs are acceptable, and the company therefore considers manufacturing intermediate for a development product on Line A1. The estimated duration of the production is 12 full-capacity weeks with standard changeover time, release time & ramp-up and delivery in week 84 on the horizon. The problem was solved including the development product which returned an infeasible solution showing that not all product inventories can be kept above their safety levels when the clinical production is introduced. Before continuing to step 3 in the framework, it is relevant to know how far from feasibility the solution is since small deviations from safety stock might be accepted by the DM. Therefore, the model was reformulated with an objective of minimizing the maximum slack value (objective function O2). The reformulated mathematical program (Scenario II) was solved giving an objective value of 0.06 which shows that the lowest expected inventory level is 0.44. If this number is not accepted by the DM, step 3 in the framework can be explored. As can be seen, this example shows that capacity planning is relevant to both account for changing demands for current products but also to handle new product introductions (NPIs).

4.3. Capacity analysis

In step 3, an example of the indirect approach to capacity changes is presented and applied to all legacy products on Line A1. The goal is to determine the minimum amount of capacity required to avoid safety violations since these could result in patients not getting access to life-saving medicines. The reformulation from step 2 was modified by replacing objective function O2 with O3, capacity variables were fixed at 0 for all product-line combinations except for the legacy products (A, B) on line A1, slack variables were fixed at 0 and the program (Scenario III) was solved giving an objective value of 0.16 which shows that a 16% capacity increase on Line A1 is required for both products A and B.

4.4. Supply decision analysis

In step 4 the result from step 3 is used to solve the original problem from step 2 with a 16% increased capacity on Line A1 for legacy products at the beginning of the horizon (Scenario IV). The results (Table 2) show that no safety stock violations occur and the remaining KPIs are close to the baseline (Scenario I). Based on the results, it is recommended to investigate how to achieve the capacity increase and accept the clinical production contingent on finding a solution. If the increase cannot be achieved, the DM will decide if additional information is needed before the clinical production is accepted.

Table 2: Data and results (KPIs) for the 4 scenarios

	Line A1 Capacity	Clinical Demand	KPIs		
			Changeovers	Lowest inventory	Final inventories
Scenario I	Baseline	No	20	0.5	2.64
Scenario II	Baseline	Yes	N/A	0.44	N/A
Scenario III	+16%	Yes	N/A	N/A	N/A
Scenario IV	+16%	Yes	22	0.5	2.72

5. Conclusions & Perspectives

A framework is described for handling mid-term capacity and production planning applied to the pharmaceutical industry. It allows the DM to evaluate current supply capabilities, identify capacity requirements and analyze the effect of specific changes to the system. This knowledge can be used to handle capacity proactively and increase collaboration across the supply chain from development through manufacturing and distribution. The framework was applied to a case study abstracted from industrial API production to give an example of its use. Future research will focus on expanding the framework with specific methods for each step and to add uncertainty to both framework and models. Plan robustness will be added to ensure feasibility subject to certain levels of uncertainty which will reduce volatility and firefighting in the planning processes as requested by industrial practitioners. Additionally, methods for multi-objective optimization will be added and model application will be explored including triggers/frequencies for updating data and applying the framework.

Acknowledgements

The authors would like to acknowledge the financial support from Novo Nordisk A/S, Product Supply Active Pharmaceutical Ingredient (PS API) Manufacturing for the industrial PhD program in manufacturing development.

References

- C. M. Marques, S. Moniz, J. P. de Sousa, A. P. Barbosa-Póvoa and G. Reklaitis, Decision-support challenges in the chemical-pharmaceutical industry: Findings and future research directions, *Computers & Chemical Engineering*, 134, 1-36
- C. T. Maravelias and C. Sung, 2009, Integration of production planning and scheduling: Overview, challenges and opportunities, *Computers & Chemical Engineering*, **33**, 12, 1919-1930.
- Z. Hu and G. Hu, A two-stage stochastic programming model for lot-sizing and scheduling under uncertainty, 2016, *International Journal of Production Economics*, 180, 198-207
- D. Smirnov, W. van Jaarsveld, Z. Atan and T. de Kok, Long-term resource planning in the high-tech industry: Capacity or inventory?, 2021, *European Journal of Operational Research*, 293, 3, 926-940
- E. Kondili, C. C. Pantelides, R. W. H. Sargent, 1993. A general algorithm for short-term scheduling of batch operations– I. MILP formulation. *Computers & Chemical Engineering* 17, 2, 211–227

Towards Resilience in Next-Generation Vaccines and Therapeutics Supply Chains

Miriam Sarkis^a, Kyungjae Tak^a, Benoit Chachuat^a, Nilay Shah^a and Maria M. Papathanasiou^{a,*}

^a*The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, United Kingdom*

**maria.papathanasiou1@imperial.ac.uk*

Abstract

Recent clinical outcomes of Advanced Therapy Medicinal Products (ATMPs) highlight promising opportunities in the prevention and cure of life threatening diseases. ATMP manufacturers are asked to tackle engineering product and process-related challenges, whilst scaling up production under demand uncertainty; this highlights the need for tools supporting supply chain planning under uncertainty. This study presents a computer-aided modelling and optimisation framework for viral vector supply chains. A methodology for the characterisation of process-related uncertainties is presented; the impact of input demand and process bottlenecks on optimal supply chain configurations and capacity allocations is assessed. A trade-off between cost and scalability emerges, larger costs incurring at higher input demands, whilst ensuring improved flexibility under demand uncertainty. Furthermore, bottlenecks uncertainty drives the optimisation to alternative strategic decisions, highlighting the need for a systematic integration within the framework.

Keywords: mathematical programming, supply chain optimisation, advanced therapeutics, pharmaceutical manufacturing

1. Introduction

Advanced Therapy Medicinal products (ATMPs) form a novel class of therapeutics with promising outcomes in the prevention and treatment of life-threatening diseases. ATMPs mostly rely on the delivery of genetic components to the target patient; this is achieved either via insertion of a vector carrying the genetic payload (*in vivo*) or by transplanting cells that been engineered to produce therapeutic proteins or factors (*ex vivo*). The gene therapy field is rapidly growing, with the U.S. Food & Drug Administration (FDA) forecasting 10-20 gene therapies approvals per year by 2025 (FDA, 2019). Viral vectors are currently at the forefront of next-generation vaccines and therapeutics. Adenoviral vectors have demonstrated their suitability and flexibility as vaccine carriers, with the commercial approval of Vaxzevria (AstraZeneca), the Janssen and the Sputnik COVID-19 vaccines (EMA, 2021). On the other hand, adeno-associated viruses have shown to be suitable for *in vivo* gene therapy with approvals of Luxturna and Zolgensma, lentiviruses (LV) find broad application in *ex vivo* gene therapy and utilised in the manufacturing of cell-based gene therapy products, such as autologous CAR-T cell therapy (Papathanasiou et al., 2020). As more advanced therapeutics and vaccine platforms reach clinical trials and commercialisation, the field is experiencing a need for innovative solutions for process development, manufacturing and infrastructure to ensure clinical availability and patient accessibility. Similar to product development and manufacturing, ATMPs distribution faces challenges that are often present across pharmaceutical supply chains (SC) (Fig.1). These include capacity planning under uncertainty of demands, clinical trial

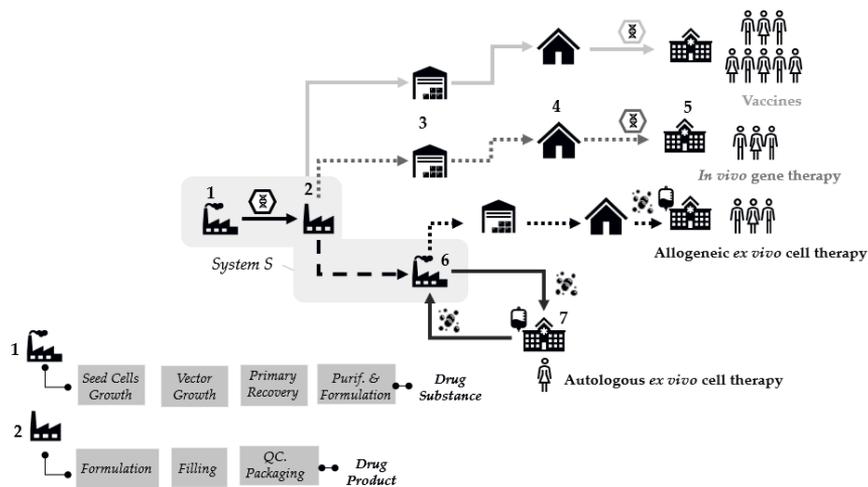


Figure 1: Generalised SCs for viral vectors: (1) primary and (2) secondary manufacturing, (3-4) regional-district storage, (5) hospitals/administration points and (7) leukapheresis sites, (6) cell therapy manufacturing facilities.

outcomes, process capabilities and dosage requirements (i.e. *in-risk* manufacturing and distribution). This highlights the need for informed decision-support tools for ATMPs SCs to improve strategic and operational planning (Sarkis et al., 2021). Several optimisation-based frameworks for scheduling of pharmaceutical plants and SCs under uncertainty have been proposed in recent works (Sousa et al., 2011; Siganporia et al., 2014; Vieria et al., 2019). This study presents (i) a novel methodology for the identification of underlying manufacturing uncertainties in emerging biopharmaceutical industries, focusing on LV vectors SCs (system *S*, Fig. 1) and (ii) an assessment of how these may impact optimal SCs configurations from clinical to commercial scale.

2. Materials & Methods

2.1. Techno-economic modelling

A primary and a secondary manufacturing platform for LV vectors was modelled using SuperPro Designer (Intelligen) to (i) allow for techno-economic assessment of the overall manufacturing process and (ii) evaluate how underlying uncertainties may propagate to key performance indicators of interest of the SC. SuperPro relies on built-in sets of algebraic and differential equations to calculate material and energy balances, sizes equipment, schedules operations and computes CapEx and OpEx. Information regarding production processes, resources costs was obtained from literature (Comisel et al., 2021; Perry and Rayat, 2021) and input ranges were recorded (Table 1). Additional inputs were sourced from SuperPro's industrial-based databases and were assumed to vary within a range of $\pm 50\%$ of their nominal value, preserving order of magnitude and allowing assessment of input significance. The uncertainty analysis was conducted using a SobolGSA-MatLab-Component Object Model (COM)-SuperPro interface, with the primary manufacturing model simulated for 1023 quasi-random combinations of inputs within the pre-defined range.

2.2. Optimisation framework

A mixed-integer linear programming (MILP) problem was developed (Table 2); given a range of production scales, manufacturing and logistics costs, the optimisation identifies (i) candidate

Table 1: Key input ranges for uncertainty analysis

Input	Nominal	LB	UB	Units	Section	Step
Target viral titre (main product)	1×10^7	1×10^6	1×10^8	TU mL ⁻¹	USP	Vector growth
Raw material (medium) cost	33.1	16.6	49.6	\$ kg ⁻¹	USP	Seed cells; vector growth
Reaction task lengths	11.0	8.0	14.0	days	USP	Seed cells; vector growth
Buffer utilisation	150.5	75.3	225.8	kg batch ⁻¹	DSP	Purification
Recovery (%)	90.0	70.0	99.0	-	DSP	Purification (dead-end filtration)
DNA removal task length	60	30	90	min	DSP	Purification (nuclease)
Filtration task length	240	120	360	min	DSP	Purification (dead-end filtration)
Chromatography task length	75	37.5	112.5	min	DSP	Purification (chromatography)
Elution yield (%)	41.0	29.0	90.0	-	DSP	Purification (chromatography)
Raw material (buffer) cost	43.5	21.8	65.2	\$ kg ⁻¹	DSP	Purification (nuclease)
Concentration factor	3	1.5	4.5	-	DSP	Purification (dead-end filtration)
Volume of Eluant	2	1	3	BV	DSP	Purification (chromatography)
Recovery (%)	99.0	70.0	99.0	-	DSP	Purification (ultra-filtration)
Rejection coefficient (%)	1.0	1.0	2.0	-	DSP	Purification (sterile filtration)

SC structures and allocated capacities, (ii) operational plans and (iii) transportation flows. The aim of the optimisation problem is to maximise the total network cost (Eq.1), comprising CapEx, OpEx and transportation costs (Eq 2-6). For manufacturing nodes j and f , the operational cost is subdivided into a labour-dependent component, a variable (batch-size dependent) a facility-dependent cost, accounting for equipment depreciation and maintenance (Eq. 4). A set of logic constraints ensure that feasible connections are established only if nodes are built (Eq. 7-14), whilst imposing 1 size s is selected for every node. Additionally product allocations occur only in installed equipment (Eq. 15-16). Material balances at each node ensure balanced flow between nodes, only if a match is established (Eq. 17, 22-26). In nodes j and f , batches are manufactured in campaigns, with the first batch being collected after α days and subsequent ones every r days (i.e. cycle time) (Eq. 18-21); this holds for USP, DSP and secondary manufacturing process sections.

2.3. Case study: viral vectors

The LV vector SC considered in this study comprises 3 candidate locations for primary and secondary manufacturing respectively, 3 storage nodes and 5 demand zones (i.e. cell therapy facilities). The impact of input demand on optimal SC configuration is assessed via 3 demand scenarios, which capture (i) a Phase I-II clinical demand (1,000 doses y⁻¹ on aggregate), (ii) Phase III-commercial demand (Comisel et al., 2021) and a (iii) commercial demand (Papathanasiou et al., 2020), with a commercial CAR-T cell therapy facility in $L5$. The impact of cycle time r^u and USP processing time α^u uncertainty on SC structures and capacity allocations is also assessed (Table 3). Finally, the resilience of the network is tested by fixing strategic decisions (i.e. scale and number of parallel lines), whilst allowing operational decisions to adjust accordingly, thereby identifying a SC *breakpoint*.

3. Results & Discussion

3.1. Techno-economic modelling and uncertainty analysis

The results of the proposed uncertainty analysis can be decoupled into throughput-related outputs (Fig. 2a-e) and cost-related (Fig. 2f-k) outputs. Process tasks length uncertainty is expected to impact upstream and downstream process (USP, DSP) times, with USP processing time remaining within 40-58 days, as opposed to 1-2 days for the DSP. With cycle times between 12-15 days, USP remains a bottleneck in the process, due to time-intensive batch cell cultures. Overall process capability is impacted (Fig. 2e), as an integer number of batches must be completed within

Table 2: Mathematical formulation of the supply chain optimisation problem

Formulation	Equations	Index
Objective	$z = TC_j^{CAP} + TC_f^{CAP} + TC_k^{CAP} + TC_j^{OP} + TC_f^{OP} + TC_k^{OP} + TC^T$	1
Capital Costs	$TC_j^{CAP} = \sum_{jst} C_s^{CAPjl} Z_{jst}; \quad TC_k^{CAP} = \sum_{ks} C_s^{CAPsl} Y_{ks}$	2-3
Operational Costs	$TC_j^{OP} = \sum_{jstl} C_{LABjl} B_{jstl} + C_s^{VARjli} P_{jstl} + C_s^{LABjl} Z_{jst}; \quad TC_k^{OP} = \sum_{ki} C^{OPki} Q_{kli}$	4-5
	$TC^T = t_c \sum_{ijf} D_{jf} Q_{ijf} + t_c \sum_{ifk} D_{fk} Q_{ifk} + t_c \sum_{ikl} D_{kl} Q_{ikl}$	6
Logic Constraints	$X_{jf} \leq \sum_s Y_{js} \quad \forall j, f; \quad X_{fk} \leq \sum_s Y_{fs} \quad \forall f, k; \quad X_{kl} \leq \sum_s Y_{ks} \quad \forall k, l$	7-9
	$\sum_j X_{jf} \geq \sum_s Y_{fs} \quad \forall f; \quad \sum_f X_{fk} \geq \sum_s Y_{ks} \quad \forall k; \quad \sum_k X_{kl} \geq 1 \quad \forall l$	10-12
	$Z_{jst} \leq Y_{js} \quad \forall j, s, l; \quad Z_{fsl} \leq Y_{fs} \quad \forall f, s, l$	13-14
	$W_{jstl} \leq Z_{jst} \quad \forall j, s, l, i; \quad W_{fsl} \leq Z_{fsl} \quad \forall f, s, l, i$	15-16
Sample Flow Constraints	$Q_{min} X_{jf} \leq \sum_l Q_{ijf} \leq Q_{max} X_{jf} \quad \forall j, f$	17
Manufacturing Constraints	$T_{min} W_{jstl} \leq T_{jstl} \leq T_{max} W_{jstl} \quad \forall j, s, l, i; \quad U_{jst} = \sum_i T_{jstl} \quad \forall j, s, l;$	18-19
Sample Node Balances	$B_{jstl} = W_{jstl} + 1/r(T_{jstl} - \alpha W_{jstl}) \quad \forall j, s, l, i; \quad P_{jstl} = x_{st} B_{jstl} \quad \forall j, s, l, i$	20-21
Network Balances	$\sum_l P_{jstl} = \sum_f Q_{ijf} \quad \forall i, j; \quad \sum_j Q_{ijf} = \sum_l P_{fjstl} \quad \forall i, f; \quad \sum_l P_{fjstl} = \sum_k Q_{ifk} \quad \forall i, f$	22-24
	$\sum_f Q_{ifk} = \sum_k Q_{ikl} \quad \forall i, k; \quad \sum_k Q_{ikl} \geq D_{il} \quad \forall i, l$	25-26

Table 3: Summary of scenarios considered

Scenario	Location-specific product demand (doses y^{-1})					Capability (d b^{-1})	
	L1	L2	L3	L4	L5	α^u	r^u
A	200	200	200	200	200	40	12
B	1,000	1,000	1,000	1,000	10,000	40	12
C	1,000	1,000	1,000	1,000	40,000	40	12
D	200	200	200	200	200	58	15
E	1,000	1,000	1,000	1,000	10,000	58	15
F	1,000	1,000	1,000	1,000	40,000	58	15

the annual operating days (335 days y^{-1}). Uncertainty in USP titers, DSP recovery and material consumption per unit product determines a range of batch sizes (Fig. 2d). Consequently, a varying number and size of output batches impact labour-related OpEx and variable OpEx respectively (Fig. 2h-k). CapEx output ranges instead depend on variability in stream volumetric flows, resulting from uncertainty in medium/buffer consumption rates per unit product produced.

3.2. Supply chain optimisation

The major contributor to SC cost is primary manufacturing accounting for 69-89% of costs across all scenarios considered (Fig. 3b-d). In scenario A, a clinical demand (1000 doses y^{-1}) is fulfilled by a single primary manufacturing site at 500 L, comprising a single USP and DSP line (Table 4). As the demand increases in L5 to commercial-scale (B-C), the framework selects to build the facility at 2000 L, with the addition of 1-3 parallel USP lines. Economies of scale are expected to be a key driver in lowering the unit production cost. At the same time, the installation of a parallel USP line increases the throughput of the facility and the utilisation of the DSP line in J1. Similar trends are seen in scenarios D-F, which consider longer cycle and processing times. Whereas D and E yield the same optimal capacity allocations, in scenario F the different process capability drives the optimisation to install an additional facility in J3, with an extra USP and DSP line respectively compared to C. As shown in Fig. 3a-c, the demand-optimised scenarios B and E

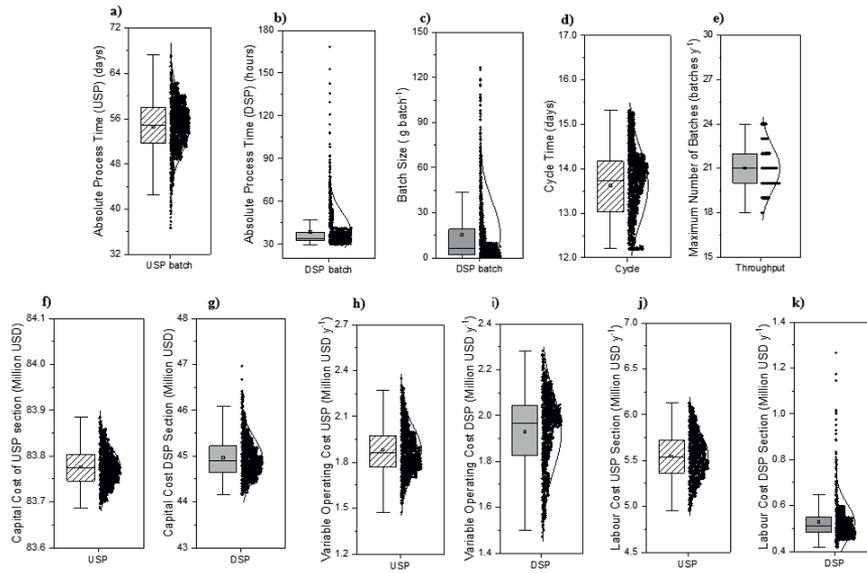


Figure 2: Results of uncertainty analysis for (a-e) throughput-related quantities and (f-k) cost-related quantities for 2000 L scale primary manufacturing process.

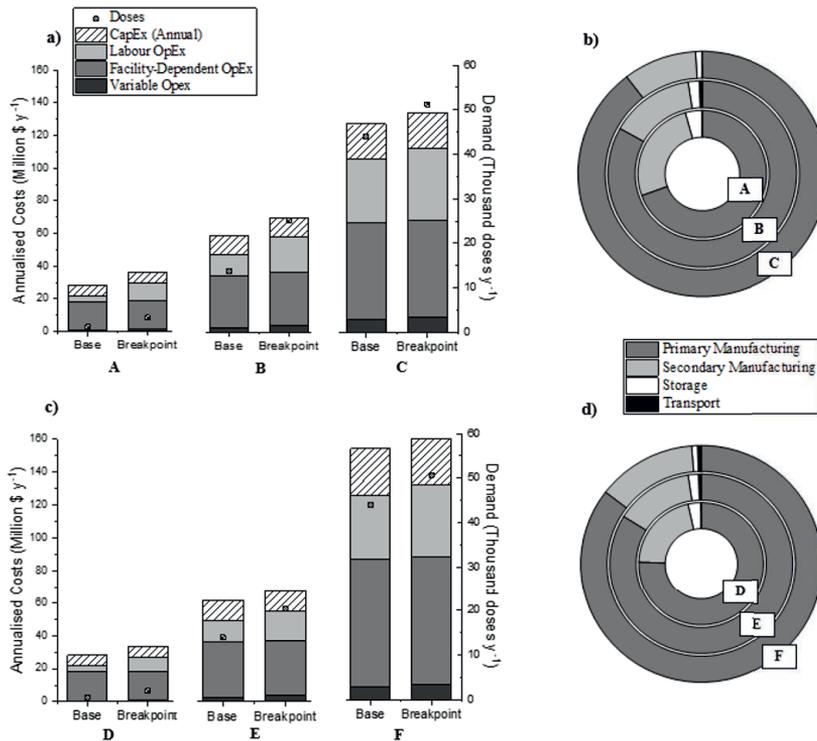


Figure 3: Cost-scalability analysis and SC costs for (a-b) $r^H = 12 \text{ d b}^{-1}$ and for (c-d) $r^H = 15 \text{ b}^{-1}$

Table 4: Summary of capacity allocations by scenario

	Scenario					
	A	B	C	D	E	F
Scale (L)	500	2000	2000	500	2000	2000
Primary Manuf.	$J1$	$J1$	$J1$	$J1$	$J1$	$J1+J3$
USP lines	1	2	4	1	2	2+3
DSP lines	1	1	1	1	1	2
Secondary Manuf.	$F1$	$F1$	$F1$	$F1$	$F1$	$F1+F3$
Scale (vials min^{-1})	50	50	50	50	50	50
F&F lines	1	1	1	1	1	1+1

require the same expenditure to fulfill the input demand, as cost is computed as a function of scale, number of batches and batch size. When assessing scalability, scenario B allows for a scale-up up to 25,000 doses y^{-1} , thanks to the shorter cycle times, whereas the SC network in scenario E saturates at 20,000 doses y^{-1} . Similarly, for scenarios C and F, the two SC networks can scale-up to 50,000 doses y^{-1} , albeit with 20 million USD y^{-1} gap in estimated costs and alternative capacity allocations and transportation links. This could result from a minimisation of transportation cost, which favours in this case a decentralised network as opposed to centralised manufacturing. This highlights the need to (i) carry out strategic decisions whilst considering bottlenecks that can emerge at operational level and therefore (ii) integrate an evaluation of the input uncertainty space as part of the optimisation. The latter is translated into considering uncertainty ranges for all throughput-related and cost-related inputs.

4. Conclusions

This study presents a computer-aided modelling and optimisation framework for vector-based ATMPs manufacturing and distribution networks. The developed tool identifies cost-effective SC structures and capacity allocations for a range of clinical to commercial demand scenarios and quantifies SC scalability under demand uncertainty. A methodology to characterise key process-related uncertainties is proposed; hence, the impact of cycle time uncertainty on SC network performance is demonstrated. The need for augmenting such tool to systematically enumerate the input uncertainty space as part of the optimisation has been highlighted. The identification of a unified resilient solution to realisations of uncertainty outcomes is expected to bring benefits to the emerging ATMPs industry, as it transitions from clinical to commercial operations.

References

- R. Comisel, B. Kara, F. H. Fiesser, S. S. Farid, 2021. Lentiviral vector bioprocess economics for cell and gene therapy commercialization. *Biochemical Engineering Journal* 167, 107868.
- EMA, 2021. COVID-19 vaccines.
- FDA, 2019. Statement from FDA Commissioner.
- M. M. Papathanasiou, C. Stamatis, M. Lakelin, S. S. Farid, N. Titchener-Hooker, N. Shah, 2020. Autologous CAR T-cell therapies supply chain: challenges and opportunities? *Cancer Gene Therapy* 27, 799–809.
- C. Perry, A. C. M. E. Rayat, 2021. Lentiviral vector bioprocessing. *Viruses* 13, 268.
- M. Sarkis, A. Bernardi, N. Shah, M. M. Papathanasiou, 2021. Decision support tools for next-generation vaccines and advanced therapy medicinal products: present and future. *Current Opinion in Chemical Engineering* 32, 100689.
- C. Siganporia, S. Ghosh, T. Daszkowski, L. Papageorgiou, S. S. Farid, 2014. Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnology Progress* 30(3), 594–506.
- R. T. Sousa, L. G. Papageorgiou, N. Shah, 2011. Global supply chain planning for pharmaceuticals. *Chemical Engineering Research and Design* 89, 2396–2409.
- M. Viera, T. Pinto-Varela, A. P. Barbosa-Póvoa, 2019. A model-based decision support framework for the optimisation of production planning in the biopharmaceutical industry. *Computers & Chemical Engineering* 121, 212–223.

Network-Based Analysis of Electrified Chemical Processing with Renewable Energy Sources

Ioannis Giannikopoulos^a, Alkiviadis Skouteris^a, David T. Allen^{a,b}, Michael Baldea^{a,c} and Mark A. Stadtherr^{a,*}

^a*McKetta Department of Chemical Engineering, The University of Texas at Austin, 200 East Dean Keeton Street, Austin, Texas 78712-1589, USA*

^b*Center for Energy and Environmental Resources, University of Texas, Austin, TX 78758, USA*

^c*Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA*

**markst@che.utexas.edu*

Abstract

In an effort to decarbonize and electrify the chemical industry, the incorporation of renewable energy sources in the energy input portfolio of chemical processes is a key step. To this end, we present a study of the effects of connecting a wind farm to a small-scale chemical process network comprising a shale gas processing plant and steam cracker. The steam cracker is assumed to be able to utilize a hybrid energy stream to support the endothermic cracking reactions, whereby conventional (fossil fuel) heat generation can be replaced with electricity whenever the latter is readily available from the wind farm. We probe the impact of employing this energy blend, as well as the network economics and CO₂ emissions, showcasing the potential benefits of incorporating renewable energy sources in lieu of/in addition to conventional ones to support chemical manufacturing.

Keywords: Decarbonization, Network Modeling, Electrification, Wind Power, Chemical Manufacturing

1. Introduction

The energy use of the industrial sector accounted for 33% of the total U.S. energy consumption in 2020, with almost 60% of this amount being related to chemicals and petroleum and coal products (U.S. Energy Information Administration, 2021c). While the majority of this energy need is met by fossil fuels, wind power contributed up to 8.4% of total U.S. utility-scale electricity generation (U.S. Energy Information Administration, 2021d). However, cases of under-utilization of available wind power are common, with the main issues being the transmission capacity of the wind generation facilities to the main electricity load centers and the inherent variability in generation rates. Infrastructure additions that would only be operational at full capacity when wind power is at its peak are uneconomical. More effective use of wind power could entail expanded local use of electricity at times of peak wind power generation, or finding ways to store energy at times of peak wind power production. In the case of chemical production (which is a form of expanded use of electricity and energy storage), the main advantage would be to (locally) produce chemicals that would be readily compatible with existing fuel transport systems. In this work we investigate modalities for utilizing (peaks in) wind power generation to support the decarbonization and electrification of chemical processing. We utilize a small-scale process network model that represents

a shale gas processing plant, an ethane cracking facility and a wind farm. We consider the limiting case where the endothermic cracker can utilize a hybrid energy stream, comprising an infinitely variable supply of conventional (fossil) fuels and electricity. Depending on the electricity production level, production of the ethane cracker may be fully electrified or fully conventional, or anywhere in between. We choose to maximize overall network profitability, which enables us to study the energy supply with the most cost-efficient use of available wind power. Furthermore, CO₂ emissions are calculated and presented.

2. Background

In recent years, there has been significant research interest in electrification and decarbonization in the chemicals and energy sectors. For example, Chen et al. (2019) evaluated the effect of direct and indirect electrification on the chemical industry by using a methanol process as a case study. Electrification is achieved by changing the energy mix required by the process (direct electrification) or by utilizing alternative feedstocks (indirect electrification). van Rooij et al. (2017) studied CO₂ reduction in the chemical industry by designing a plasma process that would transform CO₂ to CO. In this case, decarbonization is achieved by converting CO₂ into a different and useful molecule. Such approaches are further discussed by Schiffer and Manthiram (2017), who study the electrification and decarbonization of the chemical industry via electrochemistry, focusing on the largest production volume and energy consumption chemicals. Papadis and Tsatsaronis (2020) focus on the energy sector and discuss the challenges and steps that need to be considered on environmental, economic, technical, social and political levels in order to limit the adverse effects on the environment. Even though these researchers, and the many others working in this field, focus on different aspects of electrification and decarbonization, there is a consensus that the integration and efficient use of renewable resources are necessary and will be the main drivers of these efforts.

3. Problem Definition

In order to study the impact of electrifying chemical processing, we consider a canonical process network comprising a natural gas processing plant and ethane cracker, supported by a wind farm whose generation capacity varies throughout the day. The choice of gas processing plant/cracker is motivated by the need to process distributed fossil fuel sources, such as natural gas liquids (NGLs) obtained from shale gas fields. Such distributed processing lends itself naturally to electrification via locally generated renewable energy. In the scenario considered, the gas processing plant utilizes shale gas feedstock and produces natural gas that is ready for use, and NGLs (ethane, propane, butanes, and smaller amounts of heavier hydrocarbons), with the latter used to produce ethylene via steam cracking. In order to obtain a realistic representation of the process economics, we utilize data (process stoichiometries, production costs, material costs, selling prices) from the IHS Markit (2012) Process Economic Yearbook. Data on electricity prices throughout the day are obtained from the Electric Reliability Council of Texas (2021). The capacity of the gas processing plant is 150 million ft³, which is the average capacity for gas processing plants, according to the U.S. Energy Information Administration (2019), and the assumed wind farm capacity is 250 MW.

We represent the network as a directed graph, and describe here a network model in general terms. The set of network nodes is denoted as N , which is comprised of sets of supply nodes (denoted by S), sink nodes (K), conventional unit manufacturing nodes (U), electrified unit manufacturing nodes (Z) and a wind farm node (E). Therefore, $N = S \cup K \cup U \cup Z \cup E$. The set of all the materials and utilities (e.g., natural gas, steam, electricity) in the model is denoted by I and the set of all the technologies used in the manufacturing nodes are denoted as J . A set of inputs $I_{in,m}$ is associated with each node $m \in U \cup K \cup Z$ and a set of outputs $I_{out,n}$ is associated with each node $n \in U \cup S \cup Z \cup E$. Potential flows from node $n \in N$ to node $m \in N$ are represented with a directed edge (n,m) in the graph. The set of all edges representing potential flows of $i \in I$ is then given

by $G_i = \{(n, m) \mid i \in I_{out,n} \cap I_{in,m}\}$. Chemical manufacturing nodes (sets U and Z) are assumed to operate continuously, while the wind farm (set E) may operate intermittently.

The model serves as the basis for formulating a multiperiod linear program (LP). For the case study presented here, the objective function is maximization of the daily total profit:

$$\max P = \sum_{n \in N} \sum_t (R_{n,t} - C_{n,t}) - W \quad (1)$$

where P is the total daily profit from the network, $R_{n,t}$ is the revenue from node n during time period t , $C_{n,t}$ is the cost incurred at node n during time period t , and W is a daily fixed cost (i.e. not dependent on production rate) associated with the wind farm (which incurs fixed costs even when not producing power). Time periods in the model are one hour in length. The revenue term is given by:

$$R_{n,t} = b_{n,t} P_{n,t} \quad (2)$$

where $b_{n,t}$ is the unit selling price of n at time period t and $P_{n,t}$ is the amount of n produced during time period t . The cost term is given by:

$$C_{n,t} = a_{n,t} P_{n,t} \quad (3)$$

where $a_{n,t}$ is the unit cost of producing n at time period t . The fixed charge W for wind farm capital costs, operation and maintenance was estimated as \$118,500/day based on published data (U.S. Energy Information Administration, 2021a; Reuters Events, 2017). Note that wind farm costs may be overestimated because there are financial incentives for the installation of wind farms and renewable energy production that are not accounted for here.

An important aspect of this case study is that the steam cracker can operate on a (hypothetical) hybrid energy source: conventional fossil-based energy, electricity or some combination of the two. In the model the cracker is represented as two different nodes, one in set U and one in set Z ; however, there would physically be only a single hybrid cracker unit. The fraction of energy obtained from electricity is η_e , and from conventional sources $1 - \eta_e$. Here η_e is a continuous variable and $\eta_e \in [0, 1]$.

The constraints in the optimization problem are as follows:

1. There are material and utility balances governing flows in the network. These are formulated in terms of the network flow variables as shown by Giannikopoulos et al. (2021).
2. There may be supply constraints on the raw materials entering the network and demand constraints on the products exiting the network.
3. There may be limits on the flow of electricity to and from the grid, due to transmission constraints.
4. There are constraints on the availability of wind power, based on the variability of wind. This is quantified using the nominal capacity of the wind farm (specified as 250 MW) and multiplying it by a wind capacity factor (WCF) which is a function of time. The wind capacity factor is a representation of the wind energy performance which is assumed proportional to wind speed. For a given time period the wind speed is determined by random sampling from a Weibull probability distribution:

$$f_v = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad (4)$$

where v is the wind speed (m/s), k is the Weibull shape factor, and c is the Weibull scale parameter (m/s), which is a measure of the characteristic wind speed of the distribution. The WCF is then set

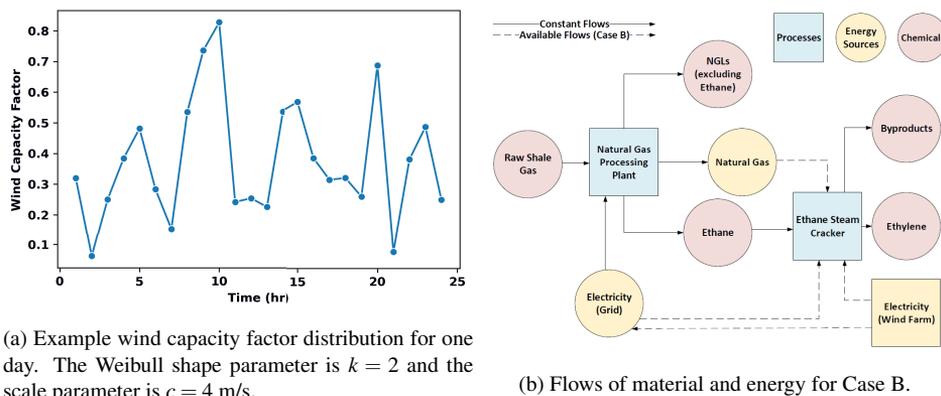


Figure 1: Wind capacity factor distribution and flow diagram for Case B.

to $WCF = \min\{0.1v, 1\}$, with the proportionality factor of 0.1 chosen to obtain an average WCF of about 0.35 (corresponding to the U.S. average). An example plot of the WCF for a 24-hour period is shown in Figure 1a.

Finally, the model accounts for CO_2 emissions from the production of the utilities used (steam, natural gas, electricity), as well as the emissions from the combustion of natural gas. Information about the CO_2 equivalent emissions from production of the utilities is based on data from GREET Argonne National Lab (2021), and information about the emissions from natural gas combustion are based on data from the U.S. Energy Information Administration (2021b). Based on these data, we are able to estimate the CO_2 equivalent emissions from the cracking process, and how these change as the mix of hybrid energy varies between electricity and fossil fuel. Note that these emissions data are not currently included in the objective function for the optimization model; however, this can be easily done (alternatively emissions could be treated as a separate objective in a multiobjective optimization analysis).

4. Case Studies and Results

We are interested in probing the effect of a renewable energy source, in this case the wind farm, in contributing energy to a small-scale process network involving a hypothetical steam cracker that can operate on a hybrid energy source (conventional fossil fuels and/or electricity). As two simple initial case studies, we will first solve the optimization-based model described above for a scenario with no wind farm (Case A), and then solve it for a case in which the wind farm energy is available (Case B). A simplified representation of the network flows for Case B is presented in Figure 1b. In each case, we will be interested in determining the total profit of the network, the total CO_2 emissions and the fraction η_e of electricity-powered operation in each time period. In both cases, there are no raw material supply constraints, so the production levels in the gas plant and ethane cracker are always at full capacity. In Case B, the wind farm can sell electricity to the grid at the prevailing price, but there are restrictions on transmission capacity that limit sales to 12 MWh daily, divided equally over the 24 one-hour time periods. This restriction is an upper limit; thus, the wind farm does not have to sell this amount to the grid, unless it is more profitable for the overall network system. For Case B, the model is run for 5 days, each with a different WCF distribution, in order to better account for the variability of the wind-generated electricity.

In Case B, the production levels due to electrified (production fraction η_e) and conventional operation (production fraction $1 - \eta_e$), are shown in Figure 2. This shows an average day of production,

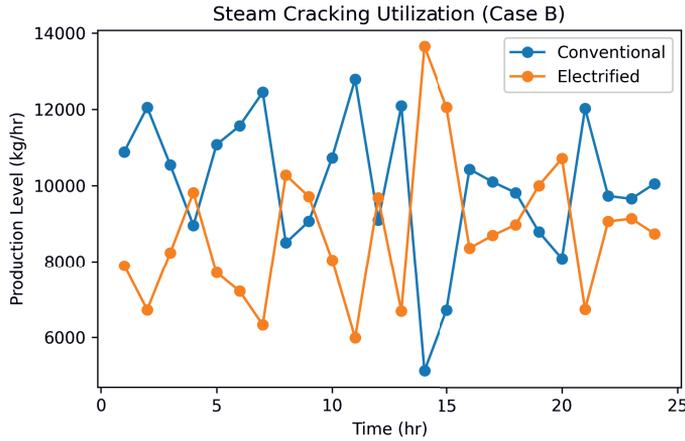


Figure 2: Steam cracking production level for each type of energy used (average at each hour over the five-day period studied).

determined by taking the average at each hour over the five-day period studied. The wind farm sells power to the grid at the maximum allowable rate, with all the remaining power used for the steam cracker. Thus, the degree of electrification depends strongly on the WCF, and tracks the amount of available electrical power closely. For the situation in this case study, both electrical and conventional fossil energy are used continuously throughout the average day. Of course, usage of electricity may be higher or lower depending on the nominal capacity assumed for the wind farm and on values specified for the wind distribution parameters c and k . It is important to note that, for this simple case study, we have assumed that an instantaneous and continuous change can be made between electrical- and fossil-powered operation of the steam cracker, which may not represent the true dynamics of the hybrid-powered operation. But, by considering this limiting case, we can still help evaluate the potential benefits of adopting such a hybrid technology, especially when paired with a renewable energy source.

The objective function in the model is the maximization of the overall network profit. The total network production levels are constant and the same in both cases; therefore, the difference in the total profit between Case A and Case B will stem only from the partial electrification of the cracker and sale of electricity to the grid in Case B, as well as the investment and other fixed costs associated with the wind farm in Case B. A summary of the maximized profit values for Case A and Case B is given in Table 1. We are also interested in the CO₂ equivalent emissions in both cases, as decarbonization is a key goal. Values for the CO₂ equivalent emissions are also shown in Table 1. Comparing Cases A and B, these values indicate a modest decrease in profit but a very substantial decrease in CO₂ emissions due to the addition of the wind farm and hybrid-power operation of the steam cracker in Case B.

Table 1: Network profit and CO₂e emissions for both case studies, with change relative to Case A.

	Case A	Case B–Day 1	Case B–Day 2	Case B–Day 3	Case B–Day 4	Case B–Day 5
Profit (1000 \$/day)	1,009	920 (−8.9%)	913 (−9.5%)	915 (−9.3%)	919 (−8.9%)	917 (−9.1%)
CO ₂ e Emissions (t/day)	812	385 (−52.6%)	482 (−40.6%)	459 (−43.5%)	395 (−51.4%)	436 (−46.3%)

5. Conclusion

In this work, we used a small-scale chemical process network comprising a shale gas processing plant and an ethane steam cracker, with connection to a wind farm, to study the potential use of a hybrid energy system to support the cracker. This hybrid system would allow conventional (fossil fuel) heat generation to be replaced with electricity whenever the latter is readily available from the wind farm. An optimization-based network model was used to develop simple case studies comparing a purely fossil-fueled case with the hybrid case, indicating, for the specific cases considered, that a large reduction in CO₂ emissions could be achieved with only a modest loss of profit. However, given the current incentives for adoption and utilization of renewable energy sources, the profit reduction could be even smaller. This suggests there is long-term benefit in adopting renewable energy resources and pairing them with chemical manufacturing, contributing to emissions reductions through electrification and decarbonization.

6. Acknowledgments

This paper is based upon work supported primarily by the National Science Foundation under Cooperative Agreement No. EEC-1647722 (CISTAR: NSF Engineering Research Center for Innovative and Strategic Transformation of Alkane Resources). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- Argonne National Lab, 2021. GREET, the Greenhouse gases, Regulated Emissions, and Energy use in Technologies Model.
URL <https://greet.es.anl.gov>
- C. Chen, Y. Lu, R. Banares-Alcantara, 2019. Direct and indirect electrification of chemical industry using methanol production as a case study. *Applied Energy* 243 (January), 71–90.
- Electric Reliability Council of Texas, 2021. Market Prices.
URL <http://www.ercot.com/mktinfo/prices>
- I. Giannikopoulos, A. Skouteris, T. F. Edgar, M. Baldea, D. T. Allen, M. A. Stadtherr, 2021. Geospatial network approach for assessing economic potential of ethylene-to-fuel technology in the marcellus shale region. *Industrial & Engineering Chemistry Research* 60 (41), 14801–14814.
- IHS Markit, 2012. IHS 2012 Process Economics Program yearbook.
URL <https://ihsmarkit.com/products/chemical-technology-pep-index.html>
- E. Papadis, G. Tsatsaronis, 2020. Challenges in the decarbonization of the energy sector. *Energy* 205, 118025.
- Reuters Events, 2017. US wind O&M costs estimated at \$48,000/MW; Falling costs create new industrial uses: IEA.
URL <https://www.reutersevents.com/renewables/wind-energy-update/us-wind-om-costs-estimated-48000mw-falling-costs-create-new-industrial-uses-iea>
- Z. J. Schiffer, K. Manthiram, 2017. Electrification and Decarbonization of the Chemical Industry. *Joule* 1 (1), 10–14.
- U.S. Energy Information Administration, 2019. U.S. natural gas processing plant capacity and throughput have increased in recent years.
URL <https://www.eia.gov/todayinenergy/detail.php?id=38592>
- U.S. Energy Information Administration, 2021a. Average U.S. construction costs for solar generation continued to fall in 2019.
URL <https://www.eia.gov/todayinenergy/detail.php?id=48736>
- U.S. Energy Information Administration, 2021b. How much carbon dioxide is produced when different fuels are burned?
URL <https://www.eia.gov/tools/faqs/faq.php?id=74&t=11>
- U.S. Energy Information Administration, 2021c. Use of energy explained: Energy use in industry.
URL <https://www.eia.gov/energyexplained/use-of-energy/industry.php>
- U.S. Energy Information Administration, 2021d. Wind explained: Electricity generation from wind.
URL <https://www.eia.gov/energyexplained/wind/electricity-generation-from-wind.php>
- G. J. van Rooij, H. N. Akse, W. A. Bongers, M. C. M. van de Sanden, 2017. Plasma for electrification of chemical industry: a case study on CO₂ reduction. *Plasma Physics and Controlled Fusion* 60 (1), 014019.

MINLP framework for systems analysis of the chemical manufacturing industry using network models

Alkiviadis Skouteris^a, Ioannis Giannikopoulos^a, David T. Allen^{a,b}, Michael Baldea^{a,c} and Mark A. Stadtherr^a

^a*McKetta Department of Chemical Engineering, The University of Texas at Austin, 200 East Dean Keeton Street, Austin, Texas 78712-1589, United States*

^b*Center for Energy and Environmental Resources, The University of Texas at Austin, 10100 Burnet Road, Bldg. 133, R7100, Austin, Texas 78758, United States*

^c*Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 E 24th St, Austin, Texas 78712, United States*
markst@che.utexas.edu

Abstract

The shale gas boom in the U.S. has had a strong effect on the U.S. petrochemicals industry. The availability of feedstocks, particularly natural gas liquids (NGLs), has significantly increased, prompting new investments and increased production, as well as research on new technologies and routes to products. Optimization-based industry network superstructure models can be used, with a cost minimization objective, to evaluate the potential of such new technologies and routes. A key aspect of such models is the approach used for propagating material prices and costs as new technology is adopted. This is needed because the extent to which a new technology is used may affect the raw material costs of processes elsewhere in the network, leading to a nonlinear model. In previous work, we proposed a cost propagation method for this situation, leading to solution of the nonlinear problem as a sequence of linear ones. Here we explore an alternative solution strategy based on use of a mixed-integer nonlinear programming (MINLP) model formulation, and compare to the previous approach.

Keywords: Optimization, Network modeling, MINLP, Chemical manufacturing, Supply chain

1. Introduction

Recent technological advancements in hydraulic fracturing and horizontal drilling have led to rapid increases in crude oil and natural gas production in the United States (U.S. Energy Information Administration, 2021). A result of this growth is the significant increase in the production of natural gas liquids (NGLs), which are often abundant in shale gas. NGLs consist mainly of ethane, propane, butanes and (in smaller amounts) C5+ compounds, molecules that constitute some of the most important building blocks of the chemical manufacturing industry. Therefore, the increase in the availability of shale gas and associated NGLs has provided a unique opportunity to expand the U.S. chemical manufacturing industry (Siirola, 2014; Yang and You, 2017).

Given the scale of capital investment involved in industry expansions, judicious decisions must be made regarding the choice of products, technologies and capacities. Network models are often at the core of such decision processes, due to their ability to capture existing and potential pathways and interconnections between different manufacturing processes (Derosa and Allen, 2015).

In recent work (Skouteris et al., 2021), we developed and used a network superstructure model involving several hundred of the highest-volume chemicals and hundreds of potential processing technologies, and considered the industry at the level of the entire United States. The model served as the basis for formulating an optimization-based industry configuration framework, seeking to minimize overall cost. A key innovation consisted of the fact that—in contrast to previous work (e.g., DeRosa and Allen, 2016)—process costs and material prices were allowed to change and respond to external perturbations in the network structure (e.g., adding a new process). To solve the resulting nonlinear problem, we used (Skouteris et al., 2021) a successive linear programming (SLP) approach that alternates between solving a constant-cost linear program (LP) and a special cost-propagation algorithm that updates process costs and material prices based on the LP results. In this paper, we explore another option for solving this nonlinear problem, based on formulating it as a mixed-integer nonlinear program (MINLP). Developing this alternative will enable comparisons to the SLP approach in terms of accuracy and computational efficiency, and is part of our strategy for evaluating the overall efficacy of the SLP scheme. The initial work reported here is focused on a simple network example, which we use for testing prior to scaling up to the larger U.S. industry model.

2. Background

Optimization-based, chemical industry network models originated with the work of Stadtherr and Rudd (1976), who initially focused on minimization of resource usage and subsequently on minimization of industry cost (Fathi-Afshar et al., 1981). Many variations and extensions of this initial work, with different applications, have appeared since, as reviewed by Skouteris et al. (2021) and Derosa and Allen (2015). Network models represent the industry as a directed graph, with nodes corresponding to manufacturing processes and edges corresponding to material and utility flows between the nodes. Each process node j is characterized by a process stoichiometry and a net production cost C_j (including raw material and utility costs plus annualized fixed capital cost and other operating costs) and each edge by a flow rate. The industry network model is then typically formulated as an LP, provided that process and materials costs are assumed to be constant and not dependent on the industry configuration (Derosa and Allen, 2015). The LP is comprised of material balance constraints for each material, as well as supply and demand constraints for primary raw materials and final end products, respectively. The objective is often chosen to be minimization of total industry cost, but many other driving forces have been explored (e.g., environmental considerations). For example, we recently developed (Giannikopoulos et al., 2022) a multi-objective industry network model that, in addition to a minimum total cost objective, also considers minimization of carbon loss (e.g., as CO₂ emissions) in the network, accounting for both feedstock carbon and fuel carbon used to provide process energy.

As established in our previous work (Skouteris et al., 2021), assuming process costs to be constant and independent of the process utilization levels is not always realistic and can lead to the model being unresponsive towards external perturbations to the network, such as introducing a new process. This was addressed by making the process costs dependent on the production levels, thereby rendering the network-level optimization problem nonlinear. To address this, we developed an SLP approach that alternates between the solution of a constant-cost LP and the use of an iterative cost-propagation procedure that updates material prices and thus process costs based on the results of the previous LP. The updated costs are then treated as constants in the next LP and the solution process continues until there are no cost changes. The price update mechanism used was based on the dominant-producer price leadership model (Scherer and Ross, 1990), in which the largest producer of a material determines its price. Complete details of the LP formulation, as well as the cost-propagation procedure used in the SLP approach, are given by Skouteris et al. (2021) and will not be repeated here.

3. MINLP Problem Formulation

As an alternative to the SLP approach, we consider here an explicit MINLP formulation of the problem, which allows for validating the proposed cost propagation/SLP framework. The MINLP formulation and solution procedure will be described here. As a preliminary step, a baseline LP minimizing total industry cost (Skouteris et al., 2021) using constant baseline process costs $C_{j,0}$ is solved, with the solution representing the state of the system prior to any external perturbations (e.g., addition of a new process):

$$\min_{X_j} C_{\text{tot}} = \sum_{j \in J} C_{j,0} X_j \quad (1)$$

$$\text{s.t. } \sum_j a_{i,j} X_j \geq D_i - S_i, \quad \forall i \in I. \quad (2)$$

Here, I and J are the sets of all materials and processes, respectively, in the model, $a_{i,j}$ is the input-output coefficient for material i in process j (negative if material i is consumed in process j , positive if i is produced; unity if i is the main product), X_j is the nonnegative production level of process j (in terms of flow rate of main product), and S_i and D_i are specified exogenous raw material supply and final product demand rates, respectively. Eq. (2) is a combination of material balance, supply and demand constraints and is expressed in terms of annual mass flow rates. From the solution of this baseline LP, the largest producer of each material i occurring as a main product is determined, and the initial price $B_{i,0}$ of i is then set equal to its production cost in this dominant process (currently no profit margin is added). Now a new process is added to the model to study its impact, and the industry network is reoptimized as an MINLP.

As in the SLP approach, the largest producer for each material i occurring as a main product has to be determined, since its cost will be used as the price for i in other processes, and, if i is a main product that is not produced at all in the optimal industry structure, then its price should remain at the initial price $B_{i,0}$. For the MINLP case, these considerations are captured via additional constraints in the optimization problem rather than as a separate calculation task. To avoid using the non-differentiable *max* function, we propose the following constraints:

$$P_i \geq X_j, \quad \forall i \in I_{\text{mp}}, \quad \forall j \in J_i \quad (3)$$

$$P_i \leq X_j + M(1 - d_{i,j}), \quad \forall i \in I_{\text{mp}}, \quad \forall j \in J_i \quad (4)$$

$$\sum_{j \in J_i} d_{i,j} = e_i, \quad \forall i \in I_{\text{mp}} \quad (5)$$

$$P_i \geq e_i \quad (6)$$

Here I_{mp} is the set of all materials produced as main products and J_i is the set of all processes yielding i as a main product. Two binary variables are used: $d_{i,j} = 1$ if process j is the largest producer of i as a main product and zero otherwise, and $e_i = 1$ if i is produced as a main product and zero if it is a main product not produced. M is a constant with a value that must be large enough to be an upper bound for all production rates X_j . When $e_i = 1$, P_i is the largest production rate of i , which occurs in the process j for which $d_{i,j} = 1$. When $e_i = 0$, $d_{i,j} = 0$ and $X_j = 0$, $\forall j \in J_i$. This means that Eqs. (3) and (4) can be satisfied for any $P_i \in [0, M]$. However, this is problematic if $P_i = 0$ since now Eq. (4) can be satisfied for either $d_{i,j} = 0$ and $e_i = 0$ or $d_{i,j} = 1$ and $e_i = 1$. To force $e_i = 0$ and thus all $d_{i,j} = 0$ in this case, we add constraint (6), which ensures that $e_i = 0$ when $P_i = 0$.

The prices B_i (in cents/lb) for each material $i \in I_{\text{mp}}$ and the net unit production costs C_j for each process $j \in J$ are given by

$$B_i = \sum_{j \in J_i} d_{i,j} C_j + B_{i,0}(1 - e_i), \quad \forall i \in I_{\text{mp}} \quad (7)$$

$$C_j = C_{j,0} + \sum_{i \in J} -a_{i,j}(B_i - B_{i,0}), \quad \forall j \in J \quad (8)$$

If there is no production of material i , then $d_{i,j} = 0, \forall j \in J_i$, and $e_i = 0$; thus, the first term in Eq. (7) becomes zero and the second term becomes $B_{i,0}$, the baseline price. Otherwise, if i is produced, then $e_i = 1$; thus, the second term in Eq. (7) becomes zero and the first term becomes the C_j for the largest producer of i ($d_{i,j} = 1$). As in the baseline case, no profit margin has been added. In Eq. (8), the notation $i \cap j$ indicates that the summation is over all materials i participating in process j . The summation represents the change in the materials cost for process j due to price changes.

Finally, since the constraints, Eqs. (3)–(6) will not force $e_i = 1$ when i is produced as a main product, the objective function used in the LP is modified for the MINLP by adding a penalty term:

$$\min_{X_j, P_i, d_{i,j}, e_i, B_i, C_j} \bar{C}_{\text{tot}} = \sum_{j \in J} C_j X_j - \sum_{i \in I_{\text{mp}}} e_i \quad (9)$$

Here the summation over e_i acts to set all e_i not already forced to zero by the constraints to have a value of one.

The MINLP defined by Eqs. (2)–(9) was implemented and solved in GAMS (GAMS, 2021) using the DICOPT solver (Duran and Grossmann, 1986), with the MILP master problems solved by using CPLEX and the NLP subproblems were solved using IPOPT.

4. Discussion

The MINLP formulation was validated on several small-scale networks (not representing physical processes used in practice) prior to use on the much larger full U.S. industry model. The results for the network depicted in Fig. 1 will be described here. In this network, before any external perturbations, there are three supply nodes (P3001, P3002, P3003) providing primary raw materials A, B and C to processes P1, P2 and P4, which in turn produce the intermediates D, E and F. Processes P3 and P5 then can produce the end product G. Overall, there are two competing pathways to produce G, one consisting of P2 and P3, and one consisting of P4 and P5 (each also requiring use of D from P1). The process costs have been set in such way that the P2-P3 pathway is much cheaper than the P4-P5 pathway, resulting in the former being selected in the baseline state.

A new process, P2005, is then introduced in existing network. P2005 produces the intermediate F and can therefore potentially replace P4. If the process costs are assumed to be constant, then any reduction in the cost of F due to use of the new technology P2005 will have no effect on the cost of operating P5, thus hampering the adoption of P2005. However, if the process costs are allowed to vary, as considered here, and P2005 can produce F more cheaply, then the lower price of F will be reflected in the cost of operating P5, thus favoring the adoption of P2005 in the network.

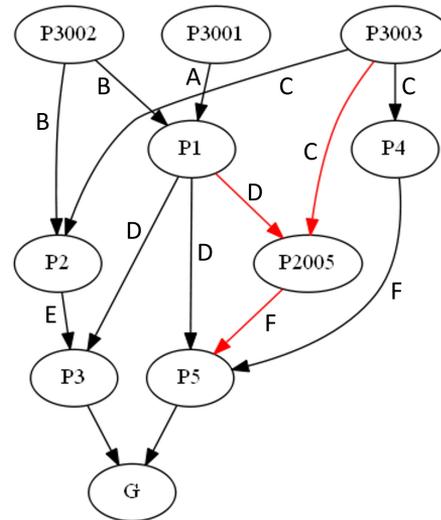


Figure 1: Graph representation of small industry network example problem. Node P2005 and the red edges represent the addition of a new process to the existing network.

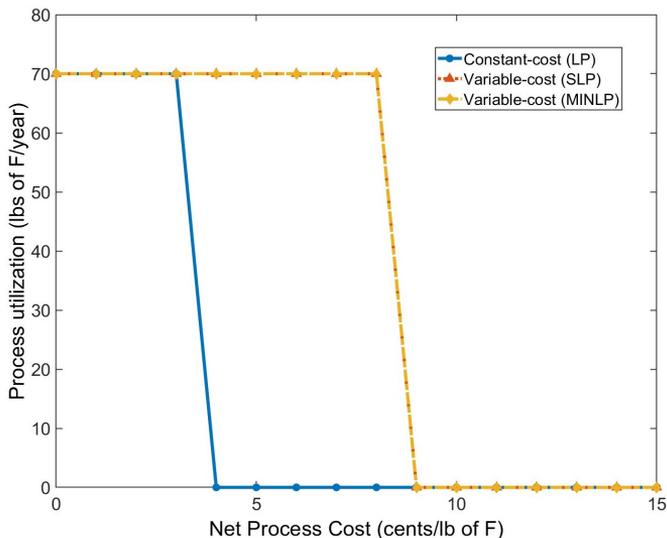


Figure 2: Process utilization for new technology P2005 as a function of its net process cost. Results for the SLP and MINLP methods are the same and overlap.

This can be seen in Fig. 2, which shows the level of adoption of P2005 (in terms of production rate) as a function of its net production cost. These results were obtained by solving the constant cost LP model, the SLP model, and the MINLP model with successively larger values of the P2005 net production cost, incremented by one cent/lb.

In the case of constant costs, where no price and costs changes are propagated throughout the network, the maximum adoption cost for the new process is just 3 cents/lb of F. For higher cost points, the new process is dropped from the optimal network configuration. This is the cost level for which the P2005-P5 pathway becomes cheaper than the P2-P3 one, even if the effect of cheaper F production is not propagated to P5. However, when this effect is propagated, the maximum adoption cost increases to 8 cents/lb of F. This is attributed to a price reduction in F, which is in turn propagated to the operating cost of P5. For example, at the insertion cost point of 5 cents/lb of F, the price of F decreases from 14 to 5 cents/lb, which leads to the cost of P5 decreasing from 13 to 6.7 cents/lb of G. This also leads to the price of G being reduced from 10 cents/lb to 6.7 cents/lb, since its largest producer switches from P3, which has a cost of 10 cents/lb of G, to P5. There is only one production level throughout the cost range where the new process is adopted and this is the same for all cases and is equal to 70 lbs of F/year. This is the amount needed to produce 100 lbs of G/year in P5 based on its specified stoichiometry, which is the value specified for the exogenous demand for G.

Comparing the results of the SLP and MINLP approaches for solving the nonlinear variable-cost problem, it is clear that both approaches yield the same solutions, in terms of adoption rates for the new process (shown in Fig. 2), the production rates of other processes, and the cost and price values (not shown). This, and experience on other small networks, indicates that the original SLP scheme employed in our previous work (Skouteris et al., 2021) was successful in accurately solving the variable-cost problem. In terms of efficiency, the SLP approach is fast and scales up well to the full industry network. The MINLP can be solved in about the same time for this small-scale network, but may not scale as well to the full network. This is a concern that will be further investigated.

5. Conclusions

The increased production of shale gas and, consequently, NGLs has led to a unique opportunity to expand the U.S. petrochemicals industry. Thus, to optimize the benefits of this expansion, it is important to have modeling and computational tools to aid the decision making that is driving this expansion. Network modeling is one such approach for assessing the impact of new technologies. In previous work using network models (Skouteris et al., 2021), we highlighted the importance of considering process costs to be variable, allowing them to respond to the addition of new processes into the industry network, and developed a new framework based on successive linear programming (SLP) to deal with this nonlinear problem. In this work, as an alternative to the SLP approach, we considered a new explicit MINLP formulation of the problem, and used it on a small-scale network example to validate the earlier SLP framework. Both approaches gave the same optimal solutions over a range of net process costs for an added process, indicating no loss of accuracy by using the SLP approach. Both approaches were very efficient computationally on the small-scale example problem, but whether the MINLP approach scales well to much larger problems is a concern. In future work, we will test the MINLP approach described here on the full-scale U.S. chemical industry network.

6. Acknowledgments

This paper is based upon work supported primarily by the National Science Foundation under Cooperative Agreement No. EEC-1647722 (CISTAR: NSF Engineering Research Center for Innovative and Strategic Transformation of Alkane Resources). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- S. E. Derosa, D. T. Allen, 2015. Impact of natural gas and natural gas liquids supplies on the United States chemical manufacturing industry: Production cost effects and identification of bottleneck intermediates. *ACS Sustainable Chemistry and Engineering* 3 (3), 451–459.
- S. E. DeRosa, D. T. Allen, 2016. Impact of New Manufacturing Technologies on the Petrochemical Industry in the United States: A Methane-to-Aromatics Case Study. *Industrial & Engineering Chemistry Research* 55 (18), 5366–5372.
- M. A. Duran, I. E. Grossmann, 1986. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming* 36 (3), 307–339.
- S. Fathi-Afshar, D. S. Maisel, D. F. Rudd, A. A. Treviño, W. W. Yuan, 1981. Advances in petrochemical technology assessment. *Chemical Engineering Science* 36 (9), 1487–1511.
- GAMS, 2021. The General Algebraic Modeling System (GAMS) Documentation. Accessed on 17 November 2021. URL <https://www.gams.com/>
- I. Giannikopoulos, A. Skouteris, D. T. Allen, M. Baldea, M. A. Stadtherr, 2022. Multi-objective optimization of production cost and carbon loss in the U.S. petrochemicals industry, To be presented at 14th International Symposium on Process Systems Engineering (PSE2021+), June 19–23, 2022, Kyoto, Japan.
- F. M. Scherer, D. Ross, 1990. *Industrial Market Structure and Economic Performance*, 3rd Edition. Houghton Mifflin.
- J. J. Siirola, 2014. The Impact of Shale Gas in the Chemical Industry. *AIChE Journal* 60, 810–819.
- A. Skouteris, I. Giannikopoulos, T. F. Edgar, M. Baldea, D. T. Allen, M. A. Stadtherr, 2021. Systems analysis of natural gas liquid resources for chemical manufacturing: Strategic utilization of ethane. *Industrial & Engineering Chemistry Research* 60 (33), 12377–12389.
- M. A. Stadtherr, D. F. Rudd, 1976. Systems study of the petrochemical industry. *Chemical Engineering Science* 31 (11), 1019–1028.
- U.S. Energy Information Administration, 2021. Natural gas explained: Where our natural gas comes from. Accessed on 15 Nov 2021. URL <https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php>
- M. Yang, F. You, 2017. Comparative techno-economic and environmental analysis of ethylene and propylene manufacturing from wet shale gas and naphtha. *Industrial & Engineering Chemistry Research* 56, 4038–4051.

A Novel Approach for Vehicle Fleet Sizing and Allocation under Uncertain Demand

Demian J. Presser^{a,b}, Vanina G. Cafaro^{a,b}, Diego C. Cafaro^{a,b}

^a*INTEC (UNL-CONICET), Güemes 3450, Santa Fe, Argentina*

^b*Fac. de Ingeniería Química (UNL), Sgo del Estero 2829, 3000 Santa Fe, Argentina*

Abstract

In this work, we propose a novel optimization model based on mixed integer linear programming (MILP) whose goal is to optimize the size and allocation of vehicle fleets under uncertain demand patterns. In contrast to previous contributions, the probability of meeting demand is a model decision and the approach is able to handle any probability distribution to characterize the number of trips required in the network. Furthermore, fleet procurement and contracting decisions for alternative types of vehicles are integrated into the model. The formulation allows to considerably reduce costs by avoiding oversizing and can be adapted to diverse demand characterizations and contracting terms from different industries. The proposed approach is able to tackle problems of real dimensions in very low computational times in comparison to previous works. A real-world case study involving maintenance operations in geographically spread assets is addressed to highlight the model capabilities and draw conclusions.

Keywords: fleet sizing, allocation, planning, optimization, uncertainty.

1. Introduction

Fleet management is a complex logistic process that involves the simultaneous planning of tactical and operational decisions. The sizing of vehicle fleets, their procurement policies and allocation to services are key decisions for modern industries that need to meet the demand of geographically distributed operations. Efficient fleet planning is an indispensable prerequisite to guarantee the supply of services, resources and equipment to keep operations running. Planning maintenance operations in offshore wind farms (Gundegjerde, 2015), sizing and allocating aircraft fleets to airline routes (Repko, 2017), organizing shipments in freight forwarding companies (Zak, 2008) or, in more general terms, optimizing transportation services through a network of origin and destination nodes using a fleet of vehicles (List, 2003) are challenging problems proposed in the literature. Fleet planning integrates short-term and long-term decisions aimed at fulfilling transportation demands that are often subject to uncertainty, usually leading planners to overestimate the number of required resources. An efficient vehicle procurement and utilization plan relies on solving critical tradeoffs between fleet acquisition and contracting costs, asset performance decay and penalties for unmet demand, among others, under a wide range of possible scenarios.

Many contributions have been made to solve this problem. Turnquist and Jordan (1986) propose a mathematical programming model to optimize the fleet size of reusable containers to match production and shipping schedules under deterministic and stochastic travel times. They conclude on the need of further treating demand uncertainty in mesh-like networks. Later on, List et al. (2003) propose a two-stage stochastic formulation for fleet sizing (first stage decision) and allocation to trips

(scenario dependent decision), accounting for uncertainty in the forecasted demand and fleet productivity. The authors highlight the tradeoff between fleet investment costs and the risk of unmet demand, which is solved through a computationally challenging approach. In turn, Wu et al. (2005) propose a linear programming model to address the rental car fleet sizing problem, deciding on the acquisition and sale of vehicles and the assignment and movement of vehicles to different locations to meet seasonal demands, at minimum total cost. Sha and Srinivasan (2016) make use of agent-based simulation to address the tank car fleet sizing problem in the chemical industry under deterministic demand. In 2020, Vanga and Venkateswaran develop an analytical model for sizing fleets of vehicles that carry generalized reusable articles. They consider cases of lost-sales and backorders, and assess the service level for different policies at a later stage. Nevertheless, resource allocation decisions are not considered. Most of the above-mentioned contributions are based on discrete probabilistic scenarios and stochastic modeling frameworks whose complexity grows rapidly with the number of possible outcomes. In most cases, service levels are evaluated after optimization and specific probability functions are used to characterize demand (Papier and Thonemann, 2008). In short, previous authors propose either analytical methods based on limiting assumptions or two-stage stochastic programming models that become intractable for rather small instances.

In contrast to previous approaches, we propose a novel MILP formulation aiming to optimize the size and allocation of vehicle fleets under uncertain demand patterns where the probability of meeting demand is a model decision. Moreover, the model is able to handle any probability distribution to characterize the number of trips required in the network. The optimal number of units to contract from different suppliers should be established according to the most convenient service level, seeking to minimize penalties for unmet demand and costs of contract agreements. The results show that an effective allocation of the contracted fleet to routes can significantly reduce the number of vehicles required for sufficiently high levels of service.

2. Problem Statement

The problem addressed in this work can be stated as follows. Given: (i) a set of nodes, (ii) random variables accounting for the total number of trips to be made between the nodes during the time horizon, (iii) the type of vehicle required in each trip, (iv) alternative contracts with vehicle suppliers, (v) unit costs for vehicle acquisition, rental, hiring/firing, and (vi) penalty costs for unsatisfied demand; we seek to optimally determine: (a) the fleet size for each type of vehicle; (b) the allocation of vehicles to trips over the network; (c) the service level to be provided, usually represented by the probability to meet demand; and (d) the plan of contracts (vehicles hiring/firing) along the time horizon, in order to minimize the total costs.

Figure 1 illustrates the main elements of the model, namely the nodes, the travel times between locations, and random variables accounting for demands, given in number of trips. For simplicity, a Gaussian distribution rectified at zero (Socci et al., 1998) is assumed in the illustrative example, although any probability distribution is eligible for parameterization. Although travel times are symmetrical in this figure, the model is able to consider different travel times according to origin-destination nodes and vehicle types. It is important to highlight that the duration of the trips is one of the critical drivers for the assignment of vehicles to routes. A good allocation will be one that takes maximum advantage of supplier time availability, reduces the number of vehicles in non-critical connections, and reinforces the level of service in the most critical paths.

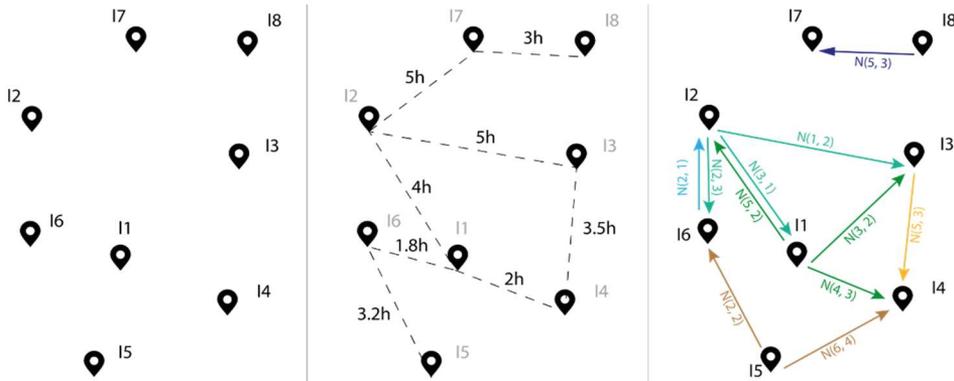


Figure 1 – Left: geographical distribution of the nodes. Middle: travel times between nodes (any direction). Right: random variables accounting for the demand of trips between nodes.

3. Mathematical Formulation

Let $i \in I$ stand for the nodes, $v \in V$ for the vehicle types and $p \in P$ for providers. Moreover, $\tau \in T$ and $k \in K$ represent daily and monthly periods, respectively. Finally, $s \in SL$ are discrete service levels. Two general contracting modes are proposed in the formulation and validated with our industry partner. The first is established on an annual basis, where the supplier makes a certain number of vehicles available for every day of the next year. On the other hand, monthly additions to the fleet are possible through flexible contracts. The objective function is given in Eq. (1), where $QF_{v,p}$ is the number of vehicles to be contracted on an annual basis, and $QV_{v,p,k}$ the number of available vehicles through flexible contracts. Unit costs $cc_{v,p}$, $ccv_{v,p}$ and $ccvt_{v,p}$ are monthly rates for vehicles of type v from supplier p accounting for fixed annual contracting, flexible contracting and flexible contract termination charges, respectively. Lastly, $CDI_{i,j,t}$ stands for the total expected cost due to daily unmet demand, under the selected contracting scheme and fleet allocation.

$$\text{Min } z = \sum_{v \in V} \sum_{p \in P} [QF_{v,p} cc_{v,p} + \sum_{k \in K} (QV_{v,p,k} ccv_{v,p} + QVT_{v,p,k} ccvt_{v,p}) \sum_{i \in I} \sum_{j \in J} \sum_{\tau \in T(k)} CDI_{i,j,\tau}] \quad (1)$$

Each supplier provides vehicles serving during a specific number of hours per day ($hdisp_{v,p}$). The number of vehicles to be procured through different contracts will define the overall number of hours available to meet demands (see RHS of Eq. 2). In turn, the service hours required to meet a given service level may be covered by allocating vehicles to trips over the network. Travel time requirements are defined by assigning vehicles to trips, as in the RHS of Eq. (2). $X_{v,i,j,k}$ is the number of trips from i to j assigned to vehicles of type v during the time period k while parameter $d_{i,j}$ is the travel time between nodes i and j . Then, Eq. (3) accounts for vehicle flow balances.

$$\sum_{i \in I} \sum_{j \in I} (d_{i,j} X_{v,i,j,k}) \leq \sum_{p \in P} (hdisp_{v,p} (QF_{v,p} + QV_{v,p,k})) \quad \forall v, k \quad (2)$$

$$\sum_{\substack{i \in I \\ i \neq j}} X_{v,i,j,k} = \sum_{\substack{i \in I \\ i \neq j}} X_{v,j,i,k} \quad \forall v, j, k \quad (3)$$

The proposed MILP is capable of handling uncertainty in travel demands between nodes for each type of vehicle. We recognize two sources of uncertainty: (a) the intrinsic variability of operations due to external phenomena (e.g. weather conditions),

roadblocks, unforeseen operating inconveniences, among others, which result in the random behavior of the demands around the expected value for each month; and (b) the prediction error, typical in all forecasting tools. Both values are usually synthesized in a dispersion parameter that is transferred as input data to the optimization model. It is important to highlight that the demand uncertainty can be modeled with any probability distribution. Figure 2 illustrates the demand profile in a given network of locations, for a given month and vehicle category. The characterization of the travel demands between nodes is shown for different pairs. The example is based on rectified normal distributions. Each directed arc has an associated mean and a standard deviation, which represent the expected number of trips required from the origin to the destination, and its variability. These parameters are used to build a probability distribution to estimate the cost of unmet demand for each possible service level, as shown in the following equations. Notice that the variability may significantly differ depending on the type of operations to develop in each node.

To compute the cost of the expected unmet demand for a given contracting plan, we illustratively show the calculations based on a normal distribution rectified at zero. Extensions are straightforward. Let $\bar{t}d_{i,j,v,k}$ and $tsd_{i,j,v,k}$ be the mean and standard deviation of the demand of trips of vehicle v between i and j , while cut_v represents the cost of missing one trip of vehicle v . Hence, the cost of expected unmet demand during period k if a service level sl is adopted, can be stated as follows:

$$cutd_{i,j,v,k,sl} = \int_{t=nt_{v,i,j,k,sl}}^{\infty} \frac{1}{tsd_{i,j,v,k}\sqrt{2\pi}} e^{-\frac{(t-\bar{t}d_{i,j,v,k})^2}{2tsd_{i,j,v,k}^2}} cut_v(t - nt_{v,i,j,k,sl}) dt \quad \forall i, j, v, k \quad (4)$$

In Eq. (4) $nt_{v,i,j,k,sl}$ accounts for the number of trips required to reach the service level sl , according to the number of standard deviations over the mean (see Eq. 5). Notice that Eq. (4) can be conservatively approximated by a finite summation as in Eq. (6), which is favored by the integer nature of variable t (i.e., number of trips). $MAXT_{v,i,j,k}$ is the maximum number of trips to achieve full coverage ($sl \sim 100\%$).

$$nt_{v,i,j,k,sl} = \bar{t}d_{i,j,v,k} + \gamma_{sl} tsd_{i,j,v,k} \quad \forall i, j, v, k, sl \quad (5)$$

$$cutd_{i,j,v,k,sl} = \sum_{t'=nt_{v,i,j,k,sl}}^{MAXT_{v,i,j,k}} \frac{1}{tsd_{i,j,v,k}\sqrt{2\pi}} e^{-\frac{(t'-\bar{t}d_{i,j,v,k})^2}{2tsd_{i,j,v,k}^2}} cut_v(t' - nt_{v,i,j,k,sl}) \quad \forall i, j, v, k \quad (6)$$

The optimization of the service level implies solving the tradeoff between the costs of contracting more vehicles to be ready for operation, and the expected costs due to unmet demand. The latter are defined by constraint (7), which imposes a minimum charge depending on the service level selected for every link $i-j$ and vehicle v during month k . That decision is made through the 0-1 variable $Y_{v,i,j,sl,k}$. In turn, Eq. (8) imposes a lower bound to the number of trips $X_{v,i,j,k}$ that are required to reach a service level of kv_{sl} standard deviations over the mean. Note that $KS_{v,i,j,k}$ is a decision variable that equals kv_{sl} if the service level sl is adopted for the link $i-j$ and vehicle v over month k (Eq. 9).

$$CDI_{v,i,j,\tau} \geq \sum_{sl \in SL} cutd_{v,i,j,sl,k} Y_{v,i,j,sl,k} \quad \forall v, i, j, k, \tau \in T(k) \quad (7)$$

$$X_{v,i,j,k} \geq \bar{t}d_{v,k,i,j} + KS_{v,i,j,k} tsd_{v,k,i,j} \quad \forall v, i, j, k \quad (8)$$

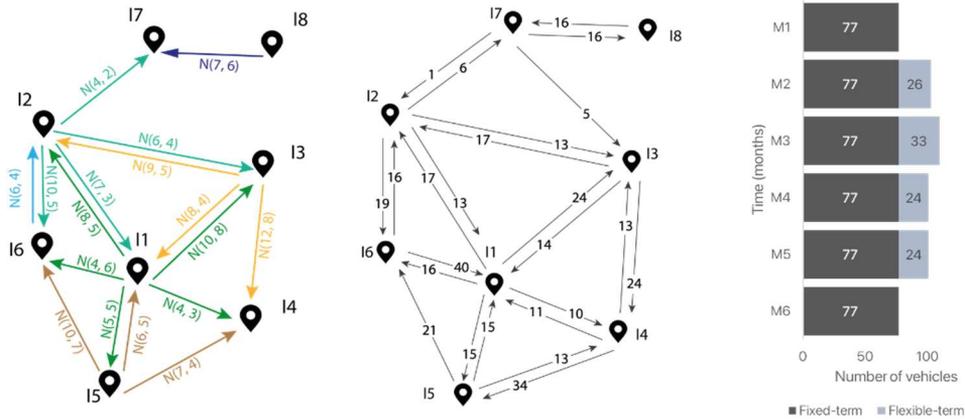


Figure 2 – Optimal solution for one of the vehicle types in the proposed case study.

$$KS_{v,i,j,k} = \sum_{sl \in SL} kv_{sl} Y_{v,i,j,sl,k} ; \sum_{sl \in SL} Y_{v,i,j,sl,k} = 1 \quad \forall v, i, j, k \quad (9)$$

4. Results and Discussion

An illustrative case is proposed in this section to show the potential of the tool. We consider two types of vehicles with weekly travel demands for each pair of nodes that are forecasted for the next 6 months. Each demand is characterized by a rectified normal distribution that captures the possible outcomes of the number of trips required for every week of that month. Figure 2 (left) shows the demand characterization for month 2, whereas the duration of the trips for both vehicles between the nodes is presented in Figure 1 (center). Decision variables to be optimized are the number of vehicles to be contracted on a fixed basis (for the 6 months), the number of vehicles to be hired on a flexible monthly basis, and the service level to be provided by assigning vehicles to every path in the network. The monthly contracting cost is set at US\$ 10,000 and US\$ 14,000 per vehicle, for fixed and flexible contracts, respectively. A penalty cost of US\$ 5,000 is considered for every unsatisfied trip. For this example with 8 nodes the MILP (after preprocessing) comprises 1619 0-1 variables, 329 continuous variables and 331 constraints. The formulation is coded in GAMS and solved using CPLEX 31.5 in a CPU time of 40 s. Figure 2 also illustrates the optimal solution obtained. At the right of Figure 2 we show the optimal contracting scheme for each month.

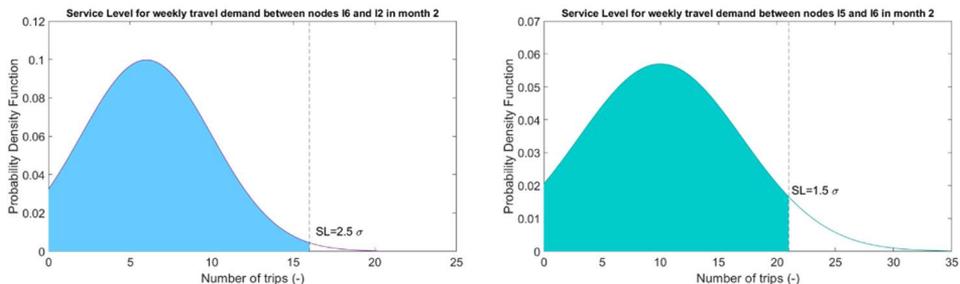


Figure 3 – Optimal service level for some of the links in the illustrative case (month 2).

A total of 77 vehicles are engaged on a fixed basis, and some additional units are hired on a variable basis during months 2, 3 and 4 in order to satisfy demands with the optimal service level. Similarly, Figure 3 illustrates the optimal service level obtained for the routes connecting nodes I6-I2 and I5-I6 during month 2, respectively. It is interesting to note that some routes may reach very high service levels due to the need to pass through those links to fulfill other services. On the other hand, the model has the capability of incorporating different suppliers for the same type of vehicle, accounting for their quality/reliability indicators, and integrating other benefits such as joint contracting (several types of vehicles from the same supplier) and quantity discounts.

5. Conclusions

We have developed a novel MILP mathematical programming model to optimize the size of vehicle fleets and their allocation to trips in a network of nodes asking for transportation services. The model accounts for demand uncertainty with no need to generate scenarios and solve large stochastic formulations. Moreover, the service levels are optimally selected to meet the demand of trips along each pair of nodes. In contrast to previous contributions, the formulation proves to be computationally efficient, flexible and capable of handling demand characterizations under any probability distribution. The model overcomes current industry practices that arbitrarily oversize fleets to avoid non-compliance with critical services. Significant reductions in logistic costs and emissions are observed, yielding a certainly better performance of the supply chain. This is achieved by optimally setting flows, minimizing idle times and avoiding unjustified trips. Since fleet sizing and allocation decisions are made in a single stage, we are currently performing further experiments to assess the quality of our solutions in comparison to two-stage stochastic models where vehicle allocations are recourse (wait-and-see) actions. Future research aims to add routing and inventory decisions to gain greater detail in the solutions and validate the model accuracy with real data.

References

- C. Gundegjerde, I.B. Halvorsen, E.E. Halvorsen-Weare, L.M. Hvattum, L.M. Nonas, 2015. A stochastic fleet size and mix model for maintenance operations at offshore wind farms, *Transp. Res. Part C, Emerging Tech.*, 52, 74-92.
- M. Repko, B.F. Santos, 2017. Scenario tree airline fleet planning for demand uncertainty, *Journal of Air Transport Management*, 65, 198-208.
- J. Zak, A. Redmer, P. Sawicki, 2008. Multiple objective optimization of the fleet sizing problem for road freight transportation. *Journal of Advanced Transportation*, 42, 379-427.
- G.F. List, B. Wood, L.K. Nozick, M.A. Turnquist, D.A. Jones, E.A. Kjeldgaard, C.R. Lawton, 2003. Robust optimization for fleet planning under uncertainty, *Transp. Res. Part E, Logistics and Transp. Review*, 39, 209-227.
- M.A., Turnquist, W.C., Jordan, 1986. Fleet Sizing under Production Cycles and Uncertain Travel Times. *Transportation Science*, 20, 227-236.
- P. Wu, J.C., Hartman, G.R., Wilson, 2005. An integrated model and solution approach for fleet sizing with heterogeneous assets. *Transportation Science*, 39, 87-103.
- M. Sha, R. Srinivasan, 2016. Fleet sizing in chemical supply chains using agent-based simulation. *Computers & Chemical Engineering*, 84, 180-198.
- R. Vanga, J. Venkateswaran, 2020. Fleet sizing of reusable articles under uncertain demand and turnaround times. *European Journal of Operational Research*, 285, 566-582.
- F. Papier, U.W. Thonemann, 2008. Queuing models for sizing and structuring rental fleets. *Transportation Science*, 42, 302-317.
- N.D. Soccia, D.D., Lee, H. Sebastian Seung, 1998. The rectified Gaussian distribution. *Advances in neural information processing systems* 10, 350-356.

Fresh vs frozen: assessing the impact of cryopreservation in personalised medicine

Niki Triantafyllou^a, Andrea Bernardi^a, Matthew Lakelin^b, Nilay Shah^a, Maria M. Papathanasiou^{a*}

^a*Sargent Centre for Process System Engineering, Imperial College London, SW72AZ, London, United Kingdom*

^b*TrakCel Limited, 10/11 Raleigh Walk, Cardiff, CF10 4LN UK*
maria.papathanasiou11@imperial.ac.uk

Abstract

Chimeric Antigen Receptor (CAR) T cell therapy is a type of patient-specific cell immunotherapy demonstrating promising results in the treatment of aggressive haematological malignancies. Autologous CAR T cell therapies are based on bespoke manufacturing lines and distribution nodes that are exclusive to the production and delivery of a single therapy. Given their patient-specific nature, they follow a 1:1 business model that challenges volumetric scale up, leading to increased manufacturing and distribution costs. Manufacturers aim to guarantee the in-time delivery and identify ways to reduce the production cost with the ultimate objective of releasing these innovative therapies to a bigger portion of the population. In this work, we investigate upstream storage to the supply chain network as means to introduce greater flexibility in the modus operandi. We formulate and assess different supply chain networks via a Mixed Integer Linear Programming model.

Keywords: CAR T cell therapy; supply chain optimisation; MILP; personalised medicine; cryopreservation

1. Introduction

Chimeric Antigen Receptor (CAR) T cell therapy is an emerging type of ex vivo autologous cell therapy, where the patient's T cells are collected and are genetically modified to express the CAR (Vormittag *et al.*, 2018). This enables them to better identify and kill cancer cells. The therapy is then administered to the patient. The unprecedented success in early clinical trials involving patients with acute lymphoblastic leukaemia (ALL) and B cell lymphomas have led to historic regulatory approvals from the U.S. Food and Drug Administration (FDA) in 2017 and the European Medicines Agency (EMA) in 2018. Five autologous CAR T cell therapies have been granted approval so far and their demand is expected to rapidly grow in the following decade (UPMC, 2021). They are currently offered at a relatively high price, potentially attributed to the 1:1 business model describing their lifecycle, that varies between \$300,000 and \$475,000 (Spink *et al.*, 2018). Presently, cell therapies are administered to approximately 5,000 patients worldwide. However, projections show that up to 50,000 patients per year will be receiving cell and gene therapies with up to 60 products in 2030 (Quinn *et al.*, 2019).

A unique feature of CAR-T therapies is that the patient's cells are used as raw material and therefore the therapy manufacturing and distribution does not allow volumetric scale-up, posing challenges towards commercialisation (Papathanasiou *et al.*, 2020).

Furthermore, the short shelf life of the therapy introduces tight time constraints challenging distribution and handling, especially under high-demand scenarios. The turnaround time of the therapies from leukapheresis to treatment administration can vary between 15 and 24 days for the commercially available treatments (Nucleus Biologics, 2021). Timely therapy delivery is of utmost importance for the patients as it can significantly impact therapy efficiency (Han *et al.*, 2021). To overcome these limitations, digital tools that utilise the principles of mathematical modelling and optimisation are used to assist decision-making by coordinating the different tasks and identifying the optimal supply chain network structures (Sarkis *et al.*, 2021a). The complexity of the CAR T cell supply chain is reflected by the product's lifecycle (Sarkis *et al.*, 2021b). The main steps of the CAR T cell therapy chain are: (a) leukapheresis, (b) manufacturing, (c) Quality Control, (d) therapy administration under hospital admission.

Given the autologous and sensitive nature of CAR-T therapies, novel, decentralised manufacturing models are considered promising alternatives to a traditional centralised approach, since small scale local manufacturing facilities offer greater flexibility and can significantly reduce transportation costs (Harrison *et al.*, 2018). On the other hand, intermediate storage of the therapies may offer greater flexibility in the operation of the overall distribution network. As far as transportation and storage are concerned, the samples can either be transported fresh (-80 °C) or cryopreserved (-180 °C). In general, cryopreservation is more expensive but allows to extend the shelf life up to 14 days, and therefore is the preferred option when more flexibility is desirable. Cryopreservation currently is conducted at the leukapheresis clinical centres with hospitals not providing this service. The cryopreserved therapy needs to be thawed before it reaches the manufacturing site, something that adds extra complexity to the logistics of the process (Avramescu *et al.*, 2021).

In this work, we present a Mixed Integer Linear Programming (MILP) model for the CAR-T cell therapy supply chain optimisation considering different randomised demand scenarios of up to 2,000 therapies per year. The main objective is to assess the impact of cryopreservation on the supply chain performance in terms of flexibility and risk of loss of the therapy due to short shelf life of non-cryopreserved products. The supply chain networks designed here are assessed with respect to two key performance indicators: (a) average production cost, and (b) average response treatment time.

2. Materials and methods

The mathematical model used in this work is a MILP model that describes the CAR T cell therapy supply chain (Bernardi *et al.*, 2022). The model assists the identification of the optimal supply chain network structure, mode of transport, assignment and sequence of the therapies to the manufacturing facilities, ensuring in-time delivery and minimum therapy cost (Table 1). The supply chain network includes 5 nodes: namely, leukapheresis site, storage, manufacturing site, quality control (QC), and hospital. The therapy lifecycle starts with the patient being assigned to a leukapheresis site, where T cells are separated from the bloodstream. Next, the leukapheresis sample is shipped fresh (-80°C) or frozen (-180 °C, cryopreserved) to the manufacturing site for manufacturing, followed by QC. Finally, the CAR T cell therapy is transported fresh to the hospital for administration to the patient.

In this study, we investigate two cases: fresh or frozen handling of the therapy upstream of the network (Figure 1). In the frozen case, the therapy is cryopreserved, and optional storage is available for up to 8 days. The upper bound of 8 days was chosen as a suitable

window not compromising patient experience. Storing therapies before manufacturing may offer greater flexibility and improved scheduling of a manufacturing facility with given capacity. Bernardi *et al.* proved essential to optimally coordinate manufacturing and distribution tasks. On the other hand, when the therapy is transported fresh (-80 °C) storage is bypassed because of the limited shelf life. The supply chain model aims at simultaneous planning and scheduling with its objective being to minimise the total cost over a long-term planning horizon (annually), while operating the supply chain in the short-term (daily). In addition, the minimisation of the turnaround time appears as a constraint in the model.

Table 1: Overview of the model formulation.

Index	Mathematical Formulation	Description
Objective function		
(1)	$\min C_{total\ cost} = C_{manufacturing} + C_{transport} + C_{storage} + C_{quality\ control}$	Total cost of therapies
Constraints		
(2)	$TRT_p = t_{delivery} - t_{start} \leq U^t$	Return time of therapy
(3)	$CAP_{m,t} = FCAP_m - \sum_p INM_{p,m,t}$	Capacity constraint
(4)	$X1_{c,m} \leq E1_m, \forall c, m, \quad X2_{m,h} \leq E1_m, \forall c, h, \quad \sum_m E1_m \leq U^M$	Network constraints
(5)	$INC_{p,c,t} = OUTC_{p,c,t} + T_{LS}, \quad \forall p, c, t$ $INS_{p,m,d,t} = OUTS_{p,m,d,t} + T_{STOR_d}, \quad \forall p, m, d, t$ $INM_{p,m,t} = OUTM_{p,m,t} + T_{MFE} + T_{QC}, \quad \forall p, m, t$	Sample balances at each node (leukapheresis, storage, manufacturing)
(6)	$LSR_{p,c,m,j,t} = LSA_{p,c,m,j,t} + TT1_j, \quad \forall p, c, m, j, t,$ $FTD_{p,m,h,j,t} = MSO_{p,m,h,j,t} + TT2_j, \quad \forall p, m, h, j, t$	Transport constraints

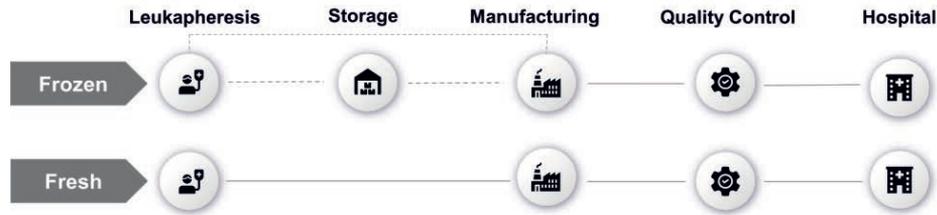


Figure 1. CAR T cell supply chain network with 5 or 4 nodes for the frozen and fresh cases respectively.

The supply chain is tested for different demand scenarios of up to 2,000 patients per year generated by an in-house bi-level decomposition algorithm. The model parameters, such as product demand scales, cost coefficients and current and planned manufacturing facilities sizes are assumed to be deterministic and were obtained through expert discussions with TrakCel Ltd, something that enables the development of industrially accurate cases. This work considers 4 leukapheresis sites and 4 hospitals in the UK and 6 manufacturing sites found in the UK, Europe and US. Following standard European and UK procedures of public or semi-public healthcare systems, it is assumed that the choice of collaborating hospitals and leukapheresis centres is not entirely under the manufacturer’s control. Hence, the two corresponding nodes are provided as model inputs and not as decision variables. For the leukapheresis sites, a capacity of 8 patients daily is assumed, considering that the leukapheresis procedure lasts 2-3 hours per patient. For the hospital site, the capacity is not a bottleneck for the administration of the therapy and therefore no upper limit is provided. The manufacturing sites have a capacity of up to 4, 10, or 31 parallel lines, and a forward-looking scenario of 7 days of manufacturing

duration is assumed. Quality control takes place in the manufacturing site and the hospital and leukapheresis centres are different facilities located in the same area. Transportation can happen either within 24 or 48 hours in case of fresh therapy and 24, 48 or 72 hours in case of frozen therapy, regardless of the transport mode (i.e. truck, rail etc). This is to align with standard practice in the cell therapy market where courier contracts are cost based on the estimated delivery time. The fresh transportation costs are 50% less than the frozen ones.

The model is implemented in Python 3.7.1 and Pyomo 6.1.2 and solved with CPLEX 12.9. All computational experiments were performed in a 24-core Xeon E5-2697 machine with 196GB. Here we present the results for the two cases (fresh and frozen) for demand profiles of increasing size. All scenarios are solved using a bi-level decomposition algorithm that is capable of reducing the computational complexity of the model and providing optimal solutions at reduced CPU times (Erdirik-Dogan and Grossmann, 2008; Terazzas-Moreno and Grossmann, 2011). The CPU time ranged between 12 s and 950 s.

3. Results and discussion

The proposed model is used for the design and assessment of the optimal supply chain network for four different demand scales (200, 500, 1,000 and 2,000 patients annually). The demand scenarios considered are assumed to be homogeneous and repeated every trimester. The model is allowed to invest in the establishment of up to two manufacturing facilities for the optimal network structure.

Figure 2 depicts the results for the four scenarios in terms of average production cost and return time per therapy. The average cost per therapy is broken down into transport cost, manufacturing cost and Quality Control cost. The storage cost is negligible and thus not presented in the graph. Figure 3 illustrates the utilisation of the manufacturing facilities for the scenario of 500 therapies per year.

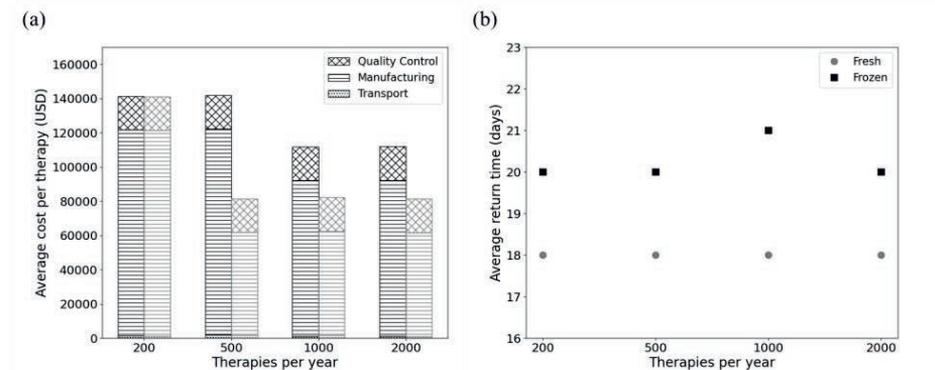


Figure 2. Comparison of fresh (black) and frozen (grey) transportation of the CAR T cell therapies based on (a) average cost per therapy (USD), where the cost is broken down in transport cost, manufacturing cost and Quality Control cost and (b) average return time per therapy.

In Figure 2a, we can observe that for the scenario of 200 patients per year the average cost per therapy is identical for fresh and frozen transportation, whereas in all the other cases the frozen transportation leads to a significantly reduced cost per therapy. This can be attributed to the incorporation of storage in the model, as it allows for improved manufacturing planning. The manufacturing cost is the biggest contributor to the total cost, with QC cost being fixed in every scenario and transportation costs being negligible.

The average return time for fresh transportation is 18 days for all demand scenarios, and 20 days for the majority of frozen transportation scenarios (Figure 2b). This is expected as in the frozen case the model can choose to store samples for up to 8 days and choose between 3 modes of transport (24, 48 or 72 hours). The cost vs return time trade-off can be observed in Figure 2; when the total return time of therapy decreases, the cost increases.

The steepest decrease of 43% of the average therapy cost takes place in the 500 therapies per year scenario. The model proposes the establishment of two manufacturing facilities (m3 and m6) in the case of fresh therapy with a total of 20 parallel lines (Figure 3a), while in the case of frozen therapy it chooses to build only facility m3 with a total of 10 parallel lines (Figure 3b). From Figure 3a we can observe that the facility m6 in the fresh case is substantially underutilised, resulting in an increased manufacturing cost per therapy. In the frozen transportation case, m3 is roughly always operating at maximum capacity leading to decreased manufacturing costs. Therefore, the reduction in the average therapy cost is attributed to the supply chain network structure and utilisation of the manufacturing facilities.

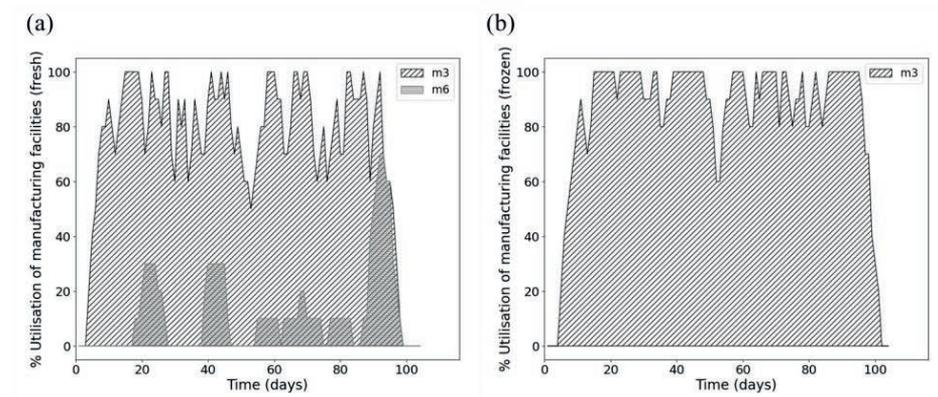


Figure 3. Utilisation of manufacturing facilities to be built for (a) fresh transportation and (b) frozen transportation in the 500 therapies per year scenario.

4. Conclusions

In this work, the impact of cryopreservation was assessed compared to fresh therapy transportation by the utilisation of a digital tool based on mathematical modelling and optimisation. The MILP model provided useful insights for the design and optimal supply chain networks in personalised therapies. Cryopreservation increases the shelf life of the samples, allowing to utilise an intermediate storage between the leukapheresis centre and the manufacturing facility. The results indicate that the extra flexibility granted by cryopreservation almost always leads to a decrease in the average therapy cost under the investigated demand scenarios, but the average vein-to-vein time slightly increases. Overall, the model identifies a trade-off between cost and average return time. The cost reduction achieved in the frozen case is linked to an improved manufacturing planning, which can potentially debottleneck the complex supply chain and logistics, and ultimately decrease the risk of loss of the therapies- especially for high-demand scenarios. Future work will aim to assess the effect of variable manufacturing time and failure rates of fresh versus frozen manufacturing on the supply chain robustness and on the average therapy cost.

5. Acknowledgements

Funding from the UK Engineering & Physical Sciences Research Council (EPSRC) for the Future Targeted Healthcare Manufacturing Hub hosted at University College London with UK university partners is gratefully acknowledged (Grant Reference: EP/P006485/1). Financial and in-kind support from the consortium of industrial users and sector organisations is also acknowledged.

References

1. FDA-approved CAR T-cell Therapies | UPMC Hillman. <https://hillman.upmc.com/mario-lemieux-center/treatment/car-t-cell-therapy/fda-approved-therapies> (accessed Nov. 05, 2021)
2. Kymriah vs. Yescarta [UPDATED] | Nucleus Biologics.” <https://nucleusbiologics.com/resources/kymriah-vs-yescarta/> (accessed Feb. 10, 2021)
3. A. Avramescu, R. Allmendinger, M. López-Ibáñez, 2021, Managing Manufacturing and Delivery of Personalised Medicine: Current and Future Models, arXiv, 2105.12699v1
4. A. Bernardi, M. Sarkis, N. Triantafyllou, M. Lakelin, N. Shah, M. M. Papathanasiou, 2022, Assessment of intermediate storage and distribution nodes in personalised medicine, *Comput. Chem. Eng.*, 157, 107582
5. M. Erdirik-Dogan, I. E. Grossmann, 2008, Simultaneous planning and scheduling of single-stage multi-product continuous plants with parallel lines, *Comput. Chem. Eng.*, 32(11), 2664–2683
6. R. P. Harrison, S. Ruck, Q. A. Rafiq, N. Medcalf, 2008, Decentralised manufacturing of cell and gene therapy products: Learning from other healthcare sectors, *Biotechnol Adv.*, 36(2), 345–357
7. D. Han, Z. Xu, Y. Zhuang, Q. Qian, 2021, Current progress in CAR-T cell therapy for hematological malignancies, *J Cancer.*, 12(2), 326–334
8. M. M. Papathanasiou, C. Stamatias, M. Lakelin, S. Farid, N. Titchener-Hooker, N. Shah., 2020, Autologous CAR T-cell therapies supply chain: challenges and opportunities?, *Cancer Gene Ther.*, 27, 799–809
9. C. Quinn, C. Young, J. Thomas, M. Trusheim, 2019, Estimating the Clinical Pipeline of Cell and Gene Therapies and Their Potential Economic Impact on the US Healthcare System, *Value in Health*, 22(6), 621–626
10. P. Vormittag, R. Gunn, S. Ghorashian, F. S. Veraitch, 2018, A guide to manufacturing CAR T cell therapies, *Curr. Opin. Biotech.*, 53, 164–181
11. M. Sarkis, A. Bernardi, N. Shah, M. M. Papathanasiou, 2021, Decision support tools for next-generation vaccines and advanced therapy medicinal products: present and future, *Curr. Opin. Chem. Eng.*, 32(1), 100689
12. M. Sarkis, A. Bernardi, N. Shah, M. M. Papathanasiou, 2021, Emerging Challenges and Opportunities in Pharmaceutical Manufacturing and Distribution, *Processes*, 9(457), 267–277
13. K. Spink, A. Steinsapir, 2018, The long road to affordability: a cost of goods analysis for an autologous CAR-T process, *Cell Gene Ther. Insights*, 4(11), 1105–1116
14. S. Terrazas-Moreno, I. E. Grossmann, 2011, A multiscale decomposition method for the optimal planning and scheduling of multi-site continuous multiproduct plants, *Chem. Eng. Sci.*, 66(19), 4307–4318

A novel hybrid algorithm for scheduling multipurpose batch plants

Dan Li,^a Dongda Zhang,^a Nan Zhang,^a Liping Zhang,^b Jie Li^a

^a*Centre for Process Integration, The Department of Chemical Engineering, The University of Manchester, Manchester M13 9PL, UK*

^b*Wuhan University of Science and Technology, Wuhan, Hubei, 430081 P. R. China*

Abstract

A novel hybrid algorithm integrating advantages of genetic algorithm (GA) on computational efficiency and sequence-based mixed-integer linear programming (MILP) model on solution optimality is proposed for industrial-scale scheduling problems of multipurpose batch plants. The computational results show that the hybrid algorithm has strong convergence and is efficient to generate the same or better schedules with less computational efforts compared to the existing methods.

Keywords: Scheduling; Multipurpose batch plants; Hybrid algorithm; Genetic algorithm

1. Introduction

Scheduling multipurpose batch plants is challenging due to the process flexibility and layout complexity. Specifically, products are produced using product-specific processing steps with batch splitting, mixing, and recycling allowed. A variety of mixed-integer linear programming (MILP) models have been proposed, including sequence-based and time-grid-based models (Harjunkoski et al., 2014). Discrete-time models confirm its capability on processes with time-varying availability but fail to guarantee data accuracy due to discretization errors. Unit-specific event-based models are prominent on solution optimality. However, they are hard to solve industrial-scale problems as the appropriate number of event points is unknown a priori. Although several decomposition methods (Nishi et al., 2010) have been attempted, generation of optimal or near-optimal solutions with low computational costs for large-scale scheduling is still an unresolved issue.

Hybrid algorithm is promising to address large-scale problems as integrating advantages of multiple approaches. Recently, a discrete-continuous algorithm (Lee and Maravelias, 2018) was developed to generate approximate solutions quickly in first stage and then refine the solution to optimality. Metaheuristics are superior in high solution quality and computational efficiency. He and Hui (2010) and Woolway and Majozi (2018) demonstrated the GAs are efficient to schedule multipurpose batch plants in short time frames due to their inherent parallelism and strong global search capability. Therefore, in this work, we take advantages of GA on low computational costs and integrate it with a sequence-based MILP model to generate near-optimal or optimal solutions for scheduling of industrial multipurpose batch plants. The computational results show that the designed algorithm can generate the same or better solutions and reduce the computational time, over one order of magnitude in some instances, compared to existing methods.

2. Genetic algorithm

A novel GA is designed to generate suboptimal or optimal solutions using short computing time. In GA, an encoding set Pop_c of individual chromosome c is defined to represent solutions for a given problem. One chromosome $c = \{c^1, c^2, c^3\}$ comprises of

three parts representing production sequence, unit assignment and task repetition, respectively. In the first part, a variable $c_n^1 = p \in \mathbf{S}^P$ represents product p going to be manufactured at n -th production step in the production. In each production step, one batch of the selected product p is produced, and one product may be produced for multiple batches to satisfy order demands. In the second part, $c_{mi}^2 = j \in \mathbf{J}_i$ is the assigned unit j for batch m of task i that have multiple feasible units. In the third part, c_o^3 is a binary variable. When $c_o^3 = 1$, the o -th performed task would be repeated. c^3 is designed to influence the sequence of tasks performed to produce different batches of products.

A decoding algorithm is designed to synthesize solutions based on above solution representation. It steps iteratively through c^1 to produce products until demand requirements are met. In each production step, the producing task $i \in \mathbf{I}_{S^P}^P$ for the product p must be performed and other tasks providing intermediate states would be performed when the state is insufficient. Unit assignments are determined by c^2 , and whether to repeat the performed task is decided by c^3 . For minimization of makespan, a heuristic rule ‘earliest starting strategy’ is adopted to start every batch of tasks as early as possible.

$$f_c = (1/MS_c)^{20} \quad (1)$$

Fitness value of individual c is calculated using Eq.(1), where MS is makespan. The roulette wheel method is adopted to select parents who would be subjected to two-point crossover and two-point mutation. A knowledge-guided search (Zheng and Wang, 2018) is incorporated to adjust production sequence and unit assignment of child chromosome based on experiential possibility. Pr_{np}^g denotes the experiential possibility of product p is produced at n -th production step (i.e. $c_n^1 = p$) at g -th generation. At each generation, the experiential possibility is updated based on memory and experiences from NF elitists in current generation, as indicated in Eq.(2), where $I_{np}^{sc} = 1$ if $c_n^1 = p$ for the sc -th selected elitist. Evolution of the GA terminates until maximum generation is reached.

$$Pr_{np}^g = (1 - a) \cdot Pr_{np}^{g-1} + a \cdot \frac{\sum_{sc=1}^{NF} I_{np}^{sc}}{NF} \quad (2)$$

3. Sequence-based MILP formulation

Sequence-based MILP formulation is developed where two binary variables $X_{imi'm'}$ and $X_{S_{imi'm's}}$ are defined to denote the sequential relations between batch m of task i and batch m' of task i' on unit $j \in (\mathbf{J}_i \cap \mathbf{J}_{i'})$ and storage $s \in (\mathbf{S}_i \cap \mathbf{S}_{i'})$, respectively. A binary variable z_{im} is defined to denote if batch m of task i is performed. A task that is allowed to be processed in multiple processing units is divided into different tasks. All batches of a task should be processed in sequence, as constrained in Eq.(3).

$$z_{im} \leq z_{i(m-1)} \quad \forall i, m > 1 \quad (3)$$

Sequence constraints are formulated as Eq.(4-6). If the batch m of task i precedes one batch m' of the task i' on a unit j , it must also precede the batch $(m' + 1)$ of task i' , as indicated in Eq.(4). Similarly, the batch m' of task i' succeeds batch $(m - 1)$ of task i if it is processed after the batch m of task i , as Eq.(5). Eq.(6) expresses the sequential relation between two batches of different tasks i and i' performed in one same unit.

$$X_{imi'(m'+1)} \geq X_{imi'm'} \quad \forall j, i \in \mathbf{I}_j, i' \in \mathbf{I}_j, i \neq i', m, m' < M \quad (4)$$

$$X_{i(m-1)i'm'} \geq X_{imi'm'} \quad \forall j, i \in \mathbf{I}_j, i' \in \mathbf{I}_j, i \neq i', m > 1, m' \quad (5)$$

$$X_{imi'm'} + X_{i'm'im} \geq z_{im} + z_{i'm} - 1 \quad \forall j, i \in \mathbf{I}_j, i' \in \mathbf{I}_j, i < i', m, m' \quad (6)$$

Batch size b_{im} is bounded by the maximum B_i^{max} and minimum B_i^{min} capacity.

$$B_i^{min} \cdot z_{im} \leq b_{im} \leq B_i^{max} \cdot z_{im} \quad \forall i, m \quad (7)$$

Eq.(8) indicates the duration constraint, where a_i and β_i are fixed and variable terms of processing time, respectively. If a state s is subject to zero-wait ($s \in \mathbf{S}^{ZW}$), the duration of its producing task ($i \in \mathbf{I}_s^P$) is exactly equal to the processing time in Eq.(9). Sequencing constraints for the same task and different tasks in a unit are given in Eq.(10) and Eq.(11).

$$T_{im}^f \geq T_{im}^b + a_i \cdot z_{im} + \beta_i \cdot b_{im} \quad \forall s \in \mathbf{S}, s \notin \mathbf{S}^{ZW}, i \in \mathbf{I}_s^P, m \quad (8)$$

$$T_{im}^f = T_{im}^b + a_i \cdot z_{im} + \beta_i \cdot b_{im} \quad \forall s \in \mathbf{S}^{ZW}, i \in \mathbf{I}_s^P, m \quad (9)$$

$$T_{i(m+1)}^b \geq T_{im}^f \quad \forall i, m < M \quad (10)$$

$$T_{i'm'}^b \geq T_{im}^f - H \cdot (1 - X_{imi'm'}) \quad \forall j, i \in \mathbf{I}_j, i' \in \mathbf{I}_j, i \neq i', m, m' \quad (11)$$

Variable T_{ims}^S denotes the transfer time of batch m of task $i \in \mathbf{I}_s$ into or out from the dedicated storage of state s . Eq.(12) enforces for batch m of task $i \in \mathbf{I}_s^P$ producing s , the transfer time T_{ims}^S equals to its finish time T_{im}^f . The start time (T_{im}^b) of batch m of task $i \in \mathbf{I}_s^C$ consuming s equals to its transfer time out from the storage in Eq.(13). Sequences on storage for batches of the same tasks and different tasks are given by Eqs.(14-15).

$$T_{ims}^S = T_{im}^f \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s^P, m \quad (12)$$

$$T_{ims}^S = T_{im}^b \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s^C, m \quad (13)$$

$$T_{i(m+1)s}^S \geq T_{ims}^S \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, m < M \quad (14)$$

$$T_{i'm's}^S \geq T_{ims}^S - H \cdot (1 - XS_{imi'm's}) \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, i' \in \mathbf{I}_s, m, m', i \neq i' \quad (15)$$

Eqs.(16-18) are formulated to enforce precedence of batches on storage, which are similar to Eqs.(4-6). When task i and i' can be processed in the same unit and related to one state, sequence relations for their batches keep consistent on the unit and storage, as Eq.(19).

$$XS_{imi'(m'+1)s} \geq XS_{imi'm's} \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, i' \in \mathbf{I}_s, i \neq i', m, m' < M \quad (16)$$

$$XS_{i(m-1)i'm's} \geq XS_{imi'm's} \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, i' \in \mathbf{I}_s, i \neq i', m > 1, m' \quad (17)$$

$$XS_{imi'm's} + XS_{i'm'ims} \geq z_{im} + z_{i'm'} - 1 \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, i' \in \mathbf{I}_s, m, m', i < i' \quad (18)$$

$$XS_{imi'm's} = X_{imi'm'} \quad \forall j, s \in \mathbf{S}^{IN}, i, i' \in \mathbf{I}_s \cap \mathbf{I}_j, m, m', i \neq i' \quad (19)$$

$CB_{imi'm's}$ is defined to monitor batches m' of task i' transferred before batch m of task i ($i \neq i'$) in Eqs.(20-22) or task ($i = i'$) by Eq.(23). It equals to $b_{i'm'}$ if batch m' of task i' is transferred before or at the same time as batch m of task i . Otherwise, it equals to 0.

$$CB_{imi'm's} \leq B_{i'}^{max} \cdot XS_{i'm'ims} \quad \forall s \in \mathbf{S}^{IN}, i, i' \in \mathbf{I}_s, m, m', i \neq i' \quad (20)$$

$$CB_{imi'm's} \geq b_{i'm'} - B_{i'}^{max} \cdot (1 - XS_{i'm'ims}) \quad \forall s \in \mathbf{S}^{IN}, i, i' \in \mathbf{I}_s, m, m', i \neq i' \quad (21)$$

$$CB_{imi'm's} \leq b_{i'm'} \quad \forall s \in \mathbf{S}^{IN}, i, i' \in \mathbf{I}_s, m, m', i \neq i' \quad (22)$$

$$CB_{imim's} = b_{im'} \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, m, m' < m \quad (23)$$

The inventory level of state s after transfer of batch m of task i is calculated by Eq.(24), being positive and less than the maximum storage capacity ST_s^{max} . Eq.(25) enforces total amount transferred must satisfy storage limitation, where In_0_s is the initial inventory.

$$0 \leq \sum_{i' \in I_s} \sum_{m': (i=i' \cap m > m') \cup (i \neq i')} \rho_{i's} C B_{imi'm's} + \rho_{i,s} b_{im} + In0_s \leq ST_s^{max} \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, m \quad (24a,b)$$

$$0 \leq \sum_{i \in I_s} \sum_m \rho_{is} \cdot b_{im} + In0_s \leq ST_s^{max} \quad \forall s \in \mathbf{S}^{IN} \quad (25a,b)$$

For the objective of minimizing makespan, demand constraints are given in Eq.(26). Makespan must exceed the finish and transfer time of all batches.

$$\sum_{i \in I_s^p} \rho_{is} \cdot \sum_m b_{im} \geq D_s \quad \forall s \in \mathbf{S}^p \quad (26)$$

$$MS \geq T_{im}^f \quad \forall i, m \quad (27)$$

$$MS \geq T_{ims}^s \quad \forall s \in \mathbf{S}^{IN}, i \in \mathbf{I}_s, m \quad (28)$$

4. Hybrid framework

A hybrid algorithm is proposed to harness GA and sequence-based MILP, as illustrated in Figure 1. There are three stages in the hybrid framework: GA is first used to generate good-quality solutions; determinations on the batch numbers of tasks and sequencing relations are extracted; the majority of the binary variables in the sequence-based MILP model are fixed based on the extracted information and the MILP model is finally solved to optimality. Multiple feasible solutions ($P_{size} \cdot 0.01$) generated by GA are optimized using the third-stage MILP model, where P_{size} is the population size of GA.

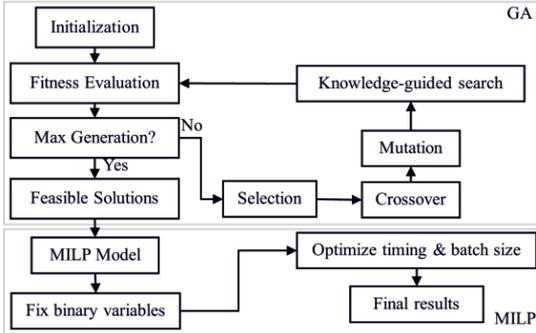


Fig.1 Illustration of hybrid algorithm

Batch size and timings would be optimized in the sequence-based MILP model. Therefore, in GA, batch sizes of tasks are assumed to be maximum and processing times are assumed to be fixed as maximum value (i.e. $\tau_i^G = \alpha_i + \beta_i \cdot B_i^{max}$). The strategies for fixing binary variables in the MILP model are described as follows. Two tasks $i \in \mathbf{I}_i^g$ and $i' \in \mathbf{I}_i^g$ in the MILP model denote one task performing on two different units. If there is no task i' belonging to \mathbf{I}_i^g , all performed batches $m \in \mathbf{M}_i^G$ of task i in the solution from GA must be processed in the MILP model. That is $z_{im} = 1 \forall i: \exists i' \in \mathbf{I}_i^g, m \in \mathbf{M}_i^G$. One batch m of a task $i \in \mathbf{I}_i^g$ starts at time TB_{im}^G . If there are consecutively idle times from any time $t \in [TB_{im}^G, TB_{im}^G + \tau_i^G/2)$ to time $t' > t + \tau_i^G/2$ on unit $j' \in \mathbf{J}_i'$, the batch m is assumed to be potential to divide into two batches, including batch m on unit j (i.e. $z_{im} = 1$) and batch $m' = \mathbf{M}_i^G + 1$ of task i' on unit j' in the MILP model; otherwise, $z_{im} = 1$ and $z_{i'm'} = 0$. The sequencing relations between batches of tasks on the same unit in the solution from GA are used to fix $X_{imi'm's}$. $XS_{imi'm's}$ is partially fixed by enforcing batch m' of task $i' \in \mathbf{I}_s^C$ to start after batch m of task $i \in \mathbf{I}_s^P$ that provides required state for batch

m' (i.e., $XS_{im'i'm's} = 1$). Also, batch m' of task $i' \in \mathbf{I}_s^C$ starts before the finish of batch m of task $i \in \mathbf{I}_s^P$ whose s is consumed by the batch m' to ensure inventory level of the state lower than the maximum storage capacity (i.e., $XS_{i'm'ims} = 1$).

5. Case studies

Five examples are solved to evaluate the performance of the proposed hybrid algorithm. Examples 1-2 are from He and Hui (2010), where instances I1 and I4 are small instances whilst the instances I2, I3 and I5 are large ones. Example 3 are the Kallrath example from Velez et al. (2015), which is an industrial-scale example. Finite intermediate storage and zero-wait policy are involved in Examples 1 to 3. Examples 4 and 5 are the Example 1 and PP solved by Lee and Maravelias (2018). GA is implemented in MATLAB 2020 and the sequence-based MILP formulation is solved using CPLEX 12.10/GAMS 33.2 on a desktop computer with AMD Ryzen™ 9 3900X 3.8 GHz and 48 GB RAM running Windows 10. We also solve these examples using the MILP models from Vooradi and Shaik (2012) and Velez et al. (2015) denoted as VS2012 and VM2015, respectively, as well as a hybrid algorithm from Lee and Maravelias (2018) denoted as LM2018.

Table 1. Computational results for examples 1 to 3

VS2012		VM2015				Hybrid algorithm			
N	MS	Gap (%)	CPU Time (s)	H	MS	CPU Time (s)	MS	CPU Time (s)	
Example 1									
I1	22	37	-	45.2	60	37	0.8	37	3.1
I2	65	109	1.4	>3600	110	108	11.0	108	61.0
I3	131	229	5.6	>3600	219	217	3298.0	217	250.3
Example 2									
I4	15($\Delta n = 1$)	37	-	11.7	60	37	0.4	37	0.2
I5	42	102	-	1328.2	120	100	1.2	100	2.7
Example 3									
I6	11	33	3.1	>3600	60	32	13.2	32	12.1
I7	12	40	-	1856.3	60	39	103.9	39	18.7
I8	23	58	17.3	>3600	60	52	93.9	52	58.2

The computational results for Example 1 to 3 are presented in Table 1. Results show that our hybrid algorithm can yield global optimality for all solved instances. Comparing to VS2012, the computational times are dramatically decreased in all instances. Also, better solutions are found for most cases (e.g. I3, I6, I7 and I8) using our algorithm. Although the proposed hybrid algorithm obtains the same optimal solution as VM2015, our algorithm requires significantly less computational efforts (< 5 minute) for the large-scale problems. For example, our algorithm leads to a reduction in computational efforts by about one order of magnitude while solving I3 of Example 1 and I7 of Example 3.

We also address the benchmark examples with complex features on processing times. The computational results are provided in Table 2. The results show the effectiveness of the proposed framework on generating optimal or near-optimal solutions within short time frames (1 minute) even for the examples with large demands (Example 4) or complex process structures (Example 5). In the proposed framework, the first-stage GA is critical

to reduce computational effort. For instance, a good solution (44) for I11 is found in 5 seconds by GA but more than 1 hour using the discrete-time model in LM2018. The high quality of the final solutions can be contributed to the integration of the MILP model.

Table 2. Computational results for examples 4 to 5

	Demand				LM2018			Hybrid algorithm			
	P1	P2	P3	P4	H	Dis. MS	MS	CPU Time (s)	GA. MS	MS	CPU Time (s)
Example 4											
I9	100	200	-	-	40	15.5	14.25	0.8	15.5	14.25	2.5
I10	300	300	-	-	60	27	25.37	2.5	27	25.37	4.2
I11	500	400	-	-	80	44	41.38	>3600	44	41.38	6.2
Example 5											
I12	23	25	24	23	120	68	63.56	13.1	68	63.85	7.9
I13	26	50	48	46	200	120	115.9	47.9	120	115.3	62.2

6. Conclusions

In this work, a hybrid algorithm is proposed for scheduling multipurpose batch plants with a wide range of processing features, such as variable conversion fraction, processing time and multiple storage policies. The computational results demonstrate the effectiveness of the proposed algorithm to solve large-size problems to optimality. More importantly, the computational times are significantly reduced to reach optima within 5 mins, leading this hybrid strategy potentially to use as an industrial operational optimisation tool.

Acknowledges

Dan Li appreciates financial support from UOM/CSC Joint Scholarship (201908130170). Jie Li would like to appreciate the financial support from EPSRC (EP/T03145X/1).

References

- Y. Han, X. Gu, 2021, Improved multipopulation discrete differential evolution algorithm for the scheduling of multipurpose batch plants, *Ind. Eng. Chem. Res.*, 60, 5530-5547.
- I. Harjunkoski, C.T. Maravelias, P. Bongers, P.M. Castro, S. Engell, I.E. Grossmann, J. Hooker, C. Mendez, G. Sand, J. Wassick, 2014, Scope for industrial applications of production scheduling models and solution methods, *Comput Chem Eng*, 62,161-193.
- Y. He, C.W. Hui, 2010, A binary coding genetic algorithm for multi-purpose process scheduling: A case study, *Chem. Eng. Sci.*, 65(16), 4816-4828.
- H. Lee, C.T. Maravelias, 2018, Combining the advantages of discrete- and continuous-time scheduling models: Part 1. Framework and mathematical formulations, *Comput Chem Eng*, 116, 176-190.
- T. Nishi, Y. Hiranaka, M. Inuiguchi, 2010, Lagrangian relaxation with cut generation for hybrid flowshop scheduling problems to minimize the total weighted tardiness, *Comput. Oper. Res.*, 37, 189-198
- S. Velez, A.F. Merchan, C.T. Maravelias, 2015, On the solution of large-scale mixed integer programming scheduling models, *Chem. Eng. Sci.*, 136(2), 139-157.
- X. Zheng, L. Wang, 2016, A knowledge-guided fruit fly optimization algorithm for dual resource constrained flexible job-shop scheduling problem, *Int. J. Prod. Res.*, 54(18), 5554-5566.

Assessment of biomass supply chain design and planning using discrete-event simulation modeling

Helena Paulo^{a,b}, Miguel Vieira^{a,c,d}, Bruno S. Gonçalves^e, Tânia Pinto-Varela^a
and Ana P. Barbosa-Póvoa^a

^a*CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal*

^b*ISEL - Instituto Superior de Engenharia de Lisboa, IPL, Lisboa, Portugal*

^c*EIGeS, Universidade Lusófona, Lisboa, Portugal*

^d*Univ Coimbra, CEMMPRE, Dept Mech Engn, Coimbra, Portugal*

^e*ESTG, Instituto Politécnico de Leiria, Leiria, Portugal*

hpaulo@deq.isel.ipl.pt

Abstract

The use of discrete-event simulation has been gathering relevance to explore dynamic solution generation approaches in the design and planning of biomass supply chains. In this work, the combination of simulation with an optimization module in a hierarchical sequence is proposed, in order to compute the decision variables of interest to validate the economic viability of the supply chain network. The results highlight the advantage to assess strategic/tactical solutions with a more precise liability of overall uncertainty factors, such as the availability of biomass sources and demand variability.

Keywords: Design and planning, biomass supply chain, discrete-event simulation.

1. Introduction

The dynamic structure of current supply chains poses significant challenges to the implementation of effective decision support systems. This complexity often entails the simultaneous assessment of different business criteria for most strategic and operational decisions, in an interdependent structure of stakeholders facing with operational uncertainty. With the advent of digitalization, the development of decision support systems is demanding an enterprise integration with faster decisions (Vieira et al., 2021). It represents a contend for the traditional modeling techniques to encompass the tradeoffs of different dimensions, with particular importance to the growing sustainability implications of economic, environmental, and social levels.

Within the modeling research of complex systems, the development of Digital Models, Twins or Shadows have been addressed in the fields of planning and control improvement. The scope often aims at high-resolution views of physical systems or focus on specified traces when fast reaction is crucial (Liebenberg et al., 2020). Due to this multi-level decision-support, discrete-event simulation (DES) can generate a digital representation with useful data information and performance indicators (e.g. lead time, makespan, resources setting, capacity utilization), allowing to model uncertainty and evaluate how the system reacts in different scenarios as a sequence of events in time.

It is recognized the growing importance in replacing fossil resources consumption by biomass to produce bioenergy and biomaterials at integrated biorefineries (Paulo et al., 2015). However, beyond its environmental contributions, attracting investment requires additional efforts to increase its competitiveness with an efficient supply chain network, gauging the diverse biomass sources to the demand of each region. The overall challenges

concerning its design and planning can include the (1) selection of biomass sources, type, and quantities; (2) storage locations/capacities and preprocessing technologies; (3) processing facilities with technology definition, capacity, and location; (4) type of products/byproducts and respective amounts; (5) demand centers, location and quantity to satisfy; or (6) logistic operations at the supply chain nodes. Moreover, these strategic decisions are greatly influenced by the uncertainty on biomass availability and quality, products demand, or the non-mature biorefineries technologies for pre/processing operations.

The biomass supply chain design under uncertainty has been an active area within the research community in recent years. Despite largely supported by analytical optimization modeling studies, DES approaches have been gaining relevance to address the behavior and performance under different conditions to describe uncertain scenarios. For example, simulation can be used to validate the long-term economic viability of a biorefining supply chain, as presented by Sukumara et al. (2015). To examine the economic performance of the supply chain over time, including supply and demand variability, a simulation model is created to determine the optimum feedstock requirement, capital investment, and operating costs of the process. Another recent contribution, Pinho et al. (2016) modeled the biomass supply chain using DES to estimate the system behavior regarding stochastic drive and chipping times to improve management and scheduling tasks. And Eriksson et al. (2017) develop a model for weather-driven analysis of forest fuel systems, with a DES approach analyzing the supply chains that account for active and passive machine activities, such as quality changes during storage. The results illustrate significant improvements concerning machine capacity requirements.

Following the recognized advantage of DES to model uncertainty, the application to decision support in the design and planning of biomass supply chains remains fairly unexplored. The aim of this work is to explore a hybrid simulation-optimization approach, with focus on the representation of the dynamics and uncertainty interactions of the real supply chain system by means of an event-based model. This preliminary study outlines the combination of an optimization-based solution to be evaluated in a virtual model that accurately simulates the characteristics of the biomass/bioproduct flows and the network performance under alternative scenarios to generate a solution. As follows, we present an overview of methodology and discuss the results in a case study example for a biomass supply chain problem.

2. Methodology

Due to the stochastic and complex nature of biomass supply chains, an adequate choice for an effective solution generation approach resumes the research for hybridization of techniques to overcome the computational burden. As noted by Figueira and Almada-Lobo (2014), the potential hybridization of simulation with optimization approaches combines the advantages of the system detail representation with the ability to optimize solutions, leaving the hard-to-model constraints to the former. Despite this hierarchical combination can be defined in one-way or iterative, here we focus on the case where analytical models can be formulated and their solutions simulated (optimization-based simulation), in order to compute some variables part of the solution generation. The analytical model generates a solution under ideal conditions, while the simulation considers the inherent variability and provides the expected realistic outcomes.

For the proposed case of a biomass supply chain, the economic viability of a solution for the strategic/tactical decisions comprises the definition of which biomass sources supply raw-material for the integrated biorefineries to satisfy demand centers within a

transportation network. However, matching the uncertain biomass seasonality throughout the given time horizon to each center demand can generate a reduced income output. Therefore, we devise the methodology composed by an optimization module to generate an initial deterministic solution, completed at the simulation module which evaluates the dynamic representation of process interconnections.

2.1. Optimization module

Based on the work of Paulo et al. (2015), this module is composed by a MILP model using GAMS/CPLEX which is stated as follows: given (a) a superstructure defining all possible locations of biomass sources (with the corresponding availability per type), biorefineries and demand centers, considering the distances between all points; (b) transportation modes and their costs, biomass acquisition costs, biorefinery installation, fixed/variable operating costs (for a set of discrete capacities); and (c) the conversion factors of the different biomass types into products and corresponding demand; determine the optimal supply chain network structure with the (d) number, location, technology and capacity of biorefineries to install, with the corresponding amounts produced and demand centers to serve; and (e) amount of each biomass type collected at each source and transportation flows between each pair of facilities; so as to (f) minimize the supply chain total cost (Eq.1).

$$\begin{aligned} \text{Total Cost} = & \sum (\text{Biomass Cost} + \text{Fixed/Variable Operating Cost at Biorefineries} \\ & + \text{Transportation Costs} + \text{Annualized Investment Cost}) \end{aligned} \quad (1)$$

2.2. Simulation module

The embedded biomass supply chain structure of operations is evaluated with a DES model implemented in SIMIO, which computes the output scenario statistics involving the occurrence of events at discrete time point using probability distributions of parameters. The main components and library of objects available at the SIMIO software are used to achieve the most accurate representation of the network nodes, which requires advanced programming of the system properties. As a general concept of the model principles, an 'entity' (e.g. demand order) will arrive, is processed by a sequence of 'servers' (e.g. biorefinery), and then departs from the system. The level of detail, uncertainty factors and problem assumptions should be considered as suitable to case objectives. Due to the scope limitations of the present work, the focus relies in the time horizon variability of biomass availability at each selected source, the technology conversion performance at the biorefineries and bioproducts prices. For simplification, it was considered the software's default number of replications to guarantee statistical significance.

The interaction of the two modules aims to generate a supply chain configuration accounting with the precise simulated biomass/bioproducts flows in order to compute a techno-economic investment viability. Adequate parameter availability, unfulfilled demand penalty and prices forecast was considered to examine the economic performance cashflows over time. To account for the time value of money, the simulation model tracks the net present value (*NPV*) of the biomass supply chains operations (Eq.2), for n time periods, CF_k as annualized value of cashflow at year k and i discount rate.

$$NPV = \sum_{k=0}^n \frac{CF_k}{(1+i)^k} \quad (2)$$

3. Case Study

The described methodology has been applied to an illustrative case study of bioethanol and biodiesel production in Portugal. The supply chain structure integrates the process of harvesting and biomass collection, biorefineries, demand centers and all the logistic activities for biomass and bioproducts transportation. Considering the 278 portuguese municipalities, 150 are considered as biomass supply sources (i) (the remaining municipalities were excluded due to low biomass availability), and 28 locations are candidate to biorefineries installation (k), to satisfy the 18 portuguese districts in biodiesel and bioethanol demands (v). Two types of biomass types are considered: Sugar/starch biomass (b1) and seeds/animal fats (b2), processed by fermentation (m1) and transesterification (m2) technologies at the biorefineries to bioethanol and biodiesel, respectively. All models run in an Intel(R) Xenon(R) CPU ES-2660v3@2.60GHz with 64GB RAM, with GAMS (25.1.1 ver.), using CPLEX (12.8.0.0 ver.) solver, and SIMIO (9.5 ver.).

3.1. Results discussion

Setting the problem input data for a 12-months' time horizon, the optimization module generates an optimal cost solution with the supply chain characteristics presented in Table 1, with the number of biomass collection sources and total biomass amounts to be transported to the installed biorefineries. The biorefineries properties, as well as the total amount of production to entirely fulfil the demand centers by each of the biorefineries, are also displayed. While the optimal solution yields a network with assumed conditions, variability in biomass availability has influence on supply chain performance over time and directly linked to demand variability. At the simulation module, these biorefineries activities that account for biomass sources transportation, conversion to final bioproducts, product distribution and sale prices, are assessed.

Table 1. Optimal supply chain characteristics

Biomass type	Biomass collection sources	Total biomass (ton)	Biorefinery	Total product. (ton)	Demand centers
b1	i5, i8, i9, i14, i16, i26, i29, i34, i35, i38, i43, i53, i69, i74, i75, i78, i87, i89, i90, i94, i100, i112, i114, i115, i123, i126, i135, i143	700 575	Location: k13 Tech: Fermentation Cap: 250 000 ton Product: Bioethanol	241 435	v2, v7, v8, v11, v14, v15
b1	i1, i2, i4, i6, i7, i18, i21, i22, i27, i39, i41, i49, i51, i52, i58, i62, i63, i67, i72, i76, i88, i96, i97, i98, i103, i108, i109, i112, i122, i127, i129, i131, i132, i134, i137, i146	598 416	Location: k16 Tech: Fermentation Cap: 250 000 ton Product: Bioethanol	206 229	v1, v5, v6, v9, v10, v12, v13, v18
b1	i11, i20, i23, i42, i46, i50, i77, i83, i86, i117, i118, i119, i138, i142, i145	217 627	Location: k26 Tech: Fermentation Cap: 75 000 ton Product: Bioethanol	75 000	v1, v3, v6, v7, v10, v11, v13, v14, v15, v16, v17, v18
b2	i1, i6, i7, i9, i14, i16, i26, i29, i34, i35, i38, i39, i43, i49, i53, i67, i72, i74, i90, i94, i98, i108, i112, i114, i115, i131, i134, i135, i143	14 383	Location: k20 Tech: Transesterification Cap: 250 000 ton Product: Biodiesel	4 770	v3, v4, v13, v16, v17

The optimal solution of this supply chain is then modeled in SIMIO to simulate its feasibility, considering the defined parameters distributions (from historical public data of energy regulatory agency) and the year seasonality of biomass availability and bioproducts demand applied to each corresponding location (Table 2). By recording the monthly flows between nodes, the DES model is run for a period of 15 years to compute the variability and cashflows of this initial scenario, considering a discount factor of 3%.

Table 2. Seasonality discretization of biomass availability and bioproducts demand

Seasons/Months		Annual biomass availability	Biodiesel annual demand	Bioethanol annual demand
Winter	(Dec. – Feb.)	10%	21%	25%
Spring	(Mar. – May)	40%	20%	23%
Summer	(Jun. – Aug)	30%	31%	26%
Fall	(Sep. – Nov.)	20%	28%	26%

The simulation results displayed in Figure 1, related to the first year of the investment period, allows to verify the virtual utilization of the optimal theoretical capacity installed and the corresponding costs breakdown, enabling a more accurate assessment of the design decisions. Likewise, in Figure 2 is possible to identify that the availability of biomass less concentrated in Fall and Winter seasons does not allow to fulfil all corresponding districts’ demand of bioproducts. Therefore, an additional scenario can be simulated to include the additional installation of a dry-storage warehouse for biomass buffer at each biorefinery that does not satisfy associated seasonal demand, k13, k16 and k20. This scenario enables to satisfy the entire country’s demand and increases the project investment NPV from –95,2 M€ to 26,3 M€.

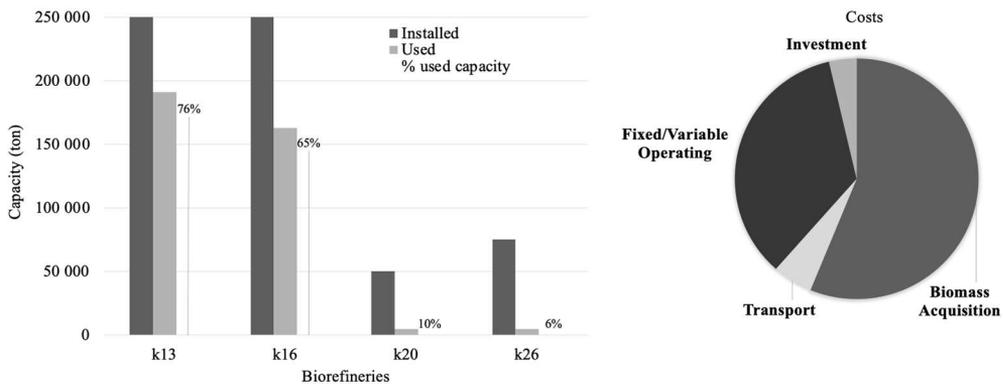


Figure 1. Results of simulated capacity utilization of each biorefinery and costs breakdown for the first year of investment project

4. Conclusions

The present work proposes the integration of a hybrid simulation-optimization approach for the design and planning of biomass supply chains, in which the role of the DES is to mimic the real-world environment to evaluate an optimized solution under dynamic and uncertain interdependent conditions. The results demonstrate that this combination offers

a method with potential to deal with such strategic/tactical decision liabilities, such as the importance of storage warehouse or pre-processing facilities to generate an efficient supply chain solution. Future work will explore the DES capabilities to assess a more robust definition of the nodes, configuration, and capacities of the supply chain by increasing the detail of the uncertainty effects in the overall operations.

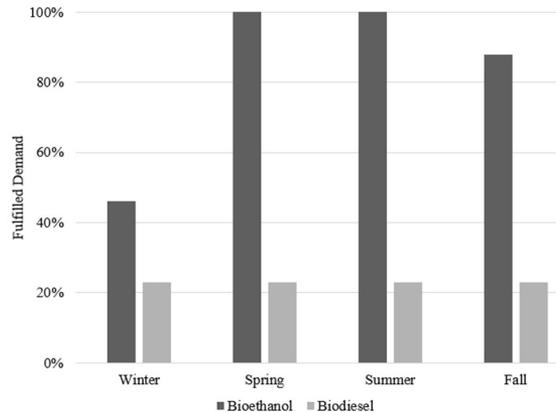


Figure 2. Percentage of total fulfilled demand of each bioproduct per season

Acknowledgments

The authors would like to acknowledge the financial support by UE/FEDER funds through program COMPETE and FCT under the project UIDB/00285/ 2020.

References

- M. Vieira et al., 2021, A two-level optimisation-simulation method for production planning and scheduling: the industrial case of a human-robot collaborative assembly line, *International Journal of Production Research*
- M. Liebenberg et al., 2020, *Information Systems Engineering with Digital Shadows: Concept and Case Studies An Exploratory Paper*, 70-84 Springer
- H. Paulo et al., 2015, Supply chain optimization of residual forestry biomass for bioenergy production: The case study of Portugal, *Biomass and Bioenergy*, 83, 245-256
- T. Pinho et al., 2016, Modelling a biomass supply chain through discrete-event simulation, *IFAC-PapersOnline*, 49, 2, 84-89
- A. Eriksson et al., 2017, Evaluation of delivery strategies for forest fuels applying a model for Weather-driven Analysis of Forest Fuel Systems, 188, 420-430
- S. Sukumara et al., 2015, A comprehensive techno-economic analysis tool to validate long-term viability of emerging biorefining processes, *Clean Technologies and Environmental Policy*, 17, 1793 – 1806
- G. Figueira. and B. Almada-Lobo, 2014, Hybrid simulation–optimization methods: A taxonomy and discussion, *Simulation Modelling Practice and Theory*, 46:118-134.

Design of a three-echelon supply chain under uncertainty in demand and CO₂ allowance prices

F. L. Garcia Castro^a, R. Ruiz-Femenia^a, R. Salcedo-Diaz^a and J. A. Caballero^a

^a*Department of Chemical Engineering, University of Alicante, Apartado de Correos 99, Alicante 03080, Spain
ruben.ruiz@ua.es*

Abstract

Nowadays there is a growing concern regarding greenhouse gas emissions and their consideration in the supply chain design. In this work we present a robust stochastic model for the design of a supply chain under uncertainty of CO₂-allowance prices and market demand. The three-echelon petrochemical supply chain superstructure consists of four production plants in Europe, storage associated with these plants and four possible markets. At each plant different products can be produced according to the available technologies. The goal is to maximize the expected net present value (ENPV), while reducing the amount of CO₂ equivalent emissions. We implemented the carbon cap and trade model from the European Union emissions Trading System, whose goal is to reduce the emission cap over time in order to achieve a climate-neutral EU by 2050. We combine the environmental LCIA data, required to determine the global warming potential, with the forecast of CO₂ allowance prices. The problem involves a multi period mixed-integer linear program (MILP) formulation, which was implemented in Python using Pyomo and solved using IBM's CPLEX algorithm. To deal with uncertainty in market demand and CO₂-allowance prices, we implemented an ARIMA model and generated multiple scenarios. Since a full discretization of the resulting probability space leads to a number of scenarios that exceeds capacities of state-of-the-art computers with ease, decomposition techniques are applied. The obtained results show an improvement of the economic performance when compared to the results from the deterministic approach that is being widely used in literature.

Keywords: CO₂ price uncertainty, stochastic model, optimum supply chain management, Augmented Lagrangian Relaxation

1. Introduction

In a globalized world the efficiency of supply chains is gaining more importance. Supply chains integrate different activities, resources and institutions to ensure effective production, storage and distribution of products. Real world uncertainties can have a large impact on the design and efficiency of a supply chain. In this paper, we not only focus on uncertainties in the market demand (which have been studied extensively in the literature), but also on uncertainties in the CO₂ prices. In this work we will consider the European Union Emissions trading system, whose goal is to achieve a climate neutral EU by reduction of the maximum amount of allowed emissions over time, c.f. European Union (2016). The study of combined uncertainty in CO₂ prices and demand require the use of a huge number of scenarios for fully discretizing the underlying probability space. Since this exceeds capacities of modern day computer with ease, we combined various techniques to being able to obtain a supply chain design that takes these combined uncertainties into account.

We first generated a large number of scenarios and clustered them according to their economic and environmental performance. In order to solve the resulting stochastic mixed-integer linear program, we applied an Augmented Lagrangian Relaxation (ALR) scheme together with an alternating direction of multipliers to ensure separability of the problem into subproblems. The resulting subproblems could then be solved in parallel to reduce computational time. For more information regarding ALR see Boyd S., Parikh N., Chu E., Peleato B., Eckstein J. (2011); Conejo A. J., Castillo E., Mínguez R., García-Bertrand R. (2006).

To the best of our knowledge, a combination of the before mentioned methods for the study of combined uncertainty in a European three-echelon petrochemical supply chain has never been used before.

2. Methodology

In this section, we briefly introduce the problem under study. Also, we review the different techniques that were combined in order to compute a stochastic supply chain model that adapts to the scenarios under study.

2.1. Problem statement

The studied three-echelon petrochemical supply chain superstructure consists of 4 production plants in Europe: Kazincbarcika (Hungary), Leuna (Germany), Mantova (Italy) and Wloclaweck (Poland). Each plant has an expansion limit between 10 and 400 kt/year and an initial capacity of 20 kt/year, c.f. Guillén-Gosálbez, G., Grossmann, I.E. (2009). All plants can produce acetaldehyde, acetone, acrylonitrile, cumene, isopropanol and phenol using one of 6 available technologies involving up to 18 different chemicals. There is a warehouse associated to each plant having the same expansion limits and initial capacities. Initially, the inventory is empty. The material flow between plant and warehouses, as well as warehouses and markets, is limited to 500 kt/year. A minimum demand satisfaction of 30% is specified. The initial maximal amount of CO₂ emissions is $2 \cdot 10^8$ kg CO₂ equivalent, with a 2.2% yearly reduction rate.

2.2. Supply chain model

The model under study is composed of environmental impact, economical assessment, capacity constraints and mass balance equation blocks. The objective is to maximize the expected net present value, while satisfying all of these equation blocks. The cap-and-trade system provides a link between economic and environmental performance. Since a full description of the model would turn out to be quite long and considering the page limit, we refer the reader to Ruiz-Femenia et al. (2013) for more details on the model. The most important equations read

$$\text{GWP}(t, s) \leq \text{MAX}_{\text{CO}_2}(t) + \text{BUY}_{\text{CO}_2}(t, s) - \text{SALES}_{\text{CO}_2}(t, s), \quad \forall t, s \quad (1)$$

$$\begin{aligned} \text{NET}_{\text{CO}_2}(t, s) &= \text{PRICE}_{\text{CO}_2}(t, s) \cdot \text{SALES}_{\text{CO}_2}(t, s) \\ &\quad - \text{COST}_{\text{CO}_2}(t, s) \cdot \text{BUY}_{\text{CO}_2}(t, s), \quad \forall t, s \end{aligned} \quad (2)$$

$$\begin{aligned} \text{NETE}(t, s) &= (1 - \text{TAX}) \left[\text{EARNINGS}(t, s) - \text{EXPENSES}(t, s) + \text{NET}_{\text{CO}_2}(t, s) \right] \\ &\quad + \text{TAX} \cdot \text{DEP}(t), \quad \forall t, s \end{aligned} \quad (3)$$

Equation (1) describes the available emissions for the supply chain at each time-step t and scenario s , and consists of the maximum amount of emissions allowed by the system in each year, as well as the amount of bought as well as sold emissions. In equation (2), the economic impact of the emissions trading on the supply chain is calculated. We call this quantity NET_{CO_2} . This term is then taken into account when calculating the net earnings for a specific time and scenario, as described in equation (3).

2.3. Uncertainty in CO₂ allowance prices and market demand

In order to make decisions on a supply chain design, oftentimes a deterministic approach is chosen in literature. Since uncertainties are not taken into account, and computations are done for a fixed deterministic setting, this can be considered a risky approach. In the design of a supply chain, there are many possible sources of uncertainty, such as CO₂ allowance prices, changes in raw material prices, energy prices, demand, etc. If these uncertainties are not taken into account, the resulting design may perform poor in unexpected future situations. In this paper, we consider uncertainties of demand and CO₂ allowance prices through combined scenarios. Each scenario represents a possible time-dependent evolution of both parameters. By taking the uncertainties into account when computing an optimal supply chain design, the result will be more robust to future changes in both parameters than the classical deterministic approach. The scenarios for the CO₂ allowance prices were generated using an autoregressive integrated moving average (ARIMA) model. ARIMA models are defined by three parameters. The first parameter determines the non-seasonal polynomial degree. The second one is the order of differentiation used to remove seasonal trends. The third parameter is the number of moving average terms, which is related to the accuracy of approximations produced by an update of the parameters from fitting the model to the historical data. In Shumway and Stoffer (2000), details can be found.

Appropriate values of the parameters were identified by fitting the historical data from 2009 to 2021 and using the Akaike and Bayesian information criteria. In our case, the best-fitting model is ARIMA (7, 1, 8).

Random time series data for the demand of each chemical at each market were generated using a normally distributed random growth with a standard deviation of 7% to a given start value at time $t = 0$. We compared the environmental and economic performance of 1000 combined scenarios, and grouped them into 20 clusters using the K-means algorithm (see fig. 1). The centroids of each cluster were then used as scenarios for the stochastic program. For an overview of clustering methods see Omran et al. (2007).

2.4. Augmented Lagrangian Relaxation

The aim of decomposition methods is to solve many easy-to-solve subproblems instead of one hard-to-solve full problem. We split the full problem into subproblems containing information only on one scenario. The main difficulty then lies in the linking of the non-scenario dependent variables from each subproblem, since these should coincide. To do so, we included as many non-anticipative constraints to the model as there are non-scenario dependent variables. These non-anticipative constraints represent the complicating constraints of the model, which have to be relaxed by introducing them as penalty terms into the objective function, together with new multipliers. Since our objective is a linear function, we added additional quadratic penalty terms to ensure convergence of the method, resulting in the so called Augmented Lagrangian Relaxation of the problem. Because of the appearing products of variables from different scenarios in the quadratic penalty terms, the resulting objective is not decomposable into subproblems for each scenario anymore. To overcome this problem, the alternating direction method of multipliers is used. This method linearizes the product of variables from different scenarios using values of previous iterations for one of the appearing variables. The result is a relaxed problem that is fully decomposable into subproblems only involving one scenario each.

A tolerance of 5% was set as stopping criteria for the ALR-loop. After each iteration, if the stopping criteria was not fulfilled, the multipliers will be updated according to the subgradient method and a new iteration is started, and we refer to Conejo A. J., Castillo E., Mínguez R., García-Bertrand R. (2006) for details. The overall procedure is explained in fig. 2

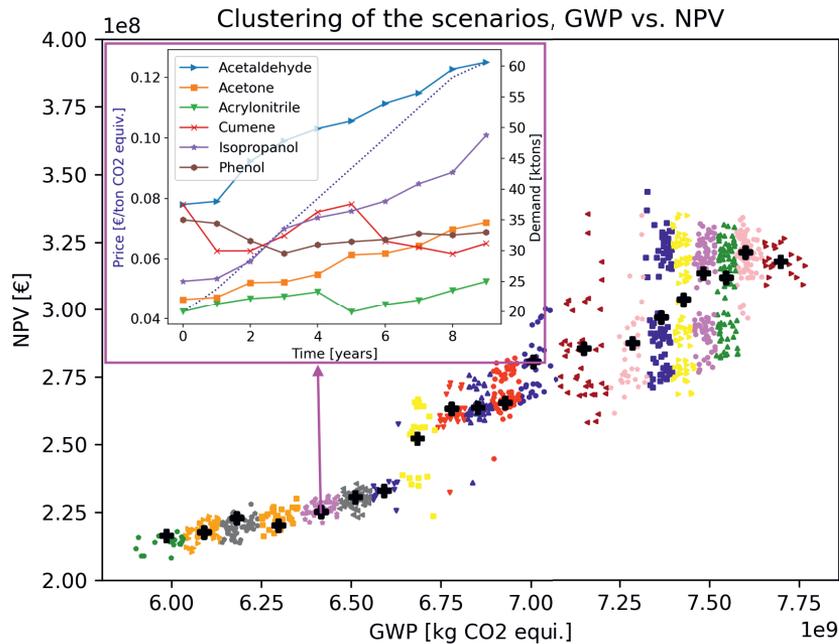


Figure 1: Economic and environmental performance of 1000 deterministic scenarios. The scenarios are clustered in different colors and the centroids are marked with a +. For the exemplary cluster centroid 5, the values of the CO₂ allowance prices (dotted blue line) and demand in the market Leuna for the 6 products over time are shown.

3. Results

Numerical experiments were conducted on a MacBook Pro model 2020 with 16 GB Ram and macOS 11.2.1. The model has been implemented in Python using Pyomo and solved using IBMs CPLEX algorithm.

Thanks to the method of alternating direction of multipliers, each subproblem can be solved independently of the others. This allows for a massive speedup through parallelization, since in the best case one processor only solves one or a few subproblems per iteration. According to the previous runtime of each subproblem, a redistribution of subproblems among the processors will be carried out between iterations. The goal is to reduce the total runtime per iteration by balancing the expected runtime of each processor. In our case (on 6 processors), we could speed up the solution process by about 80% compared to sequential solving. The problem was moreover reformulated and the number of variables per subproblem could be reduced from about 27000 to 24000, further simplifying the solution process.

We validated our method by comparing the obtained stochastic supply chain design against its deterministic counterpart from a scenario with an average investment cost, which is typically the one chosen by companies. In average, the stochastic model leads to an increase of the net present value by approx. 927600 € while reducing the global warming potential by around 225000 tons of CO₂-equivalent.

The resulting stochastic supply chain design has a net present value (NPV) of 254603229 € and

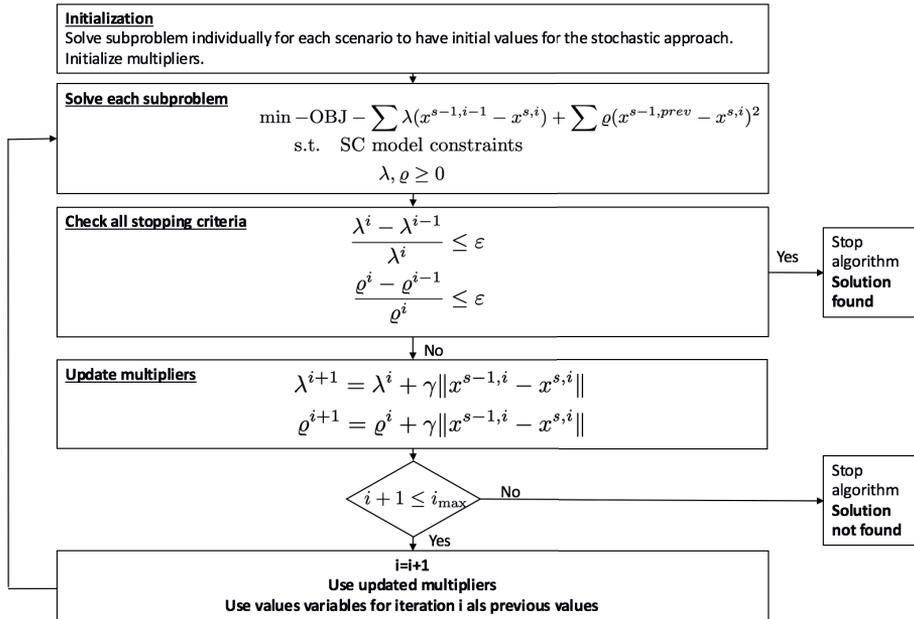


Figure 2: Schematic representation of the iterative solution algorithm. The problem has been separated into subproblems with additional coupling constraints, that appear as penalty terms in the corresponding objective function. The multiplier update uses a subgradient method.

a global warming potential (GWP) of 7624500 tons CO₂ equivalent. In average, 580 tons of CO₂ equivalent will be bought each year. In fig. 3, the capacities of the different technologies of each plant, as well as the total quantities transported between warehouses and markets in the last time-step are presented.

The major expansions have taken place at the Wloclaweck and Kzinbcarcika plants, which have the lowest production costs. At all plants, the capacity for technology 5 “Reaction of benzene and propylene” was expanded most, since cumene is not only one of the products demanded at all four markets, but also the raw material for the production of phenol and acetone (technology 6). Sines and Tarragona markets are mainly supplied by the Mantova plant, which is the closest one. These decisions reduce not only transportation costs, but also emissions.

4. Outlook and further work

We are currently working on a supply chain design, which is able to represent the penalties associated with a non-satisfaction of the demand. In this work, we studied the impact of uncertainty in market demand and CO₂ prices on the supply chain design. Nevertheless, a study for uncertainties in different parameters should be carried out, in order to recognize which parameters possibly have a higher impact on the performance. We are also working on increasing the feasible number of scenarios and assess the information loss caused by the clusterization. Considering the ALR loop, there are still many open questions regarding a robust choice and update of the multipliers for each iteration, which tend to have an enormous impact on the solvability of the stochastic model.

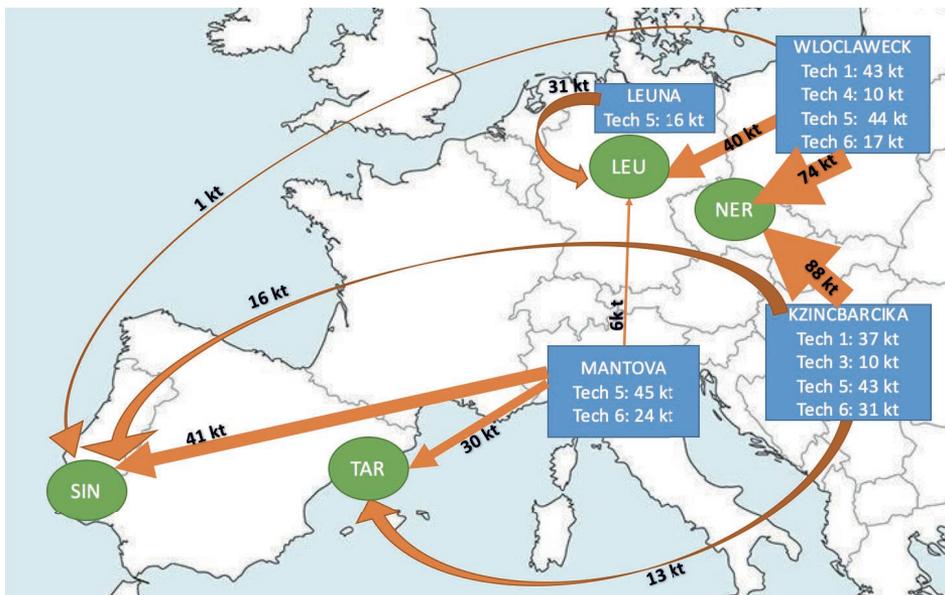


Figure 3: Stochastic supply chain design for the last time-step. The blue boxes represent the four plants and their respective warehouses. Capacity expansion information for each technology is included in the corresponding box. The four markets, Sines (SIN), Tarragona (TAR), Leuna (LEU) and Neratovice (NER) are represented with green circles. The orange arrows represent quantities transported (in kt) between each plant/warehouse and each market.

Acknowledgements

The authors gratefully acknowledge financial support to the Conselleria de Innovacion, Universidades, Ciencia y Sociedad Digital of the Generalitat Valenciana, Spain under project PROMETEO/2020/064.

References

- Boyd S., Parikh N., Chu E., Peleato B., Eckstein J., 2011. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning* 3, 1–122.
- Conejo A. J., Castillo E., Mínguez R., García-Bertrand R., 2006. *Decomposition Techniques in Mathematical Programming*. Springer-Verlag Berlin Heidelberg.
- D. Nie, T. Qu, Y. Liu, C. Li and G. Q. Huang, 2019. Improved augmented Lagrangian coordination for optimizing supply chain configuration with multiple sharing elements in industrial cluster. *Industrial Management & Data Systems*, Volume 119, Issue 4.
URL <https://www.emerald.com/insight/content/doi/10.1108/IMDS-06-2018-0253/full/html>
- European Union, 2016. The EU Emissions Trading System (EU ETS).
URL https://ec.europa.eu/clima/sites/clima/files/factsheet_ets_en.pdf
- Guillén-Gosálbez, G., Grossmann, I.E., 2009. Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE J.* 55, 99–121.
- Nyoni, Thabani and Bonga, Wellington Garikai, 2019. Prediction of CO2 Emissions in India Using ARIMA Models. *DRJ - Journal of Economics & Finance*, 4(2), pp. 01-10.
URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3346378
- M. Omran, A. Engelbrecht, A. Salman, 11 2007. An overview of clustering methods. *Intell. Data Anal.* 11, 583–605.
- R. Ruiz-Femenia, G. Guillén-Gosálbez, L. Jiménez Esteller, J. Caballero, 05 2013. Multi-objective optimization of environmentally conscious chemical supply chains under demand uncertainty. *Chemical Engineering Science* 95, 1–11.
- R. H. Shumway, D. S. Stoffer, 2000. *Time Series Regression and ARIMA Models*. Springer New York, New York, NY, pp. 89–212.

An industrial perspective on simulation versus optimisation decision-making capabilities

Robert E. Franzoi,^a Mahmoud A. Ahmednooh,^{a,b} Brenno C. Menezes,^{a,*}

^a*Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar*

^b*Division of Production Planning and Scheduling, Um Said Refinery, Qatar Energy, Doha, Qatar*

bmenezes@hbku.edu.qa

Abstract

Despite the increasing development of optimisation-based approaches over the last decades, simulation methods are still widely employed for problems that could instead be optimised. Simulation relies on achieving feasibility through a trial-and-error procedure and provides meaningful and useful resources to find quick or rigorous solutions. On the other hand, optimisation provides proper capabilities for automatic and systematic searching over the solution space. Although more computationally expensive, optimisation has been increasingly used due to the recent technological improvements and research developments in the wake of the Industry 4.0. However, industrial decision-making is still often carried out manually, which indicates an urgent need for addressing the difficulties, complexities, and issues in the implementation of automated tools. This work aims to provide guidance for industrial operations towards selecting proper decision-making capabilities. An industrial perspective on simulation and optimisation decision-making is provided, whereby highlighting the issues and challenges faced by industrial peers, and the current gaps between academic research and industrial implementation. This includes the interplay between simulation and optimisation tools, academic research, and industrial implementation. From an industrial perspective, we provide guidance for incorporating academic knowledge by the industry, exposing what are the industrial requirements to be further addressed, and assisting industrial peers towards achieving further breakthroughs.

Keywords: Optimisation, simulation, planning, scheduling, decision-making.

1. Introduction

Recent Industry 4.0 (I4) developments comprising digital transformation, big data, machinery automation, advanced analytics, and decision-making frameworks, offer opportunities for improvement in multiple applications of diverse fields. I4 technologies provide further automation of decisions with faster and more efficient operations, reduced costs, and higher product quality. There has been an urgent need for replacing outdated processes and tools more efficient and automated systems. This is fundamental towards achieving safer, more profitable, and more efficient industrial operations. In this work, we discuss simulation versus optimisation capabilities for industrial processes with particular emphasis from an industrial perspective.

Simulation methods have been historically used for decision-making. They rely on achieving feasibility through a trial-and-error procedure and provide meaningful and

useful resources to quickly find solutions for real-world problems. Simulation refers either to manual decisions (e.g., to be determined by the operations team) or to automatic computer-based algorithms (e.g., rigorous simulation) based on ranges and probabilities. Over the last decades, simulation-based software has been extensively developed for industrial applications (Yang et al., 2020).

Optimisation tools have arisen as improved alternatives for automated decisions. They provide proper capabilities to systematically find optimal (and better) solutions. Their searching algorithms are based on mathematically proven concepts in order to guarantee optimality. However, due to the increasing complexity that scales with the type and size of the formulation, there are limitations in the application of optimisation depending on the problem to be solved and the availability of computational resources. Despite such complexity issues, recent advances on decision-making algorithms and computer-aided resources have increasingly provided enhanced capabilities for properly addressing and solving industrial applications (Franzoi et al., 2018; Brunaud et al., 2020).

Such simulation versus optimisation ideology is widely known and discussed in academia. Both approaches have their advantages with highly important and meaningful applications. Nevertheless, there is a huge gap between the methodologies and strategies presented by the state-of-the-art literature and the capabilities currently used by the industry. Many industrial decision-making methods and tools are still obsolete. Common reasons include lack of knowledge, training, and qualification to employ better decision-making. This is often a result of the resistance to change from both the people and organisations. Consequently, further breakthroughs and enhancements towards improved operations are compromised or not properly exploited. Given the current economic scenario with increasing competitiveness, most companies either struggle to survive or miss significant opportunities to achieve economic and operational improvements.

In this work, we stimulate a discussion on simulation and optimisation capabilities and emphasise the importance of employing appropriate tools according to the industrial needs, limitations, and requirements for each process or application. Particular emphasis is given to the petroleum refinery industrial sector, where the decision-making process often relies on trial-and-error, manual, and simulation-based strategies and tools despite the emerging applicability of their optimisation counterparts.

2. Simulation-based versus optimisation-based industrial capabilities

Despite the increasing development and implementation of optimisation methodologies, algorithms, and tools over the last decades, the industry still heavily relies on simulation-based methods and software. In many cases, simulation is indeed more suitable or advisable. However, many organisations are still outdated in terms of software and technological capabilities, whereby not realising or exploiting opportunities to achieve increased operational efficiency and economic value. In fact, many problems and applications that could be optimised are instead handled by less efficient simulation tools. Some key differences and applications of simulation and optimisation methods are illustrated in Figure 1. Such concepts are fundamental to understand how to best employ the current industrial capabilities in a proper and efficient fashion.

A typical example of highly complex processing industry that still requires computational-based improvements, especially in terms of software and methodologies employed for the decision-making process, are petroleum refineries. In the following, we address the topic of simulation-based versus optimisation-based capabilities for

petroleum refinery operations, although such discussion and concepts are generally valid for other industries and applications as well.

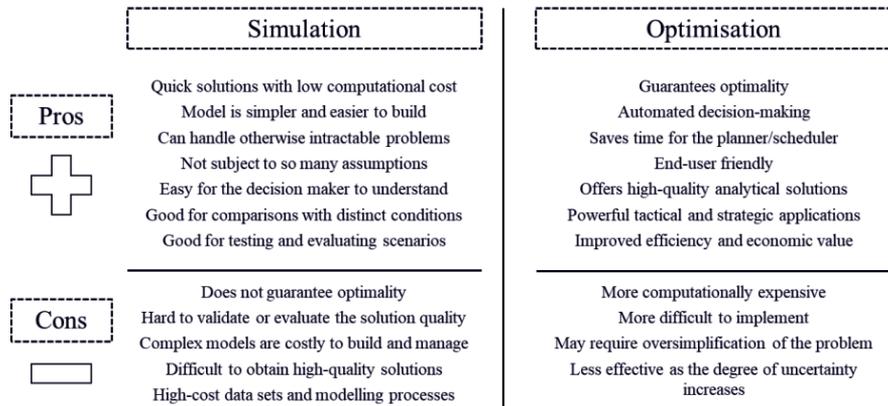


Figure 1: Simulation versus optimisation decision-making.

3. An industrial perspective on petroleum refinery decision-making

Petroleum refinery operations require determination of variables over a complex processing network, in which planning and scheduling decision-making are often performed aided by computer-based software and resources. The scheduling decision-making comprises the crude oil scheduling (feedstock edge), processing scheduling (unit-operations and streams throughout the process), and blend scheduling (product edge blending). Despite a vast amount of research and tools on scheduling optimisation, refineries worldwide still employ scheduling simulation software. That leads to feasible or suboptimal solutions, and it is typically performed based on manual or automatic trial-and-error methods (Li et al., 2020). For planning decision-making, refinery planning software are employed to determine the production throughout the plant over weeks or months. Similar to the scheduling decisions, simulation tools are still often employed to assist in the identification of financial losses or profit opportunities in the refinery through evaluation based on trial-and-error (and therefore, limited) scenarios.

There are important aspects on planning/scheduling decision-making from an industrial perspective, which comprise the interplay between simulation tools, optimisation tools, academic research, and industrial implementation, as shown in Figure 2.

Simulation versus optimisation: Simulation is still employed for problems and applications that could be better solved or handled by optimisation methods. The main reasons include insufficient knowledge to implement or develop better approaches, lack of trained or qualified personnel, fear of change, lack of investments, etc. In addition, many companies still have highly experienced planners/schedulers that manage the entire operational decision-making process, in which outdated tools are often used, and which compromises further breakthroughs and enhancements towards improved operations.

Industrial implementation of simulation tools: Trial-and-error or try-and-test simulation methodologies, either commercial or home-grown planning/scheduling, rely on simulating events to test feasibility, where the user is often responsible for manually testing different scenarios and keeping track of convoluted decision-trees that do not work

or prove feasible (Menezes et al., 2015). The simulation software is often time-consuming due to the need of continuously check and update the processing data to achieve good or proper solutions. Such procedure is not reliable, may take hours or days, and does not guarantee even feasible solutions. Furthermore, in the best-case scenario, the simulation software provides a limited number of solutions to be chosen by the planner/scheduler. This leads to exhausting efforts to model, save, and manage the numerous scenarios and to update the modelling premises and situations that constantly change. A typical outcome is that planners/schedulers abandon the solutions given by the software and return to their simpler manual spreadsheets. Hence, this translates into even less efficient decisions to be incorporated in the plant, which results in reduced operating capacity, increased process variability and uncertainty, not efficient nor reliable solutions, etc.

Industrial implementation of optimisation tools: Optimisation software can overcome some of the simulation issues often faced by planners and schedulers. However, such commercial tools are often expensive, require previous knowledge or proper training, and may be subject to resistance from the users or organisation that heavily rely on the currently used methodologies/tools and that are against or afraid of changes.

Industrial implementation of academic research: A proper discussion is required on the applicability of the methods developed in academic research into industrial processes. Planners and schedulers argue that academic research is not fully adaptable or suitable in the industrial environment, especially because of the simplifications proposed to solve complex-scope and large-sized problems. Such simplifications are often not realistic and do not provide proper capabilities for industrial implementation. Moreover, academic research that is applicable and coherent with industrial applications is often not investigated and employed by potentially interested companies.

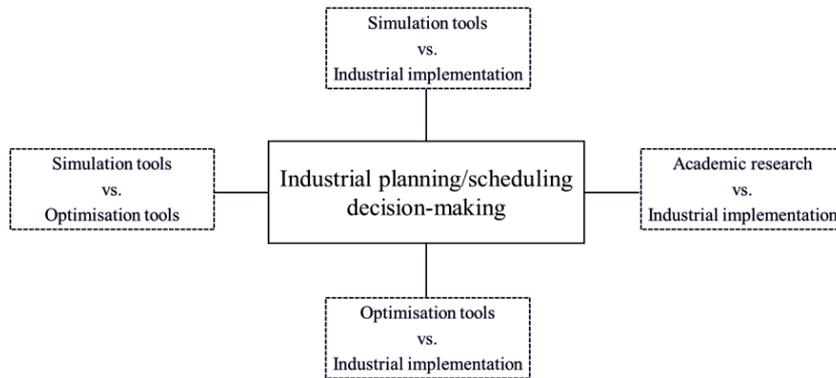


Figure 2: Interplay between simulation tools, optimisation tools, academic research, and industrial implementation.

4. Guidelines for enhanced industrial capabilities

A fundamental concept regarding the application of computational-based tools for industrial processes is that both simulation and optimisation are relevant and important, but for distinct purposes and applications. Simulation can be used to evaluate multiple distinct scenarios in order to provide information on how a given system is expected to work, i.e., system response under different conditions. There have been long joint developments in simulation tools for the industry, including strong collaboration in tailoring opportunities and solutions according to the industrial needs, graphical user

interfaces for operator interactions, and rigorous simulation software, which are widely used for controlling process conditions according to operational limitations, constraints, or requirements. Conversely, optimisation provides enhanced decision-making capabilities for identifying better solutions under specific scenarios or operating conditions. This includes finding better process conditions that lead to the optimal solution of a given problem or application.

The decision on whether to use simulation or optimisation strongly depends on the problem, application, complexity, and solution requirements. Simulation is preferred for applications in which optimisation may fail (e.g., highly degenerate, nonconvex, or complex problems) and when the operational and economic benefits of optimal solutions do not worth the requirements of resources (e.g., time, effort) for the development and implementation of optimisation methods. Simulation is indeed quite important for diverse industrial problems, but it is fundamental to highlight the importance of employing optimisation methods for applications that are optimisable, i.e., suitable to be solved by optimisation-based algorithms and tools rather than simulation-based or trial-and-error procedures. In fact, many works from the literature (Khor and Varvarezos, 2017; Franzoi et al., 2021) emphasise the importance of investigating such applications for replacing simulation by optimisation whenever the latter is suitable and applicable aiming to achieve better process efficiency and improved operations.

A key topic to be further addressed by the industry is the importance and necessity to seek better and more efficient decision-making capabilities. This involves technological and research advancements in software, tools, resources, and methodologies that provide a positive impact on the operational decision-making with increased economic value. Organisations must overcome the traditional resistance and limitations imposed by outdated tools and methods. Towards an enhanced decision-making reality within the Industry 4.0, it is fundamental to learn with the state-of-the-art technological development and to adhere to novel and improved capabilities currently available for industrial applications. Within this context, it is highly advisable to gather academic knowledge on planning and scheduling applications. This can be achieved by research and development teams, and through academic-industry collaborations. Such concepts may provide high-quality implementable research with mutual overall benefits. An illustrative example is provided as follows to exemplify industrial operations driven by simulation tools that could be better managed by their optimisation counterparts.

5. Illustrative example

The scheduling operations to be determined in a crude oil refinery are assisted by a scheduling simulation software that provides potentially feasible solutions. The software requires manual inputs and checking of information in multiple distinct simulations, which typically takes hours for an experienced scheduler to determine the solution to be implemented in the plant. Replacing it by an optimisation decision-making tool requires investments in research and software and may take months for proper development and testing. Such an optimisation capability is anticipated to provide:

- Easier and quicker rescheduling of the problem whenever needed.
- A pool of multiple optimal solutions that can be chosen by the user depending on any processing limitation or requirement.
- Possibility of continuously optimising the problem and integrating the updated solutions with the scheduling operations in an automatic and systematic fashion.

- Significant economic savings in the operations.

In addition, the solution can be checked manually to ensure high quality and consistency, and the simulation tool can be used as an alternative if the optimisation fails.

6. Conclusions

An industrial perspective on simulation and optimisation decision-making capabilities is provided herein, whereby highlighting some key gaps and challenges. Such aspects include the interplay between simulation tools, optimisation tools, academic works, and industrial implementation. Simulation is still widely employed for problems that could be rather optimised; simulation tools are often not easily or properly applicable; optimisation tools are often not available or require additional resources; and despite huge academic breakthroughs, most methodologies are not suitable for industrial implementation.

Simulation tools have been historically employed for industrial applications. Recently, optimisation decision-making has been increasingly used given the enhanced industrial capabilities (e.g., technological improvements, higher computer power) and research developments (e.g., improved algorithms and methodologies). This provides significant advancements with more efficient processes and operations. The main benefits include more flexible, adaptable, and reliable processes with higher economic value. In this work, we emphasise the importance of optimisation replacing simulation for problems and applications that are optimisable. Such concepts are widely known in the academic literature, but further improvements are still required in the industry.

The industrial decision-making is often carried out manually. Therefore, there is an urgent need to address the main difficulties, complexities, and issues on replacing simulation by optimisation from an industrial perspective. This work aims to stimulate the discussion on such meaningful topics from an industrial perspective and to provide guidance for incorporating academic knowledge by the industry, exposing the industrial requirements still unaddressed or that need further development or improvements, and assisting industrial peers on how to properly achieve such breakthroughs.

References

- B. Brunaud, H.D. Perez, S. Amaran, S. Bury, J. Wassick, I.E. Grossmann, 2020, Batch scheduling with quality-based changeovers, *Computers and Chemical Engineering*, 132, 106617.
- R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, 2018, Effective scheduling of complex process-shops using online parameter feedback in crude-oil refineries *Computers Aided Chemical Engineering*, 44, 1279-1284.
- R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, 2021, A moving horizon rescheduling framework for continuous nonlinear processes with disturbances, *Chemical Engineering Research and Design*, 174, 276-293.
- C.S. Khor, D. Varvarezos, 2017, Petroleum refinery optimisation, *Optimisation and engineering*, 18(4), 943-989.
- F. Li, M. Yang, W. Du, X. Dai, 2020, Development and challenges of planning and scheduling for petroleum and petrochemical production, *Frontiers of Engineering Management*, 1-11.
- B.C. Menezes, J.D. Kelly, I.E. Grossmann, A. Vazacopoulos, 2015, Generalized capital investment planning of oil-refineries using MILP and sequence-dependent setups *Computers and Chemical Engineering*, 80, 140-154.
- H. Yang, D.E. Bernal, R.E. Franzoi, F.G. Engineer, K. Kwon, S. Lee, I.E. Grossmann, 2020, Integration of crude-oil scheduling and refinery planning by Lagrangean decomposition, *Computers and Chemical Engineering*, 138, 106812.

Decomposition of Two-stage Stochastic Scheduling Problems via Similarity Index

Daniel Montes^{a,*}, José L. Pitarch^b and César de Prada^{a,c}

^a*Department of Systems Engineering and Automatic Control, University of Valladolid, C/ Real de Burgos s/n, Valladolid, 47011, Spain*

^b*Instituto Univ. de Automática e Informática Industrial (ai2), Universitat Politècnica de Valencia, Camino de Vera s/n, Valencia, 46022, Spain*

^c*Institute of Sustainable Processes, University of Valladolid, C/ Dr. Mergelina s/n, Valladolid, 47011, Spain*

**danielalberto.montes.lopez@uva.es*

Abstract

Two-stage stochastic scheduling problems (TSSP) are computationally demanding and challenging to solve in a monolithic fashion, as they often involve numerous discrete and continuous variables. One option is to apply decomposition methods (e.g. Benders, Progressive Hedging, etc.) with the idea of splitting the original problem into smaller ones that progressively approach the solution of the monolithic formulation. Here, we propose an alternative decomposition method based on the so-called Similarity Index (*SI*). The *SI* measures the similarity of schedules along a discretized time horizon. The non-anticipativity constraints of the TSSP are removed from the model and an *SI* is incorporated instead in the objective function to be maximized. Then, the original problem can be divided into subproblems on a scenario basis. An iterative procedure is set up so that the subproblems solutions progressively increase the global *SI*. The algorithm stops when the *SI* has reached a value of 1, i.e, the first-stage decisions are identical for all subproblems. The *SI* decomposition was successfully applied to an industrial-like TSSP of an evaporation network, using different MILP solvers: CPLEX and Gurobi. Two instances of different size and scenarios were considered. For the small one, significant improvements over the monolithic approach were not observed. However, for the large instance, our decomposition method was considerably faster.

Keywords: Online Scheduling, Decomposition Methods, Uncertainty, MILP

1. Introduction

Two-stage stochastic formulations are a popular way to take uncertainty into account in scheduling formulations. In two-stage scheduling problems (TSSP), the decision variables are split into two categories: here and now decisions (first-stage variables) which have to be taken now with the available information and cannot be changed later; and wait and see decisions (recourse variables) which are adjusted according to the realization of the uncertain parameters. This results in large-scale problems that involve both continuous and discrete variables which are computationally demanding to solve. Hence, uncertainty is often ignored in online or closed-loop scheduling, as such applications require providing solutions in real-time. One possibility to overcome this obstacle is to use decomposition methods. Unfortunately, not many ideas have arisen different to the classic approaches: the Benders Decomposition (Benders, 1962) and the Progressive Hedging (PH) algorithm (Rockafellar and Wets, 1991).

The Benders method, and its variations, is widely employed but has a high computational burden too, as each iteration generates either an optimality or a feasibility cut to reduce the feasible region of the master problem (Rahmaniani et al., 2017). The PH algorithm decomposes over scenarios by computing suitably penalties with average values of the decision variables in each iteration (Peng et al., 2019). However, the discrete nature of some first-stage decisions may provoke that the solutions obtained by the independent scenarios lead to opposite values in each iteration, so convergence of PH to a feasible solution is not guaranteed.

Here we propose a novel decomposition method based on the idea of measuring how *similar* the first-stage decisions are among the different scenarios. Such single measure is provided by the so-called similarity index. All subproblems are solved simultaneously in an iterative way, updating a weight for maximizing the similarity index. The proposal is tested in an industrial-like case study.

The concept of similarity index and the algorithm of the decomposition method are described in Section 2. Section 3 presents the case study that is used for testing our proposal. The results are presented in Section 4. Lastly, the paper closes with Section 5 where the conclusions are presented.

2. Similarity Index

The Similarity Index (SI) in a TSSP, first proposed by Palacín et al. (2018), is an aggregated measure of how similar the recourse variables are among the scenarios, i.e., an indication of the solution robustness. It is calculated by fuzzifying the discrete decisions over a specified time horizon, so that a decision taken at instant t is weighted by 1 (100%), but it also influences the nearby time instants t_n with a decreasing value that is proportional to the time difference $|t - t_n|$. See Figure 1 for a graphical representation. The SI is then computed as the total intersection area among the scenarios schedules, normalized by the maximum possible area (where all the solutions would coincide). Consequently, an SI equal to 1 means that the schedules for all the scenarios are identical.

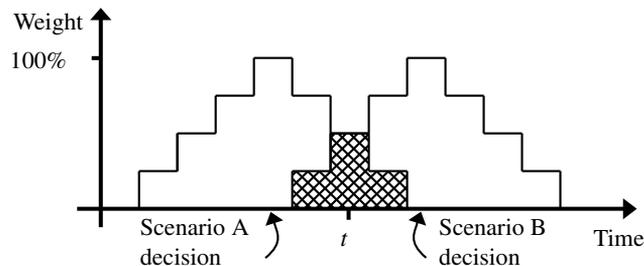


Figure 1: Example of fuzzifying the same discrete decision along seven time periods in two scenarios. The SI is computed from the intersection between both scenarios A and B and the total fuzzified area.

There is no reason to restrict the SI to the recourse variables. In a TSSP, by definition, the non-anticipation criterion enforces the first-stage variables to be identical. Therefore, if the SI were applied to them, it would necessarily take the value of 1. However, if the non-anticipativity constraints are removed from the formulation, the global model would be block-separable on a scenario basis, i.e., each scenario could be solved as an independent scheduling problem. Those solutions would not coincide in the first-stage and the SI would be lower than 1. However, an iterative procedure could be set up so that the subproblems (scenarios) are solved iteratively while pushing the SI up to the point where the solutions coincide in the first stage.

From now on, just fuzzification along the two neighboring time periods is considered to enhance the understanding of the decomposition method. The formula for computing the similarity index

for first-stage variables in this case is shown in (1).

$$SI := \sum_{t=1}^{t_R} \frac{\min_{e \in \mathcal{E}} \{y_{te} + 0.5y_{(t-1)e}|_{t>1} + 0.5y_{(t+1)e}|_{t<t_R}\}}{2t_R - 1} \quad (1)$$

Where t_R is the end of the robust horizon, \mathcal{E} is the set of uncertainty realization scenarios, and $y_{te} \in \{0, 1\}$ are the binary variables of the scheduling problem. Decision variables outside the robust horizon do not contribute towards the SI calculation.

Note that the $\min\{\cdot\}$ operation is nonlinear, so (1) cannot be used in linear scheduling formulations. Nonetheless, a lower bound for the SI can be computed using slack variables s_t and inequalities as in (2) and (3). The slack variables need to be maximized in the objective function of the optimization problem so that the lower bound on the SI is tight. Consequently, the right-hand side of (2) is incorporated in the objective function.

$$SI \geq \sum_{t=1}^{t_R} \frac{s_t}{2t_R - 1} \quad (2)$$

$$s_t \leq y_{te} + 0.5y_{(t-1)e}|_{t>1} + 0.5y_{(t+1)e}|_{t<t_R} \quad \forall e \in \mathcal{E}; t : 1, \dots, t_R \quad (3)$$

2.1. Decomposition Algorithm

Two similarity indexes need to be computed. The *Local Similarity Index* (SI_e) is computed as part of the optimization subproblems. The SI_e are maximized individually in each subproblem so that the similarity of the local solution with respect to the other scenarios is as high as possible. Hence, the computation of the local SI_e requires information from the other scenarios. For the subproblems to be separable, a previous solution from the other subproblems needs to be used. Such *reference* solution is the one corresponding to the scenario that yielded the worst local SI_e in the previous iteration. Using all the scenarios solutions can lead to competitive behavior among subproblems, so the global problem may get stuck in an infeasible solution and convergence is never reached. A *Global Similarity Index* (SI) is computed using the solutions collected from the subproblems. It is used to update a parameter that progressively increases the importance of maximizing the local SI_e . The subproblems read as in (4).

$$\begin{aligned} & \underset{y_{te}, s_t}{\text{minimize}} && J_e - \lambda \sum_{t=1}^{t_R} \frac{s_t}{2t_R - 1} \\ & \text{subject to} && \text{Local model constraints} \\ & && s_t \leq \bar{y}_t + 0.5\bar{y}_{t-1}|_{t>1} + 0.5\bar{y}_{t+1}|_{t<t_R} \quad t : 1, \dots, t_R \\ & && s_t \leq y_{te} + 0.5y_{(t-1)e}|_{t>1} + 0.5y_{(t+1)e}|_{t<t_R} \quad t : 1, \dots, t_R \end{aligned} \quad (4)$$

Where J_e is the original objective function and λ is a Lagrange-like multiplier that is updated in each iteration using a formula inspired by the sub-gradient method (Shor, 1985) as in (5). The variables \bar{y}_t contain the solution of the scenario that yielded the worst SI_e in the previous iteration.

$$\lambda_{k+1} = \lambda_k - \alpha_{k+1} (SI - 1) \quad (5)$$

In (5), α is the step size, decreased in each iteration. For instance, $\alpha_{k+1} = 0.9\alpha_k$.

The proposed algorithm starts by initializing the multiplier parameter λ and the discrete decision variables in the previous iteration \bar{y}_t to zero, which implies that the local SI_e is set to zero in the first iteration. The next steps of the algorithm comprise the calculations that are repeated until the stop criteria are met. The subproblems are solved in parallel and both their solutions and their local SI_e are collected. The global SI is calculated using (1). The solution of the scenario that yielded

the worst local SI_e is assigned to the variables \bar{y}_t that will be used in the subsequent iteration to re-compute the local SI_e . The multiplier parameter λ is updated using (5) and the termination criteria is checked. If neither the global SI has reached 1 nor the maximum iterations have been reached, the process is repeated. The solution of the subproblems y_{te}^* is reported after the algorithm has converged. Algorithm 1 summarizes these steps.

Algorithm 1 Similarity Index Decomposition

Require: $\alpha_0, \text{tol}, k_{\max}$

- 1: $k \leftarrow 0, \lambda_0 \leftarrow 0, \bar{y}_t \leftarrow 0$ ▷ Initialization
- 2: **repeat**
- 3: **for** e in \mathcal{E} **do**
- 4: $y_{te}^*, s_t^* \leftarrow \arg \min_{y_{te}, s_t} J_e - \lambda_k SI_e$ ▷ Solve (4)
- 5: $SI_e^* \leftarrow SI_e(s_t^*)$
- 6: **end for**
- 7: $SI \leftarrow \text{Eq. (1) with } y_{te}^*$ ▷ Global SI computation
- 8: $\bar{y}_t \leftarrow \arg \min_{y_{te}^*} \{SI_e^*\}$ ▷ Solution with worst SI_e
- 9: $\alpha_{k+1} \leftarrow 0.9\alpha_k$ ▷ Multiplier update
- 10: $\lambda_{k+1} \leftarrow \lambda_k - \alpha_{k+1} (SI - 1)$
- 11: $k \leftarrow k + 1$
- 12: **until** $|SI - 1| \leq \text{tol} \vee k = k_{\max}$ ▷ Convergence check
- 13: **return** y_{te}^*

3. Case study

The similarity index decomposition was tested in an industrial-like case study of an evaporation network. Each evaporation plant is made up of two evaporation chambers with several heat exchangers in between. The objective is to concentrate an inlet stream of a certain product so that it is recycled to the main process of the factory. The plants need to be cleaned regularly as fouling increases steam consumption and production costs. The scheduler must assign products and loads to plants and decide when to stop the plants for cleaning.

The allowed transitions between stages are clearly defined: A plant that is under operation can go to standby or remain in operation. A plant in standby can go to a cleaning stage or remain on standby. A plant that is in a cleaning stage can go directly to operation or go to standby. And finally, a clean plant on standby can remain in such a state or start operating. Additionally, the model shall account for: a plant can only process a product at a time; only one cleaning stage can occur per day; if a product has been assigned to a plant, it cannot be changed unless the plant is cleaned; a terminal cost is established for plants that end up in an advanced fouling state at the end of the scheduling horizon, etc.

Palacín et al. (2018) formulated a two-stage scheduling problem for this case study using the pre-defined precedence approach. Here, we recall their example and apply our decomposition method to two different instances of the problem. The model equations are not displayed for brevity. Using the notation in such paper, Equation (1) for computing the global SI becomes:

$$SI = \sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{M}_R} \frac{\min_{e \in \mathcal{E}} \{E_{vtse} + 0.5E_{v(t-1)se} |_{t>1} + 0.5E_{v(t-1)se} |_{t < t_R}\}}{n_v (2t_R - 1)} \quad (6)$$

Where \mathcal{V} is the set of evaporation plants, \mathcal{S} is the set of possible stages, \mathcal{M} is the set of days in the discretized scheduling horizon, \mathcal{P} is the set of products, and \mathcal{E} is the set of uncertainty realization scenarios. The binary variables E_{vtse} specify if a plant v at day t is at stage s in the scenario e . The total number of evaporation plants is denoted by n_v .

The local similarity index is computed using (2) and (3), updated to the case study notation as shown in (7) and (8) respectively.

$$SI_e = \sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{M}_R} \frac{s_{vts}}{n_v(2t_R - 1)} \tag{7}$$

$$\begin{aligned} s_{vts} &\leq \bar{E}_{vts} + 0.5\bar{E}_{v(t-1)s|t>1} + 0.5\bar{E}_{v(t+1)s|t<t_R} \\ s_{vts} &\leq E_{vts} + 0.5E_{v(t-1)se|t>1} + 0.5E_{v(t+1)se|t<t_R} \quad \forall v \in \mathcal{V}, s \in \mathcal{S}, t \in \mathcal{M}_R \end{aligned} \tag{8}$$

Where \bar{E}_{vts} are the binary variables of the subproblem that yielded the worst SI in the previous iteration. Consequently, \bar{E}_{vts} are not decision variables of the optimization problems. Each subproblem is then formulated using the model equations in Palacín et al. (2018), except for the non-anticipativity constraints, plus inequalities (8).

4. Results

Two problem instances of the case study were analyzed, denoted by A and B. In instance A, the scheduling problem was solved for three plants, two products, and four scenarios for the products demands. The resulting problem had 18000 variables, 17280 of which were binary. The subproblems had 4320 binary variables each. Instance B consisted of a larger problem with nine plants, three products, and sixteen scenarios for the products demands. The resulting model had 196000 variables, 90% of which were binary. The sub-problems had 11025 binary variables each.

Both instances were solved in GAMS 36.2.0 using IBM ILOG CPLEX 20.1 and Gurobi 9.1. The maximum execution time was set to one hour and the relative gap tolerance to zero. The default values were used for the rest of the solvers options. The monolithic problems of both instances were solved using concurrent optimization on 32 CPU threads. The subproblems of the decomposed approach were solved using 8 threads in instance A, and 2 in instance B.

The execution time and the final relative gap for each solver are reported in Tables 1 and 2 for instances A and B respectively.

Table 1: Instance A: Three plants, two products and four scenarios.

Approach	CPLEX		Gurobi	
	Gap	Time	Gap	Time
Monolithic	0.00%	30.84s	0.00%	7.46s
Decomposed	0.00%	26.88s	0.00%	20.91s

For instance A, both solvers could reach the optimal solution in less than a minute with the monolithic approach. Our decomposition method outperformed the monolithic approach only when using CPLEX. Using Gurobi, it was around three times slower. However, this problem instance is relatively small and the advantages of using the decomposition method are only fully realized in large settings as instance B. The algorithm converged in five iterations.

Table 2: Instance B: Nine plants, three products and 16 scenarios.

Approach	CPLEX		Gurobi	
	Gap	Time	Gap	Time
Monolithic	0.76%	3600s*	0.00%	2590.12s
Decomposed	0.00%	40.58s	0.00%	42.89s

For instance B and using CPLEX, the monolithic approach only managed to get a 0.76% gap solution within the specified time limit. The decomposed approach greatly outperformed it by arriving to the optimal solution in only 40.58 sec. Furthermore, when the time limit was increased to one day, CPLEX could only get a 0.44% gap solution. It is clear that the solver struggles to improve the optimality gap after finding a feasible solution. Unlike CPLEX, Gurobi managed to get an optimal solution within the time limit with the monolithic approach, specifically 2590.12 seconds. However, our decomposition algorithm was around sixty times faster, arriving to the optimum in just 42.89 seconds. The decomposed approach only needed two iterations to converge.

5. Conclusions

This work proposed a novel decomposition method that uses a single indicator to measure the similarity between discrete-time schedules. The non-anticipativity constraints are removed from the formulation, which enables a scenario-based decomposition. Each scenario is solved as an independent problem that includes the maximization of the similarity among the other scenarios. An iterative procedure is repeated until the first-stage variables of all the subproblems coincide. This setting enables parallelization beyond the native capabilities of the optimization solvers.

Note that the algorithm only has two tuning parameters: the step-size decrease rate α and the length of the time horizon along the discrete decisions are fuzzified for computing the similarity index. The proposed method was tested in two problem instances of different sizes. In the smaller one, significant improvements were not observed when using the decomposition approach. However, in the larger instance, the computation time required for achieving optimal solutions was significantly reduced with the decomposition method. When using CPLEX, the monolithic approach could not arrive to an optimal solution in one hour. In contrast, our decomposition method managed to arrive to an optimal solution in less than a minute. When using Gurobi, the optimal solution was obtained around sixty times faster. In conclusion, the potential of our proposal is fully realized as the number of scenarios increases.

The convergence properties of the similarity index decomposition have not been studied yet and future work will focus on it. We also aim to extend the method to non-linear mixed-integer (MINLP) scheduling problems and continuous-time scheduling formulations.

6. Acknowledgements

These results are part of the InCO4In (PGC2018-099312-B-C31) and LOCPU (PID2020-116585GB-I00) research projects that are supported by the EU with FEDER funds, as well as by the Spanish MICINN and AEI. The first author has received financial support from the 2020 call of the pre-doctoral contracts of the University of Valladolid, co-financed by Banco Santander.

References

- J. Benders, 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4, 238–252.
- C. G. Palacín, J. L. Pitarch, C. Jasch, C. A. Méndez, C. de Prada, 2018. Robust integrated production-maintenance scheduling for an evaporation network. *Computers and Chemical Engineering* 110, 140–151.
- Z. Peng, Y. Zhang, Y. Feng, G. Rong, H. Su, 2019. A Progressive Hedging-Based Solution Approach for Integrated Planning and Scheduling Problems under Demand Uncertainty. *Industrial and Engineering Chemistry Research* 58 (32), 14880–14896.
- R. Rahmaniani, T. G. Crainic, M. Gendreau, W. Rei, 2017. The benders decomposition algorithm: A literature review. *European Journal of Operational Research* 259, 801–817.
- R. T. Rockafellar, R. J.-B. Wets, 1991. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research* 16, 119–147.
- N. Z. Shor, 1985. *Minimization Methods for Non-Differentiable Functions*, 1st Edition. Springer, Berlin.

A New MILP Formulation for Scheduling of Crude Oil Operations

Pelin Dologlu,^a Sena Kurban,^a İrem Marttin,^{a,b} Nazlı Ataman,^a Gizem Kuşoğlu Kaya,^a Funda İşeri,^a Özgür Kabak,^b Y. Ilker Topcu^b

^a*Turkish Petroleum Refinery, 41790, Körfez, Kocaeli, Turkey*

^b*Industrial Engineering Department, Faculty of Management, Istanbul Technical University, 34367, Istanbul, Turkey*
pelindologlu@gmail.com

Abstract

Scheduling of crude oil operations is one of the most critical and challenging problems in petroleum refineries. Crude oil operations have more impact on the profitability of the whole refinery as these operations are the first steps of refinery operations. Crude oil operations scheduling is a complex problem as many complicated relationships and many conditions must be considered. In a refinery, a planner tries to schedule all crude oil operations by using a wide variety of complex and very dense data sets. Daily planning is mainly based on manual scheduling by trial and error method and operational experiences. Therefore, as a planner, it is almost impossible to reach the optimal solution or even a satisfactory feasible solution by handling all the data sets in a short period. Unfortunately, there is no efficient and adaptive tool or software that can be used for different refinery scheduling cases. In this study, an efficient scheduling and allocation model for an oil refinery located in İzmir is developed for unloading crudes into multiple storage tanks from the vessels arriving at various times, blending crudes, and feeding the crude distillation units from these tanks at various rates over time. To tackle the scheduling problem effectively, a Mixed Integer Linear Program (MILP) is formulated. The model is solved for the refinery, and satisfactory solutions are gathered.

Keywords: crude oil operations, short-term scheduling, MILP, optimization.

1. Introduction

Scheduling of crude oil operations is one of the most critical and challenging problems in oil refineries. Crude oil operations have a high impact on the continuity and profitability of the refinery, as it is the first step of the all refinery operation process. Besides, crude oil operations scheduling is very complex due to continuous, discrete, and binary variable optimization, and it is known as NP-hard (Li-ping & Nai-qi, 2011).

Scheduling tool provides some utilities such as (1) crude storage optimization, (2) the best evaluation of the monthly plan implementation, (3) guiding the operations, (4) optimal sustainable operations even with unexpected changes, and (5) comparison actual operating results with planning objectives. It also enables to use of superior capacity, increases plant efficiency and improves visibility and control throughout the supply chain. Therefore, the necessity of computer aided methods to optimize refinery operations should be considered in quick response to problems and opportunities and making a decision (Reddy et al., 2004a). The optimization of production plans is a hot topic for both industry and academia due to its advantages.

Several mathematical methods such as linear programming (LP), mixed-integer linear programming (MILP), and mixed-integer nonlinear programming (MINLP) are used to solve scheduling problems. Also, numerous modeling approaches were suggested to handle various scheduling problems related to different production environments, process constraints, and features. Especially, time representation is one of the most important features of modeling approaches taken in models, as it determines many aspects consisting of modeling flexibility, size, and constraints types (Lee et al., 2019). All present formulations of timing can be categorized into two main types, namely discrete and continuous-time formulations. The first attempts in scheduling problems are based on the discrete-time approach in which the time horizon is split into a count of time slots of uniform durations and events (Shah, 1996; Reddy et al., 2004b and Saharidis et al., 2009). Unlike the discrete-time approach, the time slots are not uniform in the continuous-time approach so that events can start and end at any point in the continuous-time domain (Yadav et al., 2012; Chen et al., 2012 and Karuppiyah et al., 2008).

In this paper, a new MILP model is formulated for the refinery scheduling problem involving crude oil unloading from vessels, transferring crude between tanks, and charging the crude distillation units (CDUs). The long-term scheduling problem of an oil refinery located in İzmir is addressed to find the optimum schedule. Compared to other refinery structures in the literature, the tanks are not categorized as charging and storage tanks in this refinery. The tanks here are used for both charging and storage. Also, around 50 types of crude oil can mix in the tanks. Several specific operational features make refinery management as a complex problem. There is no effective computer tool or study in the literature for such a complex oil refinery to our best knowledge. The motivation in the study is that the model formulation reflects real characteristics of the refinery structure and operations, so it is valuable for the theory to take shape in an industrial applications. The discrete-time model is preferred in this study as it allows handling the constraints more easily than continuous models.

2. Materials and Methods

In this section, the process to be analyzed is described, and then the MILP formulation to optimize the process is presented.

2.1. Process Description

In this paper, the scheduling of crude oil operations is conducted for an oil refinery located in İzmir, Turkey. Since, the planning department makes decisions for very complex crude oil operations manually, it is impossible to reach the optimal solution or even a feasible solution by handling all the data sets in a short period. Thus, it is crucial to develop an optimization tool for an effective management of crude oil operations.

Crude oil operations of the refinery are difficult to determine manually with the ever-increasing constraints. Operations in the İzmir Refinery consist of 1 docking station, 12 crude oil tanks, and 2 CDUs. An overview of crude oil operations is shown in Figure 1. Various crude oil types are shipped to the refinery and processed according to their yield. There are about 50 different types of crude oil that have been received and processed by the refinery up to now, and more different types of crude oil may also be shipped. It is very hard to process varying types of crude oil because of immiscibility and precipitation problems. Since it is hard to manage to have one storage tank for each 50 or many crude oil types; crude oil storage, blending and charging operations are handled in the same tank. Also, blending may occur in the pipelines if multiple tanks are opened. Besides, some of the crude oil types have unique properties and must be stored in particular tanks.

Therefore, tanks are labeled as national and non-national, and also with a pump and without a pump. Domestic crude oils should be stored at specific tanks because of government regulations. Also, specific types of crude oils are stored in tanks that have pumping infrastructure installed to feed continuously for economic reasons. However, this brings out other constraints related to pumps. Furthermore, operations on tanks are complex because of the piping system and measuring tank content. As unloading from vessels to tanks and transfers between tanks are using the same piping system, these operations cannot be done simultaneously. Also, only unloading or charging operations for one tank are allowed to measure tank content correctly. Operability to complex tank processes is the main concern in the refinery, and thus, there are too many constraints related to tank operations. Having too many crude oil types causes operational burdens in CDUs. Each crude oil has widely varying properties, processability and product yields, so they have minimum and maximum rates to be processed in each CDUs. Most of these constraints are unique to the İzmir Refinery, and there is limited information in the literature.

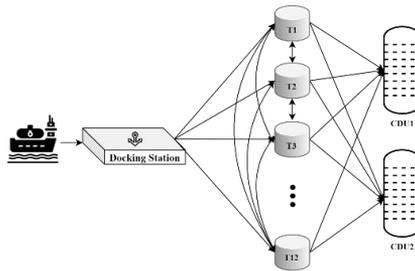


Figure 1. An overview of crude oil operations

2.2. Proposed MILP

In this section, the proposed MILP model is presented that corresponds to the scheduling problem described above. The model uses a discrete-time formulation that includes 24-hour intervals over a scheduling horizon of 60 days. The following assumptions for the model are considered: (1) Perfect mixing occurs at each tank. (2) Remaining crude oil in the pipeline during the unloading is neglected. (3) Resting times in tanks are ignored.

The modeled system has $t \in T$ days, $i \in I$ crude oil tanks, $j \in J$ CDUs, $h \in H$ crude oil types, and there are four operations $o \in O$ which are unloading, transfer-sending, transfer-receiving, and charging. Crude oils have $p \in P$ set of distinguishing properties such as American Petroleum Institute (API) gravity and sulfur. Tanks, pumps, and CDUs have $m \in M$ set of minimum and maximum limits in terms of these features and volume limits. Lastly, $k \in K$ vessels arrive at the system over the planning horizon.

The parameters of the model are set as follows: the initial volumes of tanks, $VI_{h,i}$ (m^3); API-sulphur contents and blending properties of crude oil types, $PR_{h,p}$ and $TB_{h,hh}$, respectively; minimum and maximum capacity of tanks, LC_i and UC_i ; whether the tank has a pumping system, E_i ; unloading limits to each tank, UL_i ; tolerance limit to obtain whether a crude oil type is going to be considered for blending, To_i ; maximum transfer capacity between tanks, $FL_{i,ii}$. Vessels have three descriptive parameters which are arriving times $KT_{k,t}$, volume and crude oil type of each vessel, KV_k , and $KH_{k,h}$ respectively. Also, the parameters for the CDU characteristics are total processing volume, $TD_{i,j}$; maximum charging capacity from each tank to each CDU, $DL_{i,j}$; min-max limits of processing each crude oil in each CDU, $MR_{i,h,j,m}$; API and Sulphur content limits of each CDU, LA_j , UA_j and US_j , respectively. Finally, $card^I$ shows the count of the total number of elements within the set I .

While scheduling, unloading, transfer, and charging amounts and the volume of tanks in terms of crude oil types for each time period are determined, so the variables for them are defined as $u_{t,i,k}$, $tr_{t,ii,h}$, $c_{t,i,j,h}$, and $v_{i,t,h}$ respectively. The API ($a_{i,j}$) and Sulphur ($s_{i,j}$) values of CDUs are needed to be known for each CDU. To ensure operating conditions, some binary variables are defined. Let $ub_{t,k}$ and $cb_{t,i,j}$, show whether there is any unloading and charging operations, respectively. $tk_{i,t,o}$ shows whether a tank is assigned to an operation. Lastly, $y_{i,i,h}$ represents whether the crude oil types exist in the tanks.

The planning process is required to be managed with minimum transfer operation as it restricts the other crucial operations. Thus, the objective function is the minimization of total transfer operations, Eq. (1). The first set of constraints include material balance, pump, and operational limitations. Material balances are given in Eq. (2). Eq (3) states that the volume $v_{i,t,h}$ must be within the capacity limits of tanks. Eqs. (4-6) ensures that the amount of flow in unloading, transfer, and charging (if there is any) operations should not exceed the pump limits. There are 2 CDUs, so there can be two charging operations for each tank per time period, Eq. (7). The operational limits in charging for pumping tanks are given Eqs. (8a-8c). Only one transfer or unloading operation is allowed at the same time due to the line structure of the system, given in Eqs. (9a-10). There cannot be simultaneous inflow and outflow to the tank. However, there can be two outflows, Eq. (11). As a result of this condition, simultaneous unloading and transfer operations are not allowed for a tank, Eq. (12).

$$\min Z = \sum_t \sum_i \sum_{ii} \sum_h (tr_{t,ii,h}) \quad (1)$$

$$v_{i,t,h} = v_{i,t-1,h} + \sum_k (u_{t,i,k} * KH_{k,h}) + \sum_{ii} (tr_{t,ii,i,h}) - \sum_{ii} (tr_{t,ii,i,h}) - \sum_j (c_{t,i,j,h}) \forall i, h, t \quad (2)$$

$$LC_i \leq \sum_h (v_{i,t,h}) \leq UC_i \forall i, t \quad (3)$$

$$\sum_k (u_{t,i,k}) \leq tk_{i,t,1} * UL_i \forall i, t \quad (4)$$

$$\sum_{ii} \sum_h (tr_{t,ii,i,h}) \leq tk_{i,t,2} * FL_{i,ii} \forall i, t \quad (5a,5b) \quad \sum_{ii} \sum_h (tr_{t,ii,i,h}) \leq tk_{i,t,3} * FL_{i,ii} \forall i, t$$

$$\sum_h (c_{t,i,j,h}) \leq cb_{t,i,j} * DL_{i,j} \forall t, i, j \quad (6)$$

$$\sum_j (cb_{t,i,j}) \leq tk_{i,t,4} * 2 \forall i, t \quad (7)$$

$$\sum_i (cb_{t,i,j} * E_i) \leq 2 \forall t, j \quad \sum_i \sum_j (cb_{t,i,j} * E_i) \leq 3 \forall t \quad \sum_i \sum_j (cb_{t,i,j} * (1 - E_i)) \leq 2 \forall t \quad (8a,8b,8c)$$

$$\sum_i (tk_{i,t,2}) \leq 1 \forall t \quad \sum_i (tk_{i,t,3}) \leq 1 \forall t \quad (9a,9b)$$

$$\sum_k (ub_{t,k}) \leq 1 \forall t \quad (10)$$

$$tk_{i,t,1} + (0.5 * tk_{i,t,2}) + tk_{i,t,3} + (0.5 * tk_{i,t,4}) \leq 1 \forall t, i \quad (11)$$

$$card^i * (1 - \sum_k (ub_{t,k})) \geq \sum_i (tk_{i,t,2}) \forall t \quad (12)$$

Constraints related to unloading operations are as follows: Eq. (13) ensures that all the volume of that vessel should be unloaded. However, the whole volume of vessel may not be unloaded in just one time period. Thus, the sum of volume unloaded from one vessel in each time period must be equal to the total volume of that vessel, Eq. (14). Vessels should be unloaded sequentially, Eq. (15).

$$\sum_t \sum_i (u_{t,i,k}) = KV_k \forall k \quad (13)$$

$$ub_{t,k} * KV_k \geq \sum_i (u_{t,i,k}) \forall t, k \tag{14}$$

$$ub_{t,k} + ub_{t,t,k-1} \leq 1 \forall k, t < tt \tag{15}$$

The total charging amount may change on daily basis. Eq. (16) states that the total charging amount must be equal to the input value. Same conditions are also applied to the ratio of crude oil types that are going to be charged into CDUs, Eq. (17).

$$\sum_i \sum_h (c_{t,i,j,h}) = TD_{t,j} * 24 \forall t, j \mid TD(t, j) \neq 0 \tag{16}$$

$$\sum_i (c_{t,i,j,h}) \leq TD_{t,j} * 24 * 0.01 * MR_{t,h,j,max} \forall t, h, j \mid MR_{t,h,j,max} \neq 0 \wedge MR_{t,h,j,max} \neq 100 \tag{17}$$

The calculation of API and Sulphur values belonging to the CDUs are stated in Eqs. (18-19). The API and Sulphur contents must then be within given bounds, Eqs. (20-21)

$$a_{t,j} = \left(\sum_i \sum_h (PR_{h,API} * c_{t,i,j,h}) \right) / (TD_{t,j} * 24) \forall t, j \tag{18}$$

$$s_{t,j} = \left(\sum_i \sum_h (PR_{h,SULPHUR} * c_{t,i,j,h}) \right) / (TD_{t,j} * 24) \forall t, j \tag{19}$$

$$LA_j \leq a_{t,j} \leq UA_j \forall t, j \tag{20}$$

$$s_{t,j} \leq US_j \forall t, j \tag{21}$$

Finally, blending constraint is handled in Eq. (22). Eq. (23) states that if the volume of *h* crude oil type is above the tolerance limit, then the binary variable takes the value 1 as it should be considered in the blending constraint.

$$y_{t,i,h} + y_{t,i,h,h} \leq 1 \forall i, t \geq 2, h > hh \mid TB_{h,h,h} = 0 \tag{22}$$

$$v_{i,t,h} - TO_i \leq y_{t,i,h} * UC_i \forall t, i, h \tag{23}$$

3. Results and Discussions

In line with the needs of the planning department, the time horizon is chosen as 60 days. The time interval of 24 hours is used. The selected industrial case study contains 20 vessels shipped 9 different types of crude oil, 2 CDUs, 12 tanks (2 tanks are in maintenance), including 13 crude oil types. The volume of vessels varies between 40,000 and 175,000 m³ and total volume of 20 vessels is about 1.9 million m³. The final model includes 93,398 variables, of which are 10,524 binary, and 28,829 constraints. The scheduling problem is solved by CPLEX in GAMS 28.2.0 using an Intel Xeon Gold 6252 CPU @2.10 GHz with 12.0 GB of RAM. The optimum solution is obtained within 1191.09 seconds of CPU time. The MILP model results are shown in Figure 2 for the first 10 days of the planning period.

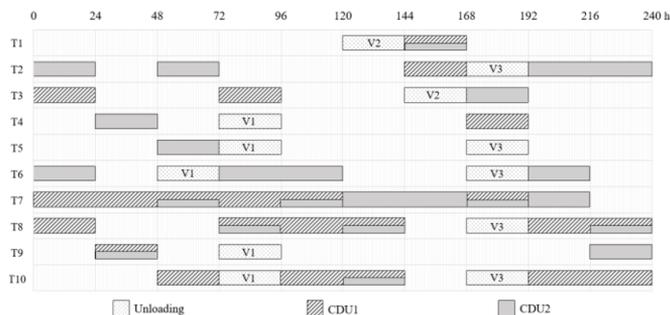


Figure 2. Gantt chart of optimal solution for first 10 days -of planning period

The results of MILP is compared with the planning department's schedule in Table 1. According to the results, there are no transfer operations in MILP results, because the objective function of the MILP is the minimization of total transfer operations. It is also observed that the computation time of the MILP model is shorter than the working hours of the experienced planner and selected tanks, charge compositions based on actual results are compatible with the MILP model results. On the other hand, more tanks are used to unload each vessel's operations, and the total number of switching tanks for charging to CDU is more in MILP result than the planning department's schedule. Therefore, we may conclude that the results of MILP are promising and satisfactory for the given case. However, there is still room for improvement.

Table 1. Comparison of MILP and planning department's schedule

	MILP Model	Planning Department
Computation time	1191.09 sec. (0.33 h)	7200 sec. (2 h)
Total number of transfer operations / Duration	0	3 / 20 h
Penalty time of vessels	24 h for 3 rd vessel (V3)	-
Total duration of unloading operations	120 h	58 h
Total number of used tanks for unloading operations	12	7
Total number of switching tanks for charging to CDU	54	26

4. Conclusions

The industrial applications of optimization-based scheduling methods are highly challenging tasks. The proposed discrete MILP model supports optimal decision making in operational planning processes of the refinery and reflects real-life scheduling problems. The model is promising in terms of support in decision making of production planning. As a further research, a multiple objective programming can be developed for taking account the other criteria given in Table 1. Furthermore, continuous-time modeling can be considered as a better optional technique to identify the planning conditions, and the model can be extended to the other refinery cases.

References

- B. Li-ping, and W. Nai-qi, 2011, Short-term scheduling of crude oil operations and its complexity, *Industrial Engineering Journal*, 14(1), 67.
- G.K. Saharidis, M. Minoux, Y. Dallery, 2009, scheduling of loading and unloading of crude oil in a refinery using event-based discrete time formulation, *Computers & Chemical Engineering*, 33(8), 1413-1426.
- H. Lee, C.T. Maravelias, 2019, Combining the advantages of discrete-and continuous-time scheduling models: Part 2. systematic methods for determining model parameters, *Computers & Chemical Engineering*, 128, 557-573.
- N. Shah, 1996, Mathematical programming techniques for crude oil scheduling, *Computers & Chemical Engineering*, 20, 1227-1232.
- P.C.P. Reddy, I.A. Karimi, R. Srinivasan, 2004a, A new continuous-time formulation for scheduling crude oil operations, *Chemical Engineering Science*, 59(6), 1325-1341.
- P.C.P. Reddy, I.A. Karimi, R. Srinivasan, 2004b, Novel solution approach for optimizing crude oil operations, *AIChE Journal*, 50(6), 1177-1197.
- R. Karuppiah, K.C. Furman, I.E. Grossmann, 2008, Global optimization for scheduling refinery crude oil operations, 32(11), 2745-2766.
- S. Yadav, M.A. Shaik, 2012, Short-term scheduling of refinery crude oil operations, *Industrial & Engineering Chemistry Research*, 51(27), 9287-9299.
- X. Chen, I. Grossmann, L. Zheng, 2012, A comparative study of continuous-time models for scheduling of crude oil operations in inland refineries, *Computers & Chemical Engineering*, 44, 141-167.

Multiperiod optimization model for CO₂ capture, utilization and storage, Colombian case study

Alexandra Duarte^{a,b}, Javier D. Angarita^b, Juan P. Espinosa-Cárdenas^b,
Javier Lizcano^b, Rafael C. García-Saravia^b, Ariel Uribe-Rodríguez^{c,*}

^a*Caldas University, Engineering Department, 170004 Manizales, Colombia.*

^b*TIP Colombia, Consultant Engineer, 687547 Girón, Colombia.*

^{c,*}*Colombian Petroleum Institute, ECOPEPETROL, 681011 Piedecuesta, Colombia.*

ariel.uribe@ecopetrol.com.co

Abstract

Colombia is formulating policies to accelerate the energy transition, decarbonize its industries and decrease its dependency on fossil fuels. In this work we present a novel optimization framework for the carbon capture, utilization, and storage (CCUS) design applied to the Colombian case. The model maximizes the net present value considering technologies for carbon capture, multimodal transportation, CO₂ utilization for enhanced oil recovery (CO₂-EOR) and geological storage to meet a given target on greenhouse gases (GHG). The novelty of this work is the integration of the CCUS design with the oil field development for CO₂-EOR. Thus, the optimal petroleum production campaign (i.e., production wells for CO₂ injection) is computed based on the oil production profiles, petrophysical properties and economic data. Moreover, this approach provides insights regarding investments on carbon capture technologies such as timing, sizing, location, and type of technology. The model also determines the optimal transportation mode considering on-shore (pipelines, railcars, and trucks) and off-shore (pipelines and ships). We tested our approach on a nationwide long-term planning horizon (20 years), considering 30 CO₂ emissions sources distributed across 12 regions in the country, 16 carbon capture technologies, 3 different plant sizes (small, medium, and large), 6 potential oil fields for CO₂-EOR, 2 sinks for geological storage, and 4 multimodal transportation modes. Results shown a carbon capture cost of 99.8 USD/t, with a cost breakdown of 45%, 30% and 25% allocated into the carbon capture processes, CO₂ transportation and CO₂-EOR utilization, respectively.

Keywords: supply chain design, carbon capture, utilization, and storage, CCUS, enhanced oil recovery, CO₂-EOR

1. Introduction

Global CO₂ emissions declined to 31.5 Gt CO₂ in 2020 in comparison to 33.4 Gt CO₂ in 2019, this partial recovery is entirely related to restrictions on transport activity in 2021 due to the Covid-19 pandemic. However, the emissions projected future growth to 33 Gt CO₂ in 2021 (IEA, 2021). Even with an increase in CO₂ emissions from oil of over 650 Mt CO₂ in 2021, these should remain 500 Mt CO₂ below 2019 levels. This reduction in emissions is achieved due to the global concern in mitigating fossil fuels use and the strategies of CCUS. In Colombia, the target is to reduce greenhouse gas emissions by 50% in 2030 and is much more closely aligned with the country's objective of achieving carbon neutrality by 2050 (UPME, 2019). To achieve these goals, is necessary to consider key aspects like increase the use of new renewable energy sources, efficient use of energy, and analysis of CCUS supply chain.

Herein, we present an optimization framework for the CCUS supply chain design. Our approach can be distinguished from previous studies in the following aspects: firstly, a complete CO₂-EOR model that involves evaluating the real considerations in the field such as petrophysical properties, dynamic oil and CO₂ production profiles and some operative's conditions are integrated into the CCUS model. Secondly, a transport patterns considered by Zhang et al., (2020) with free movements between sources and reservoirs are incorporated. And thirdly, surrogate models to evaluate the capital expenditure (CAPEX) and operational expenses (OPEX) proposed by Hasan et al., (2015) are incorporated into the CCUS model.

2. Model formulation

The main components of a CCUS supply chain include: i) stationary CO₂ sources, ii) carbon capture facilities, iii) transport network and, iv) geological storage and CO₂-EOR as utilization. The proposed mathematical model is based on the multiperiod model for CCUS published by Han et al., (2012). In this model, a superstructure composed of regions (g) enclosing different types and sizes of capture, sequestration and utilization facilities is formulated as a large-scale mixed-integer linear programming (MILP). The CO₂ captured from an emission source in a region g , according to an established target, can be sent to sequestration or utilization facilities within the same region g or to another region g' depending on its capacity and availability. Regions are connected to each other by different potential transportation modes (pipeline, truck, railcar, and ship) and different types and sizes of intermediate storage are also considered in the regions to adequately supply the non-pipeline transportation modes.

Han et al., (2012) considered utilization facilities to produce biobutanol and green polymers in South Korea. For the Colombian case, we have assumed EOR-CO₂ as the potentially most important use of CO₂. The EOR-CO₂ has been modeled considering the typical petrophysical properties of oil reservoirs and the dynamic oil and CO₂ production profiles. The oil and CO₂ production rates depend on the CO₂ injection rate, where the CO₂ injection and production rates are normalized by hydrocarbon pore volume ($HCPV_m$, [m³]) of each reservoir m . According to Eq. 1, $HCPV_m$ is a function of reservoir properties as area (A_m , [m²]), net pay thickness (h_m , [m]), average porosity (ϕ_m , [-]), and the initial water saturation ($S_{wi,m}$, [-]).

$$HCPV_m = A_m h_m \phi_m (1 - S_{wi,m}) \quad \forall m \quad (1)$$

The oil production is normalized by the original oil in place ($OOIP_m$, [m³]) which is defined following the Eq. 2.

$$OOIP_m = \frac{HCPV_m}{\beta_{oi,m}} \quad \forall m \quad (2)$$

Where $\beta_{oi,m}$ represent the initial oil formation volume factor. Each reservoir is characterized by a unique $HCPV_m$, and it determines the oil and CO₂ production rates of the field. The oil and CO₂ production profiles, considering high, medium, and low oil production, are calculated using a piecewise linearization strategy of the dimensionless curve representing the incremental recovery factor for oil and the cumulative CO₂ production (both as a function of cumulative CO₂ injection as shown in Figure 1).

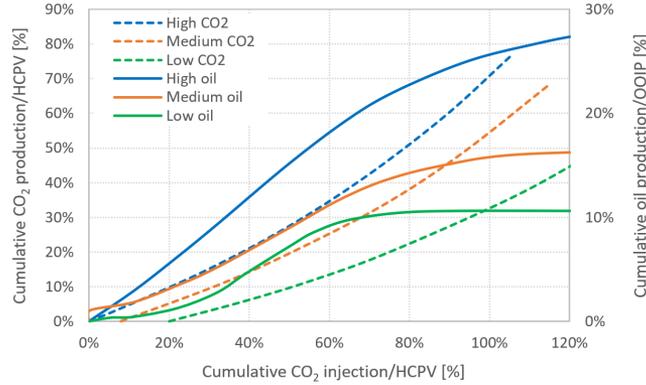


Figure 1. Cumulative production of oil (solid) and CO₂ (dash) for high, medium, and low oil recovery and CO₂ production (Adapted from Kolster et al., 2017)

The rate of CO₂ injected each period of time t per field ($CO2INJ_{m,t}$, [Mt/y]) is limited by both the maximum CO₂ injection rate per well ($qCO2_m^{max}$, [Mt/(y·well)]) and the number of injection wells available in the field ($WI_{m,t}$, [well]). According to Eqs. 3 and 4, $qCO2_m^{max}$ is assumed equal as the historical water injection rate in the field and $WI_{m,t}$ is limited to the oil fields drilling campaign ($wt_{m,t}^{max}$, [well]).

$$CO2INJ_{m,t} \leq \sum_{td \leq t} qCO2_m^{max} * WI_{m,td} \quad \forall m, t \quad (3)$$

$$\sum_{td \leq t} WI_{m,td} \leq wt_{m,t}^{max} \quad \forall m, t \quad (4)$$

The $CO2INJ_{m,t}$ is normalized with the $HCPV_m$ to calculate the cumulative CO₂ injection ($CO2INJCUMH_{m,t,h}$, [-]), as seen in Eq. 5; where ΔT is the time between two consecutive time periods.

$$\sum_h CO2INJCUMH_{m,t,h} = \Delta T * \sum_{td \leq t} \left(\frac{CO2INJ_{m,td}}{HCPV_m} \right) \quad \forall m, t \quad (5)$$

Each linear segment of the production curves h is limited by lower ($loinj_{m,h}$, [-]) and upper ($upinj_{m,h}$, [-]) limits (see Eq. 6) and each linear segment is selected introducing a binary variable ($BinH_{m,t,h}$) that can active only one of the linear segments as shown in Eq. 7.

$$loinj_{m,h} * BinH_{m,t,h} \leq CO2INJCUMH_{m,t,h} \leq upinj_{m,h} * BinH_{m,t,h} \quad \forall m, t, h \quad (6)$$

$$\sum_h BinH_{m,t,h} \leq 1 \quad \forall m, t \quad (7)$$

The cumulative oil and CO₂ produced ($CO2PRODCUM_{m,t}$, [-], and $OILPRODCUM_{m,t}$, [-], respectively) are calculated with the Eqs. 8 and 9. Where $aco2_{m,h}$ and $\beta co2_{m,h}$ represent the slope and the intercept of the linear segment h of the CO₂ breakthrough curve, respectively; $aoil_{m,s}$ and $\beta oil_{m,s}$ represent the slope and the intercept for each segment s of the oil performance curve.

$$CO2PRODCUM_{m,t} = \sum_h (\alpha_{co2,m,h} * CO2INJCUMH_{m,t,h} + \beta_{co2,m,h} * BinH_{m,t,h}) \quad \forall m, t \quad (8)$$

$$OILPRODCUM_{m,t} = \sum_s (\alpha_{oil,m,s} * CO2INJUMS_{m,t,s} + \beta_{oil,m,s} * BinS_{m,t,s}) \quad \forall m, t \quad (9)$$

Eq. 10 represents a CO₂ mass balance at the inlet of the injection well in function of the fresh CO₂ for EOR ($CO2PURC_{m,t}$, [Mt/y]), the recycled CO₂ ($CO2RECY_{m,t}$, [Mt/y]) and the injected CO₂. In our model, the fresh CO₂ rate is used to integrate the EOR-CO₂ module with the CCUS model proposed by Han et al. (2012).

$$CO2PURC_{m,t} + CO2RECY_{m,t} = CO2INJ_{m,t} \quad \forall m, t \quad (10)$$

The economic model in this framework is adapted from the one proposed by Calderón & Pekney, (2020). In our model we maximize the net present value (NPV , [MUSD]), which is function of the global CCUS supply chain profit after taxes and penalties ($CASHFLOW_t$, [MUSD]), the global capital costs of the supply chain ($CPXCCUS_t$, [MUSD]) and the interest rate (ir , [-]), as shown in Eq. 11.

$$NPV = \frac{CASHFLOW_t - CPXCCUS_t}{(1 + ir)^{t-1}} \quad (11)$$

3. Results

The mathematical model was formulated as a MILP that maximizes the net present value of the CCUS supply chain. The optimization problem was implemented in GAMS 33.2.0 and was solved using CPLEX v12.10.

Table 1. Selected capture technology for the CCUS model

Source	Capture technology	Size	Facilities No.
ET={2,...,7}		small	6
R1-H2-2	Compressor	small	1
F-4-NatCO ₂		small	1
F-5-NatCO ₂		small	1
PG-G = {1,7,8}	Abs-MEA	small	4
R1-FCC={1,2}		small	2
R1-CHP-1, R1-HDT-2	PSA-MVY*	small	5
R2-FCC-1, R2-CHP-6		small	2
R1-HDT-2		medium	1
Cem={1,3,7}	PSA-WEI*	small	5

*Zeolite-based physical adsorption of CO₂ for Pressure Swing Adsorption (PSA) technologies. MVY (Mobil composition of matter-seVentY) and WEI (Weinebeneite) are zeolites considered by Han et al., (2015).

The main considerations about the components of the CCUS supply chain analyzed in this study are the following: **(a)** the data for CO₂ sources and sinks was taken from Yáñez et al., (2020), as seen Figure 3. Comprising emissions from natural sources (F-4-NatCO₂ and F-5-NatCO₂), oil (R1 and R2, aggregated into 14 emission sources), cement (Cem-1, ..., Cem-7), power generation (PG-G-1, ..., PG-G-7), and bioethanol (ET2, ..., ET7) industries. Nodes A, B, H, K, M, and N are candidate oil fields for CO₂-EOR process, and nodes C1 and C2 are candidates for geological storage. **(b)** surrogate models for investment and operating costs developed by Hasan et al., (2015) were considered for modelling CO₂ capture technologies. **(c)** finally, three transport patterns (source-sink, source-source and sink- sink matching) with free movements between sources and sink,

as summarized by Zhang et al., (2020), were assumed for the multimodal-network transportation. The other parameters for transport and intermediate storage were taken from Han et al., (2012).

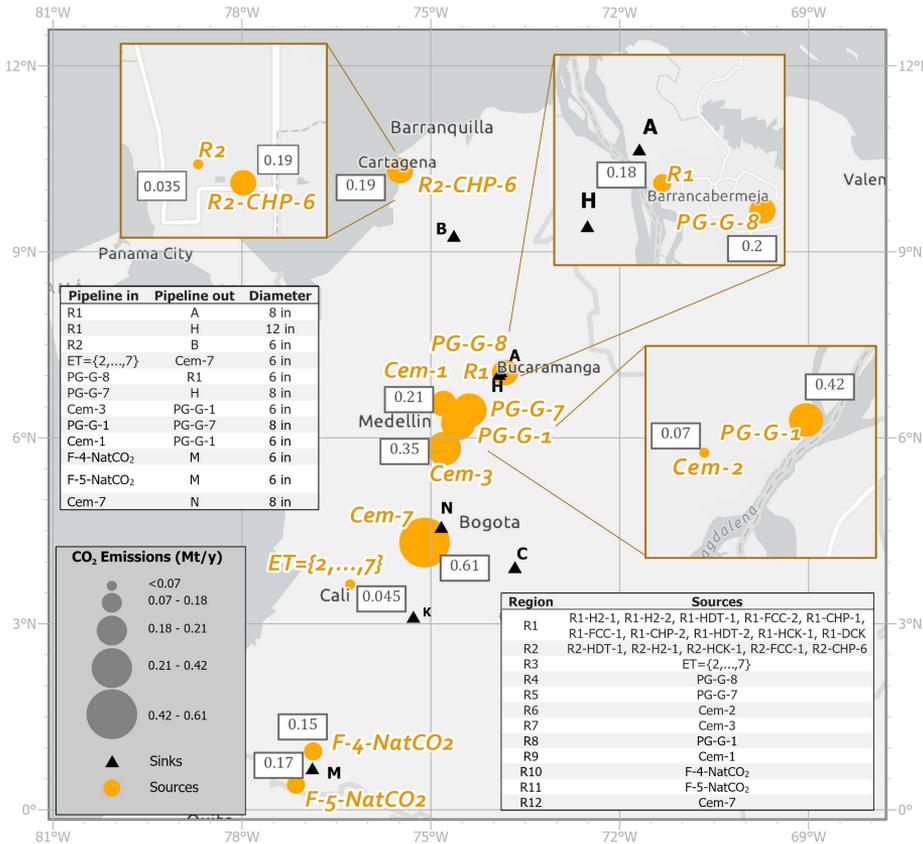


Figure 3. CO₂ sources, storage, and utilization projects in Colombia.

Absorption based on monoethanolamine (Abs-MEA) is selected for sources with a CO₂ concentration as low as 4% (PG-G-1, PG-G-7 and PG-G-8), since Abs-MEA is the only one that works in this concentration range. On the contrary, compression is selected for sources with high CO₂ concentration (ET={2,...,7}, R1-H2-2, F-4-NatCO₂ and F-5-NatCO₂). For refineries, with concentrations of 7%, the technology with the lowest investment and operational costs is PSA-MVY. Finally, when the concentration is as high as 47% as in the case of the sources: Cem-1, Cem-3 and Cem-7, the model uses the PSA-WEI capture technology, which is the one with the lowest costs for this CO₂ concentration (see Table 1). Even though the model considers 2 sinks for geological storage, the optimal solution chooses EOR as the main use for CO₂ because it represents economic benefits as opposed to the geological storage. Results of transport infrastructure are consistent, since it minimizes the costs of the network by choosing pipelines with the smallest possible diameters (6 and 8 in) to connect most of the sources with the sinks. Although it also selects 12 in pipelines to connect sources with high capture capacity, such as the R1 source. Results shown a carbon capture cost of 99.8 USD/t, with a cost breakdown of

45%, 30% and 25% allocated into the carbon capture processes, CO₂ transportation and CO₂-EOR utilization, respectively.

4. Conclusions

The motivation for this study is to present an optimization framework to support decision making for a CCUS supply chain. For this purpose, we formulated a mathematical model that considers: **(1)** the integration between EOR and CCUS in a supply chain configuration that includes petrophysical properties, dynamic oil and CO₂ production profiles and operating conditions, **(2)** transport patterns with free movements between sources and sinks, and **(3)** a surrogate model which calculates the investment and operating cost for carbon capture technologies taking into account different plant sizes and CO₂ concentrations. The model was tested with representative information from the Colombian context for the capture and use of CO₂ for EOR. The optimal solution is coherent and presents the number and size of capture technologies necessary for each source and the configuration of the transport network based on pipelines. Future research may focus on considering a more detailed model representation for designing the CO₂ pipeline transportation network.

Acknowledgments

The authors would like to acknowledge the financial support from the Centre for Innovation and Technology Colombian Petroleum Institute (ICP-Ecopetrol S.A.), and geologist Leily Candela (TIP) for her support with the GIS data. The views expressed in this paper do not necessarily reflect those of Ecopetrol S.A.

References

- Calderón, A. J., & Pekney, N. J. (2020). Optimization of enhanced oil recovery operations in unconventional reservoirs. *Applied Energy*, 258, 114072.
- Han, J.-H., Lee, J.-U., & Lee, I.-B. (2012). Development of a Multiperiod Model for Planning CO₂ Disposal and Utilization Infrastructure. *Industrial & Engineering Chemistry Research*, 51(7), 2983–2996. <https://doi.org/10.1021/ie201460v>
- Hasan, M. M. F., First, E. L., Boukouvala, F., & Floudas, C. A. (2015). A multi-scale framework for CO₂ capture, utilization, and sequestration: CCUS and CCU. *Computers and Chemical Engineering*, 81, 2–21. <https://doi.org/10.1016/j.compchemeng.2015.04.034>
- IEA. (2021). *Global Energy Review 2021*. <https://www.iea.org/reports/global-energy-review-2021/co2-emissions>
- Kolster, C., Masnadi, M. S., Krevor, S., Mac Dowell, N., & Brandt, A. R. (2017). CO₂ enhanced oil recovery: a catalyst for gigatonne-scale carbon capture and storage deployment? *Energy & Environmental Science*, 10(12), 2594–2608.
- UPME. (2019). *Plan Energético Nacional 2020-2050*. https://www1.upme.gov.co/DemandaEnergetica/PEN_documento_para_consulta.pdf
- Yáñez, E., Ramírez, A., Núñez-López, V., Castillo, E., & Faaij, A. (2020). Exploring the potential of carbon capture and storage-enhanced oil recovery as a mitigation strategy in the Colombian oil industry. *International Journal of Greenhouse Gas Control*, 94, 102938.
- Zhang, S., Zhuang, Y., Tao, R., Liu, L., Zhang, L., & Du, J. (2020). Multi-objective optimization for the deployment of carbon capture utilization and storage supply chain considering economic and environmental performance. *Journal of Cleaner Production*, 270, 122481. <https://doi.org/10.1016/j.jclepro.2020.122481>

Towards improved scheduling: an analysis on time-steps, time-horizon, and rescheduling

Robert E. Franzoi,^a Brenno C. Menezes,^{a,*}

^a*Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar*

bmenezes@hbku.edu.qa

Abstract

Decision-making capabilities and solution approaches for industrial scheduling operations are often inefficient. Both the academic literature and industry require further investigation of scheduling strategies and methodologies to achieve further breakthroughs, in which a proper selection of modelling and optimisation parameters is imperative. This work investigates the impact of time-step size and time-horizon length on multiple scheduling applications. Small time-steps provide better management of feedstocks, more efficient production, and higher profit, whereas long time-horizon can leverage spot market opportunities and achieve a more comprehensive formulation. Moreover, such parameters are especially meaningful in combination with continuous rescheduling strategies. A blend scheduling case study is employed in this work, whereby the operational scheduling and economic value of the process are investigated over distinct conditions in three examples. The interplay between rescheduling optimisation and the aforementioned scheduling parameters is highlighted. The insights drawn are useful for diverse problems and scenarios, and can significantly provide further operational and economic enhancements for industrial applications.

Keywords: Scheduling, optimisation, mathematical programming, rescheduling, industrial processes, decision-making capabilities.

1. Introduction

Computer-aided tools have been increasingly employed for scheduling operations, in which recent advances on decision-making modelling and optimisation, solving algorithms, and computer-aided resources (Franzoi et al., 2018) have provided enhanced capabilities for properly addressing and solving large-scale scheduling applications. Such mathematical representations require the pre-determination of scheduling parameters such as time-step size and time-horizon length, whose importance and impact are often neglected in the literature on the topic. This work addresses the topic of scheduling optimisation based on discrete-time formulation, whereby providing a relevant discussion on scheduling parameters used in the modelling and optimisation. These parameters are determined according to the type of problem to be solved or to the required solution in terms of accuracy or timeliness. Common choices for industrial applications are in the order of seconds or minutes (e.g., real time optimisation and control), hours or days (e.g., scheduling), months or years (e.g., procurement planning and process design).

The time-horizon length should be chosen according to the desired application and the computational resources available. Large time-horizon lengths incorporate future information into the formulation/optimisation decision-making, which is helpful for

adjusting the scheduling operations according to the known or expected scenario in the plant. In addition, the time-horizon plays a key role on the interplay between the scheduling solution to be implemented in the process and the market/sales decision-making through spot market opportunities.

Large time discretisation may cause lack of accuracy and may not fully exploit scheduling opportunities in terms of better economic value and operations (Lee and Maravelias, 2018). Small time-steps overcome such issues but lead to large formulations with potentially intractable models. However, this has been recently overcome by the recent technological advances, which provide proper capabilities for breakthrough research on large-scale optimisation. Small time-steps are closely related to online scheduling (also referred to as rescheduling) applications. Online scheduling concepts are especially important because a continuous cycle of improvement is required to reduce the deviation between the model predictions and the actual values in the plant. Operational data used in the scheduling are typically out of date or not integrated with the production, which leads to inconsistencies in the prediction throughout the process. According to Gupta et al. (2016a), rescheduling facilitates the adaptation of the schedule under uncertainties and unforeseen events, in addition to considering new information as soon as possible, which aims to increase the economic value of the process. In previous work, Franzoi et al. (2021) discuss and illustrate the importance of rescheduling for handling data uncertainty and disturbances, maintaining feasibility, restoring optimality, and providing more reliable and effective capabilities for industrial applications. Importantly, such benefits can only be achieved through a continuous cycle of updates and re-optimisations.

Although the literature on the topic indicates the importance of addressing small time-steps and large time-horizon for improved scheduling operations, to the best of our knowledge there are no works that provide comprehensive quantitative analyses and results on the impact of these features on the scheduling economics and operations. In this work, we investigate the influence of the time-step and time-horizon lengths in the scheduling solutions aiming to achieve a better understanding on the role and importance of these elements for scheduling solutions. Such discussion can be especially important to establish guidelines for future research on the topic aiming to develop more efficient methodologies and to achieve enhanced industrial operations.

2. Scheduling parameters: time-step and time-horizon

There is not standard practices or common understanding about how to properly define time-steps, time-horizon, and time-periods within scheduling formulations. On one hand, the academic literature lacks more comprehensive and conclusive studies and analyses on the impact of such parameters towards improved scheduling optimisation. On the other hand, industrial organisations worldwide do not have a standardised or optimum approach to address that question. The scheduling operations are often not properly optimised, in which outdated manual/simulation capabilities are still employed; or the optimisation tools are not sufficiently accurate or applicable. Even under the utilisation of proper and effective optimisation, the approaches, strategies, and methodologies used within the decision-making capabilities are often not appropriate. That requires an accurate representation of the problem with an integrated mathematical modelling environment (e.g., detailed operations, inclusion of operational constraints, limitations, and requirements), consideration of external factors (e.g., noises, disturbances, and uncertainties), determination of scheduling parameters, proper computational balance and tractability, etc. An important factor within this context concerns the proper selection of

scheduling parameters to be used in the mathematical formulation and optimisation. Previous literature on scheduling optimisation indicates the lack of standards to address such a question. For example, Panda and Ramteke (2019) uses time-steps of 8-hours and time-horizon of 14 days; Hou et al. (2015) considers 10-days time-horizon and 1-day time-step; Dai et al. (2021) employs 7-days time-horizon and 1-day time-step of; and Kelly et al. (2017) proposes time-horizon of 7-days and time-step of 2-hours. Ideally, the time-horizon should be as large as possible and the time-step should be as small as possible, but proper balance is fundamental to overcome computational limitations.

A previous study (Franzoi et al., 2021) addressed the importance of a continuous cycle of scheduling updating for industrial operations. In this work, we investigate how scheduling parameters, namely, time-step size and time-horizon length, affect the scheduling optimisation and its solution, whereas aiming to achieve proper balance between model complexity and computational effort. To the best of our knowledge, such analyses are missing in the state-of-the-art literature and may provide meaningful insights for improved scheduling optimisation capabilities in future works on the topic. Small time-steps enlarge the search space in the optimisation, which provides additional degrees of freedom and potentially better solutions. Large time-horizon provides further information that may significantly improve the scheduling decision-making towards better solutions. Assuming the formulation is suitable to be solved in acceptable computational time, both these concepts are meaningful and present potential for enhanced scheduling capabilities. Importantly, they need to be carefully analysed for each application to achieve proper balance between computational tractability and solution quality.

3. Problem statement

To investigate how the time-step and time-horizon sizes affect the scheduling optimisation and its solution, a blend scheduling optimisation problem for the production of diesel is addressed. Figure 1 illustrates the proposed network. There are two incoming diesel streams D1 and D2 with distinct qualities in terms of specific gravity and sulphur content. They are connected to four intermediate storage tanks S1 to S4 to be further blender in the BLENDER unit. The outgoing mixture is sent to final storage tanks F1 to F6 to be further distributed according to the demand of product P1. The data are presented in Table 1, including the properties of intermediate and final streams (i.e., specific gravity and sulphur content), their estimated market value (negative for costs of feedstocks and positive for selling prices of products), and the maximum incoming/outgoing flowrates.

Table 1: Data for the blend scheduling optimisation problem.

Unit	D1	D2	P1
Specific gravity (g/mL)	0.80	0.95	0.88
Sulphur content specification (g/g)	0.90	1.20	1.00
Maximum incoming/outgoing flowrate	100	100	500

The demands of product P1 over the time-horizon is randomly generated between 40 and 100 bbl/day, which are represented as hard bound constraints (e.g., contractual demands) and must be met. It is assumed that each daily demand needs to be supplied by a single tank F. The scheduling optimisation aims to continuously supply the demand of P1 meeting the product specifications to maximise the refinery profit.

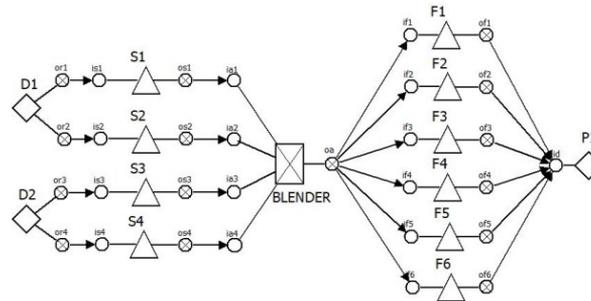


Figure 1: Blend scheduling problem flowsheet.

4. Examples and results

Three examples are proposed to evaluate how scheduling parameters affect the computational effort and solution quality (operations and economic value). They respectively address the impact of the time-step size, time-horizon length, and rescheduling combined with enhanced scheduling parameters. The examples are built as mixed-integer nonlinear programming (MINLP) problems and solved by a two-stage optimisation approach. The mathematical model and the optimisation framework are omitted for the sake of simplicity; for further details, see Franzoi et al. (2021).

4.1. Example 1: Analysis on the time-step size

Example 1 investigates the impact of the time-step size on the scheduling. This is especially important to manage operations with tanks and flows throughout the process. Small time-steps are helpful to exploit additional degrees of freedom whereby achieving more efficient operations. For example, formulations involving tanks typically consider that simultaneous incoming and outgoing flows are prohibited in the same tank. Thus, large time-steps may impose difficulties either to obtain a good schedule or to implement the solution in the plant. Two scenarios are considered:

- Scenario 1a optimises the future 20-days scheduling considering time-steps of 2-days each in a single optimisation with 10 time-periods.
- Scenario 1b optimises the future 20-days scheduling considering time-steps of 1-day each in a single optimisation with 20 time-periods.

Table 2 provides information on the scheduling solution for each scenario, which illustrates the importance of small time-step sizes. Smaller time-steps provide 6% profit increase associated with better feedstock management and less product giveaway.

Table 2: Scheduling solution for Example 1.

	Scenario 1a	Scenario 1b	Difference
Contractual demand met (%)	100	100	-
Refinery profit (k\$)	66	70	4
Computational time per optimisation (s)	1	2	1

4.2. Example 2: Analysis on the time-horizon length

Example 2 explores the impact of the time-horizon length on the scheduling. Since the beginning of the event horizon (Day1), there is a known spot market demand of 500 bbl of product P1 available at Day20, with market price 10% above the average and that can be either partially or fully met. Two scenarios are proposed:

- Scenario 2a optimises the future 20-days scheduling considering time-steps of 1-day each in a single optimisation with 20 time-periods.
- Scenario 2b performs sequential 10-days optimisations within a moving horizon rescheduling strategy (see Franzoi et al. (2021) for further information) with 1-day time-steps. There are 10 optimisations, from Day1 to Day10 until Day11 to Day20.

Table 3 presents important aspects of the scheduling solution for each scenario, which illustrates the importance of employing a sufficiently large time-horizon length.

Table 3: Scheduling solution for Example 2.

	Scenario 2a	Scenario 2b	Difference
Contractual demand met (%)	100	100	-
Spot market demand met (%)	45	100	55
Refinery profit (k\$)	121	105	16
Computational time per optimisation (s)	10	2	8

Due to the larger time-horizon in Scenario 2a, the spot market opportunity is considered in the formulation/optimisation since Day1, which provides flexibility for the plant to adjust its production accordingly. In addition, the optimisation capabilities are especially useful to ensure whether the scheduling is able to comply with the demand requirements in terms of amount and qualities. Conversely, although such spot market information is known since Day1, Scenario 2b includes it in the formulation/optimisation only at Day11, so that the scheduling does not have enough time/resources to adjust the production and misses the opportunity of selling additional product at a higher price.

4.3. Example 3: Analysis on the role of continuous rescheduling

Example 3 investigates the importance of large time-horizon length within continuous rescheduling optimisation. A spot market opportunity becomes available for Day20, in which there is an additional demand of 500 bbl of P1 with market price 10% above average and that can be either partially or fully met. However, such information is only known since Day5. Three scenarios are proposed. Information on the scheduling solution is shown in Table 4.

- Scenario 3a optimises the future 40-days scheduling considering time-steps of 1-day each in a single optimisation with 20 time-periods. The spot market opportunity is considered in the formulation.
- Scenario 3b optimises the future 15-days scheduling considering time-steps of 1-day each in two optimisations with 15 time-periods each (Day1 to Day15 followed by Day16 to Day30). The spot market opportunity is considered only in the second scheduling formulation, from Day16.
- Scenario 3c optimises the future 15-days within a moving horizon rescheduling strategy with 1-day time-steps in a total of 15 optimisations. The spot market opportunity is considered in the formulation since Day15.

Table 4: Scheduling solution for Example 3.

	Scenario 3a	Scenario 3b	Scenario 3c
Contractual demand met (%)	100	100	100
Spot market demand met (%)	0	32	100
Refinery profit (k\$)	146	155	175
Computational time per optimisation (s)	30	5	5

The large time-horizon in Scenario 3a does not capture the spot market information. There is no safety stock available and the production could not exploit such additional opportunity because. Scenario 3b considers the spot market demand only in the second optimisation run, which does not provide enough time/resources to properly adjust the production. On the other hand, higher flexibility is provided in Scenario 3c due to the continuous rescheduling operations, which allow to include the updated information into the model as soon as it becomes available. This illustrates the importance of such interplay of scheduling parameters with rescheduling strategies towards improved scheduling operations. It is worth noting that although such example emphasises spot market opportunities which are common in petroleum refineries, other fields or industries have similar cases involving peaks of demand (e.g., seasonal markets, peak of sales, etc.).

5. Final remarks and future guidelines

There is an important interconnection between rescheduling strategies and the selection of scheduling parameters. Despite highly relevant for industrial applications, such topic has been rarely discussed in the literature. In this work, we address scheduling parameters such as time-step size and time-horizon length and highlight their particular importance in combination with rescheduling strategies. The results indicate that a) proper choice of time-step size provides better selection and management of feedstocks, more efficient production, less giveaways, and higher profit; b) spot market opportunities can be better exploited by large time-horizon lengths, which is especially useful to account for future information that might be available; c) rescheduling strategies are fundamental for enhanced scheduling operations, especially when combined with proper selection of the aforementioned parameters.

Seeking enhanced industrial capabilities, it is fundamental to investigate scheduling parameters through extensive testing to identify better strategies and achieve improved solutions. The analyses carried out in this work are useful for diverse problems, applications, and conditions, and can significantly provide further operational and economic enhancements. The examples focus on the petroleum refinery industry, but the ideas and discussion provided are valid for other industrial applications as well.

References

- X. Dai, L. Zhao, Z. Li, W. Du, W. Zhong, R. He, F. Qian, 2021, A data-driven approach for crude oil scheduling optimisation under product yield uncertainty, *Chemical Engineering Science*, 246, 116971.
- D. Gupta, C.T. Maravelias, 2016, On deterministic online scheduling: major considerations, paradoxes and remedies, *Computers and Chemical Engineering*, 94, 312-330.
- R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, 2018, Effective scheduling of complex process-shops using online parameter feedback in crude-oil refineries, *Computer Aided Chemical Engineering*, 44, 1279-1284.
- R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, 2021, A moving horizon rescheduling framework for continuous nonlinear processes with disturbances, *Chemical Engineering Research and Design*, 174, 276-293.
- Y. Hou, N. Wu, M. Zhou, Z. Li, 2015, Pareto-optimisation for scheduling of crude oil operations in refinery via genetic algorithm. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(3), 517-530.
- J.D. Kelly, B.C. Menezes, F. Engineer, I.E. Grossmann, 2017, Crude-oil blend scheduling optimisation of an industrial-sized refinery: a discrete-time benchmark, In *Foundations of Computer Aided Process Operations*, FOCAPO, Tucson, AR, United States, 10-13 January.
- H. Lee, C.T. Maravelias, 2018, Combining the advantages of discrete-and continuous-time scheduling models: Part 1. Framework and mathematical formulations, *Computers and Chemical Engineering*, 116, 176-190.
- D. Panda, M. Ramteke, 2019, Preventive crude oil scheduling under demand uncertainty using structure adapted genetic algorithm, *Applied Energy*, 235, 68-82.

Pharmaceutical Industry Supply Chains: Planning Vaccines' Distribution

Inês Duarte^a, Bruna Mota, Ana Paula Barbosa-Póvoa

Centre for Management Studies of IST (CEG-IST)

Av Rovisco Pais, 1, 1049-001 Lisbon, Portugal

a) inesduarte97@tecnico.ulisboa.pt

Abstract

Pharmaceutical supply chains tend to be complex, and its management encounters several challenges. In this work, a decision-support tool based on a multi-objective mixed integer linear programming model is proposed, aiming to integrate several strategical-tactical decisions while considering the three pillars of sustainability, which are addressed as objective functions. The economic assessment is performed through the Net Present Value (NPV). The environmental impact assessment follows the Life Cycle Analysis (LCA) methodology. Accessibility of pharmaceutical products is the major focus for the social assessment, aiming to provide an equal distribution based on the burden of diseases which is made through a DALY-based metric (Disability-Adjusted Life Year). The model is applied to a representative case study aiming to discuss different optimization scenarios and allowing to understand the effect of decisions on each performance indicator. Important results were obtained that show that sustainability issues influence the supply chain design and planning.

Keywords: Pharmaceutical Industry, Sustainable supply chains, Equity, MILP

1. Introduction

The 2030 Agenda of Sustainable Development comprises several goals intended to drive global and national policies in the direction of a sustainable development of our societies. Within these goals, one seeks to support well-being at all ages and guaranteeing healthy lives, while another fosters inclusive and sustainable economic growth, employment, and proper work for all (Johnston, 2016).

Pharmaceutical companies represent a group of healthcare companies that have been facing strengthened regulations concerning economic, environmental, and social issues, driving them towards more sustainable supply chains. Moreover, pharmaceutical companies dealing with vaccines need to cope with particularities regarding the product itself, such as storage temperature and shelf-life time, influencing the resources needed to be allocated to each facility (Lemmens *et al.*, 2016). This sector plays a critical role in the healthcare structure of each country by providing medicines and vaccines with direct impact on population's quality of life. Undoubtedly, medicines are responsible for preventing and treating diseases, enhancing, or preserving health, and to avoid exacerbation of existing illnesses (Pfizer, 2019). Accordingly, along with the direct benefits for population, medicines and vaccines also contribute towards a significant cost reduction in the total healthcare costs of each country by decreasing the need for long-term care services and/or costly surgeries. Hence, addressing equity in access when designing and planning pharmaceutical supply chains has proven to be essential and helps driving this sector in the direction of a more socially sustainable industry.

Bearing in mind the described scenario for pharmaceutical supply chains, particularly when dealing with vaccines, the present work aims to study how to make strategic and tactical decisions to help attaining sustainability objectives, where the integration of social concerns is the main focus.

2. Problem Definition and mathematical formulation

The proposed model follows the work developed by Mota *et al.* (2018), where a decision-support tool for the design and planning of sustainable supply chains focusing on strategic-tactical challenges is suggested. Adaptations were made to better shape demanding characteristics of the pharmaceutical industry, such as storage technologies needed at warehouses. The generic supply chain representation considers a four-echelon structure (see Figure 1), where the raw materials flow from suppliers to factories to be transformed into final products. At factories, production technology selection is possible. Once the final products are obtained, they can either flow to warehouses or directly to markets to be sold. At warehouses, storage technology selection is possible. Moreover, transshipment between warehouses is allowed and transportation between different entities may be done by either unimodal or intermodal transportation, by road, air and sea.



Figure 1. Modelled supply chain structure

The three pillars of sustainability are introduced as objective functions. The economic objective function is obtained from the maximization of the Net Present Value (NPV) by calculating the sum of the discounted cash flows of each time-period, at an interest rate. These cash flows are obtained through the net earnings (NEt), which are given by the difference between incomes and overall costs, where the former corresponds to the amount of products sold at a certain price, and the latter by the following costs:

- raw material costs (first term)- number of products purchased from suppliers times the unit raw material cost (rmc_{mi});
- production operating costs (second term) - amount of final products produced (P_{mgit}) times the unitary operating costs of each production technology (opc_g);
- storage costs (third term)- amount of final products stored with technology g (S_{mgit}) times the unitary operating cost of storage technology (opc_g);
- transportation costs (fourth term)- flow of products transported through transportation mode a ($X_{mai jt}$) times the transportation cost per kg.km (tc_a), weight of each unit of product transported (pw_m) and distance traveled (d_{ij});
- hub handling costs (fifth term)- flow of products through the hub terminals at the airports or seaports times the unit handling costs at these terminals (hhc_j);
- airline/freighter contracted costs - contracted costs (cfp_i) for the allocated transportation capacity and/or for hub terminal use per time period (sixth term);
- inventory costs (seventh term)- amount of products in stock (S_{mgit}) times the unitary stock cost, (sc_m);

- labor costs at entities (eight and night terms) and labor costs for technologies' use (tenth term), which vary with the fixed (w_i), the variable ($wpsq$) number of workers at each entity and the number of workers needed for each technology (w_g), respectively. The labor cost at each location (lc_i), the weekly working hours (wwh) and the number of weeks per time period (wpt) are also considered, The last term describes the depreciation of the capital invested (Dpt) with tr being the tax rate.

$$\begin{aligned}
 NE_t = (1 - tr) & \left[\sum_{\substack{(m,i,j) \in F_{INCFP} \\ (a,m,i,j) \in NetP}} psu_m X_{mai jt} \right. \\
 & - \left(\sum_{\substack{(m,i,j) \in F_{OUTSUPRM} \\ (a,m,i,j) \in NetP}} rmc_{mi} X_{mai jt} + \sum_{\substack{(m,g) \in H_{prod} \\ i \in I_f}} opc_g P_{mgit} \right. \\
 & + \sum_{\substack{(m,g) \in H_{stor} \\ i \in (I_f \cup I_w)}} opc_g S_{mgit} + \sum_{\substack{(a,m,i,j) \in NetP \\ a \in (A_{plane} \cup A_{boat} \cup A_{trucks})}} tc_a \cdot pw_m \cdot d_{ij} \cdot X_{mai jt} \\
 & + \sum_{\substack{(a,m,i,j) \in NetP \\ (j \in I_{plane} \wedge i \notin I_{plane}) \cup (j \in I_{boat} \wedge i \notin I_{boat})}} hhc_j \cdot X_{mai jt} + \sum_{i \in (I_{plane} \cup I_{boat})} cfp_i \cdot Y_i \\
 & + \sum_{\substack{(m,g) \in H_{stor} \\ (m,i) \in V}} sc_m S_{mgit} + \sum_{i \in (I_f \cup I_w)} w_i \cdot lc_i \cdot wwh \cdot wpt \cdot Y_i \\
 & + \sum_{i \in (I_f \cup I_w)} wpsq \cdot lc_i \cdot wwh \cdot wpt \cdot YC_i + \sum_{\substack{(m,g) \in H \\ i \in I_f}} w_g \cdot lc_i \cdot wwh \cdot wpt \cdot Z_{gmi} \left. \right) \\
 & + tr \cdot DP_t \left. \right] \quad (1)
 \end{aligned}$$

The environmental objective function is obtained through the minimization of the environmental impact, where the environmental impact of production, storage, transportation, and entity installation is calculated for each midpoint category c , summed and normalized as represented in equation (2).

$$\begin{aligned}
 \min EnvImpact = \sum_c \eta_c & \left(\sum_{\substack{t \in T \\ (m,g) \in H}} ei_{mgc} pw_m P_{mgit} + \sum_{\substack{t \in T \\ (m,g) \in H}} ei_{mgc} vpu_m S_{mgit} \right. \\
 & + \sum_{\substack{t \in T \\ (a,m,i,j) \in H}} ei_{ac} pw_m d_{ij} X_{mai jt} + \sum_{\substack{t \in T \\ (a,m,i,j) \in H}} ei_{ic} YC_i \left. \right) \quad (2)
 \end{aligned}$$

Availability and affordability are considered by the Access to Medicine Index (AtMI), as crucial when addressing equity in access to medicines, and within these two concepts, AtMI considers that countries with highest disease burden and less ability to pay for medicines need to be prioritized. The burden of disease of a country can be measured through the metric DALYs (Disability-Adjusted Life Years), which reflects the sum of mortality and morbidity, providing a more encompassing view on health status of a population. Henceforth, the above-mentioned metric (DALY as a rate per 100,000 population) is incorporated into an objective function allowing the maximization of pharmaceutical accessibility, as represented in equation (3). In this equation, the higher the disease burden, the higher will be the value of the social factor of location i (e_i^{DALY}), thus prioritizing the location of entity i (through the decision variable Y_i) in countries with higher disease burden, as well as countries with lower levels of health expenditure.

$$\max \text{PharmaAccess} = \left(\sum_{i \in (I_f \cup I_w)} e_i^{DALY} \cdot Y_i \right) \quad (3)$$

The model additionally considers constraints that account for material balances, supply capacity, flows of products at entities, inventory levels, production, and storage technology capacities, as well as transportation capacity constraints.

3. Case-Study

The developed model is applied to a real based case-study of a supply chain of meningococcal meningitis' vaccine. Input data include the **markets which demand has to be satisfied** (U.S, Europe, Latin America, Eurasia, Asia, Africa, Middle East, Canada and Australia), **suppliers and manufacturing sites**, which include the ones already established in U.S. (Pennsylvania), Canada (Toronto), Europe (France), and Asia (India), and a fifth possible location for a factory in Africa (Kenya) is proposed so as to cover the "African meningitis belt". Moreover, **warehouses**' locations include U.S. (Pennsylvania), Canada (Toronto), Europe (France), Asia (India) and Latin America (Brazil). Possible new warehouse locations are included in the case-study: two in Africa, one in Kenya and one in Nigeria, one in Middle East (Israel), Eurasia (Russia) and Australia. The **product** considered in the study needs to be refrigerated (stored at 2° to 8°C) and each product stock keeping unit (SKU) contains 5 single-doses of meningitis vaccines.

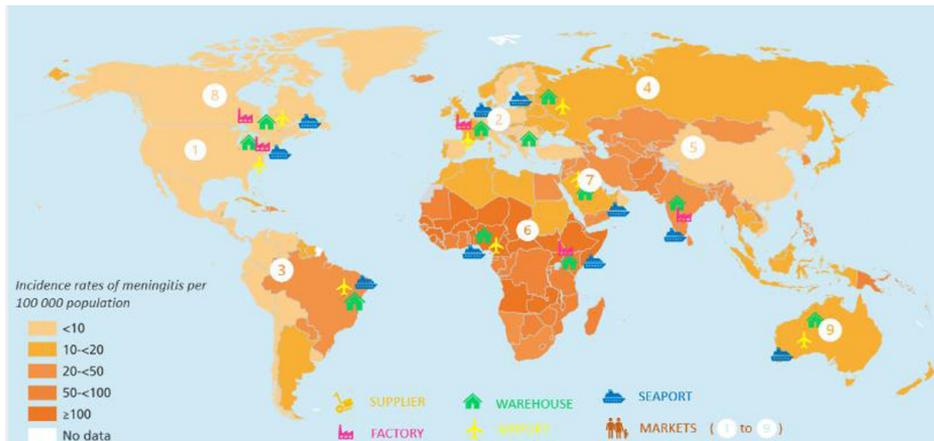


Figure 2. Superstructure considered for the case-study

Decisions were made on facility location (suppliers, factories, warehouses, hub terminals, and markets), production and storage levels as well as transportation network established.

4. Results and Discussion

Different cases are studied using a lexicographic optimization so as to obtain non-dominated solutions (see Table 1):

Case A: optimum economic performance.

Case B: optimum social performance.

Case C: optimum environmental performance.

Table 1. Summary of results considering the different cases being studied

Decisions	Cases		
	A	B	C
Entity locations	Suppliers and factories: Asia and Africa Warehouses: India and Kenya	Suppliers and factories: Asia, Europe, Canada, US, Africa Warehouses: Europe, US, Asia, Canada, Africa	Suppliers and factories: Asia, Europe, Canada, US, Africa Warehouses: Europe, US, Asia, Canada, Africa
Production levels	Most production in Asia (94%)	Most production is in Asia (38%) and Europe (32%)	Most production is in Europe (33%), Canada (26%) and Africa (20%)
Storage technology	Refrigeration	Refrigeration	None
Transportation network	Trucks of big capacity Intercontinental transportation by plane and boat	Trucks of big capacity Intercontinental transportation by boat	Trucks of big capacity Transportation by plane and boat were highly required

The most profitable solution is obtained in case A, which is not the scenario with the worst social and environmental performances. In case B, the best social performance is obtained, however this is also the case where environmental indicator perform its worst, at a cost of a 3% increase, approximately, on the environmental impact when compared with optimal performance obtained in case C. As of case C, the minimum environmental impact is achieved at a cost of approximately 1% reduction in the NPV over the same period of 10 years and social performance achieves its worst value when achieving the greenest solution, with a 67% decrease in equal access of medicines comparing with the best solution obtained in case B. Thus, the greener solution has both the worst economic and social performances.

In case A, factory in Asia is the one with higher production levels. Regarding case B, the major production of vaccines is done in Asia and Europe and in case C the installed capacity is more uniformly distributed between the five factories installed. Regarding warehouses' installed capacity, one can note that, while in case A only two warehouses are considered, located in India and Kenya, in case B seven warehouses are included in the supply chain. This can be explained by the more socially beneficial case being analyzed in situation B and being the maximization of economic objective function the second indicator considered, thus allowing for less profitable structures when comparing with case A (which corresponds to the maximization of NPV as the first indicator being maximized). Finally, in case C, no warehouses are installed, which can be due to the more evenly distribution of production across the installed factories, leading to the less need of keeping inventory in warehouses. Moreover, installation of warehouses has an impact to the environment associated, which is aimed to be minimized, as well as further costs.

The smaller variation in economic performance among the three cases can be justified by the significant contribution of production to the total costs. Moreover, by assuring a total

demand satisfaction, the cost of production can only be minimized up to a certain point, for instance, by locating a factory in a country with lower construction and labor costs. The small variation in the environmental impact across the different cases can also be justified by the major contribution of production technology to the total environmental impact. Hence, it would be important to address different options of greener production so as to minimize the impact for the environment. Recovery and remanufacturing of pharmaceutical products are extremely challenging, not only due to their limited shelf-life, but also because of their hazardousness for the environment, humans, and animals. According to some researchers on this topic, such as Amaro and Barbosa-Póvoa (2008), outdated vaccines should be properly collected to recycle, remanufacture or to be destroyed at incineration centers. These activities can potentially reduce negative environmental impacts caused by production activities of the pharmaceutical industry.

5. Conclusions and future work

The presented work proposes a decision-support tool for the design and planning of a pharmaceutical sustainable supply chain, allowing to study and comprehend the effect of each decision on the performance indicators. The application of the developed model allows for the comprehension of connections among different supply chain activities, providing an opportunity to better understand the performance of different sustainability indicators across the supply chain.

Future work is foreseen regarding recovery of pharmaceutical products, use of alternative technologies and other transportation modes (such as rail transportation). Moreover, the developed model will still be improved so as to explore different social indicators, as well as to consider uncertainty, such as uncertainty in supply, demand, production times, and also regarding the transportation modes used.

Acknowledgements

The authors acknowledge the support provided by FCT under the projects PTDC/EME-SIS/6019/2020 and FCT and P2020 PTDC/EGE-OGE/28071/2017, Lisboa -01.0145-Feder-28071.

References

- Amaro, A. C. S. and Barbosa-Póvoa, A. P. F. D. (2008) 'Planning and scheduling of industrial supply chains with reverse flows: A real pharmaceutical case study', *Computers and Chemical Engineering*, 32(11), pp. 2606–2625. doi: 10.1016/j.compchemeng.2008.03.006.
- Johnston, R. B. (2016) 'Arsenic and the 2030 Agenda for sustainable development', *Arsenic Research and Global Sustainability - Proceedings of the 6th International Congress on Arsenic in the Environment, AS 2016*, pp. 12–14. doi: 10.1201/b20466-7.
- Lemmens, S. et al. (2016) 'A review of integrated supply chain network design models: Key issues for vaccine supply chains', *Chemical Engineering Research and Design*, 109, pp. 366–384. doi: 10.1016/j.cherd.2016.02.015.
- Mota, B. et al. (2018) 'Sustainable supply chains: An integrated modeling approach under uncertainty R', *Omega*, 77, pp. 32–57. doi: 10.1016/j.omega.2017.05.006.
- Pfizer (2019) 'The Value of Vaccines in Disease Prevention: A Global Perspective', 298(18), pp. 2155–2163. Available at: <https://pfe-pfizercom-d8-prod.s3.amazonaws.com/health/VOM-Vaccines-Global-JUN2019.PDF>.
- Social, C. (2020) 'Corporate Social Responsibility - Chapter 4 of 2020 Document d'Enregistrement Universel - Sanofi'.

Plantwide control strategy for a biodiesel production process from *Jatropha curcas* oil with variable composition

Osiris Martínez-Sánchez,^a Fernando Israel Gómez-Castro,^{a*} Nelly Ramírez-Corona^b

^a *Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato, Noria Alta S/N, Col. Noria Alta, Guanajuato, Guanajuato, 36050, México. fgomez@ugto.mx*

^b *Departamento de Ingeniería Química, Alimentos y Ambiental, Universidad de las Américas de Puebla, Sta. Catarina Mártir, Puebla, Puebla, 72810, México*

Abstract

Jatropha curcas and castor oils are non-edible materials that can be used to produce biodiesel. When designing a biodiesel production process, an important aspect is that each raw material has a different lipid profile and different free fatty acids content. This may affect the performance of the production process. Nevertheless, with a proper design-control strategy, a given process can handle the variations on the composition of the raw material, maintaining the quality of the product. In this work, a strategy for the design and control of a process to produce biodiesel from *Jatropha curcas* oil with variations in triglycerides' composition of the raw material. The process is first designed in Aspen Plus V. 8.8. Then, control loops are established in Aspen Dynamics, applying disturbances of $\pm 10\%$ to the feed flowrate to test the proposed loops and the parameters of the controllers. The compositions on the feed streams are varied, maintaining the same structural design for the equipment on the process. The effect of these variations on the process operation is assessed, and the capability of the proposed control structure to stabilize the process while maintaining the product quality is verified. In general terms, the proposed control structure allows keeping the product with the desired quality, stabilizing the process after variations in the composition of the *Jatropha curcas* oil.

Keywords: biodiesel, controllability, feed composition disturbance.

1. Introduction

The use of low-quality vegetable oils helps reducing the production costs of biodiesel associated with the raw material. *Jatropha* oil is an example of this type of raw material, having a high content of free fatty acids and water. Biodiesel is mainly produced through the process of transesterification of triglycerides, with alcohol and alkaline catalysts. One of the main difficulties when working with this type of raw material is that, when synthesizing biodiesel through transesterification, alkaline catalysts could easily cause a saponification reaction that leads to the formation of soaps, which makes difficult biodiesel's downstream purification. It has been determined that to obtain good yields in the transesterification reaction, the raw material must have a percentage of free fatty acids lower than 1% (Rani et al., 2013). Berchmans and Hirata (2008) proposed producing biodiesel using *Jatropha* oil through 2 routes: using 1 reactive stage or with 2 reactive stages. For the one-stage approach, it has been observed that yields of up to 80% are

obtained using excess methanol in the reaction; this is due to the high amount of free fatty acids present in the *Jatropha* oil. On the other hand, an acid-catalyzed oil pretreatment stage allows reducing the concentration of the free fatty acids, and the presence of secondary reactions such as saponification is avoided. With this two-stages strategy, there are higher yields in the production of biodiesel, amounting to 90% under optimal conditions. Likewise, the use of different alcohols applied in the transesterification of triglycerides to produce biodiesel has been reported, where methanol and ethanol are the most used, although isopropanol cannot be ignored. It has been verified that by means of the synthesis of biodiesel using methanol and ethanol, a biofuel with characteristics that comply with the quality standards of the American standard ASTM D6571 and the European standard EN 14214 is obtained (Castillo Ospina et al., 2011; Degfie et al., 2019). Castillo Ospina et al. (2011) proposed a design for biodiesel production process with ethanol, using *Jatropha* oil with oleic acid and linoleic acid content. The process has an acid pretreatment stage followed by the alkaline transesterification of triglycerides, resulting in a high performance in the process, with a final product complying with the American standard.

The plantwide control strategy is highly recommended for its application in processes that have recycle streams, heat integration, and the interconnection of several unit operations, which is the case for a biodiesel production process through homogeneous catalysis. da Silva et al. (2019) conducted a study using the Aspen Dynamics software where different control structures were evaluated for the plantwide control strategy used in a biodiesel production process. Other works have developed decentralized control structures for biodiesel production processes with disturbances in the production (Sheng et al., 2011) or the feed flowrate (Cheng et al., 2014) In the last decade, the plantwide control strategy has been increasingly implemented in biofuel production and studied in complex processes. The processes that are fed with vegetable oils considered of low quality, such as *Jatropha* oil, have not been the exception. In order to have the capacity to satisfy a specific biofuel demand, applications of plantwide control strategies have been reported in which annual biodiesel production is maintained while respecting the quality of the biofuel produced. The plantwide control strategy is used for tuning the control loops following the corresponding heuristics proposed by different authors, and the methodologies proposed for the analysis of the perturbations applied to the system (da Silva et al., 2019; Luyben, 2013). In this study, a plantwide control strategy is proposed and assessed for a biodiesel production process, using *Jatropha* oil as raw material. The effect of composition variations in raw material on dynamic performance is analyzed, and the potential of the control structure to reject such perturbations is assessed.

2. Case study

The biodiesel production process presented in this work operates with *Jatropha* oil feed. *Jatropha* plant has a high potential for cultivation in Mexico. Table 1 reports the lipid composition of the oil used as a basis for the design of the process. This oil, composed of 4.5% free fatty acids, was characterized in Sinaloa, Mexico (Araiza Lizarde et al., 2015). In Mexico, the production of *Jatropha curcas* is favored by the different climatic and ground conditions throughout the country. According to a study carried out by INIFAP (2012), in Mexico there are at least 3,000,000 hectares with optimal conditions for the harvest of *Jatropha*. In addition, it is reported that up to 1,000,000 plants can be harvested per hectare, and a *Jatropha* plant provides approximately 3 seeds, and each of these produces 6.25×10^{-5} liters of oil (Hooda and Rawat, 2006). Based on the aforementioned

information, a process feeding of 3,500 kg/h of vegetable oil was proposed, since Mexico could produce this amount annually.

Table 1. Lipid composition of *Jatropha* oil.

Fatty acids (4.5% wt)	Triglycerides (95.5% wt)
Palmitic acid (4.1% wt)	Trypalmitin (4.1% wt)
Oleic acid (46.87% wt)	Triolein (46.87% wt)
Linoleic acid (45.18% wt)	Trilinolein (45.18% wt)
Linolenic acid (3.85% wt)	Trilinolenin (3.85% wt)

Figure 1 shows the process proposed to produce biodiesel. The *Jatropha* oil pretreatment stage is implemented; this step is catalyzed by an acid. The second reactive stage is where the transesterification of triglycerides to ethyl esters takes place. Likewise, downstream of the esterification and transesterification reactors, there is a neutralization reactor, where the catalysts are neutralized, and the resulting salts are separated to improve the efficiency of the process. Furthermore, with the help of a decanter, the biodiesel phase and the ethanol-glycerol-water phase are separated; the oily phase goes to a series of equipment where the biodiesel is washed and separated from the impurities to finally obtain the refined biofuel. On the other hand, the aqueous phase is directed to two distillation columns where most of the ethanol and glycerol present in the stream are recovered. Sensitivity analysis has been applied to the process equipment with available degrees of freedom, mainly the reactors and distillation columns, to obtain the best operating variables, aiming to keep costs down and maintain product quality. The NRTL thermodynamic model was used in the proposed process; however, in the phase separation stages with the presence of ethyl esters with glycerol, such as the decanter and the biodiesel washing column, the UNIFAC-LL model is used (Berchmans and Hirata, 2008).

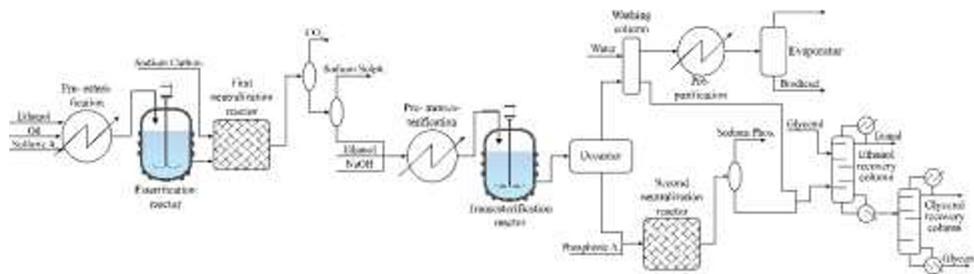


Figure 1. Process flowsheet.

3. Control strategy

The plantwide control methodology allows covering the studied system, given by a chemical process with different stages and many variables. This can be done by decentralizing the process. For processes with no recycling of matter and energy integration, it is possible to perform a decomposition based on process units to be addressed in individual problems (Luyben 2013). Control objectives have been defined in each of the process units used to produce biodiesel. These control objectives are mainly the operating variables that have the greatest influence on the operation of the equipment; this is easily detectable with the knowledge of the operation of the process. In addition, an open-loop analysis of the process has been carried out, where perturbations of $\pm 10\%$ have been applied in the oil feed flow.

Taking into consideration the proposed control objectives, the control ties proposed in the control strategy are established and shown in Table 2. To determine the location of the temperature controllers in the distillation columns, an analysis was carried out to observe the variations that the temperature presented between one stage to another, selecting the stage with the greatest variation in temperature. The controllers used are of the proportional-integral (PI) type; these are enough to control pressure and level due to the zero complexity in these control loops. In addition, to control the temperature in non-extreme systems, the PI controller is a good option (Luyben, 2013). The level and pressure controllers were tuned following the heuristics reported in Luyben (2013); however, to tune the temperature controllers, the Ziegler and Nichols method was used, supported by the tuning tool included in the Aspen Dynamics software.

Table 2. Proposed control loops.

Equipment	Controlled variable	Manipulated variable
Pre-esterification	Inlet stream temperature	Heat duty
Esterification reactor	Temperature	Heat duty
	Level	Output stream flow
Pre-transesterification	Output stream temperature	Heat duty
Transesterification reactor	Temperature	Heat duty
	Level	Output stream flow
Decanter	Liquid 1 level	Output liquid 1 flow
	Liquid 2 level	Output liquid 2 flow
Ethanol recovery columna	Pressure	Condenser heat duty
	Reflux tank level	Distilled flow
	Reboiler tank level	Condensate flow
	Stage 15 temperature	Reboiler heat duty
Glycerol recovery column	Pressure	Condenser heat duty
	Reflux tank level	Distilled flow
	Reboiler tank level	Condensate flow
	Stage 3 temperature	Reboiler heat duty

4. Results

The operation of the process has been considered for 8,160 hours of annual production (340 days), based on that, the process has the capacity to produce 30,000 tons of biodiesel per year. The steady-state process reduces the amount of free fatty acids present in the oil, with the pretreatment stage operating at 70°C. This temperature is important since having higher temperatures would decrease the performance of the reactor and elevation in costs. Likewise, in the transesterification reactor, high yields are obtained due to the lower presence of free fatty acids, in addition, with the conditions proposed such as operating at 70°C, a global conversion of triglycerides to ethyl esters above 99% was obtained. The characteristics of the obtained biodiesel have been compared with those presented in the EN 14214 standard, being satisfactory in all areas. In addition, the distillation columns allow the recovery of a good amount of ethanol and glycerol, which could be reused through the implementation of recycling.

The control strategy is evaluated in different ways. First, perturbations of $\pm 10\%$ were applied to the mass flowrate of the oil fed to the process, analyzing the behavior of the controlled variables throughout the process. In both cases, the behavior in the esterification reactor and the transesterification reactor stabilizes relatively easily. The

controllers in the distillation columns stabilize after a couple of hours, and the greatest variations are observed in the temperature controller. The purity of ethanol and glycerol remains above 99.5% in both cases, and the foregoing is shown in Figure 2. In addition, the biodiesel obtained, although in greater quantity, resembles its characteristics to those recommended by the European standard.

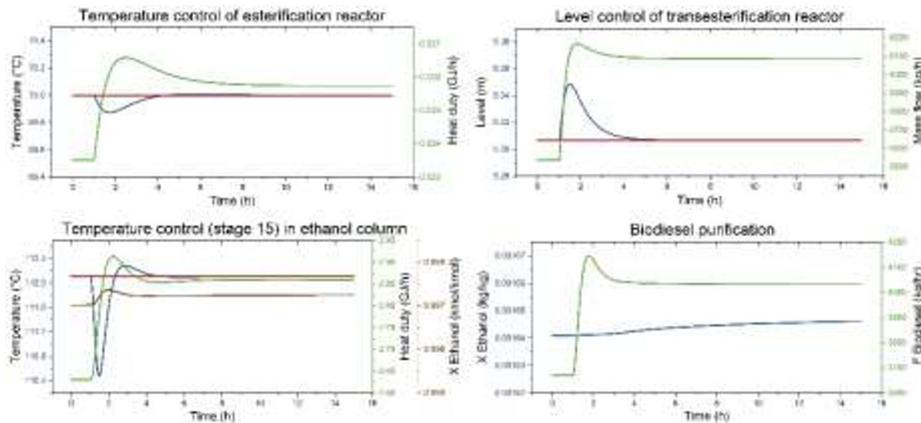


Figure 2. + 10% variation in oil feed mass flowrate.

Because of the composition of *Jatropha curcas* varies depending on its harvest origin, the capacity of the process was evaluated in the face of variations in the composition of the oil fed to the process, based on typical compositions previously reported (Araiza Lizarde et al., 2015; Guevara-Fefer et al., 2016). In these cases, the tank levels of the important equipment in the process are affected by the variation in the molar composition of the oil, however, these are less than the variation in the mass flowrate. In the distillation columns, the temperature controllers are the ones that present the greatest variations, this mainly due to the significant changes that the molar composition of the same type of oil can present due to being harvested under different conditions.

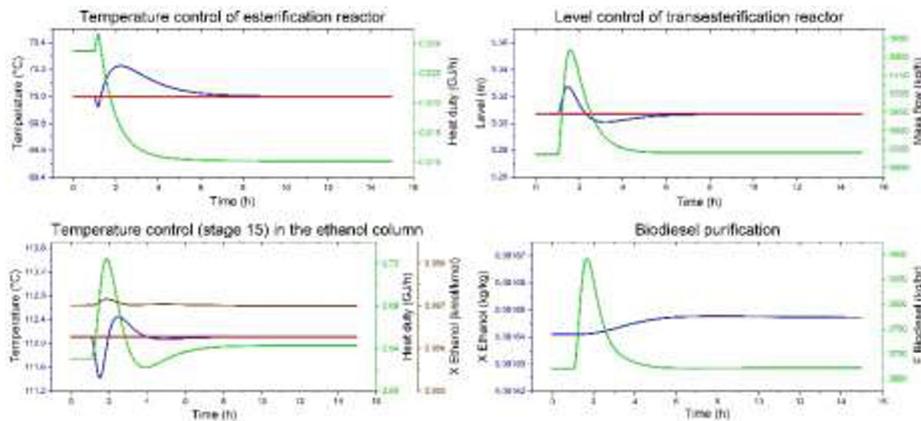


Figure 3. Composition variation in *Jatropha* oil fed.

In these cases, the purity of the ethanol and glycerol undergo small changes, while the purity and quality of the biodiesel obtained are satisfactorily maintained, as can be seen in Figure 3.

5. Conclusions

The design for a biodiesel production process has been carried out in a simulation environment, allowing the analysis for the complete process to be carried out. The operating conditions of the most influential stages in the process were analyzed through sensitivity analysis corresponding to each equipment, which proved to be enough to obtain a product capable of complying with the international regulations to which biofuels are subjected.

The proposed control strategy proved to be appropriated for the requirements sought in the work, which are the production of biodiesel from *Jatropha* oil, considering the variability of its composition or its availability. However, there are still areas of opportunity, among which it can be mentioned that the method used to tune the drivers was the Ziegler-Nichols method, which tunes the driver empirically. This can be improved by implementing a methodology to optimize these parameters, or else complement the closed-loop strategy with optimization. On the other hand, in the control loops proposed in the project, although the application with the cases presented was correct, a methodology was applied to close these loops sequentially and without considering the dead time in the controllers. This was mainly done because the proposed biodiesel production process does not contain material recycles or energy integration.

References

- N. Araiza Lizarde, L. Alcaraz Melendez, M.A. Angulo Escalante, T. Reynoso Granados, P. Cruz Hernandez, M.M. Ortega Nieblas, 2015, Propiedades fisicoquímicas del aceite de semillas de *Jatropha curcas* de poblaciones silvestres en México, *Revista de la Facultad de Ciencias Agrarias*, 47, 1, 127-137.
- H.J. Berchmans, S. Hirata, 2008, Biodiesel production from crude *Jatropha curcas* seed oil with a high content of free fatty acids, *Bioresource Technology*, 99, 6, 1716-1721.
- A.M. Castillo Ospina, P. Cuartas, J.A. Velásquez Jimenez, 2011, Obtención de biodiésel a partir de aceite de *Jatropha curcas* por transesterificación etanólica, *Investigaciones aplicadas*, 5, 1, 34-41.
- J.-K. Cheng, C.-C. Chao, J.D. Ward, I.-L. Chien, 2014, Design and control of a biodiesel production process using sugar catalyst for oil feedstock with different free fatty acid concentrations, *Journal of the Taiwan Institute of Chemical Engineers*, 45, 1, 76-84.
- B.F. da Silva, J.E. Schmitz, I.C. Franco, F.V. da Silva, 2019, Plantwide control systems design and evaluation applied to biodiesel production, *Biofuels*, 12, 10, 1199-1207.
- T.A. Degfie, T.T. Mamo, Y.S. Mekonnen, 2019, Optimized biodiesel production from waste cooking oil (WCO) using calcium oxide (CaO) nano-catalyst, *Scientific Reports*, 9, 18982.
- P. Guevara-Fefer, N. Niño-García, Y.A. De-Jesús-Romero, G. Sánchez-Ramos, 2016, *Jatropha sotoi-nunyezii* y *Jatropha curcas*, especies de Tamaulipas: una comparación desde la perspectiva de los biocombustibles, *CienciaUAT*, 11, 1, 91-100.
- N. Hooda, V.R.S. Rawat, 2006, Role of bio-energy plantations for carbon-dioxide mitigation with special reference to India, *Mitigation and Adaptation Strategies for Global Change*, 11, 445-467.
- INIFAP, 2012, Potencial productivo de especies agrícolas de importancia socioeconómica en México. <https://www.cmdrs.gob.mx/sites/default/files/cmdrs/sesion/2018/09/17/1474/materiales/inifap-estudio.pdf>. Last consulted November 16, 2021.
- W.L. Luyben, 2013, *Distillation Design and Control using Aspen Simulation*, Jhon Wiley & Sons.
- K.N.P. Rani, T.P. Kumar, T.S.V.R. Neeharika, B. Satyavathi, R.B.N. Prasad, 2013, Kinetic studies on the esterification of free fatty acids in *Jatropha* oil, *European Journal of Lipid Science and Technology*, 115, 6, 691-697.
- Y.H. Shen, J.K. Cheng, J.K. Ward, C.C. Yu, 2011, Design and control of biodiesel production processes with phase split and recycle in the reactor system, *Journal of the Taiwan Institute of Chemical Engineering*, 42, 5, 741-750.

Cascade fuzzy control of a tubular chemical reactor

Anna Vasičkaninová, Monika Bakošová, Alajos Mészáros

*Slovak University of Technology in Bratislava, Institute of Information Engineering,
Automation, and Mathematics, Radlinského 9, 81237 Bratislava, Slovak Republic
anna.vasickaninova@stuba.sk*

Abstract

Cascade control is a multi-loop control structure often used in industrial applications, which offers a possibility for applying advanced controllers. This paper compares cascade control with type-1 fuzzy controllers, type-2 fuzzy controllers, and PID controllers on the case study of a tubular chemical reactor. The primary controllers are type-1 fuzzy PID and PD controllers, type-2 fuzzy PID and PD controllers, or conventional PID controllers. The secondary controllers are type-1 fuzzy P, type-2 fuzzy P or conventional P controllers. Simulation results demonstrate that cascade control with both types of fuzzy controllers can assure better values of followed performance indices and higher energy savings measured by the coolant consumption during control of the tubular chemical reactor.

Keywords: cascade control, type-1 fuzzy control, type-2 fuzzy control, PID control, tubular chemical reactor.

1. Introduction

Between advanced control strategies, fuzzy logic control is often found in applications where conventional closed loop control does not assure satisfactory results because of non-linearity, asymmetric dynamics, or uncertainties in the controlled processes. Fuzzy logic control is based on the theory of fuzzy sets pioneered by Zadeh (1965). Zadeh (1975) also introduced the concept of the type-2 fuzzy logic. History, application, and possible future of fuzzy control are summarized in Guerra et al. (2015). Mendel (2018) introduced rule-based systems from type-1, interval type-2 and general type-2 fuzzy systems. Mittal et al. (2020) offered overview of past, present, and future trends of type-2 fuzzy logic applications including theoretical and practical implications.

Cascade control is a multiloop control strategy that enables using of advanced controllers. Meng and Hou (2011) designed cascade control with main fuzzy PID controller and auxiliary PID controller for hydro-viscous drive speed regulating start. Kumbasar and Hagra (2013) proposed a cascade control architecture, which includes the inner and outer control loops for the path tracking control of mobile robots in presence of uncertainty. García et al. (2007) designed fuzzy logic controller with intermediate variable as an alternative for cascade control with fuzzy controllers and compared both strategies. Xie and Liu (2017) formed fuzzy cascade control based on known control history for superheated temperature.

Despite intensive research and promising applications in various fields, there is a lack of studies devoted to implementation of type-1 fuzzy logic controllers (T1FLCs) and type-2 fuzzy logic controllers (T2FLCs) to tubular chemical reactors and T2FLCs in cascade control. The main goal of this paper is to show that cascade control (CC) with T2FLCs can guarantee energy savings and better performance compared to CC with conventional

PID controllers and is alternative to CC with T1FLCs when controlling systems with uncertainties, asymmetric dynamics or nonlinear systems, as tubular chemical reactors.

2. Cascade fuzzy control

2.1. Cascade control

Cascade control (CC) (Figure 1) is a multi-loop control structure used in process industry to improve control under immeasurable disturbances (Bequette, 2003). In Figure 1, C_1 is the primary (main) controller, C_2 is the secondary (auxiliary) controller, P_1 is the primary controlled system, and P_2 is the secondary controlled system. Signals r_1 and r_2 represent reference values, y_1 and y_2 are controlled outputs, e_1 and e_2 are errors, u_2 is the manipulated variable that results from the control input calculated by C_2 influenced by the disturbance d_2 . The disturbance d_1 influences the primary controlled output y_1 . Both controllers in the CC can be fuzzy controllers.

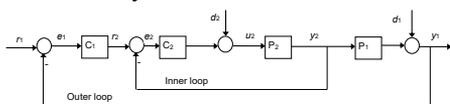


Figure 1: Scheme of a cascade control system

2.2. Type-1 fuzzy control and interval type-2 fuzzy control

The structure of T1FLC is represented in Figure 2. The crisp inputs to the dynamic controller can be errors, derivatives of errors, integrals of errors or previous values of measurements backward in time. Fuzzifier converts input data to degrees of membership by a lookup in one or several membership functions. Rule base includes various empirical rules. Defuzzifier converts the resulting fuzzy set to numbers that enter the process as control inputs.

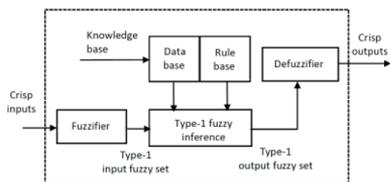


Figure 2: Type-1 fuzzy controller

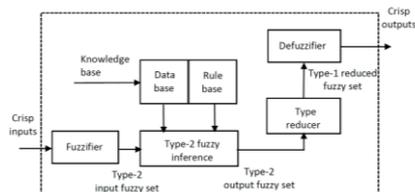


Figure 3: Interval type-2 fuzzy controller

Figure 3 represents structure of T2FLC. The rule base for T2FLC remains the same as for T1FLC, but its membership functions are type-2 interval fuzzy sets, and a reducer must be used prior defuzzification (Kumbasar, 2014). The advantage of using type-2 fuzzy logic (FL) compared to type-1 FL is that type-2 FL can handle uncertainty in control, which may be due to noise, dynamic changes in the environment, or imprecision in the models (Mittal et al., 2020).

3. Case study

The case study from chemical engineering domain is devoted to a tubular chemical reactor (TCR) with exothermic consecutive reactions $A \xrightarrow{k_1} B \xrightarrow{k_2} C$ in the liquid phase and with the co-current cooling (Dostal et al., 2015). Vasičkaninová et al. (2019) did steady-state analysis and step-response based identification of the TCR and based on presented results, TCR is the nonlinear system with asymmetric dynamics and can be treated as a system with uncertainty. As B is the main product and C is the side product, it is necessary to

keep the concentration c_B at the reference value. In CC of TCR, the concentration c_B is the primary controlled output and the temperature of reaction mixture T is the secondary controlled output. The manipulated variable is the flow rate of coolant q .

3.1. *Cascade control of the tubular chemical reactor using conventional PID controllers*
 The transfer function of the PID controller has the form (Mikleš and Fikar, 2007)

$$C = k_p \left(1 + \frac{1}{t_i s} + t_d s \right) \tag{1}$$

where k_p is the proportional gain, t_i is the integral time, t_d is the derivative time. The secondary controller was tuned experimentally as a P controller. The primary PID controller was tuned using the Rivera-Morari method (PID-RM) and the primary PI controller was tuned using the Cohen-Coon method (PI-CC) (Bequette, 2003). Two primary controllers assuring best simulation results were chosen from several designed controllers and no fine-tuning was done. Table 1 presents the controller parameters.

Table 1: PID controller parameters

Controller parameters	Primary controllers		Secondary controller
	PID-RM	PI-CC	P
k_p	1.95	2.18	-0.4
t_i	14.30	3.86	
t_d	0.67		

3.2. *Cascade control of the tubular chemical reactor using fuzzy controllers*

3.2.1. *Secondary type-1 fuzzy P controller and secondary type-2 fuzzy P controller*

Both, the secondary type-1 fuzzy P controller (P-T1FLC) and the secondary type-2 fuzzy P controller (P-T2FLC) were designed as Sugeno-type fuzzy inference systems (FISs), each with 2 rules

$$\text{If } e \text{ is } A_i \text{ Then } f_i = p_i e + q_i \tag{2}$$

where e is the error, p_i, q_i are the consequent parameters presented in Table 2 together with the antecedent parameters A_i and the parameters of the symmetric Gaussian membership function σ_i, c_i (Zhao and Bose, 2002) used for fuzzification of inputs.

Table 2: Parameters of symmetric Gaussian functions, antecedent and consequent parameters

Rule	σ_i	c_i	A_i	p_i	q_i
1	5.93	-14.79	A_1	-0.029	0.23
2	5.93	-0.81	A_2	-0.031	0.24

3.2.2. *Primary type-1 fuzzy PD controller and primary type-2 fuzzy PD controller*

Both, the primary type-1 fuzzy PD controller (PD-T1FLC) and the primary type-2 fuzzy PD controller (PD-T2FLC) were designed as the Sugeno-type FISs with 6 rules:

$$\text{If } e \text{ is } A_i \text{ and } \frac{de}{dt} \text{ is } B_i \text{ Then } f_i = p_i e + q_i \frac{de}{dt} + r_i \tag{3}$$

where e is the error, de/dt is the derivative of error, A_i, B_i are the antecedent parameters and p_i, q_i, r_i are the consequent parameters, which are presented in Table 3.

Sugeno-type FISs were generated using the subtractive clustering method. Triangular membership functions (Zhao and Bose, 2002) were used for fuzzification of inputs and Table 4 presents parameters of used triangular membership functions.

Table 3: Antecedent parameters and consequent parameters

Rule	A_i	B_i	p_i	q_i	r_i
1	A_1	B_1	121.29	-16.33	316.04
2	A_1	B_2	-22.92	-68.84	329.37
3	A_2	B_1	87.13	-192.30	201.02
4	A_2	B_2	30.52	128.56	370.74
5	A_3	B_1	103.65	-28.07	142.95
6	A_1	B_3	84.70	-4.02	178.70

Table 4: Parameters of triangular membership functions

e			de/dt		
a_i	b_i	c_i	a_i	b_i	c_i
-0.96	-0.05	0.86	-0.59	-0.28	0.006
-0.01	0.86	1.78	-0.28	0.02	0.32
0.87	1.78	2.70			

3.2.3. Primary type-1 fuzzy PID controller and primary type-2 fuzzy PID controller

Both, the primary type-1 fuzzy PID controller (PID-T1FLC) and the primary type-2 fuzzy PID controller (PID-T2FLC) were designed as the Sugeno-type FISs with 8 rules:

$$\text{If } e \text{ is } A_i \text{ and } \frac{de}{dt} \text{ is } B_i \text{ and } \int e dt \text{ is } C_i \text{ Then } f_i = p_i e + q_i \frac{de}{dt} + r_i \int e dt + s_i \quad (4)$$

where e is the error, de/dt is the derivative of error, $\int e dt$ is the integral of error, p_i, q_i, r_i, s_i are the consequent parameters. Table 5 presents the antecedent and consequent parameters. Table 6 shows the parameters of the symmetric Gaussian membership functions (Zhao and Bose, 2002) used for the fuzzification of inputs.

Table 5: Antecedent and consequent parameters

Rule	A_i	B_i	C_i	p_i	q_i	r_i	s_i
1	A_1	B_1	C_1	52.44	-0.41	132.53	45.48
2	A_1	B_1	C_2	25.95	11.54	31.96	52.56
3	A_1	B_2	C_1	27.66	-0.40	0.18	330.32
4	A_1	B_2	C_2	-2.44	0.07	0.15	330.53
5	A_2	B_1	C_1	129.66	91.08	21.09	111.06
6	A_2	B_1	C_2	141.61	-18.23	6.13	66.79
7	A_2	B_2	C_1	144.14	-2.60	11.26	85.82
8	A_2	B_2	C_2	13.47	-2.26	50.51	16.70

Table 6: Parameters of symmetric Gaussian membership functions

e		de/dt		$\int e dt$	
σ_i	c_i	σ_i	c_i	σ_i	c_i
0.82	-0.04	0.06	-0.26	3.12	0.17
0.59	1.83	0.05	0.07	3.07	7.29

4. Simulation results

The MATLAB/Simulink R2021b programming environment was exploited for simulations using CPU i7-11700 2.50 GHz, 32 GB RAM. The simulation results for six scenarios in reference tracking and disturbance rejection are presented in Figures 4 and 5.

The primary reference value was the desired value of the product concentration $c_B = 2.15 \text{ kmol m}^{-3}$. The disturbances were represented by increasing the flow rate of the reaction mixture from 0.2 to $0.36 \text{ m}^3\text{s}^{-1}$ at time 100 s and then by decreasing to $0.1 \text{ m}^3\text{s}^{-1}$ at time 200 s . The results were compared numerically assessing the total consumption of cooling water V during control, the integral performance index IAE (integrated absolute error), and ISE (integrated squared error) defined e. g. in Mikleš and Fikar (2007). Table 7 summarizes these numerical results.

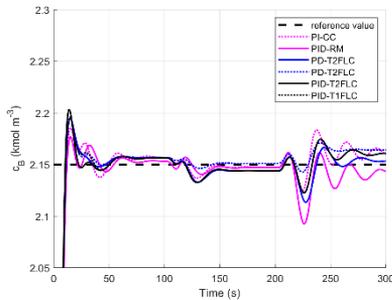


Figure 4: Control responses of the product concentration c_B in six scenarios

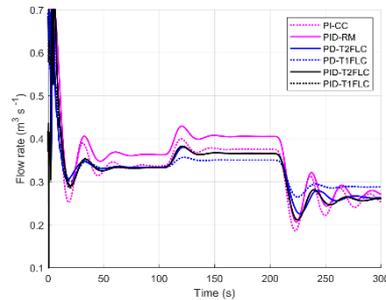


Figure 5: Trajectories of the manipulated variable in six scenarios

Table 7: Values of V , IAE, and ISE

Scenario	Primary controller	Secondary controller	V (m^3)	IAE ($\text{kmol m}^{-3} \text{ s}$)	ISE ($\text{kmol}^2 \text{ m}^{-6} \text{ s}$)
1	PI-CC	P	100.4396	9.9652	8.4881
2	PID-RM	P	108.3361	9.7175	8.5058
3	PD-T1FLC	P-T1FLC	99.7704	9.5105	8.4505
4	PD-T2FLC	P-T2FLC	99.2400	9.6808	8.4572
5	PID-T1FLC	P-T1FLC	98.5960	9.9505	8.4576
6	PID-T2FLC	P-T2FLC	98.5958	9.7361	8.3604

CC with PID-T2FLC and P-T2FLC guaranteed the lowest coolant consumption and the coolant consumption using CC with PID-T1FLC and P-T1FLC was almost the same. The coolant consumption increased by 0.65 % if CC with PD-T2FLC and PD-T2FLC was used and by 1.19 % for CC with PD-T1FLC and P-T1FLC. CC with conventional PID-RM and P assured the highest coolant consumption. CC with PD-T1FLC and P-T1FLC reached the lowest value of IAE and the second best according to IAE was CC with PD-T2FLC and P-T2FLC. The worst cascade control according to IAE achieved conventional CC with PI-CC and P controllers. The IAE increased in this CC by 4.78 %. The best value of ISE assured CC with PID-T2FLC and P-T2FLC. The second best was the CC with PD-T1FLC and P-T1FLC with the ISE value greater by 1.08%. CC with PID-RM and P controller was the worst with the ISE value higher by 1.74% compared to the best CC with PID-T2FLC and P-T2FLC. Comparing the coolant consumption, the ISE and IAE values, the CC with the primary PID-T2FLC and the secondary P-T2FLC was the best CC scenario.

5. Conclusions

CC with conventional controllers, type-1 fuzzy controllers, and type-2 fuzzy controllers was studied on TCR. CC with the primary PID-T2FLC and the secondary P-T2FLC

assured the most efficient operation of TCR. This scenario assured the lowest coolant consumption and the lowest value of the ISE performance index. According to the IAE performance index, CC with the primary PD-T1FLC and the secondary P-T1FLC was the best. The second best was CC with the primary PD-T2FLC and the secondary P-T2FLC. Based on the comparison of all results, it can be stated that both types of FLCs can be used successfully in cascade control for reaching the goals of control. Application of more complicated fuzzy type-2 controllers helped to improve the energetic efficiency of the studied TCR measured by coolant consumption. Further intensive research in this field will continue in the future.

Acknowledgement

The authors acknowledge the contribution of the Scientific Grant Agency of the Slovak Republic under the grants 1/0545/20 and 1/0691/21.

References

- B.W. Bequette, 2003. *Process Control: Modelling, Design and Simulation*, Prentice Hall, Upper Saddle River, New Jersey.
- P. Dostal, J. Vojtesek, V. Bobal, 2015. One Approach to Adaptive Control of a Nonlinear Distributed Parameters Process, In: R. Silhavy et al. (eds), *Intelligent Systems in Cybernetics and Automation Theory, CSOC 2015, Advances in Intelligent Systems and Computing* 348, Springer, Cham, 259–274.
- Y. Garcia, O. Camacho, M. Sanjuan, E. Iglesias, C.A. Smith, 2007. A new approach of cascade control based on fuzzy logic, 9th WSEAS International Conference on Automatic Control, Modeling & Simulation, 36–41.
- T.M. Guerra, A. Sala, K. Tanaka, 2015. Fuzzy control turns 50: 10 years later, *Fuzzy Sets and Systems* 281, 168–182.
- T. Kumbasar, 2014. A simple design method for interval type-2 fuzzy PID controllers, *Soft Computing* 18, 1293–1304.
- T. Kumbasar, H. Hagraş, 2013. A Type-2 Fuzzy Cascade Control Architecture for Mobile Robots, *IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK, 3226–3231.
- J.M. Mendel, 2018. *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Second Edition, Springer, Cham, Switzerland.
- Q.-R. Meng, Y.-F. Hou, 2011. A Fuzzy PID Cascade Control System of a Hydro-viscous Drive Speed Regulating Start, *Third International Conference on Measuring Technology and Mechatronics Automation*, Shanghai, China, 52–57.
- J. Mikleš, M. Fikar, 2007. *Process Modelling, Identification, and Control*, Springer Verlag, Berlin, Germany.
- K. Mittal, A. Jain, K.S. Vaisla, O. Castillo, J. Kacprzyk, 2020. A comprehensive review on type 2 fuzzy logic applications: Past, present and future, *Engineering Applications of Artificial Intelligence* 95, 103916.
- A. Vasičkaninová, M. Bakošová, J. Oravec, A. Meszáros, 2019. Model predictive control of a tubular chemical reactor, *22nd International Conference on Process Control*, 228–233.
- Y.-H. Xie, X.-Q. Liu, 2017. Fuzzy Cascade Control Research of Superheated Temperature, *International Conference on Manufacturing Engineering and Intelligent Materials*, 406–411.
- L.A. Zadeh, 1965. Fuzzy sets, *Information and Control* 8, 338–353.
- L.A. Zadeh, 1975. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning–1, *Information Sciences* 8, 199–249.
- J. Zhao, B. K. Bose, 2002. Evaluation of membership functions for fuzzy logic controlled induction motor drive, *IEEE 28th Annual Conference of the Industrial Electronics Society*, 1, 229–234.

Design and Implementation of an Optimal Control Framework for Post-combustion CO₂ Capture Process

Gaurav Mirlekar, Lars O. Nord

Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Abstract

The high penetration of renewable sources into fossil fuel-based thermal power plants demand for operational flexibility with Carbon Capture and Storage (CCS) technologies. Post-combustion CO₂ Capture (PCC) processes using chemical absorption of CO₂ in flue gas need to cope with flexible operation and the CO₂ capture rate must be controlled under fluctuating flue gas conditions. To fill this gap, in this work, an optimal control framework is proposed and implemented on PCC with monoethanolamine (MEA)-based CO₂ capture process simulation. A two-input-two-output control structure is selected from PCC that consist of flue gas flow rate and lean MEA flow rate as the input/manipulated variables while CO₂ capture rate and reboiler duty are considered as the output/controlled variables. Open-loop simulations are performed in which simulated step tests are designed by individually moving the input variables as steps and collecting the resulting data for the output variables. The classical autoregressive model with exogenous inputs (ARX) method is used for deriving the data-driven simplified dynamic model that can be embedded inside the Biologically Inspired Optimal Control Strategy (BIO-CS) casted as Model Predictive Control (MPC) to compute control moves for simultaneous control of both output variables in the dynamic simulation. The results are compared to the standalone Proportional-Integral-Derivative (PID) controller existed in the simulation in terms of the time required to reach new steady state and output tracking error. The proposed approach improves the overall performance of the process resulting in faster and flexible setpoint tracking during ramp decrease in the flue gas flow rate case study and thus providing a promising alternative.

Keywords: Carbon Capture and Storage (CCS), Process Control, Dynamic Simulation.

1. Introduction

The global Greenhouse Gas (GHG) emissions need to peak now and be reduced to net-zero by 2050 to fulfil the Paris agreement (Figueres et al. 2017). Fossil fuel-based thermal power plants accounts for the majority of the GHG, such as CO₂ release. Therefore, these energy processes can be integrated with Carbon Capture and Storage (CCS) technologies to mitigate CO₂ emissions and produce cleaner electricity (Dutta et al. 2017a). In addition, hydrogen can, as an energy carrier, enable decarbonization and increased deployment of renewable energy through sector coupling of transport, industry, buildings, as well as the power sector. Norway has the potential to produce hydrogen from the abundant natural gas resources with CCS and provide Europe with blue hydrogen required for the long-term plans (Skar et al. 2018). For such clean energy, post-combustion CO₂ capture (PCC) processes using chemical absorption of CO₂ in flue gas from power plants are the most near-term carbon capture technologies (Dutta et al.

2017a). However, the regeneration of chemical solvents is a highly energy intensive process that require particular attention. The excess heat requirement at the reboiler in the stripper lead to an increase in the overall energy requirements (Dutta et al. 2017b). Additionally, the CO₂ capture rate needs to be maintained at a desired level during power plant load variations that changes flue gas flow rates. In the past, PCC plant behavior has been analyzed by variations in steam extraction from the power plant and plantwide control methods have been studied for suitable control structures using dynamic simulations (Dutta et al. 2017a). Few studies have been performed to examine the effect of controlling the CO₂ capture rate when the PCC plants operate under fluctuating flue gas conditions. In industry, decentralized control structures and open-loop responses for load changes have been experimentally studied for flexible operations (R. Montanes et al. 2018). Additionally, the applications of nonlinear Model Predictive Control (MPC) have been demonstrated on PCC pilot plants considering CO₂ capture ratio as a controlled variable and reboiler duty as a constraint (Hauger et al. 2019). However, the simultaneous control of CO₂ capture rate and reboiler duty has not been analyzed yet. Therefore, a control strategy capable of mitigating the effects of fluctuations and improve overall performance of the system during simultaneous control with minimal human interference must be implemented.

In recent years, a Biologically Inspired Optimal Control Strategy (BIO-CS) is developed and implemented on different processes. The applications of BIO-CS include (i) a nonlinear first principle-based fermentation process model; and (ii) a simulation of a subsystem from the Integrated Gasification Combined Cycle (IGCC) that involves absorption column associated with the Acid Gas Removal (AGR) using physical solvent selexol (Mirlekar et al. 2018a, 2018b). The results of the implementation show that the BIO-CS can tackle various challenges such as, nonlinearities in the process models, multivariable control structures, performance improvement in terms of setpoint tracking error and response time when compared with classical Proportional-Integral-Derivative (PID) controller. Specifically, BIO-CS casted as MPC has displayed faster computational time advantage over standalone BIO-CS. Hence, to address challenges in CCS technologies combined with power plants, in this work, a design of an optimal control framework is presented and its implementation on PCC with monoethanolamine (MEA)-based CO₂ capture dynamic process model simulation is illustrated. The BIO-CS as MPC approach is at the core of this optimal control framework with inclusion of steps that needs to be considered during the implantation of the framework on PCC simulation. The application of the developed framework addresses problems related to control stricture with multiple variables, connection between two software platforms, maintaining required CO₂ capture rate during variations in the flue gas flow rate and increasing energy efficiency.

The paper is organized as follows: section 2 describes the PCC process; section 3 discusses the design of an optimal control framework with details on implementation steps; section 4 consists of framework implementation case study and closed-loop simulation results; section 5 concludes the paper.

2. Post-combustion CO₂ Capture (PCC) Process

The post-combustion CO₂ capture process considered for application in this paper is a dynamic simulation of an amine-based chemical absorption process in Aspen HYSYS[®] V10 (Aspen HYSYS[®] software). As depicted in Figure 1, this process consists of two cylindrical columns representing absorber and regenerator/stripper respectively. The flue gas stream containing nitrogen, water and CO₂ enters the absorber column from

the bottom. The lean solvent stream consisting of a mixture of fresh MEA, CO₂ and water is connected at the top of the absorber. In the flowsheet, the flow rate of the lean MEA is regulated via PID controller for desired CO₂ capture rate. The absorber and stripper are represented by equilibrium-based stage models and random packing material. After the absorption process, the treated gas stream at the top obtained after desired CO₂ capture with trace amount of CO₂ is released into the atmosphere.

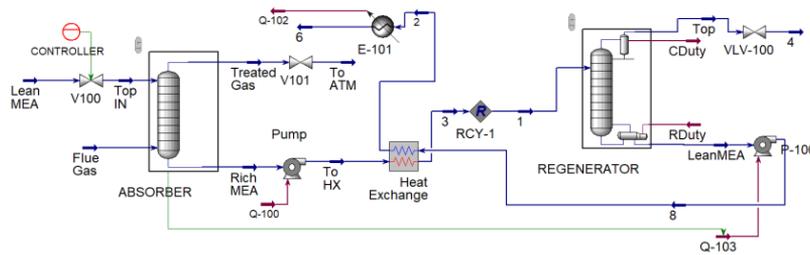


Figure 1: Schematic of the dynamic simulation of the PCC process in Aspen HYSYS[®]. The rich solvent stream coming from the bottom of the absorber is pumped through a heat-exchanger and sent at the top of the regenerator. The heat exchanger employed for heat transfer between rich and lean solvents is modeled using plate-fin configuration and the thermal properties are estimated from the steady-state conditions. CO₂ is separated from the rich solvent in the regenerator with condenser and reboiler modeled as volumes with suitable heat duty. The separated CO₂ stream at the condenser is sent for storage or compression while the lean solvent stream at the bottom of the regenerator is pumped backed into the heat-exchanger. Pumps are modeled by specifying their efficiencies and duty. Few assumptions are also made in this dynamic simulation, for example, the piping is ignored, and the corresponding pressure drop is included in the neighboring equipment. The dynamic simulation model of PCC process is validated using data from the pilot plant available in the literature (Dutta et al. 2017a). The fluctuations in the flue gas flow rate caused by ramp down or ramp up activities performed during load variations affect CO₂ capture rate and energy requirements when such PCC model is connected to a power plant. This issue in a dynamic operation of the system is addressed using the optimal control framework in this work and discussed in the next section.

3. Optimal Control Framework

- The design of an optimal control framework consists of three steps.
- i. Data-driven PCC model development for controller design
 - ii. Formulation of BIO-CS casted as MPC
 - iii. Implementation of control laws on the simulation plant

These steps are summarized in a block diagram depicted in Figure 2. The data-driven dynamic process model of PCC system is developed as a first step in the design of an optimal control framework. Initially, input/manipulated and output/controlled variables are selected based on the previous knowledge of the PCC process. For the implementation case study considered in this paper, flue gas flow rate and the valve opening that regulates lean MEA flow rate are chosen as the input variables while CO₂ capture rate and reboiler duty are considered as the output/controlled variables. Next, the open-loop simulation is performed by moving each input variable as a step change to the process. The obtained data, plotted in Figure 3, is used in the classical autoregressive model with exogenous inputs (ARX) method for deriving the simplified dynamic model. After the data processing, the dynamic characteristics of the PCC system are represented in a typical continuous-time state-space model format to be employed in the control strategy. The

second crucial step in this framework is the design of a controller based on previously developed BIO-CS that mimics ant's rule of pursuit phenomenon. In this case, BIO-CS casted as MPC configuration is used for implementation purposes due to its computation time advantage (Mirlekar et al. 2018b).

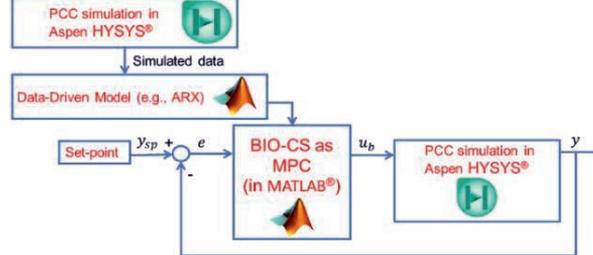


Figure 2: Schematic representation of the optimal control framework

In BIO-CS as MPC formulation, the optimal control problem associated with each agent is solved considering a fixed setpoint trajectory of an output variable for error minimization. *dynopt* (a freely available optimal control toolbox) is employed to solve the optimal control problems associated with BIO-CS as MPC and the optimal input trajectories are computed (MATLAB[®] software, *dynopt* library). Specifically, *dynopt* is a gradient-based solver that utilizes the orthogonal collocation on finite elements method in conjunction with the constrained nonlinear minimization function *fmincon* to determine optimal control trajectories.

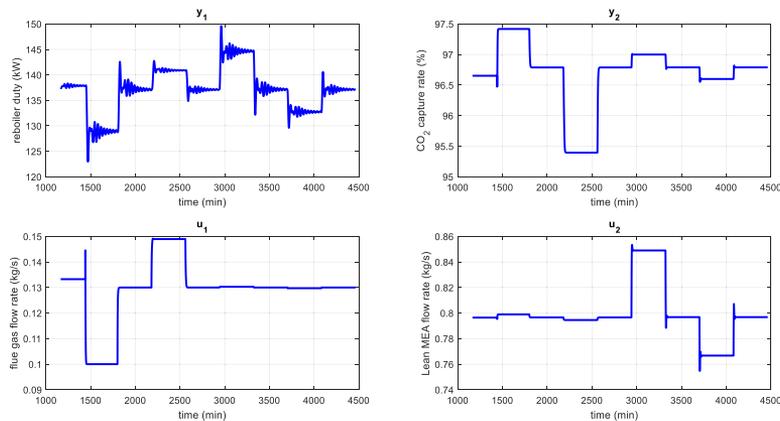


Figure 3: Open-loop simulation data used in the model development

The third important step deals with the implementation of control moves on the PCC simulation plant in an online manner. For this purpose, a freely available MATLAB[®] - Aspen HYSYS[®] link is employed for communicating the controller and the simulation plant model. As shown in Figure 2, the BIO-CS as MPC developed as a script can compute control trajectory (u_b) over a sample time horizon and call simulation for implementation using link. The script/function is paused until the implementation is being performed via simulation/integration for a specified time/horizon. Once this specified time is reached at a feedback sample time, the values of output variables (y) are sent back from simulation to controller and the loop is closed where controller subroutine is called for calculation of control trajectory (u_b) for the consecutive time horizon. The computational time for control moves calculation is about 2 seconds. In the next section, the analysis of the implementation results is presented.

4. Implementation Results

The PCC system model and the optimal control framework described in previous sections are employed for implementation purposes. Figure 4 shows the closed-loop simulation results of optimal control framework in comparison with standalone PID controller implementation for setpoint tracking case study. These results are assessed in terms of time to reach steady-state and the output setpoint tracking error. The case study mimics a real-world scenario in which variations in flue gas flow rate are introduced based on the ramp down operations of a power plant.

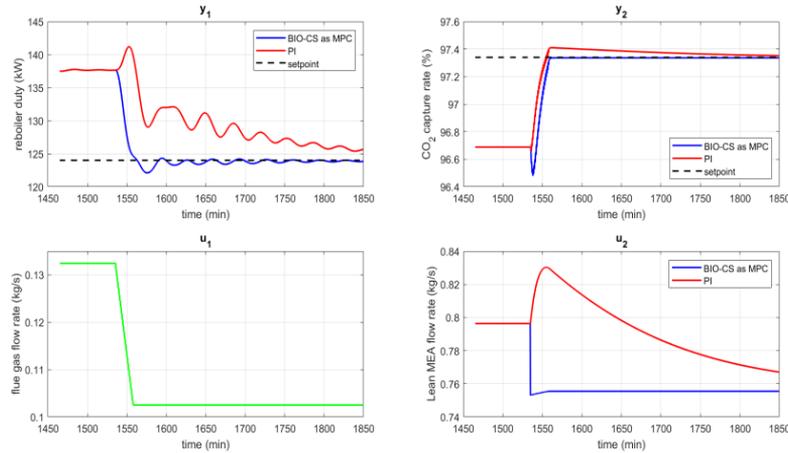


Figure 4: Closed-loop simulation results: setpoint tracking case study

For such simulation, a transient trajectory replicating a 10% ramp decrease in flue gas flow rate is applied over a period of 24 minutes followed by a constant flue gas flow rate profile for next five hours as an input to the process and represented by green solid line. Such dynamic operation is controlled by manipulating the valve opening that regulates the lean MEA flow rate for setpoint tracking scenario. For PI controller application, a fixed setpoint value of 97.34% is given for CO₂ capture rate with controller gains as $k_c = 0.1$ and $\tau_c = 0.25$. Note that the reboiler duty cannot be controlled by PI controller. The standard Aspen HYSYS[®] algorithm type is chosen for PI and tuning parameters are selected from the literature (Dutta et al. 2017b) and extensive simulation runs. In the case of BIO-CS as MPC implementation, prediction horizon/pursuit time (Δ) of 5 minutes and the feedback sample time of 1 minute is selected. A constant setpoint trajectory of 97.34% and 124kW shown by black dashed line is supplied for CO₂ capture rate and reboiler duty, respectively. PID controller is disabled, and the valve opening calculated by BIO-CS as MPC with the objective of minimizing the difference between the setpoints and the output variables is supplied to the system. The formulation of BIO-CS as MPC allows simultaneous control of both the outputs giving major advantage over standalone PI controller. As depicted in Figure 4, the corresponding output data associated with CO₂ capture rate and reboiler duty as well as the computed input trajectory of lean MEA flow rate is recorded during both control strategy implementation. During the implementation of the developed approach, the reboiler duty and CO₂ capture rate reach the desired setpoint in approximately 60 and 25 minutes from the time of introduction of flue gas flow rate ramp down profile, respectively. The settling time and the output tracking error with respect to the setpoint for both the outputs for PI execution are relatively large when compared with the BIO-CS as MPC performance. This translates into reasonably high

reboiler duty and higher energy demand to achieve required CO₂ capture during PI implementation. Thus, these results illustrate the benefits of employing the optimal control framework.

5. Conclusions

In this article, the implementation of an optimal control framework on PCC process simulation was demonstrated successfully. The goal of the control strategy design was to track the setpoints of the desired variables, in this case CO₂ capture rate and reboiler duty, during ramp decrease in the flue gas flow rate. Following objectives were accomplished in this paper: (i) A systematic design of the proposed methodology with BIO-CS as MPC at the core and steps that need to be followed during its implementation; (ii) address challenges associated with the interaction of multiple variables representing PCC control structure; (iii) communicate two software platforms and enhancing the capabilities of the process simulator. The results showed that the developed approach improves the time required to reach new setpoints for multiple outputs and the setpoint tracking error when compared with traditional control approaches. Therefore, such advancements increase energy efficiency and reduce GHG emissions in the decarbonized economy where integrations of CCS technologies into the thermal power plants is necessary. The studies associated with disturbance rejection cases and simulation of more challenging scenarios will be the subject of investigation in the future. The developed framework thus provides a promising methodology for advanced control of energy systems with CCS technologies addressing flexible operations.

6. Acknowledgement

This publication has been funded by the CleanExport project - Planning Clean Energy Export from Norway to Europe, 308811. The authors gratefully acknowledge the financial support from the Research Council of Norway and the user partners Agder Energi, Air Liquide, Equinor Energy, Gassco and Total E&P Norge.

References

- Aspen HYSYS[®] (Version 10) [Software]. <https://www.aspentech.com/en>.
- C. Figueres et al., Three years to safeguard our climate, *Nature* 546, no. 7660, 593-595, 2017.
- C. Skar et al., Norway's role as a flexibility provider in a renewable Europe, *FME CenSES*, 2018.
- Dynopt document and code library. <https://www.uiam.sk/~fekar/research/dynopt/dynopt.htm>.
- G. Mirlekar et al., A biologically-inspired approach for adaptive control of advanced energy systems, *Computers & Chemical Engineering* 117, 378-390, 2018a.
- G. Mirlekar et al., Biomimetic model-based advanced control strategy integrated with multi-agent optimization for nonlinear chemical processes, *Chemical Engineering Research & Design* 140, 229-240, 2018b.
- MATLAB[®] R2020b [Software]. <https://www.mathworks.com/products/matlab.html>.
- R. Dutta et al., Prospects of using equilibrium-based column models in dynamic process simulation of post-combustion CO₂ capture for coal-fired power plant, *Fuel* 202, 85-97, 2017a.
- R. Dutta et al., Selection and design of post-combustion CO₂ capture process for 600 MW natural gas fueled thermal power plant based on operability, *Energy* 121, 643-656, 2017b.
- R. Montanes et al., Experimental results of transient testing at the amine plant at Technology Centre Mongstad: Open-loop responses and performance of decentralized control structures for load changes, *International Journal of Greenhouse Gas Control* 73, 42-59, 2018.
- S. Hauger et al., Demonstration of non-linear model predictive control of post-combustion CO₂ capture processes, *Computers & Chemical Engineering* 123, 184-195, 2019.

Data Driven Process Monitoring, from Fault Detection and Diagnosis Points of View, in Industry 4.0 Context

Hamed Ardakani, Hadiseh Hemmati, Manuel Rodriguez Hernandez ^a

^a*Chemical Engineering Department, Universidad Politecnica de Madrid, Madrid 28040, Spain*

Abstract

Non-upgraded process monitoring approaches rationally and methodically cannot cope with the vast, diverse, and heterogeneous so-called "Big Data" produced rapidly due to Industry 4.0 applications. They require radical enhancements in the computational capacity, required CPU time, precision, and flexibility to emulate newly emerged technologies. The Industry 4.0 context, also known as networked information-based technology, is the most current level of industrial innovation that significantly improves the quality of processes and products.

This paper reviews the Fault Detection and Diagnosis (FDD) scheme as a requisite component of the process monitoring systems, considering Industry 4.0 requirements. Furthermore, to represent the adequate Industry 4.0 FDD framework, different applied technologies associated with Industry 4.0 are studied concisely.

Keywords: Industry 4.0, Process Monitoring, Big Data Analytic.

1. Introduction

The Industry 4.0 context outlines an evolution towards innovative products and the digital world. Through it, industrial automation technologies are at the beginning of a new modern period. With the intention of process monitoring, data with distinct and complex characteristics must be analyzed. Internet of Things (IoT) devices and sensors used in Industry 4.0 produce Big Data that may possess various forms of uncertainty (Hariri et al., 2019). Analyzing complex uncertain data in Industry 4.0 is a time-consuming and challenging task that necessitates the use of relevant Big Data Analytics (BDA) platforms to transform Big Data into smart data to derive hidden knowledge (Zhang et al., 2010).

Process monitoring of Industry 4.0 and large-scale systems can be tricky due to the many sensors, the high correlation among the measured variables, and the complex interaction between the faults and symptoms (Lau et al., 2013). Studies prove that applying regular process monitoring systems, such as the SCADA system, begets difficulties manipulating Industry 4.0 challenges (Tan et al., 2018). Typical process monitoring methods would produce tremendous computation pressure when faced with large-scale, complex data. They are ineffective in interpreting high-speed data flowing from various sources because of their congenital deficiencies (Gokalp et al., 2017).

In the era of Industry 4.0 and Big Data, data-driven approaches find a particularly proper context for expanding their application scopes and growing in diversity and importance (Reis and Gins, 2017). Data-driven process monitoring methods do not need the process

model and the associated expert knowledge. They have become increasingly popular in recent years, especially in complex industrial processes (Ge et al., 2013), and represent one of the most reliable areas in research and practice. However, they also have some crucial shortcomings: they depend on unbiased data, and training and testing samples should be drawn from the same distribution of data (Saufi et al., 2019).

The process monitoring challenges presented by Industry 4.0 technologies must be well answered to make them provide acceptable performance. In this paper, process monitoring requirements in the Industry 4.0 context, mainly from the data-driven FDD points of view, are studied. In this regard, fundamental components of the Industry 4.0 context are addressed, too.

2. Industry 4.0 and Associated Technologies

The transition from Industry 3.0 to Industry 4.0 requires inclusive analysis to develop a strategic plan (Zhou et al., 2016). Industry 4.0 adoption demands considerable capital investments; therefore, its characteristics for prospective transformation should be principally comprehended. Industry 4.0 is driven by several emerging concepts and technologies, including Cyber-Physical Systems (CPS), the IoT, data analytics, cloud systems, etc. For any process to be regarded as Industry 4.0, constant connectivity, human assistance, and decentralized decision-making are absolute necessities (Muhuri et al., 2019).

The introduction of CPS is one of the most remarkable changes and the core foundation of Industry 4.0. It is based on the idea of integrating the physical and virtual worlds. It removes the boundaries between these two worlds and promotes the interconnection of physical and cyber elements (Angelopoulos et al., 2020). “Things” in CPS can communicate with each other through an IoT communication platform to detect their environment, interpret available data, and act on the physical world. The IoT offers promising transformational solutions that enable the real-time interconnection of different sensors, actuators, machines, etc., safely (Xu et al., 2018). It is one of the most significant contributions to Industry 4.0 and intelligent manufacturing. The industrial IoT is a subset of the IoT and follows the exact core definition, but the “things” and goals of the industrial IoT are usually different (Thames and Schaefer, 2016).

One of the relevant concepts of Industry 4.0 is Big Data. The main characteristics of Big Data include volume, velocity, variety, veracity, and value (Yan et al., 2017). In general, Big Data refers to datasets that cannot be perceived, acquired, managed, and processed by conventional tools within a tolerable time because of their size, which usually reaches Petabyte (=1024 Terabyte) or Exabyte (=1024 Petabyte) (Nguyen et al., 2020). Since usual computers may not succeed in processing Big Data, performing the respective analysis with the cloud system would be simpler and more efficient. The main objective of cloud computing is to use enormous computing and storage resources under concentrated management to provide Big Data applications with rigorous computing capacity (Chen et al., 2014).

Cloud system developments aim to minimize the cost of capital investments in infrastructure set-up, reduce maintenance efforts, and achieve efficient management. Parallel processing, virtualized resources, and data service integration with scalable storage are all advantages of cloud system technologies. Cloud computing developments offer solutions for the storage and processing of Big Data. The role of cloud systems is crucial for Big Data as they can provide any infrastructure and a variety of tools on-demand. Given that the emergence of Big Data accelerates the growth of cloud computing, the reciprocal association between them is constructive (Nguyen et al., 2020).

3. Process Monitoring within Industry 4.0 Context

In the Industry 4.0 context, thousands of devices and sensors would be installed to collect and analyze data (Henao-Hernández et al., 2019). To extract essential information from Big Data obtained from a variety of sources, it must be synthesized into structured data, which removes constraints imposed by format, dimension, and other factors (Yan et al., 2017).

Applying a well-organized data processing framework to monitoring Big Data has become a considerable challenge. Therefore, new difficulties in understanding and managing Big Data and extracting underlying knowledge have emerged (Zhang et al., 2018). The consequential considerations are to verify that the data is of high quality and that the data sources are trustworthy. The first step to monitoring Big Data is determining where and how to store it. The following step, data processing, requires rapid data loading and querying. Additionally, efficiency and adaptivity in storage space utilization and dynamic workload patterns are expected. The next step is to analyze data that applies algorithms to interpret data and extract unknown valuable patterns, relationships, and information (Zhang et al., 2018).

3.1. Big Data Analytics

BDA refers to collecting, storing, processing, analyzing, and distributing the data and providing an integrated framework that supports decision-making. Time and infrastructure are dominant factors in BDA that would be met through cloud computing systems (Londhe and Prasada Rao, 2018). BDA tools are suitable for Industry 4.0 to ease cleaning, formatting, and transforming industrial data (Santos et al., 2017). However, there are various difficulties in applying BDA in Industry 4.0. The two primary ones would be to choose the most relevant data and to develop a model capable of exploring the underlying dependence structures (Zhu et al., 2018). Moreover, visualization of the results, heterogeneous data generated by different equipment, and incomplete data are problems for real-time BDA (Gokalp et al., 2017).

To monitor Big Data, FDD systems must possess the ability to convert problems to small ones to make better decisions and reduce costs (Hariri et al., 2019). Obviously, some primary strategies have been adopted to achieve this needed attribute. Applying several parallel algorithms, called parallelization, that work simultaneously, either doing the same tasks or different ones, and using incremental (gradual) learning algorithms that learn from data step-by-step are among the most common ones (Zimányi and Kutsche, 2015). Moreover, a hybrid of these two strategies could be a solution, too. Additional regular proposed methods in the literature are divide-and-conquer, instance selection, and granular computing (Hariri et al., 2019). This problem within the Industry 4.0 context is addressed through cloud system technology that significantly assists practical processing.

3.2. Fault Detection and Diagnosis

In general, none of the FDD methods has all the desirable features that a flawless technique must possess (Ardakani et al., 2017). Hybrid frameworks are recommended to cover the weaknesses of individual strategies and propose robust FDD systems. In a hybrid framework, methods are integrated while belonging to any type (Reis and Gins, 2017).

Fault detection and fault diagnosis are both significantly essential, but it should be noted that fault detection compared with fault diagnosis is a well-developed topic. Besides, the required time for fault detection would be trivial compared with the fault diagnosis, which may take days. Therefore, fault diagnosis needs more research concentration, while process monitoring systems commonly employ both simultaneously (Reis and Gins, 2017).

The most popular supervised data-driven FDD approaches are based on classification methods that would be considered a classification problem (Venkatasubramanian et al., 2003). Without requiring explicit mathematical models, classifiers could be trained based on pattern recognition principles on historical data, including information about normal and different faulty situations. By optimizing or adjusting the parameters, the learning process enables these classifiers to extract knowledge from data. Then, the trained classifiers can be used for process supervision to detect and diagnose faults in the process output measurements (Shokry et al., 2017).

The classifiers would be categorized into multiclass classifiers and One-Class Classifiers (OCC). Multiclass classifiers for training need at least two labeled classes of samples, while OCCs need only one class for training (Chiang et al., 2004). Depending on the FDD purposes, any combination of the classifiers may be applied. Multiclass classification of Big Data is a subject of broad interest in machine learning research. It is necessary to extract process conditions between different classes that may also be imbalanced, making it challenging to differentiate faults from normal data (Sleeman and Krawczyk, 2021). Because of the possible time constraints and sample availability, the FDD algorithms would be divided into four general groups: batch learning, incremental learning, online learning, and anytime learning (Zimányi and Kutsche, 2015). For smart FDD, real-time access to Big Data is necessary to identify faults in the shortest possible time. For monitoring Big Data, the speed with which the data is processed must meet the rate at which it is received, and this is one of the critical criteria for selecting the FDD framework (Khan et al., 2017).

3.3. Advanced Techniques

For employing FDD in the Industry 4.0 context, there is a limitation to using traditional data-driven methods, although they would be implemented in some circumstances (Yan et al., 2017). Therefore, several advanced machine learning techniques for monitoring Big Data such as Deep Learning (DL), Active Learning (AL), feature learning, distributed learning, and transfer learning are proposed. Among them, DL and AL have attracted more attention. AL, categorized into semi-supervised approaches, is applicable when data is abundant while related reliable labels are scarce or expensive. Learning in this manner is a time-consuming and not always straightforward task, but AL is designed to select critical samples for labeling to achieve high accuracy. In the literature, three main AL scenarios are proposed: membership query synthesis, stream-based selective sampling, and pool-based sampling (Qiu et al., 2016).

DL is a branch of machine learning that processes data through multiple non-linear processing layers. It is defined as the modeling of neural networks, while “deep” stands for adequate layers of representations (Kotsiopoulos et al., 2021). To construct a DL model, many parameters and hyper-parameters must be determined (Dekhtiar et al., 2018), but they are considered the most promising machine learning technique in many manufacturing fields nonetheless (Dogan and Birant, 2021). In recent years, the most common DL models used extensively in FDD systems include convolutional neural networks, restricted Boltzmann machines, deep belief networks, and deep neural networks (Angelopoulos et al., 2020). The ability to handle high-dimensional and multivariate data and discover formerly unknown knowledge makes DL techniques crucial in Industry 4.0 (Kotsiopoulos et al., 2021). In the Industry 4.0 context, DL models have been extensively implemented in FDD systems, although the high performance of DL comes with challenges and costs (Saufi et al., 2019). Training DL with Big Data is tricky since iterative calculations are often complicated to be parallelized. Therefore,

there is a tremendous surge of interest in applying parallel algorithms for training deep models (Chen and Lin, 2014).

4. Conclusions

Because of the subject's importance, there has been an increasing demand for research to provide insights into the concerns, difficulties, and solutions related to the monitoring, design, implementation, and management of Industry 4.0. Still, many impediments need to be resolved to launch Industry 4.0 applications into various industrial sectors, notwithstanding all the endeavors made up to now.

Industry 4.0 enhances process monitoring performance by quickly providing more information. In addition, process monitoring systems, particularly FDD frameworks, reinforce Industry 4.0 by efficiently extracting the required information from unprocessed data and transforming it into valuable knowledge.

The data processing in the Industry 4.0 FDD framework must be investigated profoundly since the rapidly available Big Data in Industry 4.0 is quite heterogeneous and may include vast amounts of uncertainty. Moreover, because of the volume of Big Data in the Industry 4.0 context, developed FDD models must be in accordance with the cloud system and its associated techniques to prevent computational burden and delay in detecting and diagnosing abnormal situations.

References

- A. Angelopoulos, E.T. Michailidis, N. Nomikos, P. Trakadas, A. Hatziefremidis, S. Voliotis, T. Zahariadis, 2020, Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects, *Sensors (Switzerland)*, 20, 1–33.
- M.H. Ardakani, M. Askarian, G. Escudero, M. Graells, A. Espuna, 2017, Toward Online Explore of Concept Drift for Fault Detection of Chemical Processes, *Computer Aided Chemical Engineering*, 40, 1657–1662.
- M. Chen, S. Mao, Y. Liu, 2014, Big data: A survey, *Mob. Networks Appl*, 19, 171–209.
- X.W. Chen, X. Lin, 2014, Big data deep learning: Challenges and perspectives, *IEEE Access* 2, 514–525.
- L.H. Chiang, M.E. Kotanchek, A.K. Kordon, 2004, Fault diagnosis based on Fisher discriminant analysis and support vector machines, *Comput. Chem. Eng*, 28, 1389–1401.
- J. Dekhtiar, A. Durupt, M. Bricogne, B. Eynard, H. Rowson, D. Kiritsis, 2018, Deep learning for big data applications in CAD and PLM – Research review, opportunities and case study, *Comput. Ind*, 100, 227–243.
- A. Dogan, D. Birant, 2021, Machine learning and data mining in manufacturing, *Expert Syst. Appl*, 166, 114060.
- Z. Ge, Z. Song, F. Gao, 2013, Review of Recent Research on Data-Based Process Monitoring, *Am. Chem. Soc*, 52, 3543–3262.
- M.O. Gokalp, K. Kayabay, M.A. Akyol, P.E. Eren, A. Kocyigit, 2017, Big data for Industry 4.0: A conceptual framework, *Proc. - 2016 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2016*, 431–434.
- R.H. Hariri, E.M. Fredericks, K.M. Bowers, 2019, Uncertainty in big data analytics: survey opportunities and challenges, *J. Big Data*, 6.
- I. Henao-Hernández, E.L. Solano-Charris, A. Muñoz-Villamizar, J. Santos, R. Henríquez-Machado, 2019, Control and monitoring for sustainable manufacturing in the Industry 4.0: A literature review, *IFAC-PapersOnLine* 52, 195–200.
- M. Khan, X. Wu, X. Xu, W. Dou, 2017, Big data challenges and opportunities in the hype of Industry 4.0, *IEEE Int. Conf. Commun*.
- T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, D. Tzovaras, 2021, Machine Learning and Deep Learning in smart manufacturing: The Smart Grid paradigm, *Comput. Sci. Rev*, 40.

- C.K. Lau, K. Ghosh, M. a. Hussain, C.R. Che Hassan, 2013, Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS, *Chemom. Intell. Lab. Syst.* 120, 1–14.
- A. Londhe, P.V.R.D. Prasada Rao, 2018. Platforms for big data analytics: Trend towards hybrid era, 2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput. ICECDS 2017, 3235–3238.
- P.K. Muhuri, A.K. Shukla, A. Abraham, 2019, Industry 4.0: A bibliometric analysis and detailed overview, *Eng. Appl. Artif. Intell.* 78, 218–235.
- T. Nguyen, R.G. Gosine, P. Warrian, 2020, A Systematic Review of Big Data Analytics for Oil and Gas Industry 4.0, *IEEE Access* 8, 61183–61201.
- J. Qiu, Q. Wu, G. Ding, Y. Xu, S. Feng, 2016, A survey of machine learning for big data processing, *EURASIP J. Adv. Signal Process.* 2016.
- M.S. Reis, G. Gins, 2017, Industrial process monitoring in the big data/industry 4.0 era: From detection, to diagnosis, to prognosis, *Processes*, 5, 3.
- M.Y. Santos, J. Oliveira e Sá, C. Costa, J. Galvão, C. Andrade, B. Martinho, F.V. Lima, E. Costa, 2017. A big data analytics architecture for industry 4.0, *Adv. Intell. Syst. Comput.* 570, 175–184.
- S.R. Saufi, Z.A. Bin Ahmad, M.S. Leong, M.H. Lim, 2019, Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review, *IEEE Access* 7, 122644–122662.
- A. Shokry, M.H. Ardakani, G. Escudero, M. Graells, A. Espuña, 2017, Dynamic kriging based fault detection and diagnosis approach for nonlinear noisy dynamic processes. *Comput. Chem. Eng.*, 106, 758-776
- W.C. Sleeman IV, B. Krawczyk, 2021, Multi-class imbalanced big data classification on Spark, *Knowledge-Based Syst.* 212, 106598.
- D.P. Tan, L. Li, Y.L. Zhu, S. Zheng, H.J. Ruan, X.Y. Jiang, 2018, An Embedded Cloud Database Service Method for Distributed Industry Monitoring, *IEEE Trans. Ind. Informatics* 14, 2881–2893.
- L. Thames, D.Schaefer, 2016, Software-defined Cloud Manufacturing for Industry 4.0, *Procedia CIRP* 52, 12–17.
- V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, K. Yin, 2003, A review of process fault detection and diagnosis Part III: Process history based methods, *Comput. Chem. Eng.* 27, 327–346.
- L. Da Xu, E.L. Xu, L. Li, 2018, Industry 4.0: State of the art and future trends, *Int. J. Prod. Res.* 56, 2941–2962.
- J. Yan, Y. Meng, L. Lu, L. Li, 2017, Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. *IEEE Access* 5, 23484–23491.
- P. Zhang, F. Xiong, J. Gao, J.Wang, 2018, Data quality in big data processing: Issues, solutions and open problems, 2017 IEEE SmartWorld Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Cloud Big Data Comput. Internet People Smart City Innov, 1–7.
- Y. Zhang, Y. Zhang, 2010, Fault detection of non-Gaussian processes based on modified independent component analysis, *Chem. Eng. Sci.*, 65, 4630–4639.
- K. Zhou, T. Liu, L. Zhou, 2016, Industry 4.0: Towards future industrial opportunities and challenges, 12th Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD 2015, 2147–2152.
- J. Zhu, Z. Ge, Z. Song, F. Gao, 2018, Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data, *Annu. Rev. Control* 46, 107–133.
- E. Zimányi, R.D. Kutsche, 2015, Business Intelligence: 4th european summer school, eBISS 2014 Berlin, Germany, July 6-11, 2014 tutorial lectures, *Lecture Notes in Business Information Processing*.

Safe Chance Constrained Reinforcement Learning for Batch Process Optimization and Control

Max Mowbray^a, Panagiotis Petsagkourakis^b, Antonio Del Rio Chanona^b and Dongda Zhang^a

^a*Department of Chemical Engineering and Analytical Science, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

^b*Centre for Process Systems Engineering (CPSE), Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, UK.
dongda.zhang@manchester.ac.uk*

Abstract

Reinforcement Learning (RL) has received interest within the context of decision making under uncertainty in the process industries. The primary benefit of RL arises from the formulation of the control problem as a Markov decision process (MDP), meaning that it inherits the benefits of accounting for uncertainty in a closed loop feedback control framework and models dynamics very generally via conditional probability density functions. This enables RL to handle problems with various types of exogenous and endogenous uncertainties. Despite this there has been little reported uptake of RL in the process industries. This is partly due to the inability to provide optimality guarantees under the model used for learning, but more importantly due to safety concerns. This has led to the development of RL algorithms in the context of ‘Safe RL’. Here, we present an algorithm that leverages the variance prediction of Gaussian process state space models to a) handle operational constraints and b) account for mismatch between the offline process model and the real online process. The algorithm is then benchmarked on an uncertain Lutein photo-production process against nonlinear model predictive control (NMPC) and several state-of-the-art Safe RL algorithms. Through definition of key performance indicators, we demonstrate the efficacy of the method with respect to objective performance and probabilistic constraint satisfaction.

Keywords: Safe Reinforcement Learning; Optimal Control; Dynamic Optimization; Bioprocess Operation; Machine Learning

1. Introduction

The operation of nonlinear, uncertain batch processes is a well established research focus within the academic community. Both modelling and control often poses challenges. Recent focus in the domain of control (and online optimization) has considered how best to integrate data accrued from process operation to inform decision-making and account for model uncertainties. Historically, consideration of model uncertainty in the context of optimization has been handled via stochastic and robust variants of model predictive control (MPC) (Heirung et al. (2018)). However, batch processes are often characterised by multiple, nonlinear dynamical regimes, which makes model construction notoriously difficult. This has somewhat limited the efficacy of general model predictive control in application to these processes, as they are reliant upon accurate finite dimensional (and closed form) descriptions of the true system dynamics, which has led to the rise of statistical process control approaches within industry (see Yoo et al. (2021)).

Recent efforts to remedy the apparent challenges of batch processing have manifested in the development of various scenario (Maiworm et al. (2015)), distributionally robust (Zhong et al. (2021)) and learning-based MPC approaches (Hewing et al. (2020)). Another approach is provided by Reinforcement Learning (RL). Strictly speaking RL does not require a model explicitly, lending to interpretation as model-free decision making algorithms. However, the data intensity and requirement for exploration of different control decisions necessitates offline model simulation to pre-learn a control policy (via RL) before deployment to the real process. As a result of the presence of process model-mismatch between the real process and the model used for simulation, there is a potential for one to observe distributional shift when the policy is deployed to the real process. In such a case, RL would be making extrapolative control decisions via a data-driven model and may drive the process in to operationally undesirable regimes. Further to this, process operational constraints are not explicitly handled within the Markov decision process framework. Hence, for RL to be applied safely to the process industries, development of training approaches and frameworks are required. In this work, we present an algorithm that leverages the posterior predictive distribution of a Gaussian process simulation model to jointly handle process model-mismatch and joint chance constraints. The ideas we present in this work combine the concepts of pessimism and constraint tightening, which are common to the fields of batch (or offline) RL and stochastic MPC, respectively.

2. Methodology

2.1. Problem Statement

In this work we assume that the system concerned is Markovian and expresses uncertain process dynamics, such that discrete time process evolution may be described as follows:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{s}_t) \quad (1)$$

where $\mathbf{x} \in \mathbb{X} \subseteq \mathbb{R}^{n_x}$ are states; $\mathbf{u} \in \mathbb{U} \subseteq \mathbb{R}^{n_u}$ are control inputs from a given control set; $\mathbf{s} \in \mathbb{S} \subseteq \mathbb{R}^{n_s}$ are realisations of process uncertainty termed generally to describe various sources of uncertainty; and, $t \in \{0, \dots, T\}$ is a discrete time index within a discrete finite time horizon. We would like to solve the following chance constrained problem:

$$\mathcal{P}(\pi_C) := \begin{cases} \max_{\pi} J \\ \text{s.t.} \\ X_0 \sim p(\mathbf{x}_0) \\ \mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{s}_t) \\ \mathbf{u}_t \sim \pi(\mathbf{u}_t | \mathbf{x}_t) \\ \mathbf{u}_t \in \mathbb{U}; \mathbf{s}_t \in \mathbb{S} \\ \mathbb{P}(\cap_{i=0}^T \{\mathbf{x}_i \in \mathbb{X}_i\}) \geq 1 - \alpha \\ \forall t \in \{0, \dots, T\} \end{cases} \quad (2)$$

where $\pi_c(\mathbf{u}_t | \mathbf{x}_t)$ defines a conditional probability density function, that provides a distribution over controls given observation of state; $\mathbb{X}_t = \{\mathbf{x}_t \in \mathbb{R}^{n_x} : A_j^T \mathbf{x}_t - b_j \leq 0, \forall j \in \{1, \dots, n_g\}\}$ is the set of states that satisfy an affine ($A_j \in \mathbb{R}^{n_x}$ and $b_j \in \mathbb{R}$) constraint set at a given time index; $\alpha \in (0, 1]$ is the probability allowed for violation of the constraint set for all time indices; $p(\mathbf{x}_0)$ defines the initial state distribution, which is treated a random variable, X_0 ; and, $J = \mathbb{E}_{\pi}[\sum_{t=0}^{T-1} R_{t+1}]$, is the process objective, which is equivalent to the expected sum of rewards, $R_{t+1} \in \mathbb{R}$, over the discrete time horizon. The reward is provided by a function, $R : \mathbb{X} \times \mathbb{U} \times \mathbb{S} \rightarrow \mathbb{R}$, that ranks process evolution with respect to control objectives. Due to the presence of joint chance constraints and uncertain process evolution (neither of which can be expressed via finite dimensional or closed

form expressions), one is typically required to make approximations to Eq. 2 (to obtain finite dimensional and closed form surrogate expressions) (Heirung et al. (2018)). Despite the problem formulated here, RL is generally only able to find a policy which optimizes a given objective, \hat{J} . Specifically, in the paradigm of policy optimization, parameterizations of a policy are learned, such that:

$$\theta^* = \arg \max_{\theta} \hat{J}(\theta, \cdot) \quad (3)$$

where $\theta \in \mathbb{R}^{n_{\theta}}$ define the parameters of the policy. As a result, constraints are handled via a penalty function. When the constraints are subject to realizations of uncertain variables, $\mathbf{s} \in \mathbb{S}$, the specification of the penalty function is more involved. Recently, the concept of constraint tightening has been translated from the domain of stochastic MPC to RL (see. e.g. Petsagkourakis et al. (2020)). In the following we outline an approach that leverages the posterior predictive distribution of Gaussian process (GP) models to jointly identify a constraint tightening mechanism and to handle process model mismatch.

2.2. Gaussian Process State Space Models

GPs operate within a nonparametric, Bayesian inference framework. This allows us to exploit statistical relationships we assume exist within the data to identify a function, $f: \mathbb{R}^{n_z} \rightarrow \mathbb{R}$. One can make predictions, $\mathbf{f} = [f(\mathbf{z}_1), \dots, f(\mathbf{z}_N)]$, where $f(\mathbf{z}_i) \in \mathbb{R}$, by simply querying the function at given model inputs. In order to proceed we assume the availability of a dataset $\mathcal{D} = \{Z, Y\}$, expressive of discrete time process evolution, where $Z = [\mathbf{z}_1, \dots, \mathbf{z}_N]^T$, such that $\mathbf{z}_i = [\mathbf{x}_i^T, \mathbf{u}_i^T]$ and $Y = [\mathbf{y}_1, \dots, \mathbf{y}_N]^T$, $\mathbf{y} \in \mathbb{R}^{n_y}$, represents measurements of the system state. A GP is fully specified by a mean, $\mathbf{m}(\cdot)$, and covariance function, $k(\cdot, \cdot; \lambda)$, which is parameterised by some hyperparameters λ . Selection of the mean, covariance function, and the associated hyperparameters, defines the *prior*. Inference in GPs utilizes a Bayesian framework. Therefore, a variant of Bayes' rule is used to predict the discrete time evolution of a specific component of the system state. Given the training data and modeling assumptions regarding the mean and covariance functions, this allows us to identify a *posterior* predictive distribution, $p(\mathbf{f}|Y^j, Z, \lambda)$, descriptive of discrete time process evolution. Here, $Y^j \in \mathbb{R}^N$, denotes the j^{th} column of Y . Identification of λ is achieved by maximisation of the marginal log-likelihood. If a Gaussian likelihood is chosen, the *posterior* is constructed exactly as a Gaussian distribution. At a new test point \mathbf{z}^* , a posterior predictive distribution over function values, f^* , may be identified as:

$$p(f^*|Y^j, Z, \mathbf{z}^*, \lambda) = \mathcal{N}(\mu(\mathbf{z}^*; Y^j, Z, \lambda), \sigma(\mathbf{z}^*; Y^j, Z, \lambda)) \quad (4)$$

where μ and σ have closed form expression. As GPs are multiple-input, single-output models, if one has an n_x systems states to dynamically model, then n_x separate GP models can be constructed independently and then assembled to form a state space model. It is worth emphasising that this state space model represents a probabilistic forecast of discrete time process evolution i.e. is an approximation of Eq. 1.

2.3. Integration of Gaussian Process State Space Models for Safe Reinforcement Learning

In the following, we show how to use GP state space models to identify closed form expressions for the probabilistic constraints via the concept of constraint tightening. Specifically, we leverage Boole's inequality which enables us to decompose the probability of constraint violation α across the n_g constraints, such that the probability of violating constraint j , may be defined as $\iota_j = \alpha/n_g$.

Having decomposed the joint chance constraints into individual chance constraints, we then deploy the Cantelli-Chebyshev inequality to obtain robust and closed form surrogate expressions for each chance constraint. The idea here is to identify an approximate tightened constraint set $\hat{X}_t = \{\mathbf{x}_t \in$

$\mathbb{R}^{n_x}: A_j^T \bar{\mathbf{x}}_t + \boldsymbol{\varepsilon}_{j,t} - b_j \leq 0, \forall j \in \{1, \dots, n_g\}$, where $\boldsymbol{\varepsilon}_{j,t} = f(\Sigma[\mathbf{x}_t], \iota_j)$ is a constraint and time index specific backoff; $\Sigma[\mathbf{x}_t]$ is the finite variance of the state; and, $\bar{\mathbf{x}}_t$ is the nominal (mean) state - all of which can be estimated by the posterior distribution of the GP state space model. The relevance of the Cantelli-Chebyshev inequality here is that it provides means to identify a backoff in closed form (i.e. $f(\Sigma[\mathbf{x}_t], \iota_j)$). Satisfaction of this tightened constraint set can then be handled by optimization of an l_p norm penalty function, for example see Mowbray et al. (2021). Having identified means to handle the joint chance constraints, we deploy the concept of pessimism as is common to the domain of batch (or offline) RL. Specifically, we wish to reduce exploitation of regions of the GP state space model, where the predicted variance (which expresses both aleatoric and epistemic uncertainties) is high - this is achieved by adding a penalty in the objective function. In doing so, it is expected that the effects of distributional shift when the policy is transferred to the real system will be minimised. In summary, the analysis provided here enables the identification of a closed form function, $\varphi: \mathbb{X} \times \mathbb{U} \times \mathbb{X} \rightarrow \mathbb{R}$, which includes penalty for constraint violation and for high uncertainty regions of the GP state space model. Policy optimization of φ then handles both the joint chance constraints and plant-model mismatch:

$$\theta^* = \operatorname{argmax}_{\theta} \mathbb{E} \left[\sum_{t=0}^{T-1} \varphi(\bar{\mathbf{x}}_{t+1}, \Sigma[\mathbf{x}_{t+1}], \mathbf{u}_t) \right], \quad (5)$$

In practice, the Cantelli-Chebyshev backoff, $\boldsymbol{\varepsilon}_{j,t}$ often proves conservative. To compromise between constraint satisfaction and operative performance, we plan to tune $\boldsymbol{\varepsilon}_t = [\varepsilon_{1,t}, \dots, \varepsilon_{n_g,t}]$, $\forall t$ via a set of multipliers $\boldsymbol{\xi} = [\xi_1, \dots, \xi_{n_g}]$, which are identified by solving an upper-level Bayesian optimization problem. For specific information on the Bayesian optimization scheme used, see Mowbray et al. (2021). A bilevel optimization problem is formed as a result.

3. Case Study

3.1. A Fed-Batch Lutein Photo-production process

The work developed in this paper focuses on online optimization of fed-batch bioprocesses. The performance of the proposed soft-sensors is assessed in case study on a Lutein photo-production process, which was first detailed in del Rio-Chanona et al. (2017). We assume the same model structure (which is nonlinear) and kinetic parameters as detailed in this study. The model itself describes the evolution of biomass, nitrate and lutein concentration, with lutein production strongly controlled by the complexities of cell metabolism, which are dependent on nitrate availability and incident light intensity. We additionally introduce uncertainty in the form of 5% parametric uncertainty. A set of three affine constraints were imposed on the problem, all of which reflect real operational concerns, such as the preservation of biomass productivity and limits on biomass concentration for the purposes of downstream separation. The objective of process operation is to maximise productivity of lutein production, which enables us to define J so as to reward lutein concentration and penalise nitrate concentration at the end of the batch. For full details of the case study please refer to Mowbray et al. (2021).

3.2. Case Study Design and Benchmarks

The study itself is purely computational. Because of this, we leverage the availability of the uncertain, mechanistic model (that is equivalent to Eq. 1) and conceptualise that it represents the real system. The relevant procedures to demonstrate the methodology proposed then follow: a) generate a dataset (that in practice could be available from e.g. design of experiments) by sampling the uncertain mechanistic model with space filling control trajectories; b) deploy the methodology described by building a GP state space model with the dataset and identify a policy through the framework detailed by Figure 1; c) deploy the policy identified to optimize the real uncertain process (model); and, d) benchmark the performance against NMPC, first order constrained

optimisation in policy space (FOCOPS) (Zhang et al. (2020)), the model-based offline policy optimization (MOPO) algorithm (Yu et al. (2020)), and the conservative offline model based policy optimization (COMBO) algorithm (Yu et al. (2021)). The first two benchmarks represent the best case deterministic method, and a state-of-the-art constrained RL method. The latter two are offline (batch) RL methods and designed to account for mismatch. In all RL benchmarks, constraints are handled by incorporating deterministic expressions for the hard constraint sets into a penalty function within the GP state space model. In FOCOPS, the mechanism for pessimism was implemented into the objective function to mitigate process model-mismatch. The evaluation and comparison of the different algorithms' performance was based on metrics that quantify the quality of the policy identified with respect to two main operational objectives: the expected process objective, J and the probability of constraint satisfaction, $F = \mathbb{P}(\bigcap_{t=0}^T \mathbf{x}_t \in \mathbb{X}_t)$. In the algorithm proposed one can choose to set a lower bound on $F \geq 1 - \alpha$. Clearly, closed form expression is unavailable so F is instead approximated via the sample approximate, F_{SA} , and a statistically robust metric, F_{LB} , which accounts for the limitation of finite samples. All RL policies were trained via the GP state space model and then validated on the uncertain process model, via Monte Carlo (MC) simulation.

4. Results and Discussion

All algorithms observed extensive hyperparameter tuning to ensure objective comparison could be made. All tuning was conducted via Bayesian optimization via the same scheme and objective used to tune the multipliers $\xi \in \mathbb{R}^{n_g}$ proposed in Section 2.3; also, the probability of constraint violation was set $\alpha = 0.01$. Table 1 details the result. The results demonstrate the ability of the

Table 1: Results of policy transfer to the real uncertain process (i.e. the mechanistic model). SCCPO indicates the method proposed.

KPI	SCCPO	NMPC	FOCOPS	MOPO	COMBO
J	13.65 +/- 0.082	11.58 +/- 4.07	12.31 +/- 0.092	10.56 +/- 0.065	13.24 +/- 0.082
F_{SA}	1.0	0.12	1.0	1.0	1.0
F_{LB}	1.0	0.148	1.0	1.0	1.0

method proposed to handle both constraints and mismatch. This is especially reinforced by the relative performance to NMPC, where the model is exactly the same as the uncertain real process and the only difference that exists is the presence of parametric uncertainty in the real process. All RL benchmark methods (MOPO, COMBO, FOCOPS) handle constraints with the desired probability. It is thought that this arises due to the implementation of a backoff as introduced through a) the pessimism mechanism present in MOPO and FOCOPS, b) the nature of the conservative mechanisms that underpin COMBO as well as c) the utility of hyperparameter tuning utilised in the experiment. However, it should be noted that the action of these mechanisms is not specific to any given constraint, which may go some way to explaining the conservative control performance. This is reinforced by Fig. 1, which demonstrates the policies learned on the real process over 500 MCs. The figure shows the RL methods are able to well handle the process uncertainty (with all policies showing stable, low variance control profiles). This jointly ensures sufficient nitrate conditions, whilst minimising waste. The NMPC scheme shows high variance in the controls, which arises from an inability to account for mismatch in constraint satisfaction. This leads to high variance in the productivity of the batch. The SCCPO algorithm gains performance benefits because it is able to incorporate uncertainty information into constraint satisfaction, and observes a less conservative policy, e.g. it has the lowest feasible nitrate concentration at batch termination.

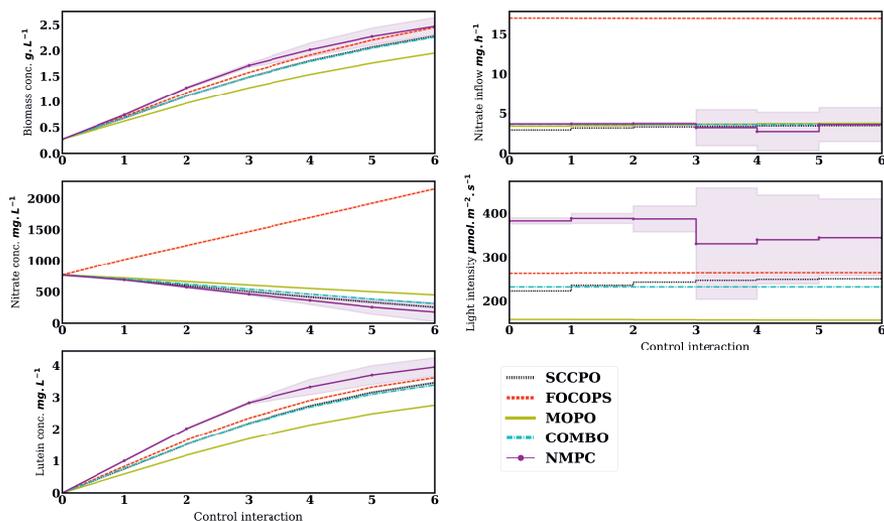


Figure 1: Final learned policies evaluated over 500 MCs on the *real* process. Plots on the left: show the state distributions (± 1 standard deviation) induced by the policies; the right shows the distribution of controls.

5. Conclusions

In this work, we have presented an algorithm that handles both operational constraints and process model-mismatch for the deployment of RL policies for the online optimization of uncertain, nonlinear fed-batch process systems. The algorithm has been benchmarked against best case deterministic method in the form of NMPC (observing a performance benefit of 13%) and state-of-the-art safe and offline RL methods (observing a performance benefit of 4%). The performance of the approach was demonstrated to be advantageous relative to the benchmarks proposed.

References

- E. A. del Rio-Chanona, N. r. Ahmed, D. Zhang, Y. Lu, K. Jing, 2017. Kinetic modeling and process analysis for *desmodesmus sp. lutein* photo-production. *AIChE Journal* 63 (7), 2546–2554.
- T. A. N. Heirung, J. A. Paulson, J. O’Leary, A. Mesbah, 2018. Stochastic model predictive control—how does it work? *Computers & Chemical Engineering* 114, 158–170.
- L. Hewing, K. P. Wabersich, M. Menner, M. N. Zeilinger, 2020. Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems* 3, 269–296.
- M. Maiworm, T. Bätghe, R. Findeisen, 2015. Scenario-based model predictive control: Recursive feasibility and stability. *IFAC-PapersOnLine* 48 (8), 50–56.
- M. Mowbray, P. Petsagkourakis, E. A. del Río Chanona, R. Smith, D. Zhang, 2021. Safe chance constrained reinforcement learning for batch process control. *arXiv preprint arXiv:2104.11706*.
- P. Petsagkourakis, I. O. Sandoval, E. Bradford, F. Galvanin, D. Zhang, E. A. del Rio-Chanona, 2020. Chance constrained policy optimization for process control and optimization. *arXiv preprint arXiv:2008.00030*.
- H. Yoo, H. E. Byun, D. Han, J. H. Lee, 2021. Reinforcement learning for batch process control: Review and perspectives. *Annual Reviews in Control*.
- T. Yu, A. Kumar, R. Rafailov, A. Rajeswaran, S. Levine, C. Finn, 2021. Combo: Conservative offline model-based policy optimization. *arXiv preprint arXiv:2102.08363*.
- T. Yu, G. Thomas, L. Yu, S. Ermon, J. Zou, S. Levine, C. Finn, T. Ma, 2020. Mopo: Model-based offline policy optimization. *arxiv preprint arxiv:2005.13239*.
- Y. Zhang, Q. Vuong, K. W. Ross, 2020. First order constrained optimization in policy space. *arXiv preprint arXiv:2002.06506*.
- Z. Zhong, E. A. del Rio-Chanona, P. Petsagkourakis, 2021. Data-driven distributionally robust mpc using the wasserstein metric. *arXiv preprint arXiv:2105.08414*.

Support Vector Machine-based Design of Multi-model Inferential Sensors

Martin Mojto^{a,*}, Karol Ľubušíky^b, Miroslav Fikar^a and Radoslav Paulen^a

^a*Slovak University of Technology in Bratislava, Bratislava 81237, Slovakia*

^b*Slovnaft, a.s., Bratislava 82412, Slovakia*

martin.mojto@stuba.sk

Abstract

There is an increasing need to monitor industrial key variables by inferential (soft) sensors. This contribution deals with the challenge of increasing the accuracy of inferential sensors yet maintaining the simple (linear) structure. In order to fulfill these opposing requirements, we design a linear multi-model inferential sensor (MIS) that switches between two models. We enhance the design of a sensor by continuous switching and optimized data labeling. The case study deals with a non-linear model of pressure-compensated temperature often used in distillation columns monitoring and control. The results show a significant accuracy improvement of MIS over a single-model sensor. The studied MIS design approaches present a great potential for practical use.

Keywords: Inferential Sensor, Monitoring, Data-based Design, Multi-model Prediction

1. Introduction

The inferential (soft) sensors (IS) find applications in all industrial and engineering fields. They provide a frequent and accurate estimation of key process variables, and therefore play a significant role in industrial monitoring (Li et al., 2021; Mojto et al., 2021; Qi et al., 2021). The principle of IS is to use the measurements from the easy-to-measure variables (e.g., temperatures, pressures) to estimate hard-to-measure variables (e.g., concentration) (Qin et al., 1997; Zhu et al., 2020).

Industrial processes usually exhibit nonlinear behavior. It significantly reduces the estimation accuracy of the simple (linear) inferential sensors. The nonlinear behavior can be compensated by the design of more complex (nonlinear) inferential sensors. Supposing the linear model structure should be maintained, it is possible to reduce the nonlinear aspect of the industrial processes by designing a so-called multi-model inferential sensor (MIS) (Khatibisepehr et al., 2012).

The standard approach for MIS design does not guarantee the continuity of the model. This can lead to significant stability issues of the (control) system that uses MIS. The further drawback is the a priori labeling of the training set without considering its impact on the MIS accuracy.

This paper investigates improvements of the standard MIS design. We enhance the approach by a continuous switching logic based on the support vector machines (SVM) methodology (Boser et al., 1992). Subsequently, this approach is extended by the optimized data labeling. We compare the effectiveness of these approaches by analyzing the performance of the single-model inferential sensor (SIS) and MIS for the nonlinear model of pressure-compensated temperature *PCT*. This

contribution deals with a special case of the sensor structure with two linear models. A possible extension with multiple models is discussed in Mojto et al. (2022).

2. Problem Statement

We aim to find a multi-model inferential sensor (MIS) with two models of the following form:

$$\hat{y} = \begin{cases} p_1^\top x + b_{p,1}, & \text{if } x \in \mathcal{R}_1, \\ p_2^\top x + b_{p,2}, & \text{if } x \in \mathcal{R}_2, \end{cases} \quad (1)$$

where $\hat{y} \in \mathbb{R}$ stands for the inferred (desired) variable, $x \in \mathbb{R}^{n_p}$ is a vector of the input data, n_p is a number of input variables, $p_i \in \mathbb{R}^{n_p}$ represents a vector of the sensor parameters in region \mathcal{R}_i and $b_{p,i}$ is a constant sensor off-set. Regions of individual model validity denoted as \mathcal{R}_i represent convex polyhedra such that $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$.

2.1. MIS-std: A Standard Approach to the Design of Multi-model Inferential Sensor (MIS)

The standard approach of the MIS design (MIS-std) involves three sequential steps:

1. A priori labeling of the training dataset. We use k -means clustering (Forgy, 1965).
2. Data classification for switching-logic design. The considered approach is support vector machines (SVM) with linear separators (Boser et al., 1992).
3. Individual sensor training. The parameters of the linear models within the MIS structure are fitted by using standard least-squares regression.

The main limitations of MIS-std are: (a) the designed models of MIS are not necessarily continuous, (b) a priori labeling of the training dataset is unaware of its impact on the MIS accuracy.

3. Design of MIS with Advanced Approaches

3.1. MIS-con: An Approach to the Design of MIS with Continuous Switching

To deal with the first limitation of MIS-std, a novel approach (MIS-con) is developed. It combines the SVM-based data classification with the individual sensor training in the optimization problem:

$$\min_{\substack{w, b_w, e \geq 0 \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^n z_i (y_i - p_1^\top x_i - b_{p,1})^2 + (1 - z_i) (y_i - p_2^\top x_i - b_{p,2})^2 + \alpha \|w\|_2^2 + \beta \|e\|_1 \quad (2a)$$

$$\text{s.t.} \quad (2z_i - 1)(w^\top x_i + b_w) \geq 1 - e_i, \quad \forall i \in \{1, 2, \dots, n\}, \quad (2b)$$

$$p_1 - p_2 - w = 0, \quad b_{p,1} - b_{p,2} - b_w = 0, \quad (2c)$$

with the number of measurements n , a normal vector and a constant off-set of the separation hyperplane, respectively, w and b_w , a vector of slack variables e , and a vector of binary parameters z that results from the data labeling procedure with $z_i = 1$ if $x_i \in \mathcal{R}_1$ and $z_i = 0$ if $x_i \in \mathcal{R}_2$. The constraints (2c) ensure continuity at the switch between the two models. This is achieved by establishing the intersection of model surfaces to coincide with the determined switching hyperplane.

As the a priori data labeling can be inappropriate for the design of a MIS with continuous switching, we allow small violations of the labeling using the slack variables e in (2b). We also consider that the user can aim at giving up some portion of model (training) accuracy for the better separation by widening the separation band. The latter feature is established by minimizing $\|w\|_2^2$ in (2a). The described features can be enforced/weakened by tuning of the positive weights α and β .

3.2. MIS-con-lab: An Approach to the Design of MIS with Optimized Data Labeling

We propose the approach MIS-con-lab to mitigate the inaccuracies caused by the a priori labeling of the training dataset. This approach searches directly for the optimal data labeling by adding z among the optimized variables in (2a). The underlying optimization problems reads as:

$$\min_{\substack{z \in \{0,1\}^n, w, b_w, e \geq 0 \\ p_1, b_{p,1}, p_2, b_{p,2}}} \sum_{i=1}^n z_i |y_i - p_1^T x_i - b_{p,1}| + (1 - z_i) |y_i - p_2^T x_i - b_{p,2}|, \text{ s.t. Eqs. (2b)–(2c)}. \quad (3)$$

We adopt the sum of absolute errors criterion in (3) to reduce the complexity of the optimization problem as this can now be transformed to mixed-integer linear program (MILP). If the sum of squared errors is used in the objective function, the optimization problem turns into mixed-integer nonlinear program (MINLP) that might be challenging especially when n is high. Moreover, the sum of squared errors criterion tends to reduce significant deviations, whereas the small deviations are neglected. This can reduce the MIS accuracy in the presence of outliers in the dataset.

MIS-con-lab can employ the regularization by penalizing the magnitudes of e and w as in the MIS-con approach to balance between the MIS accuracy and distinction of the individual models.

The problem (3) serves primarily to decide about data labels, i.e., distribution of the training data and, subsequently, of model validity regions. After fixing the values of z , the final training can be performed via solving (2). This two-step approach does not require the a priori labeling of the training set and can provide optimal MIS for the prize of an increased computational burden.

The problem (3) can be transformed to MILP via: (a) the epigraph reformulation (Milano, 2012) of the absolute value, (b) the big-M method (Griva et al., 2008) to linearize the bilinear constraints. As the variables z are binary, the big-M method does not require any new integer variables.

4. Case Study

In this contribution, the design of MIS is tested for the estimation of pressure compensated temperature PCT . This is a phenomenological variable used very often in the petrochemical industry. It is derived by combining the Antoine and Clausius-Clapeyron equations as (King, 2016):

$$\frac{1}{PCT} = \frac{R}{H_v} \ln \left(\frac{P}{P_{\text{ref}}} \right) + \frac{1}{T}, \quad \text{with } x = (T, P), \quad y = PCT, \quad (4)$$

where H_v is the heat of vaporization, R is the universal gas constant, P_{ref} is the reference pressure, P is the absolute pressure, and T is the absolute temperature. The ground-truth parameters of the PCT model are $R = 8.314 \text{ J/mol/K}$, $H_v = 55,940.550 \text{ J/mol}$, $P_{\text{ref}} = 145,325 \text{ Pa}$. The data used in this work is generated within the following intervals:

$$523.2 \text{ K} \leq T \leq 573.2 \text{ K}, \quad 2,000 \text{ Pa} \leq P \leq 20,000 \text{ Pa}, \quad 618.5 \text{ K} \leq PCT \leq 902.7 \text{ K}. \quad (5)$$

Despite the PCT model is usually used within low-pressure devices, the pressure scale used here is even lower. We use this scale in order to provide a more apparent nonlinear behavior of PCT .

We further assume that the training data of PCT is corrupted with some noise. The noise is generated as a random variable from the standard normal distribution. In order to remove the discrepancies in the variables magnitudes (P , T , PCT), we normalize the variables (P_{norm} , T_{norm} , PCT_{norm}) used in the further experiments to lie within the interval $[0, 1]$.

We study two different scenarios (clustered and uniformly distributed dataset) each with 140 synthetic measurements. The weighting parameter α is set to zero and the weighting parameter β is set to 0.01. The calculations are executed in MATLAB using Yalmip (Löfberg, 2004). The MIS approaches use Gurobi (Gurobi Optimization, 2021).

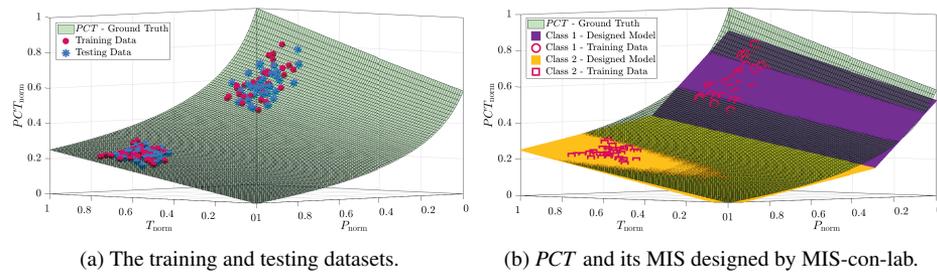
Figure 1: The PCT model on the clustered dataset.

Table 1: The median computational times (CPU time) for 100 simulation runs (clustered datasets).

	SIS	MIS-std	MIS-con	MIS-con-lab
CPU time [s]	< 0.01	0.4	0.4	5.1

4.1. Clustered Dataset

This scenario mimics the situation when the industrial data is well treated from the systematic errors. We can indicate two distinct clusters of measurements in Fig. 1a, where each cluster represents a particular operating point. The measurements are evenly divided into clusters with 50% of the measurements being randomly distributed into the training set (red points) and the rest to the testing set (blue stars). For lucidity of the results, we perform 100 runs with similar datasets.

We first analyze the performance of the trained MISs on one representative run in Fig. 1b. The ground-truth PCT is approximated by two linear models of MIS designed according to the corresponding classes of training datasets (class 1 – circles, purple model surface; class 2 – squares, yellow model surface). The accuracy of the MISs designed using all studied methods is the same. We visualize the results for MIS-con-lab as this approach has optimized the data labels.

The results in Tab. 1 indicate longer computational time of the MIS approaches compared to SIS, as expected. The results also show that MIS-std and MIS-con impose a lower computational burden than MIS-con-lab in this case. The optimized binary variables significantly increase the computational time of the MIS-con-lab approach.

A statistical evaluation of the results from 100 simulation runs is shown in Fig. 4. The accuracy of each method (via RMSE – root mean square error) on the testing set (RMSE-TS) is characterized by the blue box (25th–75th percentile, indicating variance), red line (median), and red crosses (out-

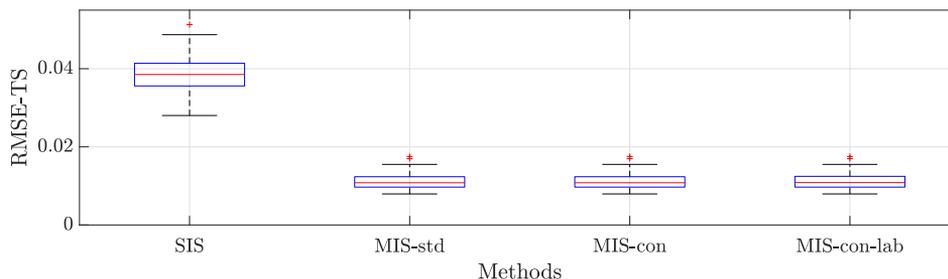


Figure 2: The prediction accuracy statistics on the testing dataset (RMSE-TS) of the designed inferential sensors from 100 simulation runs with different clustered datasets.

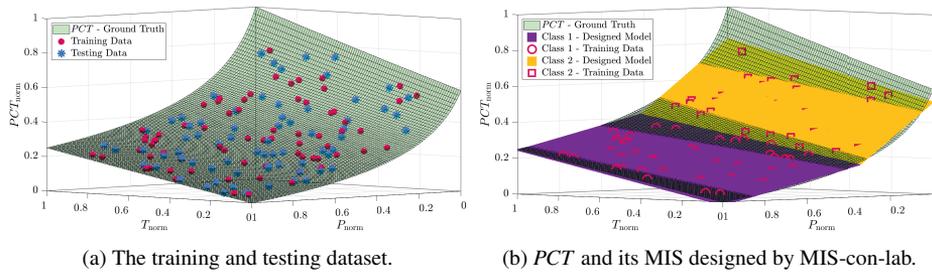
Figure 3: The PCT model on the uniformly distributed dataset.

Table 2: The median prediction accuracy on the training (RMSE-TR) and testing (RMSE-TS) datasets from 100 simulation runs (uniformly distributed datasets).

	SIS	MIS-std	MIS-con	MIS-con-lab
RMSE-TR $\times 10^3$	44.2	18.7	19.2	11.8
RMSE-TS $\times 10^3$	44.9	21.5	19.8	14.0
CPU time [s]	< 0.01	0.3	0.3	20.4

liers). We can conclude that the average/median accuracy and its variance are improved by MISs compared to SIS. The performance of all the designed MIS is the same. The MIS-std and MIS-con approaches can reach comparable accuracy to the MIS-con-lab approach as the (clustered) datasets in this scenario are appropriate for the k -means clustering (a priori labeling). The small accuracy variance confirms the good performance of these approaches.

4.2. Uniformly Distributed Dataset

This scenario involves data uniformly distributed over the whole considered range of the PCT model (see Fig. 3a). Such setting covers the cases of non-ideal industrial datasets, which usually involve this type of data. The available data is randomly and evenly distributed into the training and testing sets. These sets contain the same number of measurements. To provide accurate conclusions, we generate 100 similar datasets and statistically evaluate the results.

The results from one representative simulation run are shown in Fig. 3b. We can see that the MIS is approximated by two models (model 1 – purple surface; model 2 – yellow surface). These models are designed by the MIS-con-lab approach. We visualize the results from this approach because its accuracy is higher compared to other studied approaches.

The results in Tab. 2 prove the highest accuracy of the MIS-con-lab approach on both training and testing datasets. This method outperforms MIS-std by about 35%. Furthermore, we can indicate a slightly lower accuracy of the MIS-con approach compared to MIS-std on the training dataset. This is expected as the MIS-con approach gives up some prediction accuracy for the model-switching continuity. On the other hand, the MIS designed by MIS-con performs better on the testing set. It seems that the enforcement of continuity weakens the effects of the measurement noise in the training dataset and leads to more accurate models. The results from the computation time (CPU time) indicate a higher computational burden of the MIS-con-lab approach compared to the previous scenario (see in Tab. 1). The increased computational burden is caused by the provided dataset nature, which increases the data labeling complexity.

The results in Fig. 4 confirm the superior accuracy of the MIS-con-lab approach over SIS. The accuracy statistics of MIS-std and MIS-con involve several outliers that represent the accuracy comparable to SIS. These outliers can illustrate the low performance of the a priori labeling on

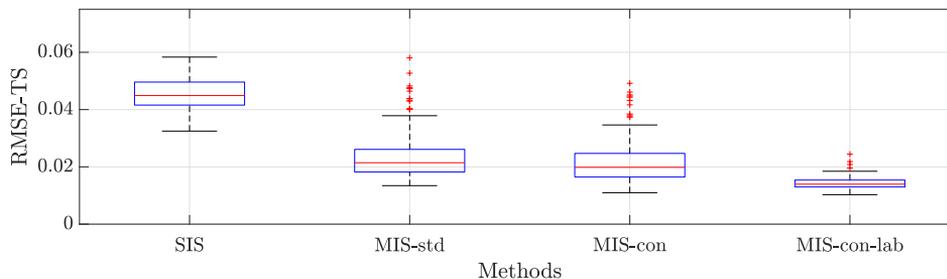


Figure 4: The prediction accuracy statistics on the testing dataset (RMSE-TS) of the designed inferential sensors from 100 simulation runs with different uniformly distributed datasets.

the uniformly distributed training dataset. The variance and number of outliers in Fig. 4 show the superiority of the MIS-con-lab approach in this respect.

5. Conclusions

This contribution is focused on the multi-model inferential sensor design using two sensor models with continuous switching and with optimized data labeling. We presented novel design approaches and compared their effectiveness against the standard design approach and against a single-model inferential sensor. The results suggest a significant accuracy improvement of the multi-model inferential sensor compared to a single-model one. The comparison of the approaches shows a very promising potential of the approach with optimized data labeling as it designs the most accurate inferential sensor and does not suffer from the limitations of the standard approach.

References

- B. E. Boser, I. M. Guyon, V. N. Vapnik, 1992. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. ACM, New York, USA, pp. 144–152.
- E. Forgy, 1965. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 21, 768–769.
- I. Griva, S. G. Nash, A. Sofer, 2008. *Linear and Nonlinear Optimization* (2. ed.). SIAM.
- Gurobi Optimization, 2021. Gurobi optimizer reference manual URL <https://www.gurobi.com>.
- S. Khatibisepehr, B. Huang, F. Xu, A. Espejo, 2012. A bayesian approach to design of adaptive multi-model inferential sensors with application in oil sand industry. *Journal of Process Control* 22 (10), 1913–1929.
- M. King, 2016. *Process Control: A Practical Approach*. Wiley.
- M. Li, T. Pan, Q. Chen, 2021. Estimation of tea quality grade using statistical identification of key variables. *Food Control* 119, 107485.
- J. Löfberg, 2004. YALMIP : A toolbox for modeling and optimization in MATLAB. In: In Proceedings of the CACSD Conference. Taipei, Taiwan.
- M. Milano, 2012. Principles and Practice of Constraint Programming - CP 2012: 18th International Conference, CP 2012, Québec City, QC, Canada, October 8–12, 2012, Proceedings. Lecture Notes in Computer Science. Springer.
- M. Mojto, M. Fikar, R. Paulen, 2022. Design of multi-model linear inferential sensors with SVM-based switching logic. In: *European Control Conference*. (submitted).
- M. Mojto, K. L'ubušký, M. Fikar, R. Paulen, 2021. Data-based design of inferential sensors for petrochemical industry. *Computers & Chemical Engineering* 153, 107437.
- J. Qi, L. Ding, S. Lim, 2021. Toward cool cities and communities: A sensitivity analysis method to identify the key planning and design variables for urban heat mitigation techniques. *Sustainable Cities and Society* 75, 103377.
- S. J. Qin, H. Yue, R. Dunia, 1997. Self-validating inferential sensors with application to air emission monitoring. *Industrial & Engineering Chemistry Research* 36 (5), 1675–1685.
- Q. Zhu, S. Joe Qin, Y. Dong, 2020. Dynamic latent variable regression for inferential sensor modeling and monitoring. *Computers & Chemical Engineering* 137, 106809.

Model predictive control for greenhouse condition adjustment and crop production prediction

Guoqing Hu^a, Fengqi You^a

^a*Cornell University, Ithaca, New York, 14853, USA*
gh429@cornell.edu

Abstract

Greenhouse cultivation targets provide protection to the plantations from severe weather and become a way to achieve controlled agricultural production. The appropriate control of the climate within the greenhouse can lead to the growth of potential yields and the extension of the growing season. However, greenhouses are usually constructed in a way without further consideration to the energy efficiency, leading to the need for an advanced automation control system to save the energy budget. Model predictive control (MPC) has been acknowledged as a suitable control strategy to solve the problem of regulating the climate and growth processes, plus MPC can also guarantee the fulfilment of the constraints specification. To develop the state-space model (SSM) required for the proposed MPC, we start by studying the dynamics of the greenhouse and crop based on the building geometry, weather dynamics, and input control signals. Afterward, we use data-driven techniques and system identification methods to construct the SSM and then incorporate the SSM into the proposed MPC framework. Based on the proposed data-driven dynamic SSM model and the data-driven uncertainty set, we propose a novel data-driven RMPC framework, which is capable of controlling the greenhouse's temperature, humidity, CO₂.

Keywords: Model Predictive Control, Greenhouse Model, Crop Model.

1. Introduction

The greenhouse is gaining popularity in recent years because of its significant contribution to the sustainable intensification of crop production in many countries (Villagran et al., 2020). Greenhouse control can be further developed into a low-cost production system, which is compatible with the shortage of natural resources and the low investment capacity of the growers (Chen and You, 2021). These agronomic and financial benefits give sufficient incentives for developing the control of greenhouse (Bennis et al., 2008). However, the control of the greenhouse is usually deemed as a complex process, regarding its multi-input multi-output (MIMO) and nonlinear characteristics. The greenhouse presents time-varying behavior and is disturbed generally by meteorologic conditions. All these difficulties adversely affect the control performances of classical controllers to the greenhouse (Fourati and Chtourou, 2007). Model predictive control (MPC) has been acknowledged as a suitable control strategy to solve the problem of regulating the climate and growth processes, handling all scheduling decisions effectively (Chu et al., 2015). Besides, MPC can also guarantee the fulfillment of the constraints specification (Camacho and Alba, 2013). MPC is a model-based control strategy that determines the optimal control sequence by solving a sequence of numerical optimization problems with constraints over a specific horizon based on the prediction

model (Agachi et al., 2016). The first input in the optimal sequence is sent into the system, and the entire computation is repetitively performed at subsequent control intervals following a receding horizon approach (Qin and Badgwell, 2003). Several efforts are conducted. In El Ghoumari et al. (2005), the authors have demonstrated the advantageous control of MPC in regulating the temperature of greenhouse compared to the PID control. In Gruber et al. (2011), the nonlinear MPC was proposed and further proved the control performance of MPC in regulating the diurnal temperature in the greenhouse around the desired reference. Despite the progress made over the years, there still exists some knowledge gap: None of the research mentioned above considers the crop model, but they only focused on the control to the greenhouse. The crop model can contribute significantly to greenhouse cultivation while being applied to the optimization work (Lentz, 1998).

In this work, we propose an MPC framework that minimizes the greenhouse's energy and natural resources consumption as well as maximizes crop production. To develop the state-space model (SSM) required for the proposed MPC, we start by studying the dynamics of the greenhouse based on the building geometry and dynamics. To take a further step, we also add the crop model and correlate it to the greenhouse model mentioned before. Afterward, we use data-driven techniques and system identification methods to construct the SSM and then incorporate the SSM into the MPC framework. Weather forecast data and measurement data about ambient temperature, relative humidity, and wind speed are gathered. Based on the proposed dynamic SSM model and forecast values, we can develop the MPC framework which can be adopted for the automation control system. This MPC framework is capable of controlling the greenhouse's temperature, humidity, CO₂, irrigation, and fertilization. The proposed framework not only is computationally efficient but also mitigates the wasted consumption issues that arise from the greenhouse power regulation.

2. State-space model construction

2.1. Greenhouse model

Vanthoor et al. (2011) has designed the greenhouse model which studies the impact of outdoor climate and greenhouse design on the indoor greenhouse climate. Weather information is comprehensively considered in this model, including ambient temperature, ambient relative humidity, global radiation, diffuse radiation, and wind speed. This model is adopted in this research regarding its fulfillment of the following three requirements: 1) this model can adequately predict the temperature, water vapor pressure, CO₂ concentration greenhouse indoor air under different greenhouse designs and regional climate; 2) this model is able to take greenhouse construction parameters and climate adjustment facilities into consideration. 3) This model is also ready to be incorporated with the tomato yield production model (Rezvani et al., 2021).

2.2. Crop model

Although Vanthoor et al. included tomato growth in the model, fertilization and irrigation were not considered. These two factors are crucial to crop production and tightly related to the quality of the crop. Therefore, we apply the VegSyst model designed by Gallardo et al. (2011). In this model, we are able to retrieve the crop growth rate by compiling the weather information data and adjusting the fertigation rate (Shang et al., 2020). Therefore, we can now develop the steady-state matrix required by the MPC framework.

2.3. System identification

Although we have the physics-based model which can describe the correlation between each state, the linear SSM is preferred in MPC computation because of computational efficiency. Therefore, we apply the system identification to the nonlinear model to identify the SSM which combines the greenhouse model and crop model. The SSM is described in the following equation:

$$x_{t+1} = Ax_t + B_u u_t + B_v v_t + B_w w_t \quad (1)$$

Where A is a 9×9 state matrix. In this SSM, 9 states are considered, including cover temperature, internal temperature, mat temperature, tray temperature, floor temperature, soil temperature, water vapor density, CO_2 density, and growth rate. B_u is a 9×7 input matrix, in which 7 control signals are considered: heater, fans, humidifier, dehumidifier, CO_2 enrichment, irrigation, and nitrogen enrichment; B_v is a 9×5 weather disturbance matrix which describes the linear correlated impact from weather data. In this research, we consider 5 weather information, including ambient temperature, ambient relative humidity, global radiation, diffuse radiation, and wind speed. B_w is a 9×3 weather forecast uncertainty matrix. Three major forecast uncertainties are considered including temperature forecast error, humidity forecast error, and wind speed forecast error (Shang et al., 2019). Figure 1 demonstrates the results of system identification. The overall mean absolute percentage error (MAPE) for the linearization is 7.06 %, indicating that this SSM is sufficiently accurate in describing the nonlinear physics correlation between each state.

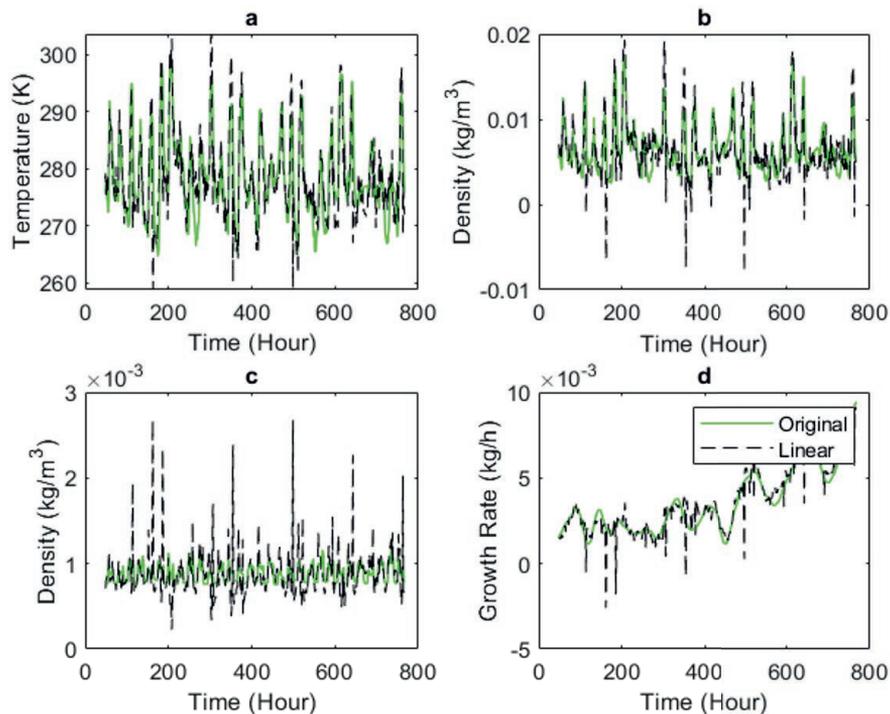


Figure 1. System identification results of: (a) indoor greenhouse temperature (b) water vapor density (c) CO_2 concentration (d) growth rate. The linear results are demonstrated as a dashed black line and original nonlinear model results are represented in a green solid line.

3. Model predictive control framework

After acquiring the SSM for the control framework. The MPC optimization problem can be formulated as below:

$$\begin{aligned} \min \quad & \sum_{i \in B_u} c_i u_i + \lambda_i^T L_i \lambda_i - \alpha \sum_{j \in X_{crop}} x_j \\ \text{s.t.} \quad & F_u u \leq f_u, \\ & F_x [Ax + B_u u + B_v v + B_w w] \leq f_x + \lambda \end{aligned} \quad (2)$$

Where F_x , F_u , f_x , f_u represent the state variable constraints matrix, control input constraints matrix, constraints for state variables, and constraints for the input. L is the weighted cost matrix that penalizes the violation to the constraints (Jia et al., 2021). λ is the slack variable that allows some extent of violation to the hard constraints. In this problem, we introduce the soft constraint to the optimization to ensure the feasibility and help alleviate the over-pessimism or infeasibility of hard constraints (Meseguer et al., 2006). Besides minimizing the cost of energy consumption and violation of the constraints, we also try to maximize crop production. α is the weighted coefficient for crop production. This coefficient is necessary because the growth rate is a much smaller value compared to the energy cost and to the constraint-violation penalty. The absence of this coefficient will lead decision impartial to the growth condition of the crop. x_{crop} is the growth rate states from the state matrix A .

4. Control profile

In this research, we simulate the greenhouse located in Ithaca, the USA from 0:00, March 1st, 2020 to 23:00, March 30th, 2020. The initial conditions are set as (Chen et al., 2021):

- For all temperature values, the starting conditions are 298.15 K
- The initial water vapor density is set as $2.5 \times 10^{-3} \text{ kg/m}^3$
- The initial CO₂ concentration is set as 0.8 kg/m³
- The initial growth rate is set as 0 kg/hr

The constraints are set as follows: For daytime (from 7:00 to 19:00), the upper limit for temperature is 302.5 K and the lower limit is 297 K; the upper limit for relative humidity is 90% and the lower limit is 80%. For nighttime (from 19:00 to next day 7:00), the upper limit for temperature is set as 297.5 K and the lower limit is 288.7 K. The relative humidity range is from 65 % to 75 %.

The control results are demonstrated in Figure 2 and Figure 3. We can observe from plots that temperature has fewer violation cases than for humidity control profile. The reason behind this is that the forecast error of temperature only occupies less than 1 % of the actual measured value (when temperature value is recorded as in kelvin), whereas the forecast error of humidity can be away from the actual measurement value up to 25 %. Therefore, the humidity forecast error can drive the control decision away from the desired constraints.

The crop growth curve is illustrated in Figure 4. We can conclude that the crop is steadily growing.

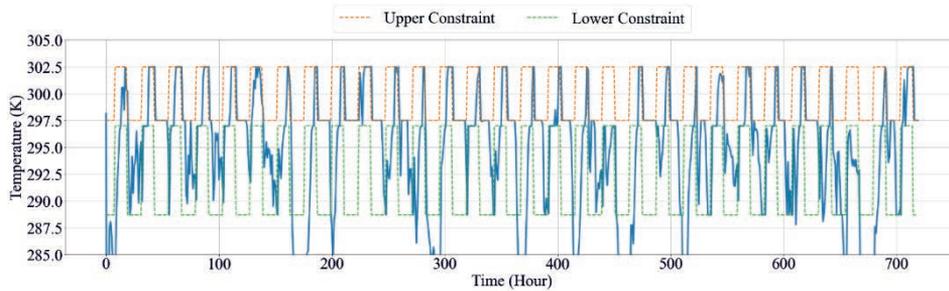


Figure 2. Temperature control profile of greenhouse from March 1st 0:00 to March 30th 0:00 in 2020. The orange dashed line indicates the upper limits for the temperature constraints while the green dash line for the lower limits.

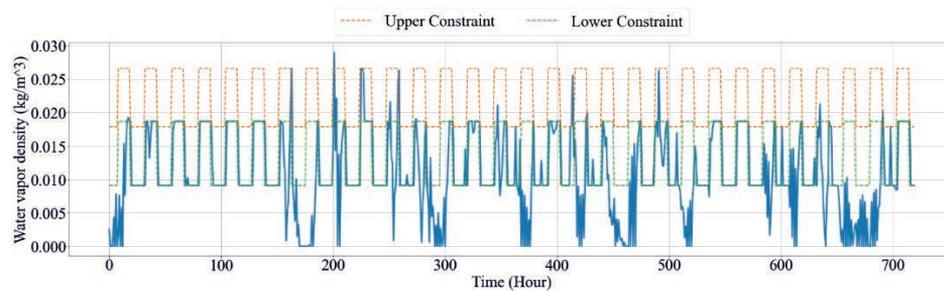


Figure 3. Humidity control profile of greenhouse from March 1st 0:00 to March 30th 0:00 in 2020. The orange dashed line indicates the upper limits for the temperature constraints while the green dash line for the lower limits.

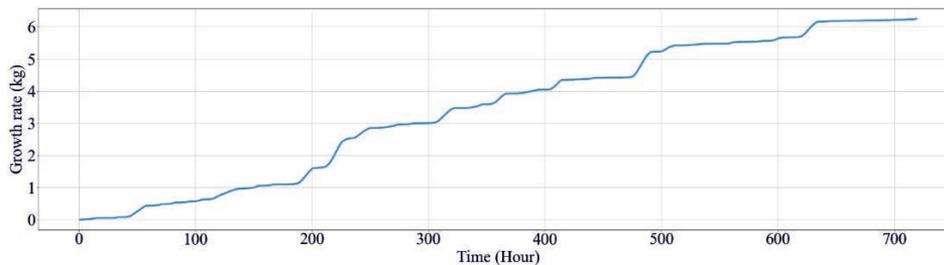


Figure 4. Gross tomato production profile from March 1st 0:00 to March 30th 0:00 in 2020.

5. Discussion

In this research, we focused on developing the MPC framework for the SSM, which combines the greenhouse model and crop model. Thanks to the linearization process, the average CPU time for each case is only 0.044 seconds, which is sufficient to finish the optimization procedure within the sampling interval of 8 hours. Plus, the soft constraint within the optimization also ensures feasibility (Lu et al., 2020). However, the drawback of this MPC framework is also noticeable. The violation of the constraints is not negligible in this control profile, as these violation cases may result in a harsh climate in the greenhouse and become harmful to the tomato growth. Therefore, Robust MPC, which takes forecast errors into consideration in obtaining optimal control inputs, should be used in this case to hedge against the forecast uncertainties.

References

- P. S. Agachi, M. V. Cristea, A. A. Csavdari & B. Szilagy, 2016, 2. Model Predictive control Advanced Process Engineering Control, De Gruyter
- N. Bennis, J. Duplaix, G. Enéa, M. Haloua, & H. Youlal, 2008, Greenhouse climate modelling and robust control. *Computers and Electronics in Agriculture*, 61(2), pp 96-107
- E. F. Camacho, C. B. Alba, 2013, Model predictive control: Springer science & business media.
- W.-H. Chen, C. Shang, S. Zhu, et al., 2021, Data-driven robust model predictive control framework for stem water potential regulation and irrigation in water management. *Control Engineering Practice*, 113, 104841.
- W.-H. Chen, F. You, 2021, Smart greenhouse control under harsh climate conditions based on data-driven robust model predictive control with principal component analysis and kernel density estimation. *Journal of Process Control*, 107, 103-113.
- W.-H. Chen, F. You, 2022, Semiclosed Greenhouse Climate Control Under Uncertainty via Machine Learning and Data-Driven Robust Model Predictive Control. *IEEE Transactions on Control Systems Technology*, DOI:10.1109/tcst.2021.3094999.
- Y. Chu, F. You, 2015, Model-based integration of control and operations: Overview, challenges, advances, and opportunities. *Computers & Chemical Engineering*, 83, 2-20.
- M. Y. El Ghomari, H. J. Tantau & J. Serrano, 2005, Non-linear constrained MPC: Real-time implementation of greenhouse air temperature control. *Computers and Electronics in Agriculture*, 49(3), pp 345 – 356
- F. Fourati & M. Chtourou, 2007, A greenhouse control with feed-forward and recurrent neural networks, *Simulation Modelling Practice and Theory*, 15(8), pp 1016-1028
- M. Gallardo, C. Giménez, C. Martínez-Gaitán, C. O. Stöckle, R. B. Thompson & M. R. Granados, 2011, Evaluation of the VegSys model with muskmelon to simulate crop growth, nitrogen uptake and evapotranspiration. *Agricultural Water Management*, 101(1), pp 107-117.
- J. K. Gruber, J. L. Guzmán, F. Rodríguez, C. Bordons, M. Berenguel & J. A. Sánchez, 2011. Nonlinear MPC based on a Volterra series model for greenhouse temperature control using natural ventilation. *Control Engineering Practice*, 19(4), pp 354-366.
- R. Jia, F. You, 2021, Multi-stage economic model predictive control for a gold cyanidation leaching process under uncertainty. *AIChE Journal*, 67, e17043.
- W. Lentz. 1998. Model application in horticulture: a review. *Scientia Horticulturae*, 74(1), pp 151-174
- S. Lu, J.H. Lee, F. You, 2020, Soft-constrained model predictive control based on data-driven distributionally robust optimization. *AIChE Journal*, 66, e16546.
- P. Meseguer, F. Rossi & T. Schiex, 2006, Chapter 9 – Soft Constraints. In: F. Rossi, P. van Beek, & T. Walsh, (eds.) *Foundations of Artificial Intelligence*, Elsevier
- S. J. Qin & T. A. Badgwell, 2003, A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7), pp 733-764
- S. M. -E. -D. Rezvani, R. R. Shamshiri, et al., 2021, Greenhouse Crop Simulation Models and Microclimate Control Systems, Next-Generation Greenhouses for Food Security, IntechOpen.
- C. Shang, W.-H. Chen, A.D. Stroock, F. You, 2020, Robust Model Predictive Control of Irrigation Systems With Active Uncertainty Learning and Data Analytics. *IEEE Transactions on Control Systems Technology*, 28, 1493-1504.
- C. Shang, F. You, 2019, A data-driven robust optimization approach to scenario-based stochastic model predictive control. *Journal of Process Control*, 75, 24-39.
- C. Shang, F. You, 2019, Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era. *Engineering*, 5, 1010-1016.
- B. H. E. Vanthoor, C. Stanghellini, E. J. van Henten & P. H. B. de Visser, 2011, A methodology for model-based greenhouse design: Part 1, a greenhouse climate model for a broad range of designs and climates, *Biosystems Engineering*, 110(4), pp 363-377
- E. Villagran, R. Ramirez, A. Rodriguez, et al. 2020, Simulation of the thermal and aerodynamic behavior of an established screenhouse under warm tropical climate conditions: A numerical approach. *International Journal of sustainable development and Planning*, 15(487-499).

Study on the noise contents of different measurements in industrial process and their impact on process monitoring

Tingting Tao, Cheng Ji, Chengyu Han, Jingde Wang*, Wei Sun*

College of Chemical Engineering, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, Beijing, 100029, China

**Corresponding Author's E-mail: sunwei@mail.buct.edu.cn*

Abstract

As a large amount of industrial data have been acquired by distributed control system (DCS), data-driven process monitoring methods have attracted a significant research interest. However, various types of noise are also collected with signal together from process instruments and data transformation, which could compromise the information of signal itself and provide a huge challenge for data analysis and modeling. Moreover, noise contents of different measurements, such as temperature measurements, pressure measurements, flow measurements, and etc., may be also varied in industrial process. Therefore, the selection of decomposition scale has a great influence on the denoising effect, while among methods reported in the literature, the decomposition scale is usually constant for all measurements in a process.

In this research, a novel standard to characterize an optimal denoising depth for each measurement is proposed. Wavelet decomposition is first selected to extract the high-frequency features of signals in different measurements for its ability to retain the local characteristics of signals in both time and frequency domains. Then information entropy is applied as a standard to characterize the depth of wavelet decomposition. An optimal decomposition scale is determined by maximizing the retention of raw signal information in each measurement. According to this standard, signal in each measurement can be processed individually according to its optimal decomposition scale. Then the processed data can be further applied to a multivariate statistics method for a more effective process monitoring. In order to verify the effectiveness of this method, the research is implemented on an industrial continuous catalytic reforming process. The results show that abnormal operational conditions can be detected earlier on the basis of the proposed method compared to other methods without considering the denoising of different variables respectively.

Keywords: wavelet denoising, process monitoring, information entropy, continuous catalytic reforming process.

1. Introduction

Production safety in chemical industry is always considered as a prerequisite, whether it is from the perspective of design or production operations. In fact, large-scale chemical production is always accompanied by a great potential safety hazard due to harsh operating conditions and various unexpected disturbances, which will cause huge losses to human lives and national economy. To prevent safety accidents, process monitoring technology has been developed in chemical processes to assist operators to detect

abnormal changes in a process plant timely. Take advantage of the real-time data collection technology provided by distributed control system (DCS), data-driven feature extraction methods have been widely adopted for fault detection. For example, as a conventional multivariate statistical method, principal component analysis (PCA) has been widely used by projecting high dimension data into a feature space and achieved good results on simulation data of Tennessee-Eastman process (MacGregor, et al., 1995). However, the actual measurement value obtained from DCS is inevitably coupled with various types of noise during data transmission. Numerous researches have been conducted to solve this problem. The widely used wavelet transform method has shown a powerful denoising effect in chemical industry by adopting both time and frequency analysis (Bakshi, 1999). Li and Wen developed a wavelet-PCA method for air-handling units monitoring, by which the advantages of wavelet transform method for data preprocessing in process monitoring has been demonstrated (Li, et al., 2014). When applying wavelet denoising, the selection of wavelet basis function and denoising depth shows a great impact on the denoising effect. Among methods reported in the literature, the decomposition scale is usually constant for all measurements in a process. However, noise distribution and contents of different measurements, such as temperature measurements, pressure measurements, flow measurements, and etc., varies in industrial process. Therefore, it is necessary to determine a proper denoising standard to achieve an optimal denoising effect for each measurement.

The aim of this paper is to extract data features for process monitoring under the premise of determining an optimal denoising depth for each type of measurement. When denoising different measurements such as temperature measurements, pressure measurements, flow measurements, and etc., it can be found that the noise contents of them are different in industrial process. On this basis, information entropy maximization, which has been proved to be a useful wavelet parameters selection standard introduced by Altay and Kalenderli (Altay, et al., 2014), is adopted to determine an optimal denoising depth for each measurement. Since signal in each measurement has been processed according to its optimal decomposition scale, the processed data are further applied to a PCA model for a more effective process monitoring. According to the results obtained from an industrial reforming heat exchanger unit, it can be seen that the fault could be earlier detected with a lower false alarm rate by the proposed method compared with other methods with only one denoising policy.

2. Methodology

In this section, the methods applied in this work are introduced.

2.1. Wavelet Threshold Denoising Method

Wavelet threshold denoising method is an effective time-frequency signal processing tool. Signals are decomposed based on a scaled and translated mother wavelet given as

$$\int_{-\infty}^{+\infty} \varphi(t) dt = 0 \quad (1)$$

In continuous wavelet transform (CWT), the decomposed coefficients are given as

$$CWT(u, s) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{s}} \varphi^* \left(\frac{t-u}{s} \right) dt \quad (2)$$

where u and s are the scaling and translation parameters. $x(t)$ is the raw signal. The discrete form of CWT is obtained by sampling the time-scale plane. In discrete wavelet transform (DWT), the decomposed coefficients are given as

$$DWT(a,b) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{a}} \phi^* \left(\frac{t-b}{a} \right) dt \quad (3)$$

where a and b are the discrete versions of u and s . De-noising of signals in the wavelet domain can be done in decomposition, thresholding and reconstruction three steps.

2.2. Shannon entropy

Shannon entropy is a concept introduced from physical systems by Shannon to estimate the amount of information (Shannon, 1948), its calculation is given as follows,

$$S = \sum_{i=1}^l p_i \log_2 \left(\frac{1}{p_i} \right) \quad (4)$$

where l is the total number and p_i is the probability of the situation i in the system. Maximum information entropy can be applied in the wavelet de-nosing process for parameter selection, the probability can be calculated as follows,

$$p_i = E_j / \sum_{j=1}^M E_j = \left(\sum_{k=1}^{K_j} w_{j,k}^2 \right) / \sum_{j=1}^M \left(\sum_{k=1}^{K_j} w_{j,k}^2 \right) \quad (5)$$

where E_j is the energy in time scale level j and M is the maximum time scale level. K_j is the number of wavelet coefficients in time scale level j . $w_{j,k}$ is the processed wavelet coefficient. Maximum separation between signal and noise can be calculated using the maximization of entropy technique calculated by Equation 6, which gives the optimal value of j and helps to determine an optimal mother wavelet.

$$S = S_{signal} + S_{noise} = \sum_{i=1}^k p_{i,signal} \log_2 \left(\frac{1}{p_{i,signal}} \right) + \sum_{i=1}^k p_{i,noise} \log_2 \left(\frac{1}{p_{i,noise}} \right) \quad (6)$$

2.3. Principal component analysis (PCA)

PCA is a classical feature extraction technique by projecting high dimensional data into a lower feature space, which has been widely used in process monitoring with T^2 and SPE statistic. Given normalized n observations of m measurement variables $X_{n \times m}$, the covariance matrix of X can be calculated as follows,

$$Cov(X) = \frac{X^T X}{n-1} \quad (7)$$

Then singular value decomposition is employed to $Cov(X)$, a score matrix T and a loading matrix P are determined by retaining the first k features that contain the most information. The raw matrix X can be decomposed as follows,

$$X = TP^T + E = t_1 p_1^T + L + t_k p_k^T + E \quad (8)$$

where p , t are the loading vector and score vector, and E is the residual matrix. For online monitoring, T^2 statistic and SPE statistic can be calculated as monitoring statistics to measure the deviation in the principal component space and residual space respectively.

3. Selection of optimal denoising parameters and its application on an industrial continuous catalytic reforming process

In this chapter, the implement procedures of proposed research on the selection of optimal denoising parameters are introduced with the case study of an industrial continuous catalytic reforming process.

3.1. Continuous catalytic reforming process and data description

The proposed Shannon entropy-based denoising strategy is applied to an industrial continuous reforming process, which contains four reactors, four furnaces and a plate heat exchanger. The pressure drop of the plate heat exchanger is of great importance because it will increase with various factors, such as environment temperature and equipment factors. However, the increase in the pressure drop is usually a process of slow change, and is difficult to be detected in time by operators. Therefore, it is necessary to build an effective process monitoring model in this process to detect the faulty deviation at its early stage. The process variables applied in this work are shown in Table 1.

Table 1 Process variables in the industrial continuous catalytic reforming process.

Variable	Description	Variable	Description
T01	Outlet temperature at cold side	PD10	Outlet pressure at cold side
T02	Inlet temperature at hot side	PD11	Pressure drop at hot side
T03	Inlet temperature at cold side	PD12	Pressure drop at cold side
T04	Outlet temperature at hot side	PD13~16	Reactor pressure drop 1
F05	Naphtha feed flow	T17~20	Furnace outlet temperature
F06	Circulating hydrogen flow	T21~24	Reactor outlet temperature
PD07	Inlet filter pressure drop at cold side	T25~28	Furnace temperature drop
PD08	Inlet pressure at cold side	PD29	Reactor inlet pressure
P09	Circulating hydrogen pressure		

3.2. Determination of the optimal denoising parameters

In this section, the selection of denoising parameters in wavelet denoising, containing mother wavelets, decomposition levels, thresholds and thresholding rules are introduced. In most existing denoising methods, all process variables are denoised under a same standard, which is called identical denoising in this work. A widely-used mother wavelet 'db4' is chosen and a decomposition level is selected properly according to the length of data. After decomposition, threshold for each layer is determined with a universal fixed threshold method. As for the thresholding rule, it has been proved that noise can be separated effectively from signal with soft threshold rule. However, considering noise content in different measurements, a novel Shannon entropy-based denoising strategy is applied to the industrial continuous reforming process. With the adjustment of mother wavelets and decomposition levels, the wavelet coefficients at different denoising depths are obtained. Then the fixed threshold and soft threshold rule are chosen for coefficient processing. Based on the processed coefficients, the signal entropy and noise entropy under different denoising parameters can be calculated. On this basis, the optimal mother wavelets and decomposition levels for different measurements can be found according to the criterion of maximum information entropy.

3.3. Process monitoring method based on new Shannon entropy denoising

A Shannon entropy-based denoising method and principal component analysis method are combined for process monitoring on an industrial continuous catalytic reforming process. The detailed flowchart of the monitoring process is shown in Figure 1. Historical data under normal operating conditions is selected for denoising. The optimal wavelet parameters suitable for the data set are determined according to the maximum information entropy criterion. Then the denoised historical data are input into the raw PCA model for

training. For the processing of real-time data, the previous denoising parameters are applied for its denoising. Then the denoised real-time data is input into the trained PCA model for testing. On the basis of this, the final monitoring results are obtained.

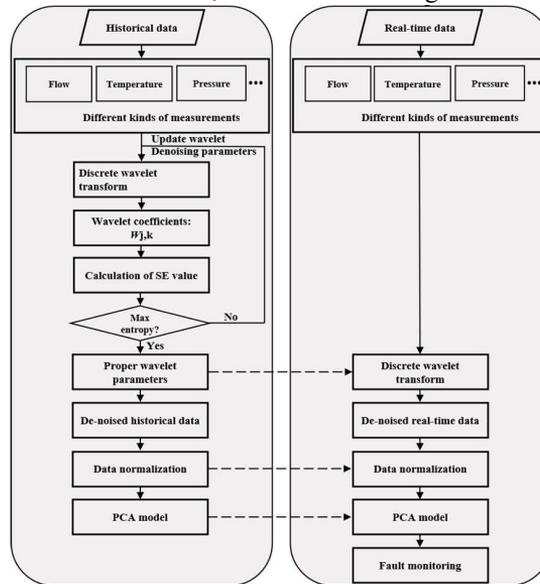


Figure 1 The detailed flowchart of the monitoring process.

4. Results and discussion

On the basis of wavelet threshold method, signals from an industrial continuous catalytic reforming process are denoised. Signal distribution of pressure variable PD07 before and after denoising is shown in Figure 2 and it can be seen that wavelet threshold method is effective for signal denoising. Three different measurements including temperature variable, flow variable and pressure variable are respectively denoised based on the maximum information entropy criterion. After normalization, the noise distribution of the variables in the time domain is shown in Figure 3. It can be seen that the noise extracted from the raw measurements conforms to the zero-mean distribution. More importantly, the result shows a higher degree of noise pollution in the pressure and flow measurement compared with the temperature measurement. Therefore, it's necessary to determine optimal denoising parameters for different measurements.

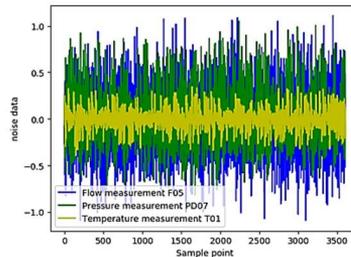
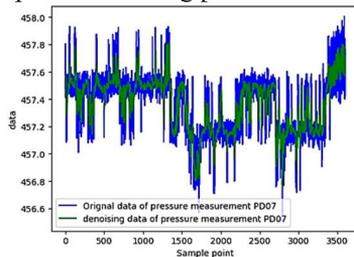


Figure 2 Signals before and after denoising Figure 3 Noise distributions of three measurements

Both historical data and real-time data are input into the PCA model in the form of a 400-length sliding window for process monitoring. The parameters of identical denoising are

introduced in Section 3.2. In separate denoising, mother wavelets and decomposition layers are determined by the maximum information entropy criterion. The parameters for temperature measurements are SYM4, 4 layers while a higher decomposition scale is recommended in pressure and flow measurements for their higher noise contents, which are SYM4, 6 and SYM8, 6. Then monitoring results based on the raw, identical denoised and respective denoised data are compared. Due to page limitations, the monitoring results without denoising have not been displayed. Based on the monitoring model constructed from raw data, the slowly changing fault can be identified at 471th sample. As is shown in Figure 4 and Figure 5, the control limit of the statistics based on the identical denoising method is exceeded at 252nd sample while the respective denoising one exceeded at 236th sample. Compared with the identical denoising method, the slowly changing fault can be detected 16 minutes earlier by the proposed method.

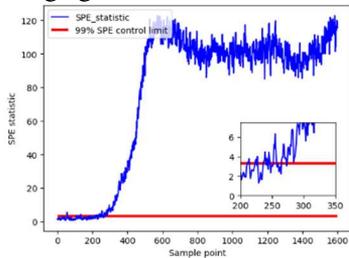


Figure 4 Result for identical denoising

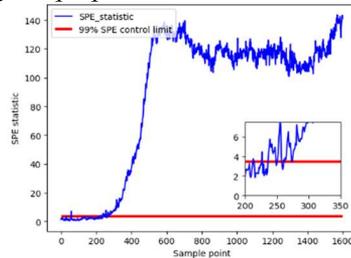


Figure 5 Result for respective denoising

5. Conclusions

In this work, a novel wavelet denoising strategy for industrial process data is proposed and applied to process monitoring with PCA. By considering the difference in noise distribution and contents of various types of variables, signal in each measurement is processed by wavelet transform individually according to its optimal decomposition scale, which is determined by information entropy maximization. Since each type of signal has been denoised to an optimal level, the signal feature could be effectively extracted by PCA, leading to a better process monitoring result. The method is applied to an industrial continuous catalytic reforming process, the slowly changing fault can be earlier detected by the proposed method compared with other methods without considering the denoising of different variables respectively. Compared with other studies with undifferentiated denoising, the determination of optimal denoising parameters on different measurements is discussed in detail in this work, which can provide a novel idea for the selection of denoising parameters in industrial practice.

References

- Ö. Altay, Ö. Kalenderli, 2014. Wavelet base selection for de-noising and extraction of partial discharge pulses in noisy environment, *Science Measurement & Technology Ict*, 9, 3, 276-284.
- B. R. Bakshi, 1999, Multiscale analysis and modeling using wavelets, *Journal of Chemometrics: A Journal of the Chemometrics Society*, 13, 3-4, 415-434.
- S. Li, W. Jin, 2014, A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform, *Energy and Buildings*, 68, 63-71.
- J. F. MacGregor, T. Kourtl, 1995, Statistical process control of multivariate processes, *Control Engineering Practice*, 3, 3, 403-414.
- C. E. Shannon, 1948, A mathematical theory of communication, *Bell System Tech. J.*, 27, 379, 623-656.

A Comparative Study on the Influence of Different Prediction Models on the Performance of Residual-based Monitoring Methods

Jiaying Ma, Yang Li, Fangyuan Ma, Jingde Wang*, Wei Sun*

College of Chemical Engineering, Beijing University of Chemical Technology, North Third Ring Road 15, Chaoyang District, Beijing, 100029, China

*Corresponding Author's E-mail: sunwei@mail.buct.edu.cn

Abstract

As a representation of the difference between observed and expected behaviors, residual has been widely used in monitoring methods based on the mechanism models, but the accurate mechanism models of certain complex processes are hard to obtain. With the rapid development of artificial intelligence, the accuracy of data-driven prediction models has been significantly improved, which makes it possible to reflect a certain status of the process by data-driven models instead of mechanism models. The residuals between status estimations from data-driven models and observations can be calculated to indicate status changes in the process. Therefore, residual-based multivariate statistical models are established for faults detection. Quite a few data-driven monitoring methods based on predicted residuals have been proposed. However, the influence of different prediction models on prediction residuals and monitoring performance hasn't been fully discussed. Referring to the above issue, various data-driven prediction models including PLS, SVM, and LSTM are applied in the Tennessee Eastman (TE) process and a real industrial case to study the influence of prediction models, evaluation indicators, distribution of prediction residuals on monitoring performance. A statistical analysis on corresponding residuals is also conducted to investigate its impact to monitoring performance.

Keywords: Multivariate linear regression, Time series modelling by LSTM, SVM (Support Vector Machine), Selection strategy of prediction models, Statistical feature of residuals

1. Introduction

Most modern chemical processes are known with large-scale, and highly complex with multiple variables (Nor, Hassan, & Hussain, 2020), which leads to the rapid growing need of process safety and reliability. To fulfill this need, process monitoring gains more and more attention recently. Many process monitoring methods have been proposed, which can be basically classified into mechanism model-based methods and data-driven methods.

An important step of mechanism model-based methods is to generate residual between measurements and estimations, and residuals are further evaluated to detect changes in their behavior caused by faults. However, the performance of mechanism model-based methods is limited by the accuracy of the model especially for certain complex processes whose mechanistic model are hard to obtain. Nowadays, data-driven methods, which solely rely on the data, are developed due to the widespread application of the instruments and distribution control system. Instead of mechanism model-based methods, data-driven

methods are introduced to reflect a certain status of the process for residual generation. Residuals obtained under normal operating conditions can be approximately considered as noise. Furthermore, the monitoring based on residuals variety can avoid the influence of data fluctuation and noise to achieve the early faults detection.

Various data-driven methods are introduced to generate residuals, which are further integrated into the residual-based process monitoring methods for better monitoring performance. Tong et.al proposed a data-driven residual generation method based on PLS for process monitoring(Tong, Lan, Yu, & Peng, 2019) to present more efficient fault detectability. Kazemi et.al used the estimated residual signal to univariate statistical control charts by developing SVM models (Kazemi, Giralt, Bengoa, Steyer, & Technology, 2020). Tao et.al proposed a LSTM-residual model on the MIT-BIH arrhythmia database to detect arrhythmia (Tao, Liu, & Liang, 2021). In the above researches, residuals are generated by linear, nonlinear and dynamic nonlinear models with the consideration of process characteristic.

Inspired by the above researches, the influence of prediction models, evaluation indicators, distribution of prediction residuals on monitoring performance are discussed in this study, where PLS, SVM and LSTM are selected as prediction models, PCA is selected as the monitoring model.

The rest of paper is organized as follows. The related methods are briefly reviewed in section2. Section 3 presents the residual-based monitoring method. The Comparison results of the Tennessee Eastman process and a real case are discussed in section 4. section 5 gives the conclusion.

2. Preliminary

In this section, the methods applied in this work are introduced.

2.1. Partial least squares regression (PLS)

Partial least squares regression (PLS) is a well-known linear prediction method by modeling latent relations between input and output variables. It was introduced by the Swedish statistician Herman O. A. Wold to predict dependent variables form independent variables and to describe the common structure underlying the two variables. As a linear method, PLS is a suitable choice for datasets that do not fit the traditional expectations demanded by ordinary regression(Pirouz, 2006).

2.2. Support vector machine (SVM)

Support vector machine (SVM) is a supervised learning model with associated learning algorithms that analyze data for prediction and classification analysis. It is developed at AT&T Bell laboratories by Vladimir Vapnik with colleagues. It can be used in regression to achieve multivariate pattern regression analysis with excellent generalization capability and high prediction accuracy(Awad & Khanna, 2015).

2.3. Long Short-Term Memory Network (LSTM)

LSTM is a deep neural network to capture the historical information of time series, and it is suitable for the prediction of long-term nonlinear series(Hochreiter & Schmidhuber, 1997). The prediction model established by LSTM can extract the nonlinear and dynamic characteristics of process data to achieve satisfactory prediction performance.

2.4. Principal components analysis (PCA)

PCA is a dimensionality reduction technology widely used in pattern recognition, image processing and feature extraction(Pearson, magazine, & science, 1901). For the purpose of process monitoring, Hotelling T^2 statistics and squared prediction error (SPE)

statistics are introduced to monitor the variation of feature space and residual space respectively.

3. Establishment of residual-based monitoring models

The residual-based monitoring method is employed to get the approximation of the local process state through prediction models, then the multivariate statistical methods are used to monitor the change of the prediction residual to achieve the early detection of faults. The complete procedure of residual-based monitoring method is presented as Figure 1.

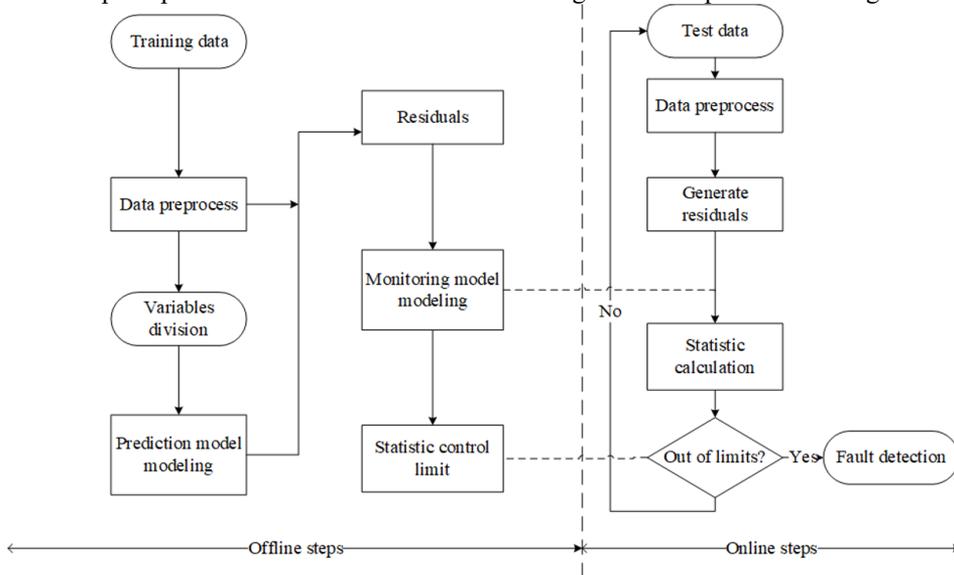


Figure 1. The detailed flowchart of the monitoring process

4. Case study

In this section, three representative prediction models are used for generating residuals, and PCA is used for residual monitoring for a fair comparison of monitoring performance. PLS-PCA, SVM-PCA and LSTM-PCA are applied in TE process and an industrial case to compare the influence of different prediction models on the performance of residual-based process monitoring methods.

4.1. Tennessee Eastman process (TE) process

Tennessee Eastman (TE) process proposed by Eastman Chemical Company is a widely used benchmark test based on a real industrial process for process monitoring. There are 33 measurement variables and 21 faults in TE process. The technological process of the TE simulation process can be briefly introduced as: the reaction raw materials A, C, D, and E are input into the reactor in gaseous form, and the main products G and H of the process are generated under the action of the catalyst. In addition, the process will produce by-products, namely by-product F.

In this case, 960 sample points are selected to establish residual-based model. It is divided into three parts: 500 sample points as training data for prediction model, 100 sample points as verifying data for prediction model and 360 sample points as training data for the monitoring model. Fault 3, fault 5 and fault 15 are hard to detect due to other

researcher's study, so 18 faults are used as testing data for the monitoring model (Chiang, Russell, & Braatz, 2000).

Table 1. Alarm point, FDR (%) and FAR (%) of 19 Faults in the TE process

Fault No.	Alarm point			FDR			FAR		
	PLS-PCA	SVM-PCA	LSTM-PCA	PLS-PCA	SVM-PCA	LSTM-PCA	PLS-PCA	SVM-PCA	LSTM-PCA
1	165	165	170	99.75	99.5	99.13	0.625	0.625	1.25
2	204	191	180	95.13	96.5	98.38	0.00	0.00	5.00
4	161	161	201	100.00	100.00	56.75	0.00	0.63	20.00
5	161	162	176	100.00	99.88	61.13	0.00	0.63	20.00
6	161	161	200	100.00	100.00	99.88	0.00	0.00	23.13
7	161	161	164	100.00	100.00	99.88	0.00	0.00	21.88
8	180	183	180	94.88	96.63	98.25	0.00	0.63	42.50
10	183	185	171	80.88	81.13	76.88	0.00	0.00	15.63
11	169	169	193	64.00	66.25	56.38	0.00	0.00	20.63
12	162	163	190	99.38	99.63	95.75	0.00	0.00	36.88
13	203	201	215	95.13	95.25	97.13	0.00	0.00	5.63
14	162	162	-	99.88	99.88	40.75	0.00	0.00	26.25
16	169	167	175	86.75	86.00	70.63	0.00	0.63	51.88
17	182	182	211	95.88	96.25	61.25	0.63	0.00	62.50
18	244	244	265	89.75	89.75	88.13	0.00	0.63	43.13
19	171	171	327	85.00	84.75	53.63	0.00	0.00	28.13
20	225	225	240	90.63	77.375	85.38	0.00	0.00	28.75
21	618	649	684	44.25	39.13	78.50	0.63	1.25	52.50
Average	-	-	-	90.06	89.33	78.76	0.10	0.28	28.09

As shown in Table 1, it is a comparison of fault alarm points, fault detection rate (FDR) and false alarm rate (FAR) of three residual-based process monitoring methods. It can be seen from Table 1 that for most faults, the monitoring performance of PLS-PCA and SVM-PCA algorithms is better than that of LSTM-PCA. Among them, the PLS-PCA algorithm has the highest average FDR for 18 faults and the lowest false alarm rate.

In residual generation part, the prediction errors of 28 variables can be obtained through their corresponding regression models. The average root-mean-square error (RMSE) of PLS, SVM and LSTM for these 28 variables are 0.465, 0.466 and 0.035, respectively. Among them, the average RMSE of LSTM is the lowest, however, the monitoring performance of residual-based LSTM-PCA model is not the best. This indicates that when the prediction model is only selected by lowest prediction error, the performance of the monitoring model is not necessarily the best.

When the process is under normal conditions, the prediction model residuals should be the noise in data, which should conform to the normal distribution. Quantile-Quantile plot is a method to compare two probability distributions by plotting their quantiles against

each other(Huang, Qiao, Liu, Liu, & Dai, 2019). Through the analysis for residuals of all variables, it shows that the probability of most residuals obtained from PLS and SVM consisting with the normal distribution is higher than the probability of residuals obtained from LSTM. Taking the variable stripper underflow (stream 11) as an example, as shown in Figure 2 (the Q-Q plots), compared to residuals from LSTM, points of residuals obtained from PLS and SVM are closer to the diagonal, which illustrates that the prediction residuals from PLS and SVM is more consistent with normal distribution.

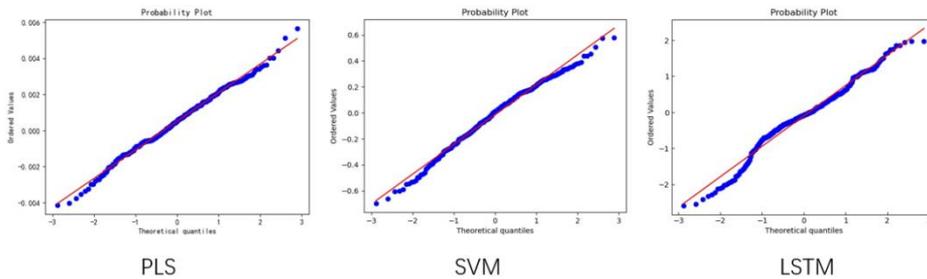


Figure 2. The Quantile-Quantile plots of the regression model in TE process

4.2. Industrial case

Data from continuous catalytic reforming unit are utilized to further illustrate the relationship between prediction model residuals and monitoring performance. In this case, 4000 sample points are selected to establish residual-based model. It is divided into three parts: 2000 sample points as training data for prediction model, 1000 sample points as verifying data for prediction model and 1000 sample points as testing data for prediction model. 2500 fault points, including a reactor temperature abnormal change, are collected as testing data for the comparison of monitoring performance.

Table 2. Alarm point in the industrial case

Methods	PLS	SVM	LSTM
Alarm point	204	191	210

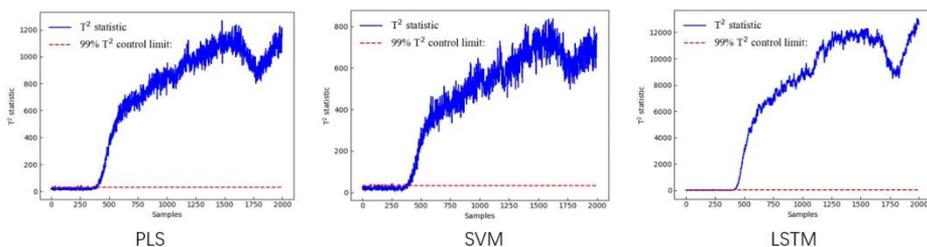


Figure 3. Monitoring results in the industrial case

As shown in Table 2 and Figure 3, it is a comparison of fault alarm points of three residual-based process monitoring methods. It can be found that SVM-PCA detects this fault at 191th sample which has a significantly earlier alarming time compared with PLS-PCA and LSTM-PCA.

Taking the cold-end outlet temperature of heat exchange as an example, as shown in Figure 4, compared to PLS and LSTM, points of residuals obtained from SVM are closer to the diagonal, which illustrates that the prediction residuals from SVM is more consistent with normal distribution.

The analysis results of two cases above states that: (1) The prediction error of the prediction model isn't the dominant factor in determining the monitoring performance, and the prediction model with the lowest prediction errors can't achieve best monitoring performance necessarily; (2) Through the research on the residual distribution of three representative prediction methods, the consistency of the residual distribution and the normal distribution has a certain corresponding relationship with the monitoring performance, which could be a factor to select the prediction model for residual-based monitoring method.

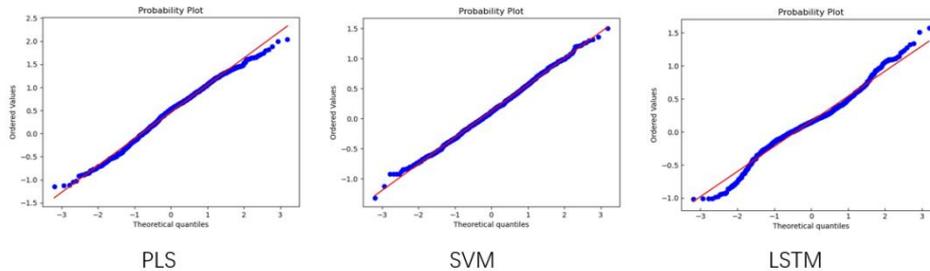


Figure 4. The Quantile-Quantile plots of the regression model in the industrial case

5. Conclusions

The influence of prediction models, evaluation indicators, distribution of prediction residuals on monitoring performance are discussed in this paper. Through the comparison of these three models, it can be seen that in addition to the traditional evaluation indicators, whether the residuals are normally distributed is also important indicators that affect the monitoring performance. The above factor should be considered in the selection and establishment of the prediction model for the residual-based monitoring method.

References

- Nor, N. M., Hassan, C. R. C., & Hussain, M. A. J. R. i. C. E, 2020, A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems, 36(4), 513-553.
- Tong, C., Lan, T., Yu, H., & Peng, X. J. J. o. P. C. 2019, Distributed partial least squares based residual generation for statistical process monitoring, 75, 77-85.
- Kazemi, P., Giralt, J., Bengoa, C., Steyer, J.-P. J. W. S., & Technology, 2020, Data-driven fault detection methods for detecting small-magnitude faults in anaerobic digestion process, 81(8), 1740-1748.
- Tao, L., Liu, B., & Liang, W. J. M. P. i. E, 2021, Automated Detection of Arrhythmia for Hybrid Neural Network of LSTM-Residual with Multi-Information Fusion, 2021.
- Pirouz, D. M. J. A. a. S, 2006, An overview of partial least squares.
- Awad, M., & Khanna, R, 2015, Support vector regression. In *Efficient learning machines* (pp. 67-80): Springer.
- Hochreiter, S., & Schmidhuber, J. J. N. c, 1997, Long short-term memory. 9(8), 1735-1780.
- Pearson, K. J. T. L., Edinburgh,, magazine, D. p., & science, j. o, 1901, LIII. On lines and planes of closest fit to systems of points in space. 2(11), 559-572.
- Y. K. Awad, M., & Khanna, R, 2015, Support vector regression. In *Efficient learning machines* (pp. 67-80): Springer.
- Chiang, L. H., Russell, E. L., & Braatz, R. D, 2000, *Fault detection and diagnosis in industrial systems*: Springer Science & Business Media.
- Huang, K.-W., Qiao, M., Liu, X., Liu, S., & Dai, M. J. a. p. a, 2019, Computer Vision and Metrics Learning for Hypothesis Testing: An Application of QQ Plot for Normality Test.

Optimal operation of parallel mini-bioreactors in bioprocess development using multi-stage MPC

Niels Krausch,^a Jong Woo Kim,^a Sergio Lucia,^b Sebastian Groß,^c Tilman Barz,^{a,d} Peter Neubauer,^a Mariano N. Cruz Bournazou,^{a,e}

^a*Technische Universität Berlin, Bioprocess Engineering, 13355 Berlin, Germany*

^b*Technische Universität Dortmund, Process Automation Systems, 44227 Dortmund, Germany*

^c*wega Informatik (Deutschland) GmbH, 79576 Weil am Rhein, Germany*

^d*AIT Austrian Institute of Technology GmbH, Center for Energy, 1210 Vienna, Austria*

^e*DataHow AG, 8600 Dübendorf, Switzerland*

Abstract

Bioprocess development is commonly characterized by long development times, especially in the early screening phase. After promising candidates have been pre-selected in screening campaigns, an optimal operating strategy has to be found and verified under conditions similar to production. In view of the large number of strains and the process conditions to be tested, high-throughput cultivation systems provide an essential tool to sample the large design space in short time. Furthermore, cultivating cells with pulse-based feeding and thus exposing them to oscillating feast and famine phases has shown to be a powerful approach to study microorganisms closer to industrial bioreactor conditions. We have recently presented a comprehensive platform, consisting of two liquid handling stations coupled with a model-based experimental design and operation framework to increase the efficiency in High Throughput bioprocess development. Using calibrated macro-kinetic growth models, the platform has been successfully used for the development of scale-down fed-batch cultivations in parallel mini-bioreactor systems. However, it has also been shown that parametric uncertainties in the models can significantly affect the prediction accuracy and thus the reliability of optimized cultivation strategies. To tackle this issue, we implemented a multi-stage Model Predictive Control (MPC) strategy to fulfill the experimental objectives under tight constraints despite the uncertainty in the parameters and the measurements. Dealing with uncertainties in the parameters is of major importance, since constraint violation would easily occur otherwise, which in turn could have adverse effects on the quality of the heterologous protein produced. Using the feedback information gained through the evolution along the tree, the control approach is significantly more robust than standard MPC approaches without being overly conservative. We show in this study that the application of multi-stage MPC can increase the number of successful experiments, by applying this methodology to a mini-bioreactor cultivation operated in parallel.

Keywords: high-throughput, model-predictive control, scale-down, multi-stage

1. Introduction

1.1. High-throughput bioprocess development

The development of a process in biotechnology is enormously time-consuming. Not only does it take a long time to find a suitable strain from a huge collection of strains, but the identification of optimal process conditions for this strain also takes a lot of time and resources. Especially due to the increased demand for resource-saving and environmentally friendly bioproducts, the requirements for faster process development are also increasing. Miniaturization and parallelization are two recent technologies that can accelerate throughput in screening processes and process development (Hemmerich et al., 2018). In particular, the use of liquid handling stations and the application of computer-aided and model-based methods could make a decisive contribution to holistic and faster process development (Hans et al., 2020). These methods have already been used for optimal experimental re-design as well as process insights and faster strain phenotyping (Anane et al., 2019).

1.2. Implications of MPC for optimal automated bioprocess development

Fed-batch is still the most widely used process strategy in bioprocessing. Here, the change of the feeding rate plays an important role. Classically, feeding sequences after the batch phase follow an exponential course, based on a previously defined growth rate. However, it can easily happen that the selected growth rate is either too low and thus the process takes a long time to reach a certain cell density or that the feeding rate was selected too high and oxygen limitation can occur. This is a problem for various processes, as it causes unwanted stress reactions and can significantly influence the yield (Baez & Shiloach, 2014). Model predictive control is a method that has been widely used in the chemical industry and has also shown good success in several simulation studies in bioprocess engineering but has had little application in real experiments (Rathore et al., 2021). This is mainly due to the fact that good estimates for the underlying parameters must be available for the model and that the inherent system dynamics in microbial cultivations with e.g. *E. coli* are much different compared to other cultivation systems like mammalian cell cultures. Finding good values for this is often time consuming and the values are subject to large uncertainties. We show here our implementation of a combination of moving horizon estimation for parameter estimation and multi-stage model predictive control. This strategy provides rough estimates for the parameters for new strains where no information is available in a short time, and multi-stage MPC is then used to compute an optimal feeding regime while explicitly taking into account the uncertainty of such parameters. As more information, in the form of measurements, is gathered, the parameter estimation can be improved and the multi-stage MPC can include this information online to improve the performance while ensuring robust process control and avoiding oxygen limitation. Since the dissolved oxygen tension (DOT) is measured with a first order delay, it is not possible to control the process optimally by means of a PID controller, since the violation of the constraint always occurs after a glucose pulse has been added. In this respect, a model-based control with prediction of whether a selected pulse satisfies the constraint is necessary. This is another reason to take the uncertainties of the parameter values into account, as this approach has proven to be significantly more robust than previous approaches (Lucia et al., 2013). By using our fully automated liquid handling station, we can quickly take many measurements to adapt our model well to the new strain and at the same time use the outputs of the MPC to optimally adjust the process online.

The multi-stage approach in particular makes it possible to still ensure robust process control even with inaccurate estimated values for the parameters.

2. Material and Methods

Initial testing experiments for the parameter estimation were conducted on our high-throughput bioprocess development platform, although the actual multi-stage experiments were performed *in silico*. The platform comprises two liquid handling stations (Freedom Evo 200, Tecan, Switzerland; Microlab Star, Hamilton, Switzerland), a mini bioreactor system (48 BioReactor, 2mag AG, Munich, Germany) and a Synergy MX microwell plate reader (BioTek Instruments GmbH, Bad Friedrichshall, Germany). With this setup, the platform is capable to perform up to 48 cultivations in parallel. DOT and pH are measured online, while the concentrations of the other state variables, i.e. glucose, acetate and biomass were measured at-line using enzymatic kits. The reader is referred to (Haby et al., 2019) for a detailed description of the facility and sampling procedure. The multi-rate sampling frequency of the different states poses another challenge for the optimization, as DOT is measured every 60 s, while the other state variables are only measured every 20 min.

The MHE part for parameter and initial state estimation can be generally formulated according to (1), subject to (2).

$$\min_{\theta, x_{0,r}} \frac{1}{2} \|x_{0,r} - x_{0,r,old}\|_{W_x}^2 + \frac{1}{2} \|\theta - \theta_{old}\|_{W_p}^2 + \sum_{k=0}^{N_{MHE}} \frac{1}{2} \|h(x_r(t), u_r(t), \theta) - y_{meas}(t)\|_{W_y}^2 \quad (1)$$

$$\text{s.t.} \\ \dot{x}_r(t) = f(x_r(t), u_r(t), \theta) \\ \theta_{min} \leq \theta \leq \theta_{max} \quad (2)$$

The estimate for the states at the initial point of the window are then denoted by $x_{0,r,old}$, the current parameter vector θ , the parameter vector from the previous horizon θ_{old} . The squared norm is applied to all calculations and the optimizer tries to minimize the deviations over the time window of length N_{MHE} of the measurements y_{meas} and the predicted outputs $h(x_r(t), u_r(t), \theta)$, considering the inputs $u_r(t)$ and current parameter vector. Each norm is further weighted by the factors W_x, W_p, W_y .

The optimization problem that is solved to compute the optimal inputs via MPC is:

$$\min_{u_r} -W_M X_r(t + N_{MPC}\Delta t) - W_L \sum_{k=0}^{N_{MPC}-1} X_r(t + k\Delta t) \quad (3)$$

$$\text{s.t.} \\ \dot{x}_r(t) = f(x_r(t), u_r(t), \hat{\theta}) \\ x_r(t_0) = \hat{x}_{0,r} \quad (4)$$

$$\text{DOT}_r(t) \geq 20\%, \quad u_r(t) \geq 3\mu\text{L}$$

Where W_v and W_L denote the weightings for the terminal- and the stage-cost, respectively. The goal was to maximize the biomass X_r in the shortest time while at the same time avoiding oxygen limitation, i.e. having a level of at least 20% dissolved oxygen in the medium. The macro-kinetic growth model used in this study hence consists of 6 ordinary differential equations (ODEs) including the measured oxygen as well as an

algebraic equation describing the actual oxygen, forming a system of differential-algebraic equations.

The problem was solved using the do-mpc software which utilizes orthogonal collocation on finite elements to discretize the system, which can then be solved using NLP optimizers (Lucia et al., 2017). One main challenge when operating this system is the non-continuous property of the inputs. Since glucose feeding is bolus-like, concentrations and volumes are changed at discrete time points. To deal with this, mass balances must be solved for the respective pulse additions, and it is mathematically more difficult to solve the system. The reader is referred to (Kim et al., 2021) for an in depth description of the underlying model and description of the optimization process. To consider the uncertainties in the parameters, a scenario tree is built. Here, each branch of the tree represents a possible combination of the uncertain parameter combinations. At each node, the optimization problem is solved again considering the current parameter uncertainty and an optimal trajectory is identified. When the parameter uncertainty is large, this approach yields much better constraint satisfaction than nominal MPC approaches. In this contribution, we tested the multi-stage approach to consider the uncertainties for a prediction horizon of 180 min and a robust control horizon of 10 min.

3. Results and discussion

Based on our previous work, where we showed the advantage of using MPC to obtain higher biomass while avoiding oxygen limitation compared to a predefined feeding regime (Krausch et al., 2020), we extended our work to account for the uncertainties in the parameters. Compared to the nominal MPC, multi-stage MPC offers the opportunity to consider the inherent uncertainties of the model parameters to be more robust. This is especially important to ensure that the defined constraints are not violated. Especially for processes with a large uncertainty in the parameters or difficult to predict dynamics, this offers a much better possibility of process control. In this setup, the framework was used to consider the uncertainties of three important model parameters, namely the maximum substrate uptake rate $q_{S,max}$, the yield coefficient of biomass per substrate excluding maintenance $Y_{XS,em}$ and the volumetric oxygen transfer coefficient $k_L a$. All these parameters can affect the oxygen consumption and hence easily lead to constraint violation. Choosing the right parameters in the scenario tree could be supported by using a subset selection method with sensitivity analysis or further data-driven approaches like PCA (Thombre et al., 2019). Many studies have shown the advantage of advanced control methods such as MPC, but have mainly been studying the nominal case for optimization (Chang et al., 2016; Ulonska et al., 2018). As depicted in Figure 1, the multi-stage MPC calculates a lower feeding rate compared to the nominal MPC, since it considers the uncertainties in the parameters. The nominal MPC can actually only ensure that the constraints are not violated if the parameters are well estimated and have very low uncertainty. The multi-stage approach, on the other hand, can identify an optimal feeding regime and account for the strength of uncertainties via a feedback loop. This is particularly advantageous for processes about which little information and measurement data are available like in early screening of new strains and therefore the parameter uncertainties are correspondingly large. Accordingly, taking these uncertainties into account leads to a lower feeding behavior in contrast to the nominal MPC, but to a much more robust process behavior. In particular, the system can react quickly in an unpredictable manner if the constraint is violated, since oxygen limitation also has a

strong influence on the expression of numerous genes and thus could strongly change the dynamics of the system.

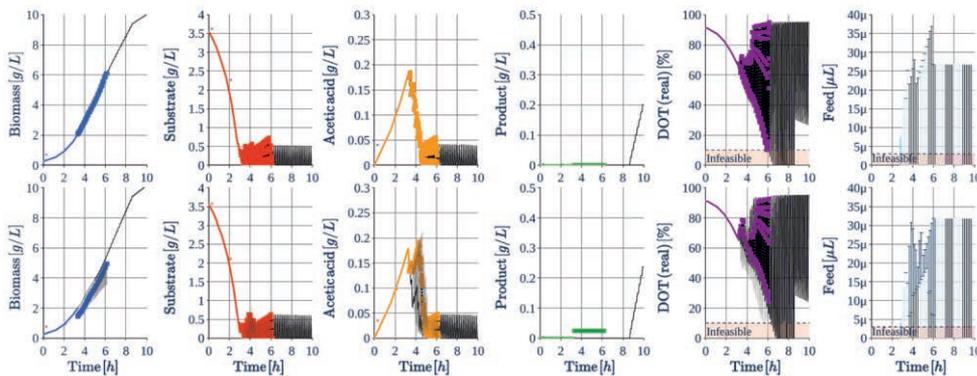


Figure 1: Comparison of nominal (upper) and multi-stage (lower) MPC at 3 h. The multi-stage approach guarantees constraint satisfaction even when the uncertainties are considered. Uncertainties for the DOT and feed are shown as shaded plots and error bars, respectively.

4. Conclusion and outlook

In this study, we have shown that a robust multi-stage MPC approach that takes into account the uncertainties in the parameters is even better than a nominal MPC. Violation of the constraint is a problem for many bioprocesses and there are various approaches to control bioprocesses. With the predictive approach, the constraints can be met well before they are violated, and the process can be guided to its optimum, so constraints are satisfied even in case of high uncertain parameters. However, the computational effort increases significantly with the number of parameters and the underlying uncertainties, so careful consideration must be given to which parameters the uncertainties should be considered for. Future work will need to deal with this problem by being more flexible in terms of what parameter uncertainties are considered and how large the prediction horizon should be, since the problem size is growing exponentially with a larger robust horizon or number of uncertain parameters to account for. Another aspect is to represent the product formation with the help of data-driven models, so that the process can also be optimized in this respect.

Acknowledgments

We thank Felix Fiedler for the development of the do-mpc software. We gratefully acknowledge the financial support of the German Federal Ministry of Education and Research (BMBF) (project no. 01QE1957C – BioProBot and 01DD20002A – 568 KIWI Biolab).

References

- Anane, E., García, Á. C., Haby, B., Hans, S., Krausch, N., Krewinkel, M., Hauptmann, P., Neubauer, P., & Cruz Bournazou, M. N [Mariano Nicolas] (2019). A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. *Biotechnology and Bioengineering*, 116(11), 2906–2918. <https://doi.org/10.1002/bit.27116>

- Baez, A., & Shiloach, J. (2014). Effect of elevated oxygen concentration on bacteria, yeasts, and cells propagated for production of biological compounds. *Microbial Cell Factories*, *13*, 181. <https://doi.org/10.1186/s12934-014-0181-5>
- Chang, L., Liu, X., & Henson, M. A. (2016). Nonlinear model predictive control of fed-batch fermentations using dynamic flux balance models. *Journal of Process Control*, *42*, 137–149. <https://doi.org/10.1016/j.jprocont.2016.04.012>
- Haby, B., Hans, S., Anane, E., Sawatzki, A., Krausch, N., Neubauer, P., & Cruz Bournazou, M. N [Mariano Nicolas] (2019). Integrated Robotic Mini Bioreactor Platform for Automated, Parallel Microbial Cultivation With Online Data Handling and Process Control. *SLAS Technology*, *24*(6), 569–582. <https://doi.org/10.1177/2472630319860775>
- Hans, S., Haby, B., Krausch, N., Barz, T., Neubauer, P., & Cruz-Bournazou, M. N. (2020). Automated Conditional Screening of Multiple Escherichia coli Strains in Parallel Adaptive Fed-Batch Cultivations. *Bioengineering*, *7*(4). <https://doi.org/10.3390/bioengineering7040145>
- Hemmerich, J., Noack, S., Wiechert, W., & Oldiges, M. (2018). Microbioreactor Systems for Accelerated Bioprocess Development. *Biotechnology Journal*, *13*(4), e1700141. <https://doi.org/10.1002/biot.201700141>
- Kim, J. W., Krausch, N., Aizpuru, J., Barz, T., Lucia, S., Martínez, E. C., Neubauer, P., & Bournazou, M. N. C. (2021). Model predictive control guided with optimal experimental design for pulse-based parallel cultivation. *ArXiv Preprint*. <http://arxiv.org/pdf/2112.10548v1>
- Krausch, N., Hans, S., Fiedler, F., Lucia, S., Neubauer, P., & Cruz Bournazou, M. N [Mariano N.]. (2020). From Screening to Production: a Holistic Approach of High-throughput Model-based Screening for Recombinant Protein Production. In *Computer Aided Chemical Engineering. 30th European Symposium on Computer Aided Process Engineering* (Vol. 48, pp. 1723–1728). Elsevier. <https://doi.org/10.1016/B978-0-12-823377-1.50288-3>
- Lucia, S., Finkler, T., & Engell, S. (2013). Multi-stage nonlinear model predictive control applied to a semi-batch polymerization reactor under uncertainty. *Journal of Process Control*, *23*(9), 1306–1319. <https://doi.org/10.1016/j.jprocont.2013.08.008>
- Lucia, S., Tătulea-Codrean, A., Schoppmeyer, C., & Engell, S. (2017). Rapid development of modular and sustainable nonlinear model predictive control solutions. *Control Engineering Practice*, *60*, 51–62. <https://doi.org/10.1016/j.conengprac.2016.12.009>
- Rathore, A. S., Mishra, S., Nikita, S., & Priyanka, P. (2021). Bioprocess Control: Current Progress and Future Perspectives. *Life*, *11*(6). <https://doi.org/10.3390/life11060557>
- Thombre, M., Krishnamoorthy, D., & Jäschke, J. (2019). Data-driven Online Adaptation of the Scenario-tree in Multistage Model Predictive Control. *IFAC-PapersOnLine*, *52*(1), 461–467. <https://doi.org/10.1016/j.ifacol.2019.06.105>
- Ułonska, S., Waldschitz, D., Kager, J., & Herwig, C. (2018). Model predictive control in comparison to elemental balance control in an E. coli fed-batch. *Chemical Engineering Science*, *191*, 459–467. <https://doi.org/10.1016/j.ces.2018.06.074>

A platform for online system identification of dynamic ultrafiltration systems

Oscar A. Prado-Rubio^{a,b}, Victor H. Grisales Diaz^c, Jakob Kjøbsted Huusom^b

^a *Departamento de Ingeniería Química, Universidad Nacional de Colombia – 170003 Manizales, Colombia*

^b *Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), DK-2800 Lyngby, Denmark*

^c *Universidad Libre Seccional Pereira, Facultad Ciencias de la Salud, Pereira, Colombia*
oaprador@unal.edu.co

Abstract

Dynamic Micro/Ultra-filtration technology has shown to be an intensified process able to provide improved selectivity and/or throughput required in challenging separations. However, there is a trade-off between operating conditions to achieve optimal operation which is time-variant and difficult to estimate. Herein, a platform for online system identification is built for a dynamic ultrafiltration system, where the Forgetting Factor Recursive Least Squared (FFRLS) method is investigated to describe transmembrane evolution at constant flux operation for a real wastewater pilot plant. The model predictions have around 4% average error, predicting relatively fast and slow system dynamics. Therefore, the model has a good permeability estimation useful to determine the membrane fouling rate in real-time.

Keywords: Online system identification, ARX model, Dynamic ultrafiltration, Membrane fouling rate modelling.

1. Introduction

Ultra and Micro-filtration are mature separation technologies widely used in water treatment, food, beverage, pharmaceutical and biotechnological industries. Membrane filtration processes have shown remarkable performance in terms of selectivity, throughput, low footprint and system modularity, operation at mild conditions, low chemical consumption, and relatively simple operation and automation. However, filtration technology is challenged to guarantee optimal stable operation under unmeasured high-variance inlet disturbances, which highly influence the system performance. The reason is two time-variant coupled phenomena referred to as concentration polarization and membrane fouling, which promote decay in the separation performance in terms of selectivity and/or throughput. To recover system performance, physical/chemical cleaning is forced, which substantially affects process operating costs and processing time (Baker, 2012).

To mitigate the adverse effect of concentration polarization and irreversible fouling, the dynamic operation has been proposed instead of conventional crossflow filtration (Prado-Rubio, O.A., 2012). This corresponds to a non-invasive physical technique where a portion of permeate is reversed using a high frequency – short duration backshock. This inverted flux can partially remove membrane fouling and disrupts the boundary layer

adjacent to the membrane surface. The latter avoids a high solute concentration at the membrane surface which reduces membrane rejection (López-Murillo *et al.*, 2021). The intensifying effect of the dynamic membrane operation has been shown in diverse applications, where overall performance can be improved significantly by reverting flux periodically (Prado-Rubio, O.A., 2012; López-Murillo *et al.*, 2021). As a drawback, process intensification can be only accomplished by a specific combination of operating conditions. Those are relatively complex to determine due to the system dynamic and unmeasured disturbances. Thus, to obtain an economically attractive operation scenario, the filtration unit must adapt to the operating conditions relaxing and tightening the driving force along with the operation. As expected, massive experimental work is required to establish a sustainable operation in real industrial conditions and sometimes operating conditions cannot be reliably estimated (Prado-Rubio, O.A., 2012).

Therefore, it is relevant to explore different modelling approaches for dynamic ultrafiltration that allows exploiting PSE tools to improve process design and operation. Recently, we have investigated off-line modelling approaches to describe dynamic membrane filtration, including scheduled linear time-invariant (LTI)-SISO and MISO autoregressive models, neural networks, and hybrid approaches (Prado-Rubio and Huusom, 2015; Prado-Rubio and Von Stosch, 2017; Grisales Díaz *et al.*, 2017). Those approaches have shown to be useful for process design but have limitations for optimal operation/control. Therefore, there is still a need for tools to determine the membrane state during real operation, thus, to have a chance to optimize the relatively high energy demanding system. The purpose of this contribution is to implement and investigate a platform to describe the dynamic filtration system based on online system identification. The platform aims to provide a real-time representation of the dynamic ultrafiltration system, focused on the fouling rate represented by the membrane permeability evolution. For testing, the platform uses a database of experimental data which are provided in real-time, and the Forgetting Factor Recursive Least Squared (FFRLS) method to correlate operating conditions when the system is operated at constant membrane flux. Thus, an estimation of the membrane permeability is performed to indicate the membrane fouling rate. This contribution corresponds to a preliminary effort to investigate the modelling approach for dynamic ultrafiltration which later can be implemented in a pilot plant.

2. Methodology

The platform has been implemented in Matlab® 2019b in three modules. The platform emulates the real system by using a database from experiments of dynamic filtration at field conditions, which have been exported from .txt (raw data from the PLC) to a .mat file. Accounting for the large sequential experiments database, continuous long-term system operation is emulated combining data from different operation scenarios. Within the algorithm, the first step is to pre-process the information for subsequent system identification including a data reconciliation strategy, and emulate real sampling time data access. Finally, the platform uses a custom-made function for Recursive Least Squared parameter estimation with forgetting factor to correlate the membrane flux and the manipulable transmembrane pressure, thus estimating the membrane permeability. Then model quality is determined and subsequently membrane permeability is predicted and analyzed.

2.1. Experimental data – produced water treatment using dynamic ultrafiltration

Produced water is wastewater in the oil and gas industry after most of the crude oil has been separated into an oil field. Produced water is a complex emulsion mainly composed of organic matter, oils, microorganisms, heavy metals, chemicals, salts, dissolved oxygen, among others. Produced water treatment is challenging accounting for the complex mixture and the time-variant nature depending on the oil well maturity and location. Membrane technology has been tested for produced water management since conventional specific gravity-based technologies have difficulties coping with the stricter environmental legislations for water disposal. Dynamic ultrafiltration experiments were performed in a real produced water treatment plant, where effluents from different sections of the processing facility were fed to the membrane system (Prado-Rubio, 2012). The membrane system was operated at constant flux using the step-up/step-down methodology (Beier & Jonsson, 2009). Under this operation strategy, the transmembrane pressure is increased during operation to compensate for the increasing transport resistance due to fouling. The pilot plant records data from flow rates, absolute pressure, and temperature every 3 seconds. Inlet and outlet water quality was characterized offline using spectrophotometry. During the long-term experiments, inlet water varied between $30 < \text{Total Suspended Solids (TSS)} < 100 \text{ mg/l}$ and $3 < \text{Total Hydrocarbons (TH)} < 1200 \text{ mg/l}$. The membrane treatment removed above 96% of TSS and 94% of TH. For this application, a subset of 26 experiments is employed to test the developed tool, those experiments contain approximately 22000 data points for every variable.

2.2. Signal pre-processing

In previous work, a signal processing algorithm has been developed which allows reconstructing frequency and amplitude of the *in-situ* cleaning strategies namely backshock and backflush (Prado-Rubio and Von Stosch, 2017). This was necessary due to the induced signal delays to avoid potential overlapping of both cleaning strategies and fast dynamics of the backshock that cannot be captured by the used sampling time in the PLC. This signal reconstruction algorithm has shown to be useful for off-line system identification and therefore, it is incorporated in the platform. However, the user can define if these reconstructed signals are used for online process identification. Additionally, in this emulated application, the user defines which experiments are used for testing. Taking advantage of the large database, different operation scenarios have been evaluated pilling nonconsecutive experiments to assess the online system identification methodology and predictions. During this investigation, scenarios from single experiments up to 7 pilled experiments were evaluated.

2.3. Online system identification

System identification tools are well-established within process control including offline and online strategies. Forgetting Factor Recursive Least Squared (FFRLS) is the most used method for system parameter identification since it is simple and stable. The forgetting factor adjusts the proportion of old and new data employed so the proportion of old data is reduced when new data is available. This algorithm has a fast convergence to the actual value (Diniz, 2013). FFRLS method is employed to correlate the evolving transmembrane pressure to operate the dynamic filtration unit at constant membrane flux, where the membrane flux is controlled through a PI control. The linear dependency of

both variables is evident in the pore-flow model, where the membrane flux (J) is proportional to the transmembrane pressure (TMP) (Baker, 2012):

$$J = L_p(TMP) \quad (1)$$

The membrane permeability (L_p - L/m² h bar) is expected to decrease during operation due to the increase in transport resistance generated by concentration polarization and fouling. Permeability evolution is a function of the operating conditions such as the inlet concentrations, pH, temperature, ionic strength, membrane, flow rates, pressures, cleaning strategies, among others. Based on the pore-flow model, it is natural to use a simple ARX model to correlate variables. ARX has been applied for offline system identification of this system previously. For preliminary modelling approach assessment, FFRLS has been implemented to estimate the TMP signal as output using the membrane flux as an inlet. In this way, the model can predict the membrane permeability evolution as an indirect indication of the membrane fouling rate.

As inputs, the user defines the model order, and the forgetting factor is a tunable parameter. The implementation was validated using the recent functions included in Matlab® for online system identification. The estimation quality is evaluated against the available experimental data and moving average method.

3. Results

As mentioned previously, different combinations of the experiments were investigated to determine the modelling approach capabilities. As an example, a particular situation is depicted in this contribution: operation under low fouling rate, followed by high fouling rate experiment and returned to low fouling rate operation. This case is particularly interesting since it contains subcritical and supercritical flux operations. Simulations were performed for a sequence of polynomials order (from 1 up to 4) and exploring the influence of the forgetting factor. First, it was evidenced that the online system identification procedure was substantially faster than the sampling time, allowing to have faster sampling time in forthcoming experiments (e.g., every second). Secondly, increasing the order of the polynomials led to a high correlation between parameters without improving the model quality. Therefore, for illustration purposes, first-order models were sufficient to match previously obtained results with higher-order MISO offline system identification (Prado-Rubio and Von Stosch, 2017).

In Figure 2, the quality of the model predictions is depicted. It should be mentioned that the estimation is performed in real-time, thus, the figure is animated. Conventionally, forgetting algorithms have shown to be suitable for tracking slow dynamics while having more difficulties describing fast changes. Interestingly, the online ARX model can reproduce most of the fast behaviour associated with the backshock and backflush, while following the slow dynamics associated with the membrane fouling. This indicates that the identified model has a high degree of experimental data reproduction under different operating strategies. Most of the data have around 4% average prediction error, while there are some points with up to 10%.

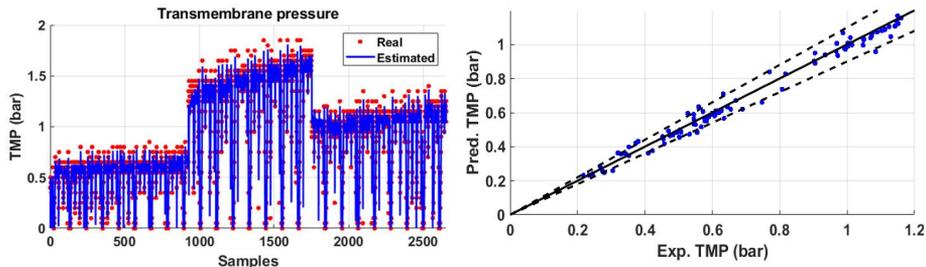


Figure 2. Simulation results of the real experimental data and model predictions. (a) full data set and (b) parity diagram for TMP using moving window average with 20 samples plus $\pm 10\%$ error.

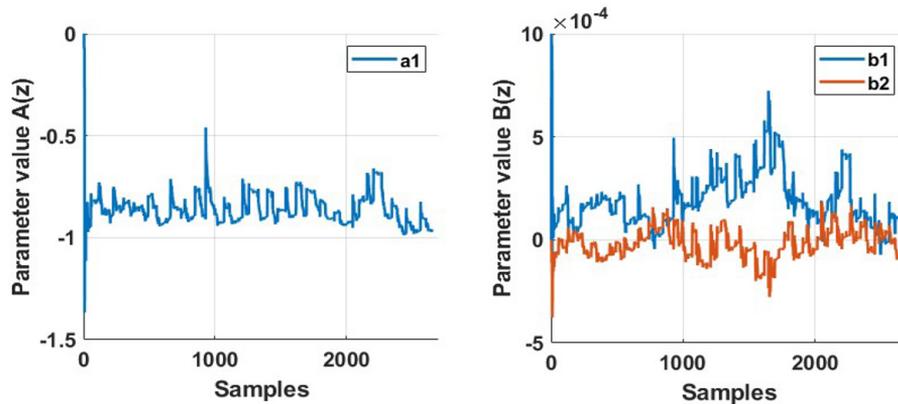


Figure 3. Adaptive parameters evolution to reproduce different fouling rate experiments.

The adaptive evolution of the parameters can be seen in Figure 3, where $A(z)$ and $B(z)$ are the polynomials for the backward shift operator z^{-1} , defining the system transfer function. As expected from the pore-flow model, the changes in the membrane permeability are tracked by B . Despite the lack of interpretability of those parameters, interestingly, $b1$ captures the increase of the fouling rate evidenced in the transmembrane pressure. Besides, the parameter $b2$ might not be entirely necessary, thus it could be omitted, and the system could be represented by even a simpler first-order discrete model.

Finally, Figure 4 depicts the predicted average permeability and the average estimated from the raw data. The online model provides a clearer representation of the membrane permeability, where it is evident the trend which indicates the fouling rate, instead of using the instantaneous estimation based on the raw data. Although *in situ* cleaning was performed, a permeability decrease was observed in all windows studied (at low or high TMP). According to the operation strategy, the controlled operation increases the transmembrane pressure to compensate for the adverse influence of fouling. At low flux operation (subcritical) the TMP increment is tolerable and associated with reversible fouling, increasing the desired flux can generate undesirable irreversible fouling. The threshold between both operation windows is challenging to estimate due to its time-variant nature, but this approach has shown that can estimate the average membrane permeability instead of the instantaneous permeability, which can be used to define an adaptive controller to avoid excessive membrane fouling during operation.

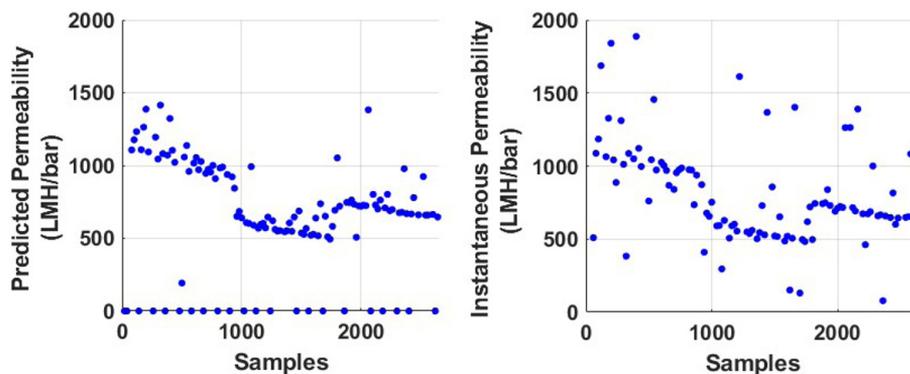


Figure 4. Comparison between predicted average permeability and average instantaneous permeability from raw data

4. Conclusions

Despite the simplicity of dynamic membrane filtration systems, its modelling has shown to be challenging accounting for the time-variant nature of the involved transport phenomena. Herein, a tool for online system identification has been implemented to allow process engineers to investigate dynamic ultrafiltration modelling. This platform is used to predict changes in transmembrane pressure (TMP) with an averaged error lower than 4%. Thus, it provides a better membrane permeability estimation compared to previous efforts with off-line system identification or using raw data to have an instantaneous calculation. This estimation provides an indirect estimate of the membrane fouling rate. These preliminary tests are encouraging to pursue an implementation in a real system to introduce adaptive control actions into the system operation, which will have a significant impact on the operation of membrane-based treatment systems.

References

- Baker, R.W. Membrane technology and applications. 3rd ed. John Wiley & Sons, Ltd, 2012.
- Beier, S.P. and Jonsson, G. Critical Flux Determination by Flux Stepping. *AIChE Journal*, 56(7), 1739-1747, 2009.
- Diniz, P.S.R. Fundamentals of adaptive filtering. *Adaptive Filtering*; Springer, pp. 13–78, 2013.
- Luis Humberto López-Murillo, Víctor Hugo Grisales-Díaz, Manuel Pinelo, Oscar Andrés Prado-Rubio. Ultrafiltration intensification by dynamic operation: insights from hybrid modelling. *Chemical Engineering & Processing: Process Intensification*. Volume 169, 108618, 1-13, 2021.
- Oscar Andrés Prado-Rubio and Moritz von Stosch. Towards Sustainable Flux Determination for Dynamic Ultrafiltration through Multivariable System Identification. *Computer-Aided Chemical Engineering*, volume 40, 2719-2724, 2017.
- Prado-Rubio, O.A. SiC Membrane Pilot Ultrafiltration Test for Produced Water Treatment. Technical report - LiqTech international S/A, 2012.
- Víctor Hugo Grisales Díaz, Oscar A. Prado-Rubio, Mark J. Willis, Moritz von Stosch. Dynamic hybrid model for ultrafiltration membrane processes. *Computer-Aided Chemical Engineering*, volume 40, 193-198, 2017.

Development of a Virtual Sensor for Real-Time Prediction of Granule Flow Properties

Rexonni B. Lagare,^a Mariana Araujo da Conceicao,^a Ariana Camille Acevedo Rosario,^a Katherine Leigh Young,^a Yan-Shu Huang,^a M. Ziyen Sheriff,^a Clairmont Clementson,^b Paul Mort,^c Zoltan Nagy,^a Gintaras V. Reklaitis^{a b}

^a*Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA*

^b*Purdue Center for Particulate Products and Processes, West Lafayette, IN 47907, USA*

^c*School of Materials Engineering, Purdue University, West Lafayette, IN 47907, USA*

**rlagare@purdue.edu*

Abstract

We report progress of an ongoing work to develop a virtual sensor for flowability, which is a critical tool for enabling real time process monitoring in a granulation line. The sensor is based on camera imaging to measure the size and shape distribution of granules produced by wet granulation. Then, statistical methods were used to correlate them with flowability measurements such as ring shear tests, drained angle of repose, dynamic angle of repose, and tapped density. The virtual sensor addresses the issue with these flowability measurements, which are based on off-line characterization methods that can take hours to perform. With a virtual sensor based on real-time measurement methods, the prediction of granule flowability become faster, allowing for timely decisions regarding process control and the supply chain.

Keywords: virtual sensor, monitoring, flowability, machine learning, size and shape distribution

1. Introduction

The manufacturing of tablets often relies on a granulation step to improve the processability of a pharmaceutical powder blend. By converting them into granules, flowability, tabletability, compressibility, and compactibility can be improved. These properties are considered to be the critical quality attributes (CQA) of the granulation unit; and in a wet granulator, the monitoring of these properties is considered critical. Unfortunately, characterization tests for these CQAs are usually off-line methods that can take hours to measure and require sample reduction procedures that can lead to significant sampling errors. It is thus important to develop faster ways to estimate the CQAs and minimize sampling error.

In a wet granulation platform employing a fluidized bed, it is possible to measure the size and shape distribution of a finished batch of granules as it discharges from the product hopper. If these real-time measurements of size and shape can be used to automatically predict the CQA of the discharging granules, decisions regarding the batch and the process could be made much faster. This time advantage could save future batches from failure, provide valuable information about the raw material, and allow optimization of

the control process parameters of downstream unit operations to match the characteristics of each batch of granules.

Particle size and shape distributions are known to be strong indicators of granule properties, so they have a great potential to be reliable predictors of a granule's CQA's. However, their measurement result in a large number of data points that are difficult to manage and process. In practice, these distributions often get reduced to 1 to 3 D-values (i.e., D10, D50, D90) prior to analysis. There has been demonstrated success in this strategy, but this practice can lead to significant loss of information from the dataset, especially when the distributions have statistical central tendencies that do not fall close to those selected D-values. An ideal solution would be to employ all available information from the size and shape distribution measurements, and then to use an appropriate data reduction technique that maximizes the relevant information from the distribution measurements. By implementing this with an appropriate feature extraction technique, the reduced dataset should maximize correlatability with properties of interest such as flowability.

Aside from the predictor variables (i.e., size and shape distribution), the need for data reduction and feature extraction is also applicable for the predicted variables (i.e., the CQA's), especially for flowability. Since there is no singular measure for it, several methods exist to characterize flowability. Often, the goal in characterization is to select the method with test conditions that can closely match the conditions to which the granules are subjected to during processing. For some applications, a single method might suffice. But if granules will be subjected to tablet pressing, they will be subjected to quasi-static flow conditions in the hopper of the tablet press, as well as dynamic flow conditions inside the feed frame. Hence, several methods are required to ensure that the granules would result in quality tablets. Furthermore, each of these methods produce multiple test result parameters that are highly correlated. This can potentially result into a large dataset that needs to be appropriately reduced to make it more manageable and maximize its predictability with real-time measurements of granule size and shape distributions.

2. Methods

2.1. Data Reduction: Principal Component Analysis (PCA)

PCA is a method that reduces the dimensionality of large datasets while retaining most of its information. This is achieved by taking an orthogonal decomposition of the covariance matrix of process variables along the directions that explain the maximum variation of the data. (Wold *et al.*, 1987) While this method gives the same number of principal components as the original variables in the dataset, it also puts maximum possible variance in the first few principal components, making it possible to drop the rest of the principal components without losing much information. With the appropriate selection of principal components, data analysis and exploration can be performed on lower number of dimensions.

2.2. Latent Variable Regression

2.2.1. Linear Regression: Partial Least Squares (PLS)

With both the predictor and the predicted variables requiring data reduction through PCA, linear regression on their projections to latent spaces can be performed (i.e., projection to their principal components). This process is known as Partial Least Squares, and it is a

widely used technique in areas such as chemometrics bioinformatics, neurosciences, and sensor development, to name a few.(Liu and Chen, 2014)

3. Materials and Equipment

3.1. Granules

The granules used in this study are made with varying compositions of lactose and microcrystalline cellulose as the excipient, acetaminophen (APAP) as the active pharmaceutical ingredient (API), and either hydroxypropyl cellulose (HPC) or polyvinylpyrrolidone as the liquid binder solution. The excipient, API, and binders are prepared in varying compositions and wet granulation process conditions to produce granules with different flowability characteristics. This work studied four types of granules labelled as: HHIU1, HHIU2, HHIU3, and HHIU4.

3.2. Granulation Equipment

The granules are produced by wet granulation using the Xelum platform manufactured by Syntegon. Xelum employs a fluidized bed, where the pharmaceutical powders are automatically dosed and pneumatically charged with the liquid binders that facilitate the formation of granules. Moreover, granulation and drying takes place in the same process chamber, which eliminates the need to transfer wet granulate and improves the system's reliability.

3.3. Size and Shape Distribution Measurement

The size and shape distribution of the granules are measured using Eyecon₂, which is a direct imaging particle analyzer developed by Innopharma Technology. By using a camera to take images of the particles at-line or inline, this tool uses image analysis algorithms to detect particle boundaries and fit an ellipse around them. The ellipse gives a major and a minor diameter, which when averaged gives a third dimension to estimate a 3D volume of the particle using the equation:

$$Volume = \frac{\pi}{6} \times D_{min} \times D_{max} \times D_{ave} \quad \text{Equation 1}$$

Using this volume, an equivalent spherical diameter is computed, and this diameter is the basis for the size distribution reported by Eyecon₂. Size distributions are reported as D-values, which are based on the cumulative size distribution. Reporting distributions in this manner fixes the number of variables for every possible form of size distributions.

The major and minor diameters of each particle are also reflective of its shape, which may be quantified as eccentricity, as shown in the following equation.

$$Eccentricity = \sqrt{1 - \left(\frac{D_{min}}{D_{max}}\right)^2} \quad \text{Equation 2}$$

Eyecon₂ inherently acquires a distribution of eccentricity/shape but reports, by default, the distribution as a mean and relative standard deviation.

3.4. Flowability Measurements

The set of flowability measurements employed in this study covers both quasi-static flow and dynamic flow. Quasi-static flow is characterized by the ring shear tester and partly by tapped density analysis, while dynamic flow is characterized by drained and dynamic angle of repose.

3.4.1. Ring Shear Tester (RST)

The Schulze ring shear tester is an essential tool for hopper design since it is mainly concerned with quasi-static flow. In this technique, powder is loaded normally to a specific bulk density and then sheared until the material begins to flow. Data is collected as yield strength as a function of normal stress. From these measurements, the flow function coefficient can be computed, which may also be referred to as flowability. Additionally, other parameters such as internal friction, wall friction, and bulk density can be determined from the Schulze RST.

3.4.2. Drained Angle of Repose and Jamming Onset

The drained angle of repose is measured using the Flodex™ tool, which essentially measures the ability of a powder to fall freely under gravity through an orifice. Initially, the powder is contained in a hopper with a flow disk at the bottom. The disk has an orifice that can be opened via a discharge valve to start the powder flow. After opening the discharge valve, not all the powder in the hopper would be able to flow out and this residual powder would remain between the edge of the orifice and the hopper walls. The angle between the surface of this residual powder and the orifice disk is called the drained angle of repose and is correlated with the flowability of the powder.

3.4.3. Dynamic Angle of Repose

The dynamic angle of repose is measured using a rotary drum developed by GranuTools called the GranuDrum™. The powder is loaded into a drum that can be rotated at a set rotating speed. As the drum is rotated from rest, the angle of the powder surface increases from horizontal until an avalanche occurs. The angle at which this happens may be referred to as the yield point and is correlated with flowability. Thereafter, the powder surface is maintained at an angle from horizontal, and this is recorded automatically using back-lit cameras as the dynamic angle of repose. As the rotation speed of the drum changes, the dynamic angle of repose also changes, revealing interesting rheological behaviors of powder during flow.

3.4.4. Tapped Density Analyzer

Tapped density analysis is performed by another tool developed by GranuTools called GranuPack™. This tool minimizes operator error during filling and volume measurements using automation and sensor technologies. Powder is loaded onto a cylinder container and its density is monitored as the container is tapped continuously. As the powder is tapped, the density increases until it asymptotically approaches a maximum. The density may be expressed as the Hausner ratio, which is basically the ratio between the tapped density and the poured density. The dynamics of the compaction during tapping is also automatically captured via the parameters characteristic number and tau. The characteristic number is the number of taps at which the density is between between the poured density and the asymptotic density (i.e., density at infinite number of taps), while Tau is another characteristic number extrapolated from an exponential model (Philippe and Bideau 2003) fitted onto the compaction curve.

4. Results and Discussion

4.1. Principal Component Analysis on Size and Shape Distributions

Size and shape distributions measurements can result in at least 24 variables as shown in the x-axis of Figure 1. Applying principal components analysis (PCA) on the dataset reduced the number of variables into just 3 principal components (PC), which can explain up to 97% of the variance in the original dataset.

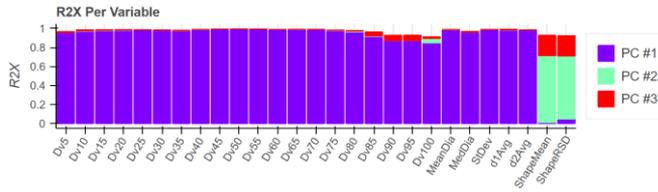


Figure 1. Explained variance per size and shape distribution variable

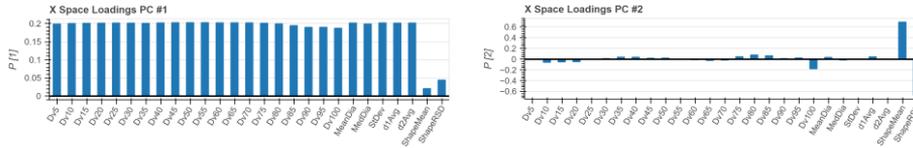


Figure 2. Loadings of size and shape distribution variables on principal component 1 (left figure) and principal component 2 (right figure)

This drastic reduction of variables suggests that many of them are highly correlated, as shown in the loadings plot in Figure 2. This is the case for the size distribution variables, as they dominate influence on the first principal component, which explains 89% of the variation in the original data. On the other hand, shape-related variables (i.e., the shape mean and relative standard deviation) have the strongest influence on the second principal component, supporting the importance of measuring shape distributions, and not just size.

4.2. Predicting Flowability Data

Measurements from the Schulze RST can lead to 9 different parameters (or variables) that are related to flowability. Similar to the Eyecon2 data, most of the variation in these parameters (up to 98%) can be explained by only three principal components. Hence, by applying partial least squares using three principal components onto the Eyecon2 data (as the predictor variables) and the RST data (as the predicted variables), the parity plot of the flow function coefficient (FFC) shown in Figure 3 show good prediction performance. Although not shown, similar performance was also observed for the rest of the variables.

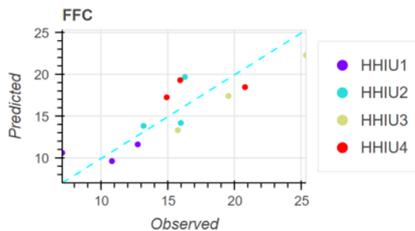


Figure 3. Predicted vs observed flow function coefficients measured from Schulze RST.

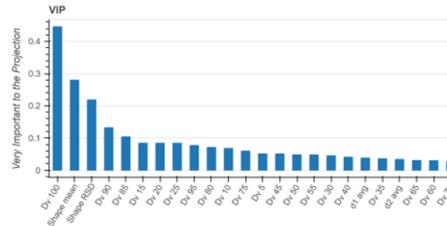


Figure 4. Ranking of variable importance to the PLS projections.

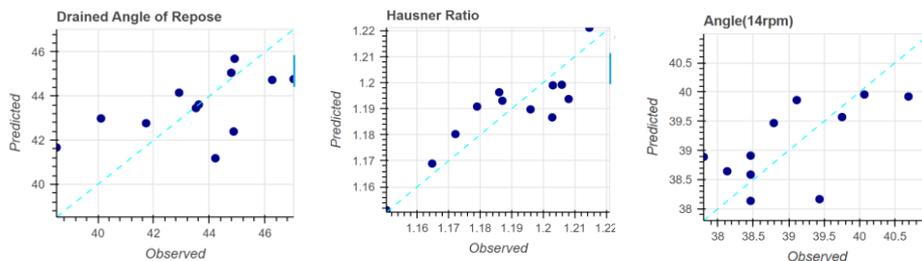


Figure 5. Parity plots for selected dynamic flow test parameters: drained angle of repose (left), Hausner ratio (middle), and dynamic angle of repose (right).

The performance of the PLS model can be attributed to the effectiveness of using all available information instead of selecting some and then ignoring the rest. Figure 4 shows how the larger D-values (e.g., D85, D90, and D100) and the shape parameters contribute the most to the PLS projections and hence its performance. This not only corroborates the importance of measuring shape distributions, but also the folly of selecting certain D-values such as D50, D10, and D90. As shown in Figure 4, those variables are not the most important. Using the aforementioned techniques, similar results were achieved from the dynamic flow tests, as shown in Figure 5 for selected parameters from the Flodex (left figure), GranuDrum (middle figure), and GranuPack (right figure) measurements.

5. Conclusions

Using PLS regression, sensor models were developed to predict flowability measurements based on size and shape distribution of granules, and parity plots show good predictability for all flowability measurements. The importance of shape measurements as well as using the complete size distribution, instead of selecting a few D-values, in the predictive performance was highlighted.

6. Acknowledgements

PCA and PLS computations were performed using the Phi toolbox (github.com/salvadorgarciamunoz/pyphi) for multivariate analysis by Dr. Sal Garcia.

This work was supported by United States Food and Drug Administration under grant 1U01FD006487-01.

7. References

- S. Wold, K. Esbensen, and P. Geladi. 1987. "Principal Component Analysis", *Chemometrics and Intelligent Laboratory Systems*, 2.1-3,37-52
- J. Liu, D.S. Chen. 2014. "Developing a Soft Sensor for an Air Separation Process Based on Variable Selection in Dynamic Partial Least Squares", *Computer Aided Chemical Engineering*, 33, 685-690
- P. Philippe, and D. Bideau. 2003. "Granular Medium under Vertical Tapping: Change of Compaction and Convection Dynamics around the Liftoff Threshold." *Physical Review Letters*, 91 (10): 104302.

A Hierarchical Approach to Monitoring Control Performance and Plant-Model Mismatch

M. Ziyen Sheriff,^a Yan-Shu Huang,^a Sunidhi Bachawala,^b Marcial Gonzelez,^{b,c}
Zoltan K. Nagy,^a Gintaras V. Reklaitis^a

^aDavidson School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, USA

^bSchool of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

^cRay W. Herrick Laboratories, Purdue University, West Lafayette, IN 47907, USA

zsheriff@purdue.edu

Abstract

Controllers are often tuned during plant commissioning, with a fixed process model. However, over time degradation can occur in the process, the process model and the controller, making it necessary to either re-tune the controller or re-identify the process model. Authors have proposed a variety of approaches to identify plant-model mismatch (PMM) and control performance degradation (CPD). While each approach may have its own advantages and disadvantages, they are generally designed to function on different timescales. The differing timescales result in the need for a multi-level hierarchical approach to monitor, detect, and manage PMM and CPD, as illustrated through a continuous pharmaceutical manufacturing application, i.e., a direct compression tablet manufacturing process. This work also highlights the requirement for index-based metrics, that enable the impact of PMM and CPD to be quantified and assessed from a control performance monitoring perspective, to aid fault diagnosis through root cause analysis to guide maintenance decisions for continuous manufacturing applications.

Keywords: control performance monitoring, plant-model mismatch, nonlinear model predictive control.

1. Introduction

The pharmaceutical manufacturing industry is being pushed to transition from batch to continuous process operation due to potential improvement in process controllability and product quality. Additional factors such as the development cost of new medicines makes it both desirable and feasible to produce smaller annual volumes of targeted dosages for smaller patient populations. Due to stringent regulations placed by regulatory bodies, the development of reliable real-time process monitoring, control and management approaches is of crucial importance, so that deviations in critical material (CMAs) and critical quality attributes (CQAs) can be minimized (Su et al., 2019). These include the need for efficient estimation and control frameworks, and algorithms to monitor these frameworks to identify and quantify plant-model mismatch (PMM) and control performance degradation (CPD). Quantification of PMM and CPD can in turn support higher level fault detection and diagnosis efforts.

Identification and management of PMM and CPD has received significant attention in the control literature. PMM can arise in the continuous manufacture of oral solid dosage for several reasons, e.g., the feeder refill step can introduce disturbances that can affect CMAs such as bulk density (Destro et al., 2021), and this can in turn result in deviations in the CQAs. A minimum variance-based assessment criterion was proposed by (Harris,

1989) to assess the condition of the working control loop but was limited to single-input-single-output (SISO) systems. More recently, partial correlation coefficient (PCC)-based and mutual information (MI)-based approaches were proposed by (Badwe et al., 2009; Chen et al., 2013) to identify PMM: both approaches are well-suited to handle cases where there is high correlation between manipulated variables. Advanced estimation and control strategies such as the moving horizon estimation-based nonlinear model predictive control (MHE-NMPC) framework have also been employed for continuous pharmaceutical manufacturing applications to handle the impact of PMM (Huang et al., 2021). While identification of PMM is important, quantifying the PMM and assessing its impact on control loop behavior will aid higher level decision making related to maintenance and safety. (Wang et al., 2012) proposed a control performance index (CPI) and loop robustness index (LRI), based on the integral absolute error and sensitivity margin to quantify PMM and CPD, respectively. Each of the methods described thus far are computed on different timescales, e.g., LRI requires identification of the transfer functions for the MIMO system and can only be carried out during regularly scheduled maintenance, while MI can be computed more frequently using closed operating data, and the MHE-NMPC framework is designed to operate on a significantly shorter timescale. Therefore, it is important to develop a multi-level hierarchical approach to utilize the quantitative information regarding PMM and CPD from different timescales, that will further support root cause diagnosis efforts and aid higher-level maintenance decisions for continuous pharmaceutical manufacturing applications.

To summarize, since the LRI which is based on the sensitivity margin is limited in applicability to SISO systems, this work seeks to extend its applicability by utilizing the disk margin proposed by (Seiler et al., 2020) for MIMO systems. This work also proposes a multi-level hierarchical framework to handle metrics that quantify PMM and CPD on different timescales to support higher level decision making related to safety and maintenance. Practical applicability will be demonstrated through an illustrative example that focuses on the continuous manufacture of oral solid dosage. The rest of this work is organized as follows. In Section 2, components of the hierarchical framework will be explained. An illustrative example using a rotary tablet press will be presented in Section 3, along with a discussion on the results. Concluding remarks will be presented in Section 4.

2. Methodology

The aim of this work is to propose an approach that enables efficient interpretation and management of quantitative information obtained from different metrics on different timescales to support higher-level fault detection and diagnosis efforts which in turn aid decision making related to maintenance and safety. The Quality-by-Control (QbC) framework proposed by (Su et al., 2019) presents a 3-level hierarchical framework for control that includes equipment-based control at Level 0, process analytical technology (PAT)-based property feedback control at Level 1, and model-based supervisory control at Level 2. This work seeks to demonstrate that the framework can incorporate quantifiable metrics to carry to enable multi-level control performance monitoring. A schematic illustration of the multi-level hierarchical framework is presented in Figure 1. Like the original QbC framework, Level 2 hosts the MHE-NMPC framework where online estimation and control is accomplished. Level 2 operates on the shortest timescale, e.g., in seconds. Metrics such as the integral absolute error (IAE), magnitude-to-product (M2P), and duration-to-reject (D2R) serve as preliminary indicators that monitor the

effectiveness of the framework. This information is utilized to determine if attention need to be paid to the metrics from Level 3. Level 3 operates on a longer timescale, e.g., in minutes or hours depending on process dynamics, where closed loop operating data is utilized to compute the MI and covariance matrix-based indices. The MI index allows the engineer to determine the number of input-output channels affected by PMM and identify these channels. The covariance matrix-based approach proposed by (Yu and Qin, 2008) is utilized to track CPD. If severe deterioration in either metric is identified at Level 3, attention needs to be paid to the metrics from Level 4. LRI for Level 4 can only be computed during scheduled maintenance as it requires the collection of open-loop data to re-identify transfer functions. The CPI for Level 4 also provides information regarding the impact of PMM on control loop behavior and can be used in combination with the LRI to determine the urgency of required maintenance. It should be noted that the proposed hierarchy enables early detection of PMM and CPD through the monitoring of metrics from different levels and time scales. Metrics obtained from all levels can be fed into an analytics platform to aid maintenance decision making and root cause analysis. However, this is out of the scope of the current work as it will require monitoring metrics from other components of the process, e.g., condition of unit operations, to demonstrate the strength of analytics platform, and will be addressed in subsequent work.

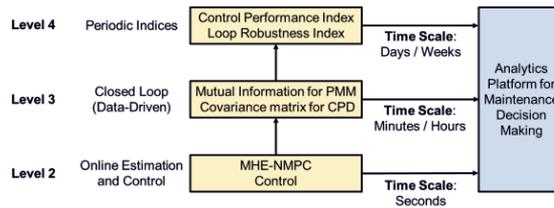


Figure 1. Schematic illustration of multi-level hierarchical framework for control performance monitoring.

3. Case Study

The case study presented in this work utilizes the process model for the rotary tablet press provided by (Huang et al., 2021) as the benchmark. The system consists of five input variables: dosing position (Dose), pre-compression thickness (Ptck), main compression thickness (Mtck), turret speed (Tret), and concentration of glidant (Csil), and four controlled variables: tablet weight (Twei), pre-compression force (Pcom), production rate (Prod), and tensile strength (Tstr). A process schematic listing the unit operations and available PAT measurements is provided in Figure 2. Model parameters for three cases of PMM (no PMM, mild PMM, and high PMM) are provided by (Huang et al., 2021).

3.1. Level 2 Monitoring

Monitoring indices for Level 2 for this case study are available in (Huang et al., 2021), where the ability to distinguish between high PMM and the other two cases was demonstrated. However, due to the effectiveness of the MHE-NMPC framework, the indices were unable to clearly distinguish between the case of no PMM and mild PMM, as PMM was effectively managed when mild. Therefore, attention needs to be paid to the indices from Level 3 to determine if the three cases of PMM can be clearly distinguished.

3.2. Level 3 Monitoring

For Level 3, MI for PMM functions by examining the correlation between the error residuals and the manipulated variables. A pseudo-binary random signal (PRBS) was utilized to provide sufficient excitation to the system to compute the MI metrics. A

summary of the results for all three cases of PMM is provided in Table 1. While the level of PMM cannot be easily visualized using the raw MI values, the percentage difference compared to the base case (no PMM) makes quantification of PMM straightforward. A threshold of 10% was set for this case study. For mild PMM, five input-output channels showed PMM, while seven input-output channels were identified for high PMM. The channels experiencing degradation are highlighted in yellow. The percentage increase is also significantly higher for high PMM. This result highlights the need for Level 3 monitoring indices due to its ability to distinguish between varying degrees of PMM.



Figure 2. Process schematic.

The covariance matrix-based assessment criterion utilized in this work was proposed by (Yu and Qin, 2008), where a generalized eigenvalue analysis is used to assess control performance. Using the case of no PMM for demonstration purposes, this case study examined three different scenarios of controller tuning, where the parameters and their values are provided in Figure 3, with the prediction horizon, control horizon, and past window of measurements in the MHE framework denoted by N_p , N_c , and N_{past} , respectively. An eigenvalue greater than 1 implies degraded performance for a particular control loop, and a value lower than 1 implies improved performance. A summary of the generalized eigenvalues, and their confidence intervals is provided in Table 2.

Table 2. Summary of generalized eigenvalue analysis for controller tuning.

Loop	Adequate Tuning				Poor Tuning			
	1	2	3	4	1	2	3	4
Eigenvalue	1.20	1.02	1.00	0.99	2.38	1.07	1.04	0.94
Lower Limit	1.09	0.93	0.91	0.90	2.17	0.97	0.94	0.85
Upper Limit	1.31	1.11	1.09	1.08	2.59	1.17	1.13	1.02

For cases of adequate and poor tuning the lower limit of the confidence interval for one loop is greater than 1, confirming degraded performance. This result is important to note as the difference between adequate and ideal tuning cannot readily be distinguished visually from the time series plots (see production rate in Figure 3 (a) and (b)).

Table 1. Summary of mutual information metrics.

No PMM									
(a) Mutual Information									
	Twei	Pcom	Prod	Tstr					
Dose	0.400	0.080	0.239	0.143					
Ptck	0.152	0.041	0.119	0.090					
Mtck	0.227	0.060	0.153	0.116					
Tret	0.119	0.041	0.213	0.089					
Csil	0.090	0.041	0.091	0.088					
Mild PMM									
(b) Mutual Information				(c) Percentage Difference (%)					
	Twei	Pcom	Prod	Tstr	Dose	Twei	Pcom	Prod	Tstr
Dose	0.408	0.087	0.229	0.143	1.804	9.835	-4.485	-0.162	
Ptck	0.149	0.047	0.127	0.099	-1.870	13.308	6.962	10.049	
Mtck	0.235	0.063	0.145	0.117	Mtck	3.218	5.710	-5.270	0.977
Tret	0.131	0.041	0.222	0.087	Tret	10.389	-0.395	4.651	-2.156
Csil	0.101	0.031	0.112	0.077	Csil	12.036	-25.955	24.193	-12.674
High PMM									
(d) Mutual Information				(e) Percentage Difference (%)					
	Twei	Pcom	Prod	Tstr	Dose	Twei	Pcom	Prod	Tstr
Dose	0.393	0.106	0.235	0.140	-1.768	32.813	-1.822	-1.765	
Ptck	0.128	0.058	0.135	0.091	-16.283	39.553	13.745	1.455	
Mtck	0.207	0.070	0.149	0.124	Mtck	-8.748	16.468	-2.808	6.372
Tret	0.101	0.041	0.223	0.086	Tret	-14.792	-0.134	4.794	-3.896
Csil	0.116	0.053	0.154	0.061	Csil	29.264	28.406	70.551	-30.248

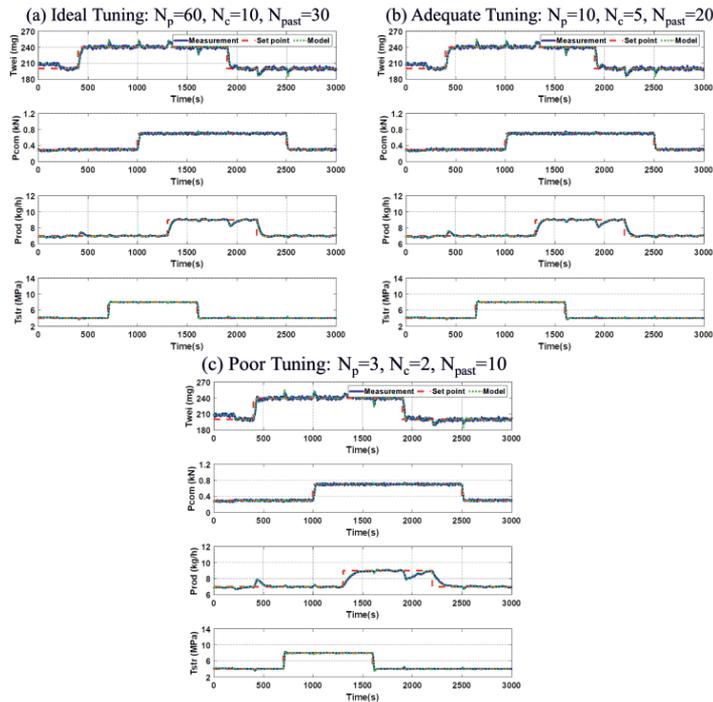


Figure 3. Time-series plots for different cases of controller tuning.

3.3. Level 4 Monitoring

Level 4 monitoring indices, CPI and LRI, were designed to be utilized together for decision making. CPI is based on the IAE and provides a means to compare the benchmark performance to current operation. While CPI is a useful indicator, the LRI provides additional information regarding the stability and robustness of the system. To enable use of the LRI for MIMO systems, this work utilizes the disk margin proposed by (Seiler et al., 2020) instead of the sensitivity margin obtained from the Nyquist plot, to make it easier to quantify and visualize the robustness of different input-output channels. Computing the LRI is a 3-step process that involves: (1) identification of the open-loop transfer functions of the MIMO system, (2) computation of disk margin for all channels, (3) computation of LRI for all channels. A summary of the CPI and LRI values for the same case study presented by (Huang et al., 2021) is provided in Table 3 and Table 4, respectively. Negative values for both the CPI and LRI indicate degradation in control performance and loop robustness. As the level of PMM increases, the CPI values become increasingly negative, indicating increased degradation in control performance. In this example, the LRI for most channels (with the exception of turret speed-tensile strength channel) are also increasingly negative in the presence PMM, implying that the robustness and stability of those channels are affected as well, requiring maintenance actions. The channels with no values were open loop unstable for both the benchmark and monitored cases. This case study demonstrated how metrics from the different timescales can be evaluated to determine if maintenance actions are required. Indices from the shortest timescale, i.e., Level 2, evaluate the feasibility of continued operation, but the indices from Levels 3 and 4 allow process engineers to periodically evaluate the urgency of required maintenance.

Table 3. Summary of CPI metrics for different cases of PMM.

	Twei	Pcom	Prod	Tstr
No PMM	0.000	0.000	0.000	0.000
Mild PMM	-0.007	-0.072	-0.421	-0.157
High PMM	-0.389	-0.214	-0.657	-0.305

Table 4. Summary of LRI metrics for different cases of PMM.

	(a) Mild PMM				(b) High PMM			
	Twei	Pcom	Prod	Tstr	Twei	Pcom	Prod	Tstr
Dose	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ptck	-0.0228	-	-1.0000	-	-0.0671	-	-1.0000	-
Mtck	-0.0029	-0.0174	-0.0681	-	-0.0089	-0.0464	-0.1101	-
Tret	0.0000	0.0000	-0.0086	0.0311	-0.0004	0.0000	-0.0160	0.0257
Csil	-	-	-	-	-	-	-	-

4. Conclusions

This work demonstrated how the QbC framework could be applied to enable multi-level control performance monitoring by incorporating indices from different timescales. Future work includes the development of a data analytics platform to aid decision making for continuous manufacturing industries, and experimental validation on the pilot plant at Purdue University.

Acknowledgement

This work was supported by the United States Food and Drug Administration under grant 1U01FD006487-01.

References

- Badwe, A.S., Gudi, R.D., Patwardhan, R.S., Shah, S.L., Patwardhan, S.C., 2009. Detection of model-plant mismatch in MPC applications. *Journal of Process Control* 19, 1305–1313.
- Chen, G., Xie, L., Zeng, J., Chu, J., Gu, Y., 2013. Detecting Model-Plant Mismatch of Nonlinear Multivariate Systems Using Mutual Information. *Industrial & Engineering Chemistry Research* 52, 1927–1938.
- Destro, F., García Muñoz, S., Bezzo, F., Barolo, M., 2021. Powder composition monitoring in continuous pharmaceutical solid-dosage form manufacturing using state estimation – Proof of concept. *International Journal of Pharmaceutics* 605, 120808.
- Harris, T.J., 1989. Assessment of control loop performance. *The Canadian Journal of Chemical Engineering* 67, 856–861.
- Huang, Y.-S., Sheriff, M.Z., Bachawala, S., Gonzalez, M., Nagy, Z.K., Reklaitis, G. v., 2021. Evaluation of a Combined MHE-NMPC Approach to Handle Plant-Model Mismatch in a Rotary Tablet Press. *Processes* 9, 1612.
- Seiler, P., Packard, A., Gahinet, P., 2020. An Introduction to Disk Margins. *IEEE Control Systems Magazine* 40, 78–95.
- Su, Q., Ganesh, S., Moreno, M., Bommireddy, Y., Gonzalez, M., Reklaitis, G. v., Nagy, Z.K., 2019. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. *Computers & Chemical Engineering* 125, 216–231.
- Wang, H., Hägglund, T., Song, Z., 2012. Quantitative Analysis of Influences of Model Plant Mismatch on Control Loop Behavior. *Industrial & Engineering Chemistry Research* 51, 15997–16006.
- Yu, J., Qin, S.J., 2008. Statistical MIMO controller performance monitoring. Part I: Data-driven covariance benchmark. *Journal of Process Control* 18, 277–296.

Optimal Control Policies of a Crystallization Process Using Inverse Reinforcement Learning

Paul Danny Anandan, Chris D. Rielly, Brahim Benyahia*

*Department of Chemical Engineering, Loughborough University, Epinal Way,
Loughborough, LE11 3TU United Kingdom*

B.Benyahia@lboro.ac.uk

Abstract

Crystallization is widely used in the pharmaceutical industry to purify reaction intermediates and final active pharmaceutical ingredients. This work presents a novel implementation of Inverse Reinforcement Learning (IRL) approach where an agent observes the expert's optimal control policies of a crystallization process and attempts to mimic its performance. In essence, an Apprenticeship Learning (AL) setup was developed where the expert demonstrates the control task to the IRL agent to help attain effective control performance when compared to the expert. This is achieved through repeated execution of "exploitation policies" that simply maximizes the rewards over the consecutive IRL training episodes. The cooling crystallization of paracetamol is used as a case study and both proportional integral derivative (PID) and Model Predictive Control (MPC) strategies were considered as expert systems. A model based IRL technique is implemented to achieve effective trajectory tracking which ensures final crystal size, considered as the critical quality attributes, by reducing the deviation from the optimal reference trajectories namely process temperature, supersaturation, and particle size. The performance of the trained IRL agent was validated against the PID and MPC and tested in presence of noisy measurements and model uncertainties.

Keywords: Apprenticeship Learning; Reinforcement Learning; Inverse Reinforcement Learning; Batch Crystallization.

1. Introduction and Background

Crystallization is one of the most widely used techniques in the pharmaceutical industry to purify intermediates and final active pharmaceutical ingredients (API) (Benyahia, 2018). It is significant that effective control strategies are required to regulate the target critical quality attributes such as crystal size, shape distributions, purity, and polymorphism. The downstream processes also depend on the ability of the crystals to flow, dry, and filter. Pharmaceutical industries are subject to stringent regulatory requirements for product quality as a small deviation from the target quality can create significant impacts on quality control, safety, efficacy, and shelf life (Lakerveld et al., 2015). This may even lead to the failure of the clinical trials during the early development stages, post-approval, and drug recalls that leads to severe economic setbacks.

Owing to the recent and significant development in the field of computer science, Artificial Intelligence (AI) has witnessed a resurgence of interest over a broad research and industrial applications. Reinforcement Learning (RL) is one among those techniques

that has been well-investigated and implemented in field of robotics and automation industries (Roveda et al., 2018). Similar investigations and developments have motivated the adoption of RL as a new control strategy for a variety of applications. RL has also been used to control chemical reactions and demonstrated its potential advantages compared to some of state-of-the-art algorithms and control techniques, such as PID (Zhou et al., 2017). The use of RL in presence of various uncertainties and measurement noise in real-world dynamic systems has also gained a lot of interest. For instance, Petsagkourakis et al., 2020 implemented a RL-based control strategy for a batch bioprocess with various uncertainties (measurement and process noises). Benyahia and coworkers, demonstrated the advantages of combining RL with a Kalman Filter (KF) for controlling batch crystallization processes for achieving robust control in presence of measurement noise and plant-model mismatch (Benyahia et al., 2021). With the continuous and successful implementations of RL in various domains, there has been a growing demand for effective training strategies. Transfer Learning (TL) is one among the most effective methods to accelerate the learning process while parallelly achieving better performance. In essence, TL follows a sequential training approach to address the continuously evolving training objectives (Taylor and Stone, 2009).

This paper outlines a novel IRL application for crystallization processes. Previously, Apprenticeship Learning (AL) approaches included the definition of a reward function that mimics the behavior of an expert controller (Mowbray et al., 2021). Such implementations rely on human expertise to capture the underlying theories of the expert controller. This paper, on the contrary, is an effort to train an agent to mimic the expert controller's actions by shadowing them. The IRL agent was associated with a KF for further TL-based trainings to deal with various levels of noise in one of the most critical measurements in crystallization; the number of crystals (number of counts). Towards this end, this paper promotes an initial IRL approach to mimic the behavior of the expert controllers, and later use TL-based training strategies with an intension to exploit and understand the potential of RL for crystallization control.

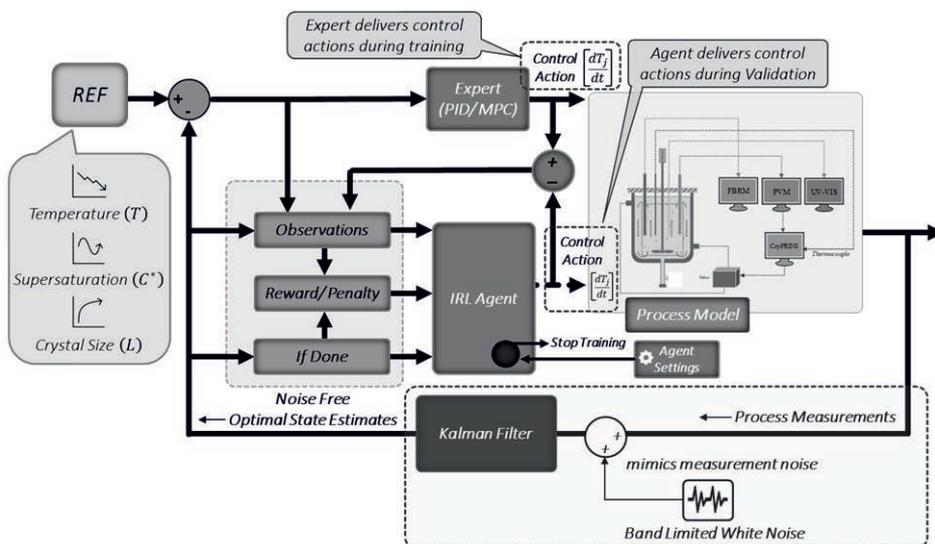


Figure 1: Overview of the IRL training and validation setup

2. Problem Formulation

The central objective of this work is to exploit the existing AI-based techniques to build robust controllers that can effectively compete with the efficiency of the traditional proportional integral derivative (PID) and Model Predictive Control (MPC) strategies. To achieve this objective, this paper presents the use of IRL approach combined with RL, TL and KF techniques to build a very responsive and efficient control strategies for batch crystallization processes that deals with various uncertainties (measurement noise and plant-model mismatch). This work makes use of three reference trajectories for process temperature, supersaturation, and particle size to ensure the product’s critical quality attributes are within safe and acceptable margins. The reference trajectories are computed using model-based open-loop dynamic optimization method illustrated by Nagy et al., 2008, where the focus was on maximising the crystal size. The case study for this work uses the batch crystallization process of Paracetamol in water. The details of the dynamic mathematical model used for this case study can be found in our previous literature (Benyahia et al., 2021). The kinetic parameters of the growth and nucleation were obtained from the literature (Nagy et al., 2008). The agent’s policy and value function definitions are based on the Deep Deterministic Policy Gradient (DDPG) technique. Please refer to our previous literature for the details regarding the agent definitions used in this work (Benyahia et al., 2021).

The setup used for this IRL training strategy is shown in **Figure 1**. During training, the expert controller (PID/ MPC) acts on the jacket temperature (manipulated variable) of the batch crystallization process. The chosen IRL approach will penalise the agent for deviations beyond a critical limit from the expert’s actions. This allows the agent to imitate the expert’s behaviour while reducing the training effort. The training makes use of a reward function that is defined based only on penalties. Therefore, the maximum theoretical reward that can be achieved in this case is zero which means that the agent’s control actions are equivalent to the expert’s (PID/ MPC) actions. In the context of IRL training, the agent will focus on minimizing the total penalties received per training episode. However, to reduce the training time, a reasonable value closer to the maximum theoretical reward was selected to stop the training. After training, the saved agent can be used to validate the performance of the IRL strategy. At the validation stage, the control loop is disconnected from the expert’s control actions and replaced by the RL agent. This ensures that the trained agent is the only control system for the batch crystallizer acting on the jacket temperature, which is used as the manipulated variable.

$$R_{t1} = \left(\underbrace{0((E_a - A_a) \leq 0.1)}_{\substack{\text{Zero penalty if the difference between the agent's and} \\ \text{the expert's control action is within an acceptable range}}} \bigvee -10(E_a - A_a) \right)$$

A proportional penalty if the difference is beyond the acceptable range

$$R_{t2} = \left(\underbrace{0(|T_e| \leq 0.1)}_{\text{Penalty for Temperature Tracking}} \bigvee -(|T_e|) \right) + \left(\underbrace{0(|S_e| \leq 0.1 \times 10^{-3})}_{\text{Penalty for Supersaturation Tracking}} \bigvee -5(|S_e|) \right) + \left(\underbrace{0(|L_e| \leq 1)}_{\text{Penalty for Crystal Size Tracking}} \bigvee -10(|L_e|) \right)$$

Figure 2: Simplified overview of the reward Function

It is important to note that this work includes IRL and RL trainings in two steps. Accordingly, **Figure 2** presents the simplified overview of the separate reward functions used in the first and second stages. The reward function for the first step (R_{I1}) focuses on the IRL approach, where the agent is forced to learn by shadowing the expert controller. The reward definition for the second step (R_{I2}) is focused on exploiting the training results obtained in the first step by combining them with RL, TL, and KF techniques. The goal was to improve the performance of the trained IRL agents with subsequent trainings targeted at more accurate and multiple trajectory tracking. The addition of KF to the trainings can help to address the issues related to various measurement noises and plant-model mismatch.

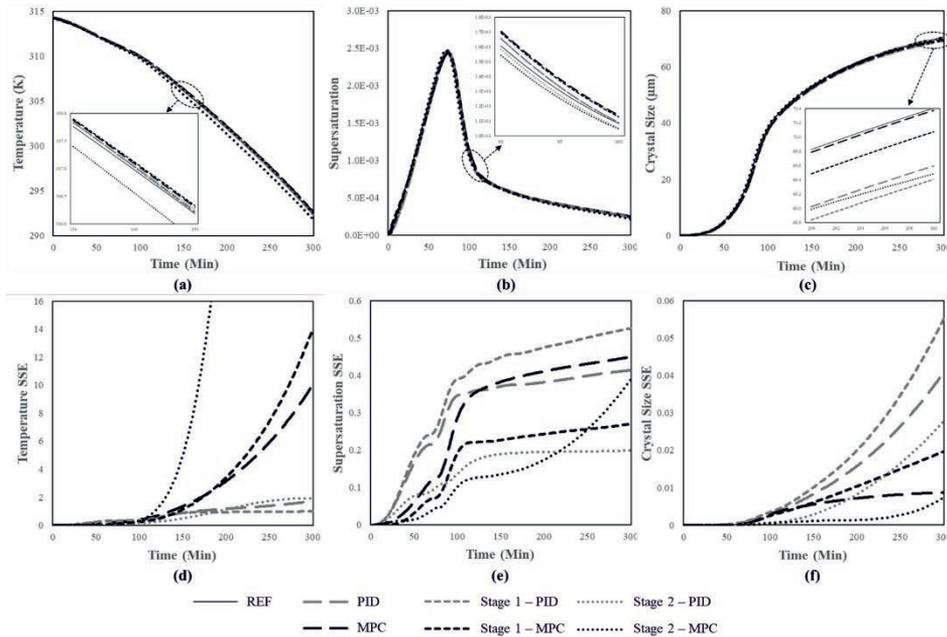


Figure 3: Overview of the validation results after stage 1 and stage 2 training.

3. Result and Discussion

The trained agents were validated in two stages in conjunction with the two-step training strategy suggested in **Figure 2**. **Figure 3** presents the comparison between the validation results captured at the end of each stage of the trainings conducted. The trajectories of the sum of the squared errors (SSE) presented in **Figure 3d, e** and **f** provide further qualitative insights about the difference in accuracy between the two stages of the trainings. The mean crystal size was considered as the target critical quality attribute and higher penalties were defined for control actions resulting with larger deviations from the reference trajectory. For a better insight, **Figure 3f** indicates that the performance of the stage1 trainings are not far off from the target expert controllers (PID/ MPC). However,

with subsequent second stage trainings, the performance of the agents has achieved an even more effective performance compared the target expert controllers (PID/ MPC).

Figure 4 is used to present the benefits of including KF during the second stage of training. The idea is to present the usefulness of KF to deal with various measurement noise and uncertainties related to plant-model mismatch. In other words, the use of KF will help achieve more effective training by providing optimal state estimates which enable more reliable and accurate evaluation of the rewards associated with the agent's control actions. The measurement noise (band-limited white noise) was introduced to the output signal related to the number of particles (μ_n) predicted by the process model. It is one of the most important real-time process measurements commonly obtained by a Focused Beam Reflectance Measurement (FBRM) technique. Similarly, the plant-model mismatch is introduced to the kinetic growth parameter (k_g) of the process model. It is important to note that separate trainings were conducted for cases with measurement noise and plant-model mismatch. The trajectories of the sum of the squared errors (SSE), presented in **Figure 4d, e, and f**, provide further qualitative insights about the difference in accuracy between the two stages of the trainings. Again, to obtain a clear conclusive performance comparison, it is necessary to pay more attention to **Figure 4f**. This is simply because of the highly weighted penalty over crystal size deviations. It is also worth noting that the results shown in **Figure 4** does not include the actual expert controllers (PID/ MPC), instead it is the comparison between the two stages of the IRL agent trainings investigating the usefulness of KF in the context of measurement noise and model uncertainties. The results indicate that the integration of KF at the second stage of the trainings is clearly beneficial, and with further TL-based trainings, the accuracy and performance of the agents can be further improved.

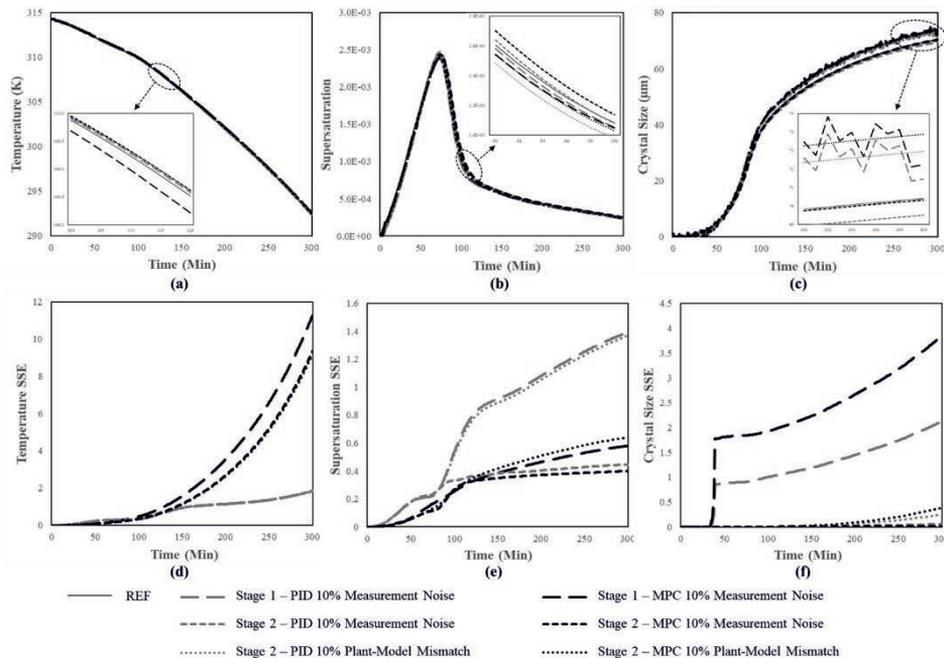


Figure 4: Overview of the results showing the benefits of integrating KF in the stage 2 training.

4. Conclusion

This paper demonstrated the use of model-based RL control policies for the trajectory tracking control of a batch crystallization process. Initially, the agent was trained using an IRL approach to mimic the behavior of the PID and MPC. A range of reward functions defined with penalties-only (no rewards) fashion were used to achieve the highest score in a reasonable amount of time, using the best compromises between exploration and exploitation. The resulting IRL agents were then validated by comparing their control performance against the expert controllers. It was shown that the IRL approach can exhibit a very close performance to the benchmark controllers. Later, the IRL agents' performances were further enhanced by conducting subsequent TL-based RL trainings. The results revealed that the IRL approach combined with TL and KF can demonstrate robust control performance in presence of noisy measurements and model uncertainties.

Acknowledgements: This work was funded by the EPSRC (EP/R032858/1) ARTICULAR project.

References

- Benyahia, B., 2018. Applications of a plant-wide dynamic model of an integrated continuous pharmaceutical plant: Design of the recycle in the case of multiple impurities, in: *Computer Aided Chemical Engineering*. <https://doi.org/10.1016/B978-0-444-63963-9.00006-3>
- Benyahia, B., Anandan, P.D., Rielly, C., 2021. Robust Model-Based Reinforcement Learning Control of a Batch Crystallization Process, in: *9th International Conference on Systems and Control (ICSC)*. IEEE, Caen, France, pp. 89–94. <https://doi.org/10.1109/ICSC50472.2021.9666494>
- Lakerveld, R., Benyahia, B., Heider, P.L., Zhang, H., Wolfe, A., Testa, C.J., Ogden, S., Hersey, D.R., Mascia, S., Evans, J.M.B., Braatz, R.D., Barton, P.I., 2015. The Application of an Automated Control Strategy for an Integrated Continuous Pharmaceutical Pilot Plant. *Organic Process Research and Development*. <https://doi.org/10.1021/op500104d>
- Mascia, S., Heider, P.L., Zhang, H., Lakerveld, R., Benyahia, B., Barton, P.I., Braatz, R.D., Cooney, C.L., Evans, J.M.B., Jamison, T.F., Jensen, K.F., Myerson, A.S., Trout, B.L., 2013. End-to-end continuous manufacturing of pharmaceuticals: Integrated synthesis, purification, and final dosage formation. *Angewandte Chemie - International Edition*. <https://doi.org/10.1002/anie.201305429>
- Mowbray, M., Smith, R., Del Rio-Chanona, E.A., Zhang, D., 2021. Using process data to generate an optimal control policy via apprenticeship and reinforcement learning. *AICHe Journal*. <https://doi.org/10.1002/aic.17306>
- Nagy, Z.K., Chew, J.W., Fujiwara, M., Braatz, R.D., 2008. Comparative performance of concentration and temperature controlled batch crystallizations. *Journal of Process Control*. <https://doi.org/10.1016/j.jprocont.2007.10.006>
- Petsagkourakis, P., Sandoval, I.O., Bradford, E., Zhang, D., del Rio-Chanona, E.A., 2020. Reinforcement learning for batch bioprocess optimization. *Computers and Chemical Engineering*. <https://doi.org/10.1016/j.compchemeng.2019.106649>
- Roveda, L., Pallucca, G., Pedrocchi, N., Braghin, F., Tosatti, L.M., 2018. Iterative Learning Procedure with Reinforcement for High-Accuracy Force Tracking in Robotized Tasks. *IEEE Transactions on Industrial Informatics*. <https://doi.org/10.1109/TII.2017.2748236>
- Taylor, M.E., Stone, P., 2009. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*.
- Zhou, Z., Li, X., Zare, R.N., 2017. Optimizing Chemical Reactions with Deep Reinforcement Learning. *ACS Central Science*. <https://doi.org/10.1021/acscentsci.7b00492>

Dynamic Real-Time Optimization with Closed-Loop Prediction for Nonlinear MPC-Controlled Plants

Daniela Dering^a and Christopher L. E. Swartz^{a*}

^a*Department of Chemical Engineering, McMaster University, 1280 Main Street West, Hamilton L8S 4L8, Canada*
swartzc@mcmaster.ca

Abstract

Current trends toward globalization and electricity market deregulation are requiring increasingly dynamic operation of chemical processes. In this paper, we develop a dynamic real-time optimization (DRTO) formulation for plants controlled by NMPC. It utilizes a prediction of the plant under the action of constrained NMPC. At every prediction time-step, an NMPC problem determines the control inputs that are applied to the dynamic process model. We show that the unconstrained NMPC problem for SISO systems affine in the inputs has a single stationary point that corresponds to the global optimum, and numerical experiments suggest that a similar property holds for the constrained problem. This allows the embedded NMPC subproblems in the DRTO formulation to be replaced by their first-order Karush-Kuhn-Tucker (KKT) conditions, yielding a single-level optimization problem. We show that the DRTO with embedded NMPC subproblems can lead to significant improvement in plant performance, and also compare the performance of NMPC to that of linear MPC.

Keywords: NMPC, DRTO, Karush-Kuhn-Tucker conditions

1. Introduction

Real-Time Optimization (RTO) utilizes a rigorous steady-state plant model to compute the set-point trajectories that optimize an economic process metric. These set-points are then provided as targets to a multivariable controller, generally model predictive control (MPC), that computes the input values applied to the plant. Because RTO utilizes a steady-state plant model, it is not executed during process transitions, but only when the plant has reached an almost steady-state operation (Darby et al., 2011). This particular limitation is averted if a dynamic plant model is used instead. This variation of RTO is commonly referred to as DRTO. In a recent development, Jamaludin and Swartz (2017) proposed to account for the MPC action in the DRTO formulation. This new variation, named CL-DRTO, models the closed-loop response of the plant, and has been shown to improve process economics compared to a DRTO formulation that utilizes an open-loop prediction of the plant dynamics (Jamaludin and Swartz, 2017). An alternative paradigm is economic model predictive control (EMPC) (Ellis et al., 2014) that replaces standard MPC by optimizing the process economics and computing the input values simultaneously. EMPC versions that utilize linear and nonlinear dynamic models have been proposed. It has been argued that the EMPC problem can not be solved as frequently as standard MPC due to the problem complexity, and that the integration of economic optimization and control in one problem lead to conflicting objectives, which may hinder process safety (Würth et al., 2011). The EMPC and DRTO approaches are representative of the so called single and two-layer control architecture, respectively.

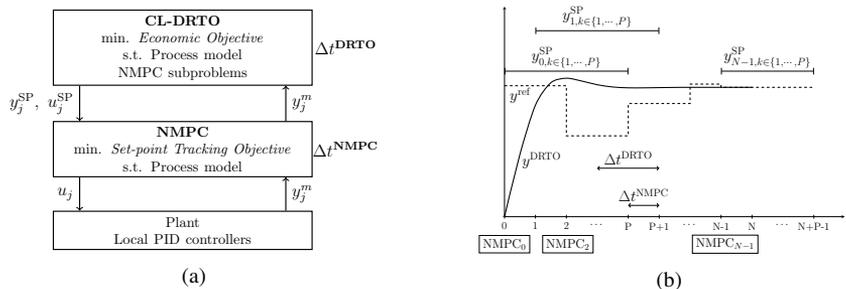


Figure 1: Schematic representation of (a) two-layer approach, and (b) optimal solution of CL-DRTO problem.

While linear MPC is the prevalent control strategy in chemical enterprises (Qin and Badgwell, 2003), it can perform poorly for large transitions and/or in highly nonlinear plants. The reason is that the built-in linear model provides a poor approximation in regions away from the steady-state operation, exacerbated by strong nonlinearity. Moreover, in multiproduct plants, multi-linear models may have to be used in order to appropriately model all the operating regions. Due to these limitations, nonlinear model predictive control (NMPC), that uses a nonlinear process model, is often the preferred choice of control strategy for highly nonlinear processes.

In this study, we propose a DRTO formulation similar to that in Jamaludin and Swartz (2017) but that accounts for the action of a nonlinear model predictive control (NMPC) instead of linear MPC. To avoid a bilevel optimization problem, we use the associated first-order Karush-Kuhn-Tucker (KKT) conditions to account for the NMPC action in the DRTO formulation. We consider in this paper systems that are affine in the control inputs, and derive conditions under which the unconstrained NMPC problem yields a single stationary point, which corresponds to the global optimum. Numerical experiments suggests that a similar property holds for the input constrained problem as well, supporting the reformulation of the embedded NMPC problems using the corresponding KKT conditions. Performance of linear and nonlinear MPC for controlling a SISO nonlinear system in the presence and absence of a DRTO layer is compared.

2. Formulation

Figure 1a provides an schematic representation of the framework for online real-time optimization and control of a chemical process implemented in this study. The CL-DRTO layer computes the reference trajectories y^{ref} and u^{ref} that optimize an economic or target tracking objective function. The set-point trajectories y_j^{SP} and u_j^{SP} are extracted from the reference trajectory and provided to the lower level NMPC that then computes the input values u_j applied to the plant. The lower-level NMPC and upper-level DRTO problems are executed at time periods Δt^{NMPC} and Δt^{DRTO} , respectively, with $\Delta t^{DRTO} / \Delta t^{NMPC} \in \mathbb{Z}^+$.

In order to account for the closed-loop behavior of the process in the CL-DRTO formulation, an NMPC problem is formulated at every time-step j along the DRTO prediction horizon to determine the input values used in the process model. Figure 1b shows a schematic representation of the response of a plant under the combined action of DRTO and NMPC. N and P are, respectively, the CL-DRTO and NMPC prediction horizon, M is the NMPC control horizon and y^{DRTO} is the output trajectory predicted by the DRTO process model. The set-point trajectories used in the NMPC subproblems are extracted from the reference trajectory y^{ref} . Note that the NMPC problem solved after time-step $N - P + 1$ sees beyond the CL-DRTO horizon N , and for that reason the reference trajectory is extended until $N + P - 1$. More detail about the CL-DRTO problem formulation is

provided in the next sections. We highlight that the general formulation and solution approach are similar to that proposed in Jamaludin and Swartz (2017), but with the linear MPC (LMPC) replaced by NMPC. We use the Python backend of CasADi (Andersson et al., 2019) and the interior point solver IPOPT (Wächter and Biegler, 2006) to implement and solve the optimization problems in this study.

2.1. Process Model

The process model includes the mathematical model of the plant, bounds on the optimization variables and process outputs, and constraints linking the process model to the NMPC subproblems. The discrete plant model is assumed to have the following form

$$x_{j+1}^{\text{DRTO}} = f^{\text{DRTO}}(x_j^{\text{DRTO}}, u_j^{\text{DRTO}}) \quad \forall j \in \mathcal{J}_0^{N-1} \quad (1a)$$

$$y_j^{\text{DRTO}} = Cx_j^{\text{DRTO}} \quad \forall j \in \mathcal{J}_1^N \quad (1b)$$

where $x_j^{\text{DRTO}} \in R^{nx}$ and $y_j^{\text{DRTO}} \in R^{ny}$ are column vectors of the predicted state and output values at time-step j , respectively. $u_j^{\text{DRTO}} \in R^{nu}$ is the vector of input values computed via solution of j^{th} NMPC subproblem, that is $u_j^{\text{DRTO}} = u_{j,0}^{\text{NMPC}}$, where $u_{j,k}^{\text{NMPC}} \in R^{nu}$. $C \in R^{ny \times nx}$ is a matrix that relates the states to the outputs, and $\mathcal{J}_a^b = \{j|a \leq j \leq b, j \in Z_0^+\}$ is a set of time-steps. $f^{\text{DRTO}}(\cdot)$ is a nonlinear function obtained via discretization of the dynamic plant model.

The optimization degrees of freedom are the reference trajectories $y^{\text{ref}} = [(y_1^{\text{ref}})^T, \dots, (y_{N+P-1}^{\text{ref}})^T]^T$ and $u^{\text{ref}} = [(u_0^{\text{ref}})^T, \dots, (u_{N+M-2}^{\text{ref}})^T]^T$ where $y_j^{\text{ref}} \in R^{ny}$ and $u_j^{\text{ref}} \in R^{nu}$. For j beyond the CL-DRTO horizon, the reference trajectories are held constant, that is $y_j^{\text{ref}} = y_N^{\text{ref}} \forall j \in \mathcal{J}_{N+1}^{N+P-1}$ and $u_j^{\text{ref}} = u_{N-1}^{\text{ref}} \forall j \in \mathcal{J}_N^{N+M-2}$. Additionally, lower and upper bounds can be imposed on the optimization variables and predicted outputs. Set-point hold constraints to force the reference trajectory to have a constant value for a specified number of time-steps could also be applied.

2.2. NMPC subproblems

The NMPC subproblems are assumed to have an equivalent formulation to the lower-level NMPC in Figure 1a:

$$\min_{u_j^{\text{NMPC}}} \Phi_j^{\text{NMPC}} = \sum_{k=1}^P (y_{j,k}^{\text{NMPC}} - y_{j,k}^{\text{SP}})^T Q (y_{j,k}^{\text{NMPC}} - y_{j,k}^{\text{SP}}) + \sum_{k=0}^{M-1} \left[\Delta u_{j,k}^T R \Delta u_{j,k} + (u_{j,k}^{\text{NMPC}} - u_{j,k}^{\text{SP}})^T S (u_{j,k}^{\text{NMPC}} - u_{j,k}^{\text{SP}}) \right] \quad (2a)$$

$$\text{subject to } x_{j,k+1}^{\text{NMPC}} = x_{j,k}^{\text{NMPC}} + \Delta t^{\text{NMPC}} f(x_{j,k}^{\text{NMPC}}, u_{j,k}^{\text{NMPC}}), \quad \forall k \in \mathcal{J}_0^{P-1} \quad (2b)$$

$$y_{j,k}^{\text{NMPC}} = Cx_{j,k}^{\text{NMPC}}, \quad \forall k \in \mathcal{J}_1^P \quad (2c)$$

$$\Delta u_{j,k} = u_{j,k}^{\text{NMPC}} - u_{j,k-1}^{\text{NMPC}}, \quad \forall k \in \mathcal{J}_0^{M-1} \quad (2d)$$

$$u_{j,k}^{\text{NMPC}} = u_{j,M-1}^{\text{NMPC}}, \quad \forall k \in \mathcal{J}_M^{P-1} \quad (2e)$$

$$u_{\min}^{\text{NMPC}} \leq u_{j,k}^{\text{NMPC}} \leq u_{\max}^{\text{NMPC}}, \quad \forall k \in \mathcal{J}_0^{M-1} \quad (2f)$$

$$u_j^{\text{NMPC}} = \left[\left(u_{j,0}^{\text{NMPC}} \right)^T, \dots, \left(u_{j,M-1}^{\text{NMPC}} \right)^T \right]^T \quad (2g)$$

where $Q \in R^{ny \times ny}$ is a positive definite matrix, and $R \in R^{nu \times nu}$ and $S \in R^{nu \times nu}$ are positive semidefinite matrices. $u_{j,k}^{\text{NMPC}} \in R^{nu}$, $x_{j,k}^{\text{NMPC}} \in R^{nx}$ and $y_{j,k}^{\text{NMPC}} \in R^{ny}$ are, respectively, the input, state and

output values for the j^{th} NMPC subproblem at time-step k . $f(\cdot) : R^{nx+nu} \rightarrow R^{nx}$ is the dynamic model of the process, here assumed to be affine in the inputs and given by

$$\frac{dx}{dt} = f(x, u) = g(x) + B(x)u \quad (3)$$

with $g(x) : R^{nx} \rightarrow R^{nx}$ and $B(x) : R^{nx} \rightarrow R^{nx \times nu}$. We assume that $f(\cdot)$ is smooth, continuous and differentiable in R^{nx+nu} . In linear MPC formulations, $f(\cdot)$ is a linear function and the resulting control problem a convex quadratic programming problem with linear constraints. Due to the convexity property, the first-order Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for global optimality (Nocedal and Wright, 2006), and they can be used instead of Equation 2 in the CL-DRTO (Jamaludin and Swartz, 2017) allowing solution as a single-level optimization problem. Because in this study $f(\cdot)$ can be nonlinear, problem 2 is potentially non-convex. Next we show some results that suggest that problem 2 with $f(\cdot)$ defined as in Equation 3 has a unique local optimum for Single-Input Single-Output (SISO) systems, that is when $ny = nu = 1$.

Consider the smaller NMPC problem,

$$\min_v \quad \Phi_R = \sum_{k=1}^P Q(y_k - y_k^{\text{SP}})^2 \quad (4a)$$

$$\text{subject to} \quad x_{k+1} = x_k + \Delta t f(x_k, u_k), \quad \forall k \in \mathcal{J}_0^{P-1} \quad (4b)$$

$$y_k = Cx_k, \quad \forall k \in \mathcal{J}_1^P \quad (4c)$$

$$v = [u_0 \dots, u_{P-1}]^T \quad (4d)$$

which retains all potential sources of nonconvexity in problem 2. For a SISO system, we have $y_k^{\text{SP}}, y_k, u_k \in R$ and $x_k \in R^{nx}$, $Q > 0$ is a scalar and $C \in R^{1 \times nx}$ a row vector. $v \in R^P$ are the optimization variables, $f(\cdot)$ is defined as in Equation 3, and Δt is the sampling time. Φ_R is ultimately a function of only x_0 , v , and $y^{\text{SP}} = [y_1^{\text{SP}}, \dots, y_P^{\text{SP}}]^T$, since the equalities in Equations 4b and 4c can be used to eliminate y_k . The first-order KKT conditions for problem 4 are simply $\nabla_v \Phi_R(v, x_0, y^{\text{SP}}) = 0$, where

$$\nabla_v \Phi_R(v, x_0, y^{\text{SP}}) = 2 \begin{bmatrix} \left((y_1 - y_1^{\text{SP}})QC + \sum_{k=2}^P (y_k - y_k^{\text{SP}})QC \frac{\partial x_k}{\partial x_1} \right) \frac{\partial x_1}{\partial u_0} \\ \left((y_2 - y_2^{\text{SP}})QC + \sum_{k=3}^P (y_k - y_k^{\text{SP}})QC \frac{\partial x_k}{\partial x_2} \right) \frac{\partial x_2}{\partial u_1} \\ \vdots \\ \left((y_{P-1} - y_{P-1}^{\text{SP}})QC + (y_P - y_P^{\text{SP}})QC \frac{\partial x_P}{\partial x_{P-1}} \right) \frac{\partial x_{P-1}}{\partial u_{P-2}} \\ \left((y_P - y_P^{\text{SP}})QC \right) \frac{\partial x_P}{\partial u_{P-1}} \end{bmatrix} \quad (5)$$

We have that for $z \in R^n$ and $w \in R^m$, $\frac{\partial z}{\partial w}$ is a $n \times m$ matrix with $(ij)^{\text{th}}$ element $\frac{\partial z_i}{\partial w_j}$. Substituting $\frac{\partial x_k}{\partial u_{k-1}} = \Delta t B(x_{k-1})$ in Equation 5, and assuming $CB(x) \in R - \{0\}$ for all $x \in R^{nx}$, we obtain that the unique stationary point of $\Phi_R(v, x_0, y^{\text{SP}})$ is given by $y_k - y_k^{\text{SP}} = 0$ for all $k \in \mathcal{J}_1^P$. This expression can be solved explicitly for the optimal value of the optimization variables v . Moreover, note that this solution corresponds to the global optimum of problem 4, since it leads to $\Phi_R(v, x_0, y^{\text{SP}}) = 0$. Therefore, the first-order KKT conditions are necessary and sufficient for problem 4. This result suggests that they may also be necessary and sufficient for problem 2 since the constraints in Equations 2d-2f are linear and the additional terms in the objective function are convex. To substantiate this conjecture, we solved the constrained problem in Equation 2 for distinct SISO systems using both the first-order KKT conditions and an optimization solver, for thousands of different initial guesses for the optimization variables. The first-order KKT conditions yield a system of nonlinear equations with complementary constraints (e.g.

$\mu_k^1(u_{j,k}^{\text{NMPC}} - u_{\min}^{\text{NMPC}}) = 0$, $\mu_k^2(u_{\max}^{\text{NMPC}} - u_{j,k}^{\text{NMPC}}) = 0$, $\forall k \in \mathcal{J}_0^{M-1}$ where $\mu_k^i \in R^+$ are Lagrange multipliers), that are handled using an exact penalty formulation (Ralph and Wright, 2004). For all the test problems, the KKT conditions and optimization solver returned the same optimal solution for all initial guesses.

3. Case Studies

We use the CL-DRTO formulation with embedded NMPC subproblems reformulated via the KKT conditions for optimization and control of an isothermal nonlinear CSTR used to produce five grades A, B, C, D and E via the reaction $3R \xrightarrow{k} P$ (Flores-Tlacuahuac and Grossmann, 2006). The dynamic model is given by

$$\frac{dC_R}{dt} = \frac{F}{V}(C_0 - C_R) - kC_R^3$$

where C_0 and C_R are the composition of the inlet and outlet stream, respectively. F is the inlet and outlet flow rate, V is the volume of liquid in the reactor, and k is the reaction rate constant. For this study, we use $C_0 = 1$ mol/L, $V = 5$ m³, $k = 2$ L²/mol²h (Flores-Tlacuahuac and Grossmann, 2006). Using the previous introduced notation, we have that $x = C_R$, $u = F$, $C = 1$, $g(x) = -2x^3$, $B(x) = 0.2(1 - x)$. We use NMPC and DRTO parameters $Q = 1$, $R = 0.2$, $S = 0$, $M = 3$, $P = 30$, $N = 50$, $\Delta t^{\text{NMPC}} = 0.5$ h, $\Delta t^{\text{DRTO}} = 2$ h and impose the bounds $0 \leq u_{j-1,k}^{\text{NMPC}} \leq 3$, $0 \leq y_j^{\text{DRTO}} \leq 0.6$, $0 \leq y_j^{\text{ref}} \leq 0.8$ for all $k \in \mathcal{J}_0^{M-1}$, $j \in \mathcal{J}_1^N$. The bounds on $u_{j,k}^{\text{NMPC}}$ are sufficient to prevent $CB(x) = 0$ from being a feasible solution, and we do not have to impose constraints on the output values in the NMPC subproblems. The NMPC and DRTO process model is discretized using the forward-Euler method with a step size of 0.1.

We consider an economic (Φ^{ECO}) and set-point tracking (Φ^{SP}) objective function formulation to be minimized:

$$\Phi^{\text{ECO}} = \sum_{j=1}^N \Delta t^{\text{NMPC}} u_{j-1} - 50R_j^1 R_j^2, \quad \Phi^{\text{SP}} = \sum_{j=1}^N (y_j - y_j^{\text{targ}})^2$$

$$R_j^1 = 0.5 \tanh\left(\gamma\left(y_j - y_j^{\text{targ}}(1 - \varepsilon)\right)\right) + 0.5, \quad R_j^2 = 0.5 \tanh\left(\gamma\left(y_j^{\text{targ}}(1 + \varepsilon) - y_j\right)\right) + 0.5$$

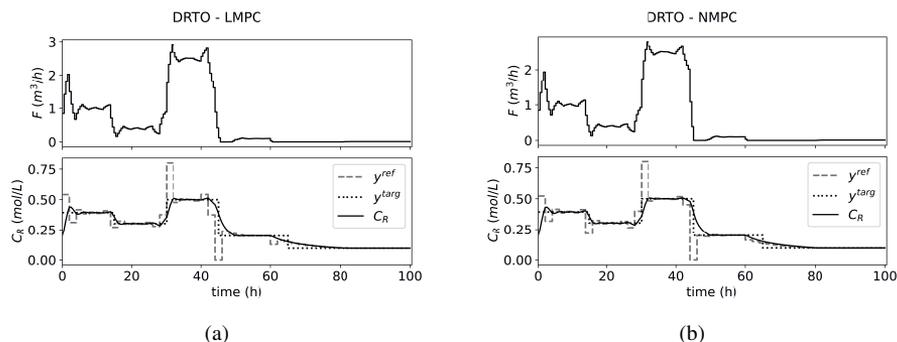
where y_j^{targ} is the specified target output concentration at time-step j , γ defines the steepness of the switching function R_j^i , and ε is the tolerance by which y_j can deviate from y_j^{targ} . We have that $R_j^1 \approx 1$ if $y_j > y_j^{\text{targ}}(1 - \varepsilon)$, and it is zero otherwise. Similarly, $R_j^2 \approx 1$ if $y_j < y_j^{\text{targ}}(1 + \varepsilon)$, and it is zero otherwise.

We compare the performance of NMPC against that of linear MPC for closed-loop control of the nonlinear CSTR, both when the CL-DRTO layer is present (Figure 1a) and when it is not. The target output trajectory y^{targ} is provided as set-points to the control layer in the latter case. The objective values at the end of the simulation horizon for these control implementations are given in Table 1. For both, LMPC and NMPC, the objective values improve considerably when the set-points are computed by the DRTO layer. The NMPC controller tends to lead to an overall better plant performance compared to LMPC. We note that by accounting for the control action, the DRTO becomes aware of the LMPC and NMPC limitations, and can therefore compute reference values that to an extent compensate for that.

The plant input and output trajectories for the case when the set-point tracking objective function is used in the CL-DRTO are shown in Figure 2. The response is quite similar for both controllers despite some small differences in the reference trajectory. We note that the value of F , while small (0.01 m³/L), is not zero after 70 hours (Flores-Tlacuahuac and Grossmann, 2006).

Table 1: Objective values for NMPC and LMPC control implementations.

	LMPC	NMPC	DRTO-LMPC	DRTO-NMPC
Φ^{ECO}	-6543	-6,592	-7,241	-7,292
Φ^{SP}	0.4388	0.4407	0.1479	0.1402

Figure 2: Input and output trajectories for the nonlinear plant under the action of CL-DRTO with embedded (a) LMPC and (b) NMPC subproblems, and Φ^{SP} as the objective function.

4. Conclusion

In this study, we proposed a DRTO formulation for plants controlled by NMPC. We showed that the unconstrained NMPC problem for SISO systems affine in the inputs has only one stationary point that is also the global optimum. Numerical experiments suggest that this property also holds for the constrained NMPC problem, making the first-order KKT conditions necessary and sufficient for global optimality. This allows us to use the associated KKT conditions to account for the NMPC control action in the DRTO formulation. At every time-step the control problem is solved to compute the input value used in the process model. The results indicate a better performance of the NMPC compared to linear MPC, and show that significant improvement in plant performance can be achieved by accounting for the control action in real-time optimization.

References

- J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, M. Diehl, 2019. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11 (1), 1–36.
- M. L. Darby, M. Nikolaou, J. Jones, D. Nicholson, 2011. RTO: An overview and assessment of current practice. *Journal of Process Control* 21 (6), 874–884.
- M. Ellis, H. Durand, P. D. Christofides, 2014. A tutorial review of economic model predictive control methods. *Journal of Process Control* 24 (8), 1156–1178.
- A. Flores-Tlacuahuac, I. E. Grossmann, 2006. Simultaneous cyclic scheduling and control of a multiproduct CSTR. *Industrial & engineering chemistry research* 45 (20), 6698–6712.
- M. Z. Jamaludin, C. L. E. Swartz, 2017. Dynamic real-time optimization with closed-loop prediction. *AIChE Journal* 63 (9), 3896–3911.
- J. Nocedal, S. Wright, 2006. *Numerical Optimization*. Springer Science & Business Media, Ch. Theory of Constrained Optimization.
- S. J. Qin, T. A. Badgwell, 2003. A survey of industrial model predictive control technology. *Control Engineering Practice* 11 (7), 733–764.
- D. Ralph, S. J. Wright, 2004. Some properties of regularization and penalization schemes for MPECs. *Optimization Methods and Software* 19 (5), 527–556.
- A. Wächter, L. T. Biegler, 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106 (1), 25–57.
- L. Würth, R. Hannemann, W. Marquardt, 2011. A two-layer architecture for economically optimal process control and operation. *Journal of Process Control* 21 (3), 311–321.

Dynamic Real-Time Optimization of a Solar Thermal Plant during Daytime

Alix Untrau,^a Sabine Sochard,^a Frédéric Marias,^a Jean-Michel Reneaume,^a Galo A.C. Le Roux,^b Sylvain Serra^a

^a*Universite de Pau et des Pays de l'Adour, E2S UPPA, LaTEP, Pau, France*

^b*Universidade de São Paulo, Escola Politécnica, São Paulo, Brasil*

Abstract

In this work, an economic Dynamic Real-Time Optimization (DRTO) of the operation of a solar thermal plant is conducted. A planning phase is used to improve the use of storage, and a storage state target is included in the objective function of the DRTO. The weight of this term in the objective function is adjusted. The trajectories for the flow rates obtained with the DRTO algorithm are used in a simulation model undergoing disturbances on the solar irradiation. This methodology is compared with a simulation following the flow rate trajectories obtained during the planning phase. The results show an improvement in the total expenses and the percentage of solar energy used to satisfy the demand, with a reasonable tracking of the storage state target, bringing good perspectives for the implementation of the method in a real environment.

Keywords: Dynamic real-time optimization, solar thermal energy, simulation.

1. Introduction

In order to achieve greenhouse gases emissions reduction, developing solar thermal plants and making the most of them appear crucial. In addition to the design, the operation of the plant can be optimized in order to reduce the operating cost, help to meet the heat demand and cut down the fossil fuel consumption in heat production. Dynamic optimization was used to optimize the operation of a solar thermal plant including a storage tank during 36 hours using weather and load forecasts (Scolan et al., 2020). Trajectories for the flow rates in the different parts of the solar plant were computed. Especially, the use of the storage tank was optimized to extend the time period where solar heat could be supplied to the consumer. More studies focusing on the dynamic optimization of Concentrated Solar Power (CSP) plants can be found in the literature. In these studies, the operation of the plant is optimized in order to maximize the income from electricity selling, with a variable electricity price and a thermal energy storage (Lizarraga-Garcia et al., 2013; Wittmann et al., 2011). Hybridization of the solar plant with a back-up fossil fuel has also been considered, leading to a larger amount of solar energy collected (Ellingwood et al., 2020; Powell et al., 2014). These studies show an improvement in the performance of the solar thermal plants, with more energy collected at a lower cost. However, it is difficult to predict the weather and the heat demand accurately several days ahead of time and dynamic optimization cannot consider disturbances, such as differences in planned weather and heat demand, to adapt the optimal strategy. The lower level in the hierarchical operation of a plant is the control. The controllers in a solar thermal plant are generally basic controllers such as PID (Proportional Integral Derivative) tracking a target value for the outlet temperature of the solar field. More complex controllers, suited for nonlinear systems with various

dynamics, have been studied for CSP plants to improve the disturbance rejection and the uncertainty handling (Camacho et al., 2007). Furthermore, it is possible to integrate an economic objective in the controller (Engell, 2007). In two studies, this approach was used to minimize the back-up fossil fuel consumption in a solar system including storage (Pintaldi et al., 2019; Serale et al., 2018). The authors suggest that a hierarchical structure, with a decoupling between the dynamic optimization and the control tasks might improve the use of storage. A Dynamic Real-Time Optimization (DRTO) could be used to compute reference trajectories for the flow rates that will be tracked by controllers (Kadam, 2002). The DRTO of a solar plant was performed in only one study (Pataro et al., 2020). The flow rate in the concentrating solar field was optimized in real-time using measurements of the ambient conditions and an economic objective function. However, the storage management was not considered. Based on this literature survey, an economic DRTO of a solar thermal plant including storage is carried out in this paper.

2. Solar thermal plant description and modelling

A typical layout for a solar thermal plant was chosen for this study and is presented in Figure 1. The production loop is composed of a solar field with 12 loops of 15 flat plate collectors each, corresponding to a panel surface of 2873 m². A glycol-water (30/70 volume) mixture is heated up in the solar field. It can by-pass the heat exchanger connecting the production circuit to the secondary circuit during the warm-up phase of the solar field, until a target temperature is reached. Once collected in the secondary loop, the solar heat can be stored in the stratified storage tank for later use or directly delivered to the consumer circuit. The storage tank can be discharged to supply heat to the consumer circuit when solar irradiance is too low to meet the demand. These different operational modes make the operation of the solar thermal plant flexible, but nevertheless complex. Variable speed pumps and control valves are used to operate the plant. A back-up gas heater, not modelled and not represented in Figure 1, is associated to the solar thermal plant to meet the load curve.

Details about the nonlinear models used for each element of the solar thermal plant can be found in (Scolan et al., 2020). In this paper, the solar field is modelled as an equivalent surface panel to speed up the calculations.

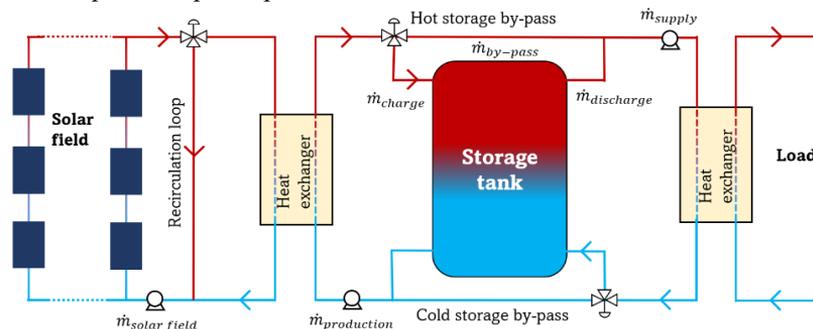


Figure 1: Architecture of the solar thermal plant

3. Dynamic Real-Time Optimization methodology

The structure of the algorithm used to optimize in real-time the solar thermal plant is presented in Figure 2. First, an offline Dynamic Optimization (DO) based on weather and load forecasts is performed for a 2-day period. The objective function for this planning phase is to minimize the costs associated with the electricity used for pumping

(130 €/MWh) and the gas (80 €/MWh) used in the backup heater. A value is given to the stored energy at the end of the time horizon because the stored heat could be valuable beyond the end of this optimization (17.5 €/MWh from (Scolan et al., 2020)). Finally, a penalty term is added to smooth the trajectories computed for the flow rates in the different parts of the plant. The planning phase is used for the storage management. Indeed, storage has slower dynamics than the rest of the plant and needs to be considered on a longer time period.

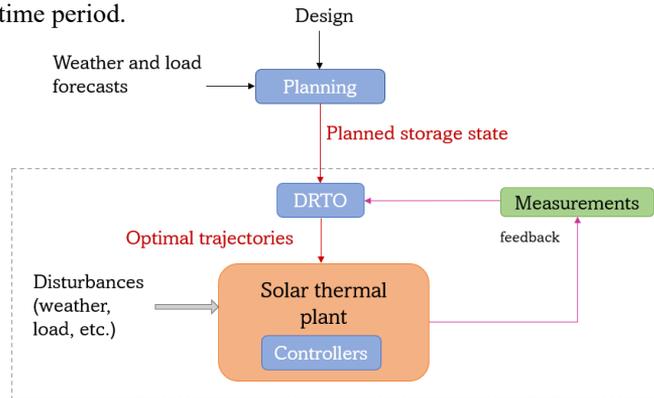


Figure 2: Optimization algorithm

The DRTO methodology is applied during daytime of the first day of the complete time horizon using the two layer approach presented in (Kadam, 2002). It starts when the warm-up phase of the solar field is completed, which corresponds to a temperature at the outlet of the solar field of 70°C. The objective function for the DRTO also minimizes the cost of gas and electricity and smooths the trajectories obtained for the flow rates. In addition, a term that minimizes the difference between the planned stored energy and the actual energy stored at the end of the day is added. A new DRTO is carried out every hour, until the end of the day. Thus, the time horizon shrinks for every DRTO (from about 12h to 1h) and the computational time also reduces from 15 minutes to less than 5 minutes. The models and the resolution method could be modified in future work to speed up the calculations. Measurements of temperatures in the different parts of the solar thermal plant are used to determine the initial state. Disturbances are also measured, and weather forecasts are updated. A simulation model is used to represent the solar thermal plant and provide the feedback measurements. The optimal trajectories are sent to the simulation model of the solar thermal plant. Perfect control is assumed in first approach, so the exact optimal trajectories are followed in spite of disturbances. The optimizations were conducted in the Gams software and the solar thermal plant was simulated in Matlab. The optimizations, solved using the simultaneous approach, are initialized with results from simulations with standard operation strategy. The complete discretized model involves up to 28,000 variables for the first DRTO with a time horizon of about 12 hours.

4. Case study

4.1. Inputs

For the weather forecasts, averaged data over the last 50 years in Pau, France, are used. For this study, two summer days are chosen: a sunny day followed by a cloudy day, as shown in Figure 3 with GHI standing for Global Horizontal Irradiance. This configuration imposes the use of storage to avoid exceeding the heat demand the first day and to supply some solar heat the second day despite the low solar irradiation.

For the real-time data during the first day, an uncertainty of p %, several values of p being tested, is applied to the entire curve of solar irradiance. Other inputs, such as ambient temperature or heat load are considered certain in this work. Weather forecasts are updated before each DRTO, as shown in Figure 4, for the hours 9 and 14 with an uncertainty of - 30 % on the GHI. This procedure is used in this first approach, but real data for weather forecasts and measurements should be used for more realistic results.

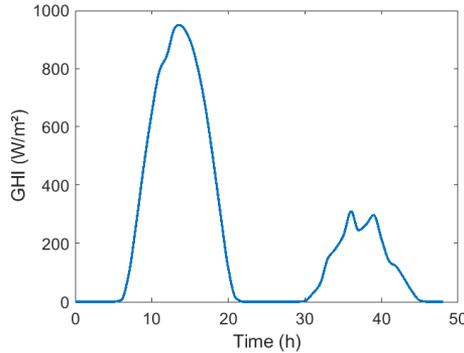


Figure 3: Solar irradiance forecast

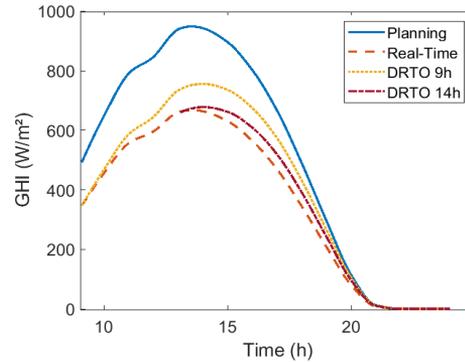


Figure 4: Real-Time solar irradiance and updated forecasts

The load curve is considered constant, with a value of 836 kW. The consumer flow enters the heat exchanger at 55°C. The storage is initially at 60°C, which represents a partially charged storage. The temperatures in the solar thermal plant are equal to the ambient temperature at the beginning of the planning phase. The initial state for the DRTO during daytime is retrieved from the planning when the temperature at the outlet of the solar field reaches 70°C.

4.2. Outputs

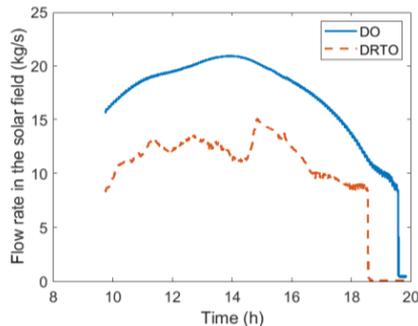


Figure 5: Comparison between DO and DRTO

Figure 5 shows an example of trajectories obtained for the flow rate in the solar field $\dot{m}_{solar\ field}$ from the planning (DO) and the DRTO with an uncertainty of - 50 % on the GHI. It can be observed that $\dot{m}_{solar\ field}$ is lower for the DRTO. In order to achieve a temperature high enough for storage and heat supply, a lower flow rate is used in the solar field associated with lower pumping cost. Similar trends can be seen for the other flow rates in the plant.

4.3. Following of the planned heat storage state

As mentioned in Section 3, the objective function of the DRTO includes a term to minimize the difference between the planned heat storage state at the end of the first day and the actual storage state. The energy difference is multiplied by the gas price (80 €/MWh) in the objective function, since the lack of stored energy will be compensated by gas in the backup heater. The cost associated with the failure to follow the planning is affected by a weight, which needs to be adjusted.

On Figure 6 two performance criteria are plotted for different weights. On the left axis, the difference between the stored energy at the end of the day and the planned target is drawn. On the right axis, the difference between the cost during the day and the lowest

cost achieved for the DRTO without any objective on storage is shown. It can be observed that, with a higher weight, the target on the storage state is better achieved but it is at a higher cost because less heat is delivered to the consumer and thus more gas is used to complete the demand. DRTO without any target on storage (weight of 0) leads to less expenses for this first day but it will cost more money on the second day where little solar heat can be produced. For the remaining part of this paper, a weight of 0.65 is chosen as it represents a good compromise between total cost for the day and tracking of the storage target.

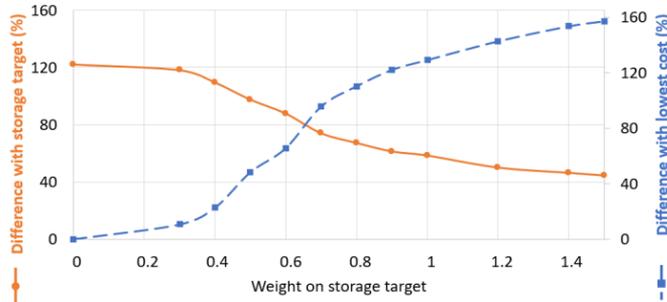


Figure 6: Effect of weight on storage target in the objective function on storage and costs

4.4. Results for different uncertainty levels on the GHI

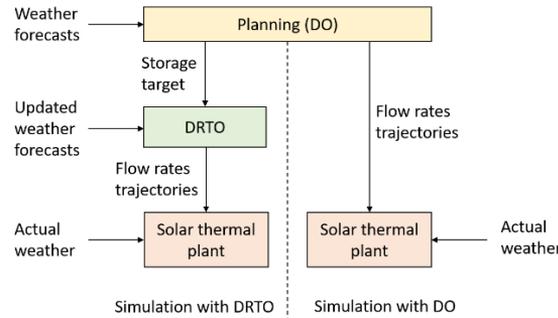


Figure 7: Comparison of simulations based on DO or DRTO

Figure 7 presents the methodology used to assess the improvements in the solar thermal plant operation obtained with DRTO. The DRTO method presented in section 3, and repeated on the left-hand side of the diagram, is compared to a simulation following the flow rates trajectories obtained during the planning phase (DO) but undergoing the disturbance on the solar irradiation.

The results of the comparison of the simulations based on offline and real-time optimizations of the solar thermal plant are presented in Table 1, for different values of disturbances on the GHI. Only negative disturbances are considered to avoid exceeding the heat demand in the simulation based on DO trajectories. Table 1 shows an improvement (bold characters) for DRTO compared to DO in the total cost (between 15 % and 44 %) and the percentage of solar energy directly used to satisfy the demand, for every case tested. However, it is at a price of a decrease in storage state monitoring. A study over a longer time period, a week or even a month, would help to assess the benefits of DRTO since the storage state at the end of the last day would be less important and the cost and solar percentage would be the only performance criteria. In Table 1, it can be noticed that for a higher uncertainty level on the GHI, the overall performance of the solar thermal plant is improved with DRTO, with a high solar energy percentage (+ 52 % compared to DO) and a small difference in storage state tracking (- 20 % compared to DO). The results observed depend on the weight chosen in the previous section. Based on these observations, a variable weight should be considered: for a small uncertainty level, the storage state target should be tracked, and the weight should be adjusted to ensure a good tracking while minimizing the cost. However, for a large

uncertainty level, the storage state target cannot be met, and it seems more suitable to minimize the cost of the day without consideration for the planned stored energy.

Table 1: Solar thermal plant performance in simulations based on DO or DRTO for different disturbances on the GHI

	GHI – 10 %		GHI – 20 %		GHI – 30 %		GHI – 40 %		GHI – 50 %	
	DO	DRTO	DO	DRTO	DO	DRTO	DO	DRTO	DO	DRTO
Difference with storage target (%)	- 17.5	- 48.0	- 35.2	- 65.9	- 54.5	- 79.3	- 75.3	- 101.9	- 97.7	- 117.3
Total expenses (€)	218.6	121.7	291.3	191.3	369.0	291.3	451.9	373.3	540.4	456.6
% Solar heat	69.3	83.0	58.5	72.5	47.0	57.4	34.7	45.3	21.6	32.8

5. Conclusion

The economic DRTO of a solar thermal plant was conducted, with a planning phase to improve storage management. The weight used for the tracking of the storage target was adjusted to make a compromise between the income achieved on the first day and the tracking of the storage target ensuring revenues the second day. Using the chosen weight, DRTO was compared with a simulation based on offline DO for various levels of uncertainty. The results indicate a general improvement in the operation of the solar thermal plant, especially for large uncertainty on the solar irradiation. Future work should focus on the testing of the methodology with real data and over a longer period of time.

References

- Camacho, E. F., Rubio, F. R., Berenguel, M., & Valenzuela, L. (2007). A survey on control schemes for distributed solar collector fields. Part II : Advanced control approaches. *Solar Energy*, 81(10), 1252-1272.
- Ellingwood, K., Mohammadi, K., & Powell, K. (2020). Dynamic optimization and economic evaluation of flexible heat integration in a hybrid concentrated solar power plant. *Applied Energy*, 276, 115513.
- Engell, S. (2007). Feedback control for optimal process operation. *Journal of Process Control*, 17(3), 203-219.
- Kadam, J. V. (2002). A Two-Level Strategy of Integrated Dynamic Optimization and Control of Industrial Processes—A Case Study. *European Symposium on Computer Aided Process Engineering - 12*, 6.
- Lizarraga-Garcia, E., Ghobeity, A., Totten, M., & Mitsos, A. (2013). Optimal operation of a solar-thermal power plant with energy storage and electricity buy-back from grid. *Energy*, 51, 61-70.
- Pataro, I. M. L., Roca, L., Sanches, J. L. G., & Berenguel, M. (2020). An economic D-RTO for thermal solar plant : Analysis and simulations based on a feedback linearization control case. *XXIII Congresso Brasileiro de Automática*, 8.
- Pintaldi, S., Li, J., Sethuvenkatraman, S., White, S., & Rosengarten, G. (2019). Model predictive control of a high efficiency solar thermal cooling system with thermal storage. *Energy and Buildings*, 196, 214-226.
- Powell, K. M., Hedengren, J. D., & Edgar, T. F. (2014). Dynamic optimization of a hybrid solar thermal and fossil fuel system. *Solar Energy*, 108, 210-218.
- Scolan, S., Serra, S., Sochard, S., Delmas, P., & Reneaume, J.-M. (2020). Dynamic optimization of the operation of a solar thermal plant. *Solar Energy*, 198, 643-657.
- Serale, G., Fiorentini, M., Capozzoli, A., Cooper, P., & Perino, M. (2018). Formulation of a model predictive control algorithm to enhance the performance of a latent heat solar thermal system. *Energy Conversion and Management*, 173, 438-449.
- Wittmann, M., Eck, M., Pitz-Paal, R., & Müller-Steinhagen, H. (2011). Methodology for optimized operation strategies of solar thermal power plants with integrated heat storage. *Solar Energy*, 85(4), 653-659.

Online Bayesian Re-design of Parallel Experiments based on Asynchronous Posterior Sampling

Martin F. Luna^a, M. Nicolás Cruz B.^{b,c}, Ernesto C. Martínez^a

^aINGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002GJC, Argentina

^bKIWI-biolab, Bioprocess Engineering, TU Berlin, Ackerstrasse 76, 13355 Berlin, Germany

^cDataHow AG, Zürichstrasse 137, 8600 Dübendorf, Switzerland
ecmarti@santafe-conicet.gov.ar

Abstract

High-throughput robotic platforms boost gathering informative data sets to support bioprocess development by resorting to on-line redesign of asynchronous parallel experiments. Due to significant uncertainty in both structure and parameters of mathematical models of bioreactors during early developmental stages, a probabilistic Bayesian approach is proposed. A novel algorithm that combines Asynchronous Posterior Sampling (APS) of model parameter distributions with chance-constrained optimization is used to bias data sampling depending on the modeling goal. As new data are available, model parameter distributions are updated using variational Bayesian inference. Myopic posterior sampling is then used online for purposefully changing cultivation conditions in parallel experiments. The proposed approach is based on a probabilistic macroscopic model, whereas the modelling goal is specified by integrating domain expertise and preferences via a reward function. A case study related to *Escherichia coli* expressing a desired product is used to demonstrate that a trade-off between improving parametric precision and biasing data gathering towards bioprocess optimization is achieved. Results obtained are encouraging for autonomous operation of robotic platforms.

Keywords: Bayesian inference, optimization, Experimental design, Probabilistic Models.

1. Introduction

The availability of high-throughput robotic platforms (Cruz Bournazou et al., 2017; Haby et al., 2019) demands a systematic methodology to efficiently explore large experimental search spaces while generating informative data sets to support model-based optimization during bioprocess development in the face of uncertainty. To this aim, online re-design of parallel experiments must explicitly pursue the goal of biasing data gathering to increase parametric precision only for the reduced region of operating conditions of interest. In the field of biopharmaceuticals, with the main goal of improving the understanding of innovative processes and reproducible end-use product quality, statistical analysis of available data highlights the importance of actively striking a balance between exploring for new knowledge and exploiting what is already known to account for scale-up effects on bioprocess development (Hernández Rodríguez, et al., 2019). The use of a Bayesian (probabilistic) approach (Martínez, et al., 2021) is appealing to account for significant prediction errors in macroscopic dynamic models of biological systems which are too shallow to comprehensively describe the full complexity of switching in metabolic pathways when responding to changes in the operating conditions or to the addition of an inducer for protein expression (Kiparissides et al., 2011).

2. Online re-design via posterior sampling

2.1. Bayesian set up

Let Ξ denote a parameter space, \mathcal{U} a decision space, and \mathcal{Y} an observation space. A probabilistic macroscopic model of a bioreactor is defined by a joint probability distribution over the following set of stochastic variables:

- x : y : the $n \times n_t$ hidden states time-series; the $p \times n_t$ observations (sampled data),
- u : the $n_u \times n_t$ redesign actions time-series,
- $\theta; \varphi$: the $n_\theta \times 1$ evolution parameters; the $n_\varphi \times 1$ observation parameters,
- α : the state noise precision (structural errors),
- σ : the measurement noise precision (analytical and sensor calibration errors).

From sample to sample in a dynamic experiment, these variables follow the equations:

$$\begin{aligned} x_t &= f(x_{t-1}, \theta, u_{t-1}) + \eta_t; \quad \eta_t = N(0, \alpha^{-1} \mathbf{I}), & \text{(hidden) state evolution} \\ y_t &= g(x_t, \varphi) + \varepsilon_t; \quad \varepsilon_t = N(0, \sigma^{-1} \mathbf{I}), & \text{observation (samples)} \end{aligned} \quad (1)$$

where f (resp. g) is the first-principles model (observation model), and η_t (observation ε_t) is the state (resp. measurement) modeling errors (noise). A probabilistic model of a bioreactor is then completed by specifying the (initial) Gaussian prior distributions for its parameters θ, φ . Also, Gamma distribution priors are defined for the precision hyperparameters α, σ .

We now consider a Bayesian setting, where at stage t a sample $\xi_t^* \in \Xi$ is drawn from a prior distribution $p_t(\xi)$ of model parameters $\xi = \{x, x_0, \theta, \varphi, \alpha, \sigma\}$. Depending on the modeling goal, a redesign action $u_t^* \in \mathcal{U}$ is chosen as if ξ_t^* were the true vector of parameters in a deterministic model, then the redesign action u_t^* is performed and the resulting (next) outcome y_{t+1} is measured. Using variational Bayesian inference (Daunizeau et al., 2014), the posterior $p(\xi_{t+1}^* | u_t^*, y_{t+1})$ is obtained, and a new stage of experimental redesign begins by using Thompson sampling (Russo et al., 2018). Thus, online re-design proceeds for n stages, resulting in a data sequence $D_n = \{(u_t^*, y_{t+1})\}, t = 1, \dots, n$, which is an ordered set of action-observation pairs in a dynamic experiment.

2.2. Goal-oriented online redesign

Let us assume that the desired redesign goal is specified via a *reward* function $r: \Xi \times \mathcal{D} \rightarrow \mathbb{R}$ that drives online decisions towards gathering informative data at the end of the experiment. Here, \mathcal{D} denotes the set of all possible data sequences that can be generated in a dynamic experiment. It is worth noting that depending on how r is defined, the data sequence D_n may be biased differently. The reward function can be related either to improving model selection, parametric precision, productivity, or a combination thereof. Alternatively, the modeling goal can be related to cumulative rewards gathered at the end of the experiment. In the latter case, online re-design decisions may require propagating beliefs from the current stage t onwards. In this work, a simple and intuitive strategy is used for online re-design aiming to maximize the estimated final reward \hat{r}_n at the end of the experiment based on sampling the posterior $p(\xi_{t+1}^* | u_t^*, y_{t+1})$ of the data D_{t+1} collected so far and exploring the redesign decision space using a deterministic model of the bioreactor dynamics with parameter ξ_t^* , by solving the mathematical program:

$$u_t^* = \operatorname{argmax}_{u \in \mathcal{U}} \hat{r}_n(\hat{y}_n | \xi_t^*, \hat{x}_t, u_t), \quad (2a)$$

$$\text{Subject to Eqs. (1) and } \mathcal{C}(x, u) \geq 0 \quad (2b)$$

where $C(x, u)$ stands for physiological constraints that prevents undesired cultivation conditions. In robotic platforms having n_{MBR} mini bioreactors, the optimization problem in Eq. (2) can be solved for each one of them. The optimal solution found u_t^* is expected to maximize the sampled value \hat{r}_n of the reward function by assuming that the value $\xi_{t,i}^*$ is the true one (unknown) vector of model parameters, \hat{x}_t is the inferred current state and \hat{y}_n is the predicted observation at the end of the experiment for a given action u_t . A schema of the procedure for online re-design using posterior sampling is shown in Fig. 1.

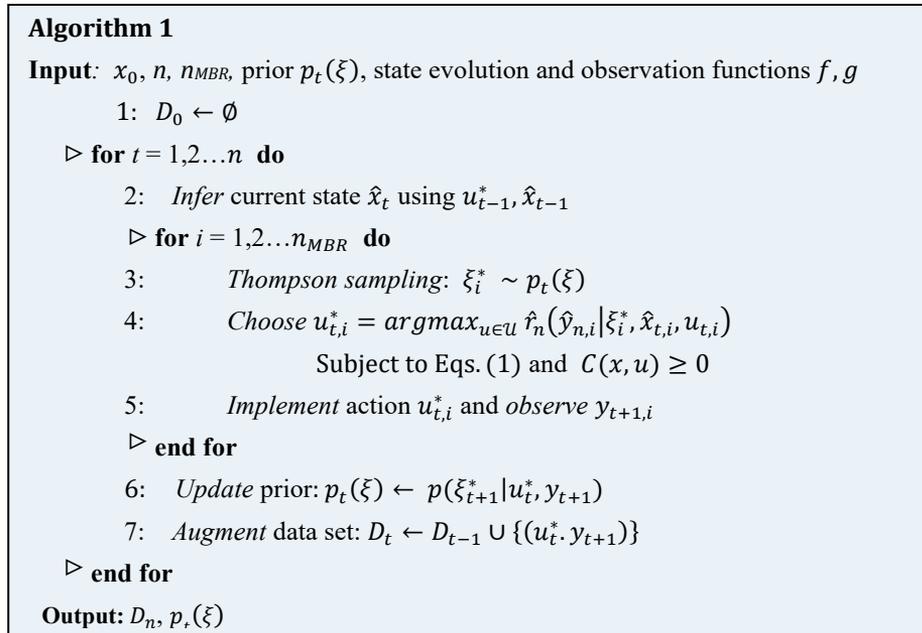


Figure 1. Online re-design strategy using Thompson sampling.

When several asynchronous parallel experiments are made, there are many possibilities for implementing the strategy in Fig. 1 to explore the search space depending on the modeling or optimization goal. The main advantage of the proposed approach is that predictions by different model parameterizations generated by posterior sampling can be experimentally tested in different mini bioreactors (MBRs), which depending on how it is done, gives rise to very different trade-offs between exploration and exploitation.

3. Case study

The proposed method was tested using artificial data generated based on the model proposed in Nickel et al. (2017), consisting of six differential equations and eighteen parameters. Eight MBRs for *Escherichia coli* fed-batch cultivations were run over a six-hour period. Samples are taken every thirty minutes and biomass (X), glucose (Glc) and acetate (Ac) concentrations are measured at-line, whereas the oxygen (O_2) is measured continuously using an online sensor. Based on new data, glucose pulses are computed following the model-based optimized feeding strategies for each MBR, and then the experiment is redesigned online. The aim of the experimental design is to find an optimal feeding profile that maximizes the amount of biomass obtained at the end of the run without allowing the oxygen concentration to drop below 20% in any of the MBRs.

The first-principles model is assumed to have the same structure as the one used to create the artificial data (no structural process-model mismatch) and 5% gaussian noise was added to the simulated values. The priors $p(\xi)$ were chosen to be wide enough to contain the real values of the parameters. The algorithm, the models and all optimization routines were implemented using Matlab 2021a. Model inversion is performed using the VBA toolbox by Daunizeau et al., (2014). Details about the toolbox can be found in the literature. Results obtained for eight MBRs runs are shown in Fig. 2. The biomass and oxygen concentrations are plotted for all bioreactors. For oxygen, only the minimum value for each sampling interval is shown for clarity. As biomass concentration increases, dissolved oxygen tension drops steadily until glucose fully depletes (not shown in Fig. 2). After that, the oxygen concentration increases until a new glucose pulse is added to the mini bioreactor, which gives rise to a drop in the dissolved oxygen tension. Once the new glucose is consumed, oxygen concentration rises again, until another pulse is fed. This pattern is repeated until the end of the experiment for all MBRs. The optimizer chooses the feeding strategy for each MBR based on the model parameterization resulting from Thompson sampling. Even when all MBRs start from the same initial condition, the different strategies give rise to a wide range of process states. However, most of the experimental datapoints are skewed towards the region of high productivity without constraint violation.

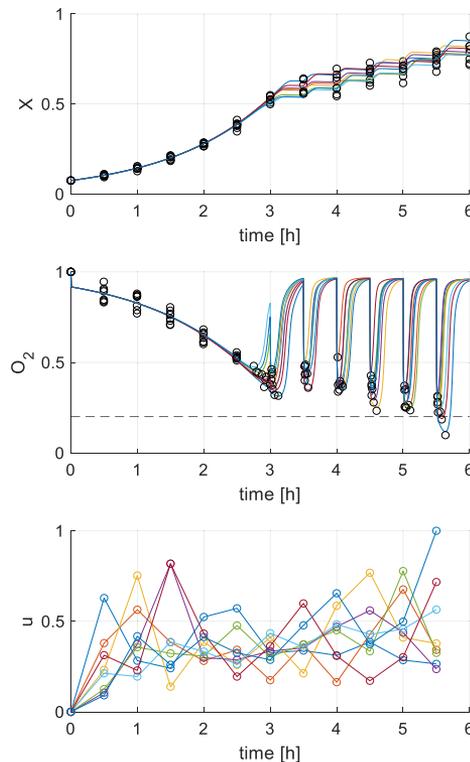


Figure 2. Normalized values for biomass and oxygen concentrations (thick lines: model predictions, circles: experimental datapoints), together with the feeding strategy for each MBR.

After each sampling time, the parameter distributions are updated with new data from all MBRs. The evolutions of some selected model parameters are presented in Fig. 3. As can be seen, the uncertainty of model parameters is significantly reduced in the first iteration and then it slowly decreases, as the mean values converge to some stationary values. This fast convergence is probably because the artificial data generator and the first-principles model have the same structure. If structural mismatch is present, convergence may take more iterations and modeling uncertainty will not be completely reduced.

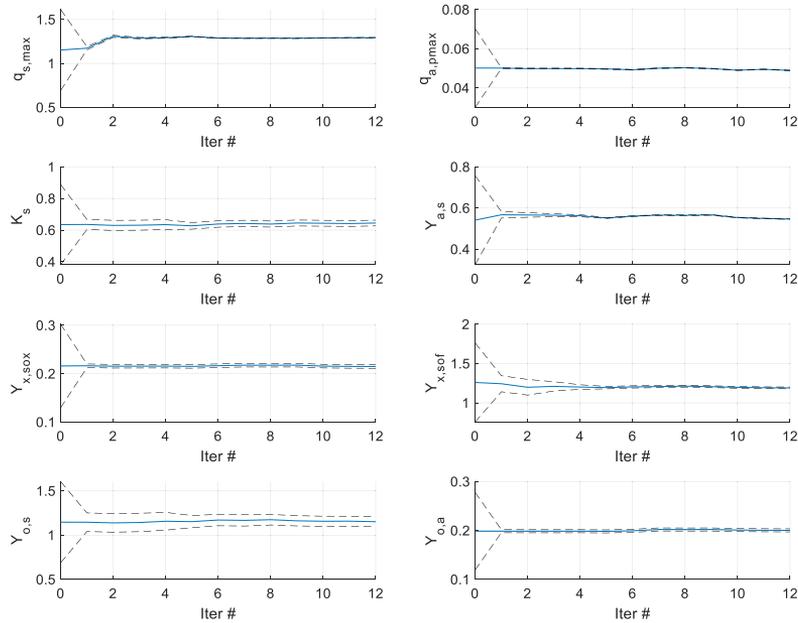


Figure 3. Evolution of selected model parameters during the experiments. Solid lines: mean values of the distributions, dashed lines: 95% confidence intervals.

Once the parallel eight experiments have been completed, the probabilistic model is updated using the complete dataset (model predictions are shown as solid lines in Fig. 2) that is biased towards improved operating conditions, which is the main objective of the proposed approach for parallel experimental redesign. The Bayesian nature of the resulting model allows for a probabilistic formulation of the optimization problem for each MBR (see Luna and Martinez, 2018):

$$u^* = \operatorname{argmax}_{u \in U} E[m_{X,end}] \quad (3.1)$$

$$\text{Subject to } \Pr(O_{2,min} > 20) > 95\% \quad (3.2)$$

where $m_{X,end}$ is the biomass amount obtained at the end of the experiment (in mg) and $O_{2,min}$ is the minimum dissolved oxygen tension (in percentage saturation). The prior and the posterior distributions of the model parameters are used to solve the chance-constrained mathematical program in Eq. (3). The resulting model-based optimized feeding strategy can be tested in a simulated experiment for a standalone bioreactor. The final amount of biomass obtained as well as the minimum dissolved oxygen concentration

during the testing run are presented in Table 1 for both model parameterizations. The uncertainty in model predictions using the prior distributions is significantly higher compared to that of the probabilistic model with the posterior distributions. As a result, the *a priori* optimized feeding profile is more cautious than the one with the probabilistic model parameterized using the updated distributions; for the later a higher amount of biomass is obtained while the dissolved oxygen tension is kept above the threshold value during the whole run. The information content of the generated data in the redesigned parallel experiments is aptly biased for solving the optimization problem in Eq. (3).

Table 1. Performance comparison of model parameterizations

	X_{end} [mg]	$O_{2,min}$ [%]
u_{prior}^*	19.7	57.59
$u_{posterior}^*$	31.6	21.20

4. Concluding remarks

A new approach for on-line redesign of parallel experiments using Thompson sampling was proposed. Results obtained for the case study demonstrate the advantages of enforcing chance-constrained physiological constraints in balancing exploration with exploitation based on sampling model parameter distributions for feed rate optimization.

References

- M. N. Cruz Bournazou, T. Barz, D. B. Nickel, D. C. Lopez Cárdenas, F. Glauche, A. Knepper, P. Neubauer, 2017, Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. *Biotechnol. Bioeng.*, 114, 610–619.
- J. Daunizeau, V. Adam, L. Rigoux, 2014, VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data, *PLoS Comput Biol* 10, 1, e1003441.
- B. Haby, S. Hans, E. Anane, A. Sawatzki, N. Krausch, P. Neubauer, M. N. Cruz Bournazou 2019, Integrated Robotic Mini Bioreactor Platform for Automated, Parallel Microbial Cultivation With Online Data Handling and Process Control, *SLAS Technology*, 24, 6, 569–582.
- T. Hernández-Rodríguez, C. Posch, J. Schmutzhard, J. Stettner, C. Weihs, R. Pörtner, B. Frahm, 2019, Predicting industrial-scale cell culture seed trains—A Bayesian framework for model fitting and parameter estimation, dealing with uncertainty in measurements and model parameters, applied to a nonlinear kinetic cell culture model, using an MCMC method, *Biotechnology and Bioengineering*, 116, 2944–2959.
- A. Kiparissides M. Koutinas, C. Kontoravdi, A. Mantalaris, E. N. Pistikopoulos, 2011, “Closing the loop” in biological systems modeling — From the in silico to the in vitro, *Automatica*, 47, 1147–1155.
- M. F. Luna & E. C. Martínez, 2018, Model-based run-to-run optimization for process development, *Brazilian J. of Chemical Engineering*, 35, 1063-1080.
- E. C. Martínez, J. W. Kim, T. Barz, M. Nicolas Cruz B., 2021, Probabilistic modeling for optimization of bioreactors using reinforcement learning with active inference, *Proc. of ESCAPE-31 2021*, 50 (Part A), 419-424.
- D. B. Nickel, M. N. Cruz-Bournazou, T. Wilms, P. Neubauer, & A. Knepper, 2017, Online bioprocess data generation, analysis, and optimization for parallel fed-batch fermentations in milliliter scale, *Engineering in the Life Sciences*, 17, 11, 1195-1201.
- D. J. Russo, B. Van Roy, A. Kazerouni, Ia. Osband, Z. Wen, 2018, A Tutorial on Thompson Sampling, *Foundations and Trends in Machine Learning*, 11, 1, 1-96.

Dynamic optimisation and comparative analysis of fed-batch and perfusion bioreactor performance for monoclonal antibody (mAb) manufacturing

Wil Jones, Dimitrios I. Gerogiorgis*

*School of Engineering (IMP), University of Edinburgh, Edinburgh, EH9 3FB, UK
D.Gerogiorgis@ed.ac.uk*

Abstract

Integrated continuous bioprocessing (ICB) holds great promise towards achieving higher efficiency, a critical target for ensuring the advent of biopharmaceutical manufacturing. Culture productivity must be drastically improved to reduce production cost projections, especially in comparison to traditional fed-batch process platforms (Pollock et al., 2017). The present paper explores the benefits of continuous bioprocessing, employing dynamic optimisation of upstream fermentation for production of monoclonal antibodies (mAb). Industry-standard cultures of hybridoma cells have been used to develop models of mAb bioreactor operation. These models track several state variable trajectories of interest, including profiles for viable cells, substrates, by-products and the target mAbs product. Dynamic optimisation of both operation modes (fed-batch and perfusion bioreactors) has been undertaken, with the objective function of maximizing the final mAb product titers through optimal feeding strategies in two reactor types for improved culture proliferation. The clear differences in terms of reactor space (capital costs) and time (operating costs) are highlighted, and a fair comparison basis is established so as to evaluate performance. For both modes of operation, a technoeconomic analysis illustrates the implications of optimal bioreactor designs, with a view to systematic decision-making in mAb ventures.

Keywords: Biopharmaceutical manufacturing; monoclonal antibodies; dynamic optimisation.

1. Introduction

Monoclonal antibodies (mAbs) are immune proteins which are genetically engineered within laboratories for the treatment of a variety of ailments, including some autoimmune diseases and cancers (Schulze-Koops et al., 2000), (Adams et al., 2005). In terms of commercial expansion, mAbs are currently the fastest growing biopharmaceutical product type in the market, representing the majority (57%) of total European Medicines Agency (EMA) biopharmaceutical approvals between 2015 and 2018 (Grilo et al., 2019).

Traditionally, biopharmaceutical process platforms have consisted of a series of batch mode operating units as seen in Figure 1. Batch operation has been favoured due to its inherently simpler configuration for the manufacture of small-scale speciality products such as (bio)pharmaceuticals (Shirahata et al., 2019). Despite this, the industrial drivers of process economics and rapid quality control generate a strong motivation to pursue semi-continuous or (preferably) continuous industrial scale biomanufacturing. The need to provide accessible therapies in developing countries, and the advent of biosimilars, means that optimising process economics is more imperative than ever in today's market. Steadily improving efficiencies of commercially available bioreactors address this need.

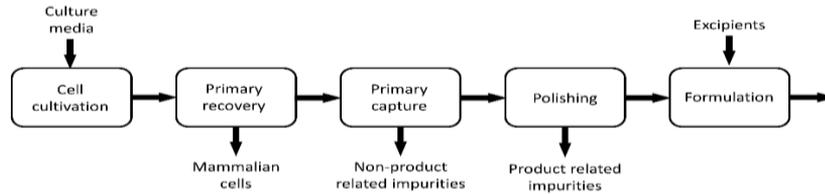


Figure 1: Generalised process platform for the manufacture of biotherapeutic proteins.

The Quality by Design (QbD) framework set out by the U.S. Food & Drug Administration (FDA) highlights the need to explore continuous processing for many therapies, and our recent paper stresses the role of parametric model analyses (Jones and Gerogiorgis, 2021). The present study capitalises on dynamic modelling and optimisation methods towards probing technoeconomic potential for fed-batch and perfusion bioreactors to manufacture mAbs. Dynamic optimisation solutions are first established in order to elucidate feeding and outlet trajectories for maximised mAb production. Following this, both bioreactor designs and cost components are studied within a technoeconomic framework, providing novel insights into operational efficiency and potential for manufacturing cost reductions.

2. Dynamic Modelling and Optimisation for mAb Production

2.1 Dynamic Model of CHO Cell Fermentation

A widely cited model published by De Tremblay et al. (1992) has been selected as a basis for the dynamic study, the state variables of which are presented in Table 1. Growth rates, consumption rates and productivities have been modelled using Monod kinetics, and all relevant kinetic parameters are detailed in the original publication. In this biosystem, hybridoma cells secrete mAbs and growth limiting metabolites (lactate and ammonia). Throughout the time domain, both designs are inoculated with glucose and glutamine substrates, to promote culture growth. In the perfusion bioreactor design, the outlet flow consists of a bleed flow to modulate cell density, and a harvest flow to recover protein.

Table 1: Dynamic model for mAb production (De Tremblay et al., 1992).

$$\frac{dV}{dt} = F_{in} - F_{out} \quad (1) \quad \frac{d[X_V]}{dt} = (\mu - k_d)X_V - \frac{(F_{in} - F_{harvest})}{V}X_V \quad (2)$$

$$\frac{d[GLC]}{dt} = (GLC_{in} - GLC) \frac{F_{in}}{V} - q_{GLC}X_V \quad (3) \quad \frac{d[GLN]}{dt} = (GLN_{in} - GLN) \frac{F_{in}}{V} - q_{GLN}X_V \quad (4)$$

$$\frac{d[LAC]}{dt} = q_{LAC}X_V - \frac{F_{in}}{V}LAC \quad (5) \quad \frac{d[AMM]}{dt} = q_{AMM}X_V - \frac{F_{in}}{V}AMM \quad (6)$$

$$\frac{d[mAb]}{dt} = q_{mAb}X_V - \frac{F_{in}}{V}mAb \quad (7)$$

2.2 Dynamic Optimisation of Fed-Batch and Perfusion Bioreactors for mAb Production

A technoeconomic comparison of both (fed-batch and perfusion) bioreactors hosting the mAb biochemical system of Table 1 should proceed on the basis of dynamic optimisation, to ensure that the two are used in the most productive (albeit diverse) regime of operation. For both cases, a heavily constrained Non-Linear Programming (NLP) problem emerges. Manipulated and state variable profiles of these ODEs have been discretised in time using orthogonal collocation of finite elements, thus facilitating NLP solver use (Biegler, 2007). The fed-batch bioreactor optimisation is solved with APOPT (Hedengren et al., 2014). The perfusion bioreactor optimisation is solved via IPOPT (Wächter and Biegler, 2006).

3. Results and Discussion

3.1 Dynamic Optimisation of a Fed-Batch Bioreactor for mAb Manufacture

Fed-batch dynamic optimisation has been conducted with the initial conditions and state variable constraints exactly as specified by De Tremblay et al. (1992), enabling the validation of our bioreactor design and optimisation results. Key design trajectories are illustrated in Figure 2. The manipulated (inlet flowrate) variable discretisation is varied between 0.05 L day^{-1} and 0.10 L day^{-1} , to analyse its effect on bioreactor performance. The bioreactor time horizon discretisation has been set to 0.5 days for both optimisations.

The trajectories of the state variables for both fed batch designs are very similar across the entire time domain, suggesting the additional flowrate set points provided by the finer (0.05 L day^{-1}) discretisation step provided no tangible benefit to bioreactor performance. Furthermore, we underline that an improved final mAb titer (157 mg L^{-1}) is achieved in comparison to the previous, lower optimised titer (147 mg L^{-1}) (De Tremblay et al., 1992).

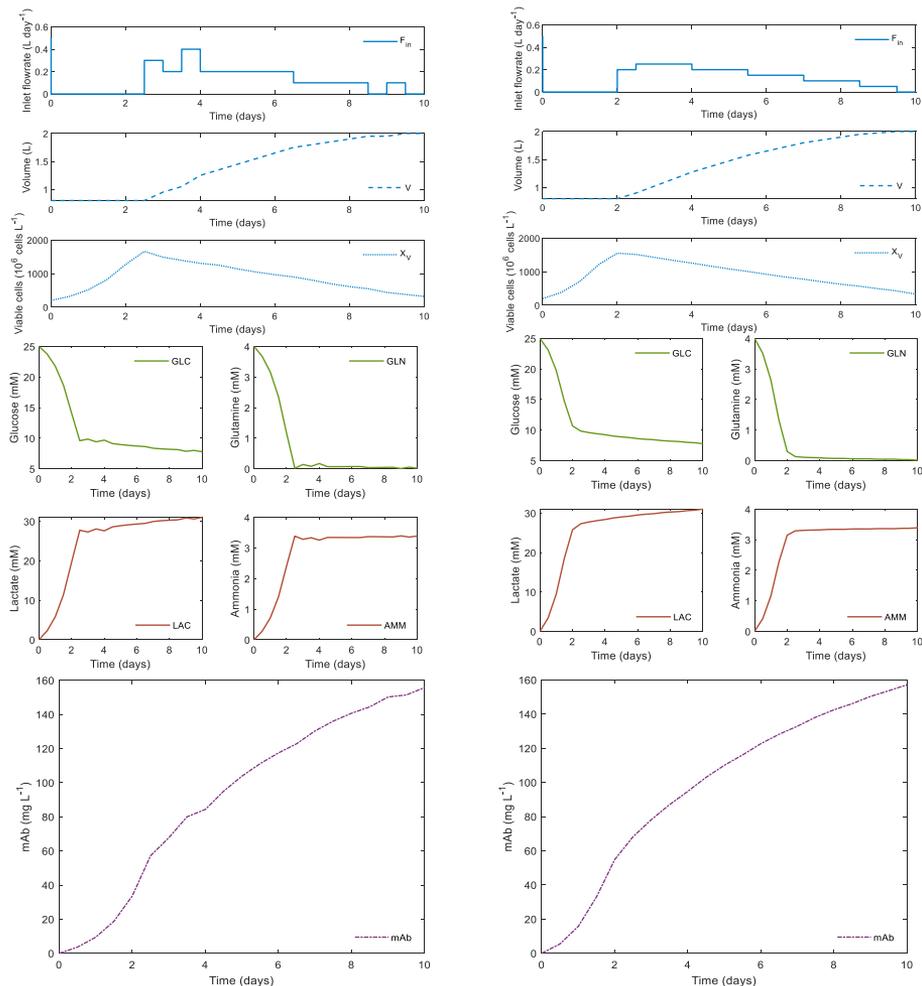


Figure 2: Optimal fed-batch bioreactor trajectories for 0.10 L day^{-1} and 0.05 L day^{-1} (R).

3.2 Dynamic Optimisation of a Perfusion Bioreactor for mAb Manufacture

Identical initial conditions and inlet flowrate constraints have been used for optimising the fed-batch and perfusion bioreactors. Manipulated variables include bleed and harvest flows, and their constraints were set equal to the inlet flowrate constraints. Key variable and mAb titer trajectories for the optimised perfusion reactor are illustrated in Figure 3.

Some fundamental features of the trajectories are summarised to analyse performance. The bleed flowrate remained fixed at zero throughout the time domain so as to retain the maximum number of cells in the reactor during operation, maximizing mAb production. A significantly larger quantity of total feed has to be provided in the optimised perfusion system, in comparison to the optimised fed-batch bioreactor (3.96 L compared to 1.20 L). The ability to harvest growth-limiting by-products of lactate and ammonia in the perfusion system has been shown to be highly beneficial to the bioreactor productivity. The importance of harvesting is evident when comparing final mAb titers for the systems, as perfusion achieves 210 mg L^{-1} , representing a 33.8% increase vs. the fed-batch case.

3.3 Comparative Technoeconomic Analysis of Fed-Batch vs. Perfusion Bioreactors

Extensive technoeconomic analysis of the two optimised bioreactors required a basis of operation and industrial scale unit sizing; reported campaign lengths vary for these, so a typical fed-batch (15 days) and perfusion time (30 days) are explored for both designs. Industrial fed-batch bioreactor sizes reported are much greater than respective perfusion units. This disparity is due to the difference in maturity of the technologies, as less than 10% of commercial biotherapeutics are manufactured via perfusion (Lindskog, 2018). Bioreactor sizes of 2000 L and 12000 L have been previously employed for perfusion and fed-batch operation modes, respectively, in the literature (Farid, 2007; Yang et al., 2019). The nominal mAb basis considered in our analysis is 200 kg annually (Klutzb et al., 2016). Figure 4 shows the annual operating cost breakdown for both bioreactor types, different initial cell densities and campaign times. Despite handling higher cell densities better, increased perfusion bioreactor culture and medium demands induce higher cost for the higher ($40 \cdot 10^6 \text{ cells mL}^{-1}$) density designs. Conversely, the perfusion bioreactor operation shows improved cost viability vs. fed-batch, for the low ($0.2 \cdot 10^6 \text{ cells mL}^{-1}$) density case.

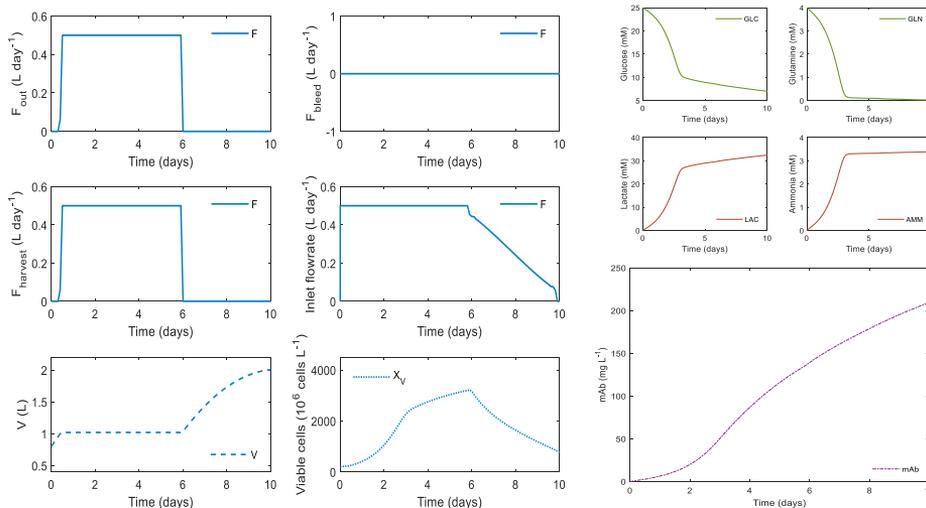


Figure 3: Optimal perfusion bioreactor trajectories of manipulated and state variables

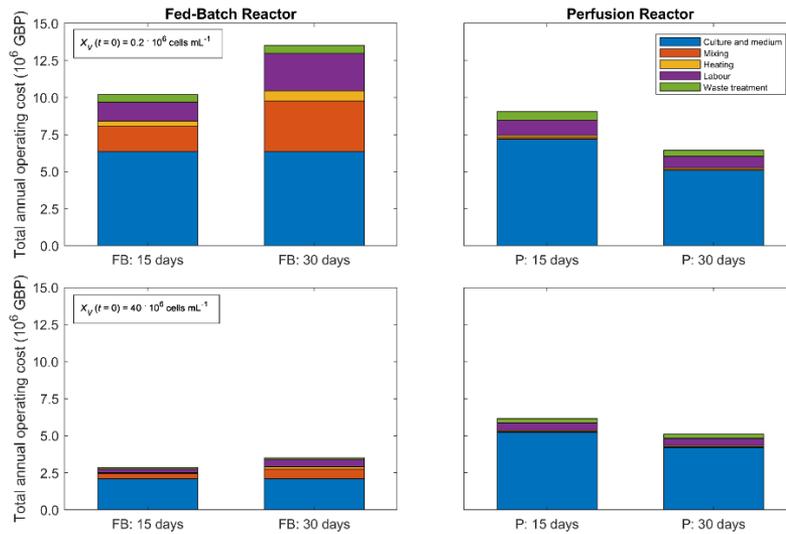


Figure 4: Operating costs of optimised (FB: fed-batch, P: perfusion) industrial designs.

A key difference concerns higher utility (heating and mixing) costs needed for fed-batch reactor vessels, even if fixed (purchase) cost per L is lower. Extended campaigns benefit perfusion bioreactor designs, but adversely affect fed-batch reactor cost-effectiveness. These observations in Fig. 4 plots agree with the established understanding that perfusion bioreactor campaigns should be longer than fed-batch bioreactor ones (Yang et al., 2019).

Finally, it is important to analyse the long-term viability of these designs for industrial implementation. Accordingly, the Cost Of Goods Sold per gram of mAb (COGS) metric is computed for the fed-batch and perfusion designs with lowest operating costs (Eq. 8).

$$\text{COGS} = \frac{\text{NPC} (\pounds)}{\text{Total mAb Production (g)}} \quad (8)$$

The Net Present Cost (NPC) is another key metric, being a function of capital expenditure (CapEx), operational expenditure (OpEx), interest rate (r) and plant lifetime (τ) (Eq. 9).

$$\text{NPC} = \text{CapEx} (\pounds) + \sum_{i=1}^{\tau} \left(\frac{\text{OpEx} (\pounds)}{(1-r)^i} \right) \quad (9)$$

For the technoeconomic analysis performed here, a range of fixed interest rates elucidated the sensitivity of the optimised designs to market pressures. The COGS results for both modes and interest rates between 1-10% (15-year plant lifetime) are shown in Figure 5.

Despite initially being marginally favourable for perfusion design (due to reduced capital expenditure from smaller vessel material), COGS value indicate that perfusion design is eventually outperformed by the fed-batch design soon afterwards (and well ahead of mAb plant lifetime), due to clear Operating Cost (OpEx) differences. At higher interest rates, perfusion mode exhibits a significant COGS value surge towards the end of plant lifetime.

For identical plant lifetime and interest rates, the fed batch design experiences a relatively small change in COGS, reinforcing the conclusion that the more established fed-batch reactor technology is still more cost-effective for mAbs in the biopharmaceutical industry.

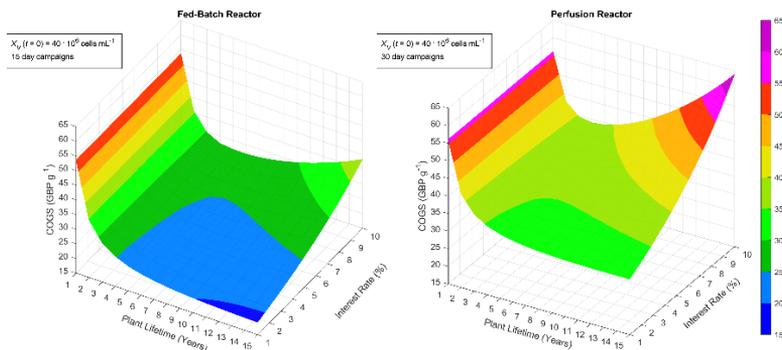


Figure 5: Cost of goods sold per gram analysis for the optimised industrial designs.

4. Conclusions

Technoeconomic analysis via dynamic optimisation of a hybridoma cell line employed for fed-batch as well as perfusion mode mAb production has been addressed in this study. The mAb titer for the optimised fed-batch design is improved compared to the previous literature value reported (De Tremblay et al., 1992). A further enhancement of the mAb titer for bench-scale design has been established for the perfusion bioreactor, highlighting the importance of reactor harvest flowrate in manipulating cell dynamics for proliferation. The economic analysis considers the cost-effectiveness of both (fed-batch and perfusion) bioreactor systems, for two initial viable cell densities and two campaign durations. Larger initial cell densities decrease Operational Cost (OpEx) for both bioreactors, while longer campaign times clearly benefit perfusion vs. fed-batch bioreactor performance. Plant lifetime analysis is conducted via the Cost Of Goods Sold (COGS) per gram metric, showing lower long-term viability of perfusion designs, especially for high interest rates.

References

- G.P. Adams, L.M. Weiner, 2005, mAb therapy of cancer, *Nat. Biotechnol.* 23(9), 1147–1157.
- L.T. Biegler, 2007, An overview of simultaneous strategies for dynamic optimization, *Chem. Eng. Process.* 46(11), 1043–1053.
- M. De Tremblay et al., 1992, Optimization of fed-batch culture of hybridoma cells using dynamic programming: single and multi feed cases, *Bioprocess Eng.* 7(5), 229–234.
- S.S. Farid, 2007, Process economics of industrial mAb manufacture, *J. Chromatogr. B* 848:8–18.
- A.L. Grilo, A. Mantalaris, 2019, The increasingly human and profitable mAb antibody market, *Trends Biotechnol.* 37:9–16.
- J.D. Hedengren et al., 2014, Nonlinear modeling, estimation and predictive control in APMonitor, *Comput. Chem. Eng.* 70:133–148.
- W. Jones, D.I. Gerogiorgis, 2021, Parametric analysis of mammalian cell culture performance for advanced mAb biopharmaceutical manufacturing, *Comput.-Aided Chem. Eng.* 50:1923–1928.
- S. Klutz et al., 2016, Cost evaluation of Ab production processes, *Chem. Eng. Sci.* 141, 63–74.
- E.K. Lindskog, 2018, The upstream process, in: *Biopharmaceutical Processing: 625–635*, Elsevier.
- J. Pollock et al., Integrated continuous bioprocessing, *Biotechnol. Prog.* 33(4), 854–866.
- H. Schulze-Koops, P. Lipsky, 2000, Anti-CD4 monoclonal antibody therapy in human autoimmune diseases, *Curr. Dir. Autoimmun.* 2, 24–49.
- H. Shirahata et al., 2019, Dynamic modelling, simulation and economic evaluation of two CHO cell-based production modes toward biopharma processes, *Chem. Eng. Res. Des.* 150, 218–233.
- A. Wächter, L.T. Biegler, 2006, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Programming* 106, 25–57.
- O. Yang et al., 2019, Comparison between batch and continuous monoclonal antibody production and economic analysis, *Ind. Eng. Chem. Res.* 58(15), 5851–5863.

Control of an industrial packed extraction column for biodiesel washing

Manuel L. Pinho^a, Catarina G. Braz^a, Henrique A. Matos^a, Carla I.C. Pinheiro^b
and José F. O. Granjo^{a,*}

^a*CERENA, DEQ, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal*

^b*CQE—Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal*

**josegranjo9@tecnico.ulisboa.pt*

Abstract

The present contribution addresses the design of feedback control in a biodiesel washing column with structured packing. The column dynamics is described with a pseudo-homogeneous dispersion model previously developed in gPROMS[®] ModelBuilder and validated against industrial data. The operation of the extraction column was analysed for scenarios where biodiesel composition changes, presence of contamination in biodiesel inlet, and change of glycerol concentration in the washing water inlet. The impact of these disturbances on the column hydrodynamics (e.g., flooding, phase inversion) and product quality was analysed. Finally, the performance in open-loop, closed-loop, and closed-loop with decouplers of the system was carefully compared.

Keywords: Biodiesel, Liquid-liquid extraction, Process Control, Dynamic simulation

1. Introduction

Biodiesel (fatty acids methyl esters or FAME) is a recognised renewable substitute of fossil diesel, and its production in Portugal has been steadily increasing with an average annual growth rate of 4.1 % since 2013 (Knoema, 2021). A recent trend in the European Union is the increasing incorporation of waste cooking oils (WCO) in biodiesel production (Grinsven et al., 2020), due to its lower cost and by new regulations limiting the production of first-generation biodiesel. This trend causes the formulations of raw oil to fluctuate significantly, requiring more process flexibility and adequate control systems to maintain the product within strict quality standards and prevent operational problems.

Existing literature devoted to process control in the biodiesel industry focuses on the transesterification reaction section, while the biodiesel washing step is described with shortcut methods (Brásio et al., 2016). However, the increase of the incorporation of lower-quality feedstocks in biodiesel processing may originate operation problems in the column due to subtle changes in the oil composition, leading to more frequent shutdowns and higher maintenance costs. As such, simplified models are of limited use to mitigate these issues and design effective control systems.

This work aims to implement a robust control system based on a real industrial biodiesel washing column model to prevent operational problems and keep the product within quality standards. The remaining of the paper is structured as follows: the implementation and tuning of the control scheme are demonstrated in Section 2; the results are presented in Section 3, followed by final remarks in Section 4.

2. Development of the control strategy for the extraction column

The dynamic model of the extraction column was developed in gPROMS[®] ModelBuilder following a detailed pseudo-homogeneous approach shown in (Pinho, 2021). The washing column is static, has structured packing and operates isothermally. The general transient mass balance equations for the continuous and dispersed phases, the boundary and initial conditions, and the hydrodynamic behaviour of the column are detailed in Mohanty (2000) and Pinho (2021) based on the most accurate state of the art. The overall model was validated with data retrieved from an industrial unit, Pinho (2021). For simplicity sake, all input/output variables will be presented in this section as deviation variables. This analysis was conducted with WCO-based biodiesel at 45 °C and a solvent-to-feed ratio (S/F) of 13.5 vol%.

Regarding extraction columns, most control systems focus on the interface level between the two phases inside the column and outlet concentrations control. If this level is not adequately stabilised, the dispersed layer can flood, causing the loss of solvent and product (Weinstein et al., 1998). Typically, the manipulated variables (MV) used to control the holdup and product composition are the dispersed-phase flow rate, the continuous-phase feed/effluent flow rate and the rotor speed (in the case of agitated Kühni columns).

2.1. Open-loop testing and system linearisation

The impact of disturbances in input variables related to the biodiesel formulation, such as interfacial tension, density, viscosity, and inlet glycerol composition (that reflects a poor phase separation) on the glycerol composition (wt%) of the biodiesel outlet stream, flooding percentage of both phases, holdup and in the phase inversion parameter (χ) was analysed. The effect of step disturbances in the inlet and outlet biodiesel and water mass flow rates over the outputs mentioned above are also shown.

To proceed with the development of a control system in MATLAB[®], it was necessary to linearise the system, which was achieved with the *System Identification* toolbox. All transfer functions had a fit to the estimation data higher than 94 % and are listed in Tables 1 to 3.

Table 1: Transfer functions for the outlet flow rates and output variables.

Transfer Functions	$Q_{\text{out water}}$ (kg/h)	$Q_{\text{out biodiesel}}$ (kg/h)
Biodiesel Glycerol Composition (wt%)	$\frac{-5.98E(-3)}{s^2+121.6s+644.9}$	$\frac{0.016s+0.135}{s^3+137.3s^2+8362s+4.40E04}$
Flooding Continuous Phase %	$\frac{359.9s^2+3.15E04s+2.84E05}{s^3+6.75E04s^2+5.98E06+5.52E07}$	$\frac{358.3s+2.50E04}{s^2+4.93E04+3.41E06}$
Flooding Dispersed Phase %	$\frac{-58.41s^2-4657s-3.88E04}{s^3+6.21E04+5.08E06+4.67E07}$	$\frac{77.49s+4937}{s^2+8.70E04s+5.34E06}$
Holdup (fraction)	$\frac{0.851s^2+14.77s+52.96}{s^3+4.47E04s^2+7.90E05+2.82E06}$	$\frac{1.664s+6.588}{s^2+5.51E04+2.18E05}$
Phase Inversion Parameter	$\frac{1.34s+18.27}{s^2+3.65E04+5.28E05}$	$\frac{2.305s+171}{s^2+4.21E04+3.11E06}$

2.2. Variable pairing

The classical methods of Relative Gain Array (RGA) and Singular Value Analysis (SVA) were used to analyse the multivariable process control problem (Seborg et al., 2011). Since manipulating both inlet and outlet flow rates for biodiesel and water would lead to an over-specified system, only the outlet streams are considered for control purposes. On the other hand, Weinstein et al. (1998) demonstrated that manipulating the outlet instead of the inlet flow rate would lead to smoother system behaviour. Consequently, the mass fraction of glycerol in biodiesel and holdup are the two controlled variables. The recommended pairings from RGA and SVA the pairings: glycerol outlet composition — water outlet flow rate, and holdup — biodiesel outlet flow rate are recommended. These pairings have a λ of 0.83 (compared to 0.82 and 1.40) and a condition

Table 2: Transfer functions for the interfacial tension and feed glycerol composition.

Transfer Functions	Interfacial Tension (N/m)	Feed Glycerol Composition (wt%)
Biodiesel Glycerol Composition (wt%)	$\frac{89.82s^2+1.37E04s+9.37E04}{s^3+158.7s^2+9404s+5.43E04}$	$e^{-0.0277s} \frac{3007s+7.68E04}{s^2+253.2s+1883}$
Flooding Continuous Phase %	$\frac{-1.26E06s^3-1.92E08s^2-1.21E10s-1.17E11}{s^4+224.5s^3+3.41E04s^2+2.16E06s+2.10E07}$	$\frac{e^{-2.03s}(2.22E05s^2+3.18E06s-1.27E07)}{s^4+411.8s^3+2.22E04s^2+5.16E05s+2.80E06}$
Flooding Dispersed Phase %	$\frac{-3.25E05s^2-2.37E07s-2.13E08}{s^3+141.5s^2+1.04E04s+9.38E04}$	$e^{-0.003s} \frac{-261.3s+2562}{s^2+74.5s+607.5}$
Holdup (fraction)	$\frac{1.313s-11.59}{s^2+90.95s+634.3}$	$\frac{2.37s-211.3}{s^2+81.64s+549.6}$
Phase Inversion Parameter	$\frac{13.08s+10.25}{s^2+95.77s+760.4}$	$\frac{-109.1s-3871}{s^2+342.9s+2579}$

Table 3: Transfer functions for biodiesel density.

Transfer Functions	Biodiesel Density (kg/m ³)
Biodiesel Glycerol Composition (wt%)	$\frac{0.754s^2+264s+2494}{188.1s^3+2.43E04s^2+1.37E06s+1.19E07}$
Flooding Continuous Phase %	$\frac{141.7s^4+4.15E04s^3+5.19E06s^2+3.56E08s+2.81E09}{276.8s^4+8.08E04s^3+1.02E07s^2+6.89E08s+5.45E09}$
Flooding Dispersed Phase %	$\frac{-13.59s^3-2046s^2-1.13E05s-8.30E05}{187.7s^3+2.80E04s^2+1.58E06s+1.18E07}$
Holdup (fraction)	$\frac{2.65E(-3)s^3+0.517s^2+41.69s+298.2}{s^3+195.5s^2+1.57E04s+1.13E05}$
Phase Inversion Parameter	$\frac{5.96E(-3)s^3+0.946s^2+59.67s+421.3}{s^3+158.5s^2+9974s+7.05E04}$

number of 3.78, whilst other pairings have a condition number of ∞ . The other possible pairings presented singular steady-state gain matrices, meaning that they are ill-conditioned. The control scheme implemented is illustrated in Figure 1. The nomenclature used in Figure 1 will refer to the controllers in the following subsection.

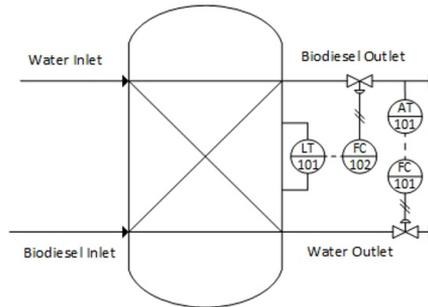


Figure 1: Control scheme for the biodiesel extraction column.

2.3. Controller tuning

Knowing which variables to pair, it is necessary to tune the controllers. The controllers tested were evaluated based on integral error criteria in set-point tracking and disturbance rejection. The three most used integral error criteria are the integral of the squared error (ISE), the integral of the absolute value of the error (IAE), and the integral of the time-weighted absolute error (ITAE). Here, the IAE was the deciding factor to select the controllers. The control system was also analysed in terms of relative stability with the concepts of Gain Margin (GM) and Phase Margin (PM) (Seborg et al., 2011). Saturation was added to the MV, -10% to 10% for biodiesel and -20% to 20% for water mass flow rates. For this process, set-point changes are not expected, so only the results for disturbance rejection are shown.

Starting with the FC101 controller, the Ziegler-Nichols (ZN) and Internal Model Control (IMC) tuning methods were tested. These methods and respective tuning procedures are detailed in Seborg et al. (2011). The controller settings for the different tuning methods are presented in Table 4, with the values of GM and PM. The controller is in parallel, and the derivative mode has a filter N . For the disturbance rejection, a disturbance of -80% in the interfacial tension was applied to the system. The PI controller tuned with IMC has the smallest value of IAE.

Table 4: Controller settings for the controller FC101 according to ZN and IMC tuning methods and Gain and Phase Margins.

Tuning Method		ZN		IMC	
Controller mode		PI	PID	PI	PID
Controller Settings	K_P	-2.05E06	-2.73E06	-2.24E06	-1.07E09
	K_I	-7.21E07	-1.60E08	-6.36E07	-1.41E11
	K_D	-	-1.16E04	-	-8.08E05
	N	-	100	-	100000
Relative Stability	GM	68	1392	67	2
	PM	34	23	37	47

Similarly to FC101, the controller FC102 was tuned with two different methods. Both ZN and IMC tuning approaches were tested; however, the IMC tuning technique (in Mathematica[®]) failed to return a viable controller. Hence, the Cohen-Coon (CC) tuning method was used (Cohen and Coon, 1953). The controller settings and relative stability margins are shown in Table 5. Analysing this table and taking the IAE criterion as the deciding factor, the PI controller tuned with CC rules has a slightly better performance when compared with the other ones.

Table 5: Controller settings for the controller FC102 according to ZN and CC tuning methods and Gain and Phase Margins.

Tuning Method		ZN		CC	
Controller mode		PI	PID	PI	PID
Controller Settings	K_P	3.02E04	4.03E04	5.76E04	5.34E04
	K_I	1.64E06	3.65E06	9.21E06	5.18E06
	K_D	-	111	-	91.7
	N	-	100	-	1000
Relative Stability	GM	∞	∞	∞	∞
	PM (°)	156	∞	125	∞

Two decouplers were added to minimise the interaction between the closed-loops, which is observable by the peaks in Figure 3, for example, in Section 3. However, analysing the transfer function of one of the decouplers, a real right-half plane pole was identified, which makes it unstable. Hence, static decouplers were used instead.

3. Results

After implementing the multiloop with decouplers, the system response was evaluated in open-loop, closed-loop, and closed-loop with static decouplers to disturbances in interfacial tension, biodiesel density and feed glycerol composition are presented in Figures 2 to 4. The response of the MV is not shown here.

It is observed that the addition of decouplers to the control system was beneficial. The control of holdup significantly improved with the reduction of settling-time and lower overshoots; the control of biodiesel glycerol composition was not ideal due to an increase in oscillation; however, the overshoots were minimised. Additionally, the closed-loop interactions were reduced by adding decouplers, which successfully reduced excessive controller action.

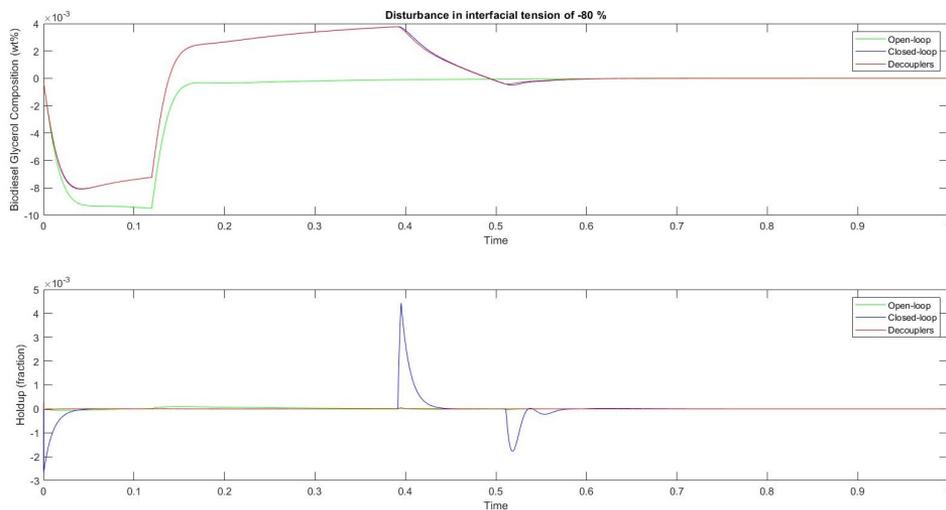


Figure 2: Evolution of the controlled variables: biodiesel glycerol composition (wt%) and holdup fraction to a disturbance in interfacial tension of -80% .

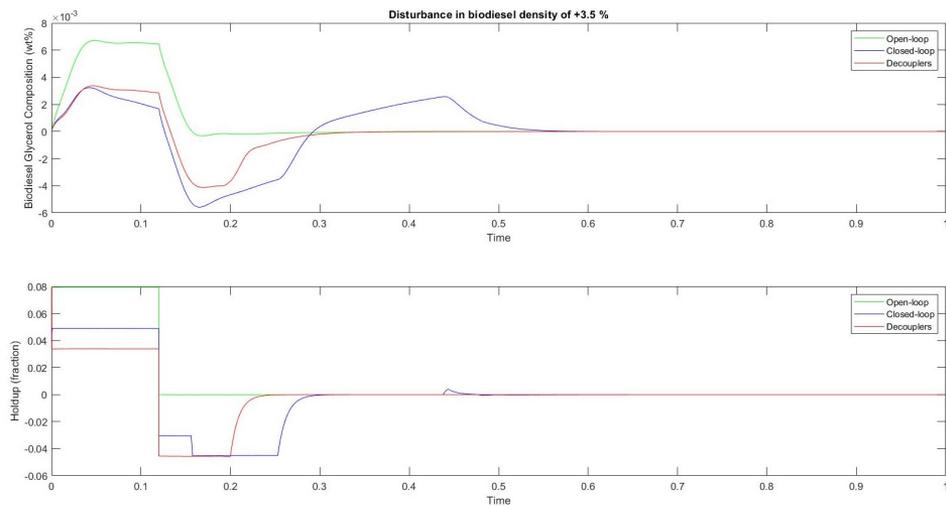


Figure 3: Evolution of the controlled variables: biodiesel glycerol composition (wt%) and holdup fraction to a disturbance in biodiesel density of 3.5% .

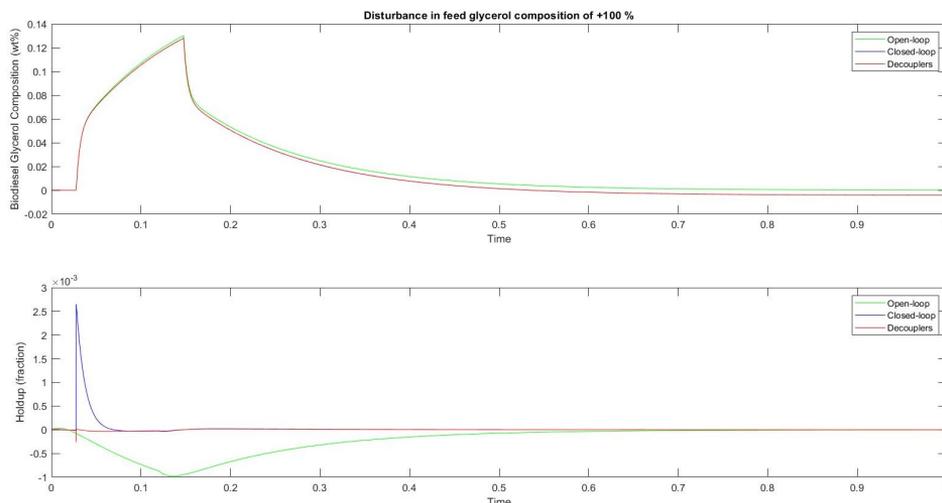


Figure 4: Evolution of the controlled variables: biodiesel glycerol composition (wt%) and holdup fraction to a disturbance in feed glycerol mass composition of 100 %.

4. Conclusion and Future Work

The present work focuses on the implementation of a 2×2 control system on a dynamic model of a biodiesel washing column. The model was developed in gPROMS[®] ModelBuilder using a rate-based approach and describing the dispersed phase as pseudo-homogeneous. In this model, mass transfer phenomena, hydrodynamics and biodiesel properties are estimated. The model was validated with industrial data provided by an industrial partner and successfully predicts the solubility of the solutes in the washed biodiesel.

The system is analysed to implement a control scheme. The best-suggested pairing is Glycerol outlet composition in the biodiesel-water outlet flow rate and holdup-biodiesel outlet flow rate. The controllers are then tuned with the IMC and CC methods, respectively. The addition of static decouplers is successful in minimising strong loop interactions. Overall, the control system has a good performance for set-point tracking and disturbance rejection.

Acknowledgements

The authors acknowledge the financial support by CERENA (strategic project FCT-UIDB/UIBD/04028/2020).



References

- A. S. Brásio, A. Romanenko, N. C. Fernandes, L. O. Santos, 2016. First principle modeling and predictive control of a continuous biodiesel plant. *J. Process Control* 47, 11–21.
- G. H. Cohen, G. A. Coon, 1953. Theoretical considerations of retarded control. *Trans. ASME* 827 (75).
- A. Grinsven, E. Toorn, R. Veen, B. Kampman, 2020. Used Cooking Oil (UCO) as biofuel feedstock in the EU. Tech. rep. Knoema, 2021. International Energy Data, Monthly Update. Accessed in 05-08-2021. URL <https://knoema.com/EIAINTL2018May/international-energy-data-monthly-update#>
- S. Mohanty, 2000. Modeling of liquid-liquid extraction column: A review. *Rev. Chem. Eng.* 16 (3), 199 – 248.
- M. L. Pinho, 2021. Dynamic simulation of biodiesel washing column. Master's thesis, Instituto Superior Técnico.
- D. Seborg, T. Edgar, D. Mellicamp, F. Doyle III, 2011. *Process Dynamics and Control*, 3rd Edition. John Wiley & Sons.
- O. Weinstein, R. Semiat, D. R. Lewin, 1998. Modeling, simulation and control of liquid-liquid extraction columns. *Chem. Eng. Sci.* 53 (2), 325–339.

Development of mechanistic reduced order models (ROMs) for glidant and lubricant effects in continuous manufacturing of pharmaceutical solid-dosage forms

Sunidhi Bachawala^a and Marcial Gonzalez^{a,b}

^a*School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA*

^b*Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN 47907, USA*
marcial-gonzalez@purdue.edu

Abstract

As the pharmaceutical industry transitions from batch to continuous manufacturing, real-time monitoring, and mechanistic model-based control are essential to conform to FDA quality standards. Glidants and lubricants are known to affect the Critical Quality Attributes (CQAs) of a tablet such as tensile strength, tablet porosity, and dissolution profile (Razavi et al., 2018; Apeji and Olowosulu, 2020). Quantitative models for predicting these effects are essential for enabling centralized control strategies of lubricant and glidant feeding and blending in direct compression tableting lines. This work presents the development of mechanistic reduced order models to capture the effects of lubricant (magnesium stearate) and glidant (silica) on CQAs and Critical Process Parameters (CPPs). A Latin Hypercube experimental campaign with thirty different mixing conditions of silica with MCC (Avicel PH200) and APAP (Acetaminophen) was carried out using a Natoli NP400 tablet press and a SOTAX AT4 tablet tester. Experiments show that the tensile strength and blend bulk density are significantly affected by the mixing conditions of silica. Similarly, adding magnesium stearate (MgSt) changes the bulk density of the blend, compaction force required to form a tablet, and tensile strength of the tablet, depending on the lubrication conditions (Mehrotra et al., 2007; Razavi et al., 2018).

Keywords: Lubricant effects, glidant effects, continuous pharmaceutical manufacturing

1. Introduction

The production of tablets in the pharmaceutical industry has predominantly been operated in batch mode, where tablets are produced using a specified amount of raw materials within a given time frame. In recent years, with the advent of process analytical technology (PAT) sensors, the transition from batch to continuous manufacturing has been made possible. However, modeling and advanced understanding of the tablet production process is essential to implement continuous manufacturing in the pharmaceutical industry. In particular, active process control using the Quality-by-Control (QbC) approach (Su et al., 2019) requires predictive and fast models of the tablet Critical Quality Attributes (CQAs). Therefore, steady-state mechanistic models which can predict Critical Process Parameters (CPPs) and CQAs of tablets are essential to the implementation of robust control strategies in the direct compression tableting line.

In continuous manufacturing, it is essential that the powder has good flowability. A glidant, such as colloidal silica (Silica), is an excipient added to improve powder flowability. A lubricant, such as magnesium stearate (MgSt), helps reduce internal friction during compaction and tablet-tooling

friction during ejection. Typically formulations use 0.25%-1% w/w of MgSt and 0-0.02% w/w of Silica. Even when added in such small amounts, these excipients significantly affect the bulk properties of the powder, such as bulk density (Mehrotra et al., 2007), and surface properties, such as the strength of solid bridges formed during compaction. This change in the properties of the blend naturally impacts the tableting process and the CQAs of the final tablet (Razavi et al., 2018; Van Veen et al., 2005), including its dissolution profile, and in turn, the bioavailability of the active pharmaceutical product (API).

It is well-known that MgSt increases the bulk density of the powder and reduces the tensile strength of the tablets (Mehrotra et al., 2007). Quantitative models to describe the effects of MgSt on the tensile strength and elastic modulus of tablets were proposed by Razavi et al. (2018). However, little research has been carried out to quantify the effects of Silica. In this paper, mechanistic reduced-order models of the entire tableting process, i.e., algebraic models based on the understanding of the underlying physical mechanisms which describe the effects of Silica and MgSt on tablet CQAs, are proposed and calibrated. These models are a means to implement moving horizon estimation-based non-linear model predictive control (MHE-NMPC) for the tablet press at Purdue's pharmaceutical continuous manufacturing pilot plant (Huang et al., 2021).

The paper is organized as follows. In Section 2, the design of experiments (DoE) to study both MgSt and Silica is described. In Section 3, the quantitative models proposed for CPPs and CQAs of the tableting process are discussed, with a special focus on the differences between MgSt and Silica effects. Section 4 discusses conclusions and directions for future work.

2. Materials and Methods

Two experimental campaigns were carried out, one to study the effects of Silica and the other to study the effects of MgSt. The materials used in this study were microcrystalline cellulose Avicel PH200 (MCC), 10% w/w acetaminophen (APAP), and Silica and MgSt at different concentrations. The DoE to study the effects of Silica was carried out using a mixture of MCC, APAP, and Silica. MCC and APAP were mixed in a Tote blender with 0-0.2% w/w Silica for 10-30 minutes. The in-die thickness was kept constant at 3.1 mm and the dosing position was chosen to be 7-11 mm, to manufacture tablets having a broad range of relative densities, i.e., 0.6-0.9. The tablet press turret speed was varied between 25-35 rpm. A Latin Hypercube sampling (Viana, 2013) of turret speed, dosing position, concentration, and mixing time for 30 experiments was created using the MATLAB function `lhsdesign`. The same procedure was repeated for the second DoE to characterize the effects of MgSt. Blends with 0-2% w/w MgSt, APAP, and MCC were prepared by mixing in a Tote blender for 11-30 minutes. The in-die thickness was also chosen to be 3.1 mm, the dosing position 9-13 mm, and the tablet press turret speed 11-20 rpm. Tablets formed with MgSt blends required higher dwell time as compared to Silica blends and hence lower turret speeds were used in the MgSt DoE. In this case, a Latin hypercube design of 20 experiments was created. Next, tablets were manufactured using a Natoli-NP400 tablet press using D-type tooling with shallow cup depth. For each run in the DoE, a SOTAX AT4 tablet tester was used to measure tablet thickness, diameter, weight, and hardness of 50 tablets under steady-state manufacturing conditions. The tablet press hopper was filled with 0.5 kg of the blend at the start of each experimental run.

3. Reduced Order Models

The bulk density of a powder ρ_b is observed to be affected by shear strain γ imparted to the powder during mixing (Mehrotra et al., 2007). As the total shear increases, the bulk density initially increases and ultimately reaches a plateau, during which no further change in the bulk density is observed. The following asymptotic relationship between the bulk density and total shear is

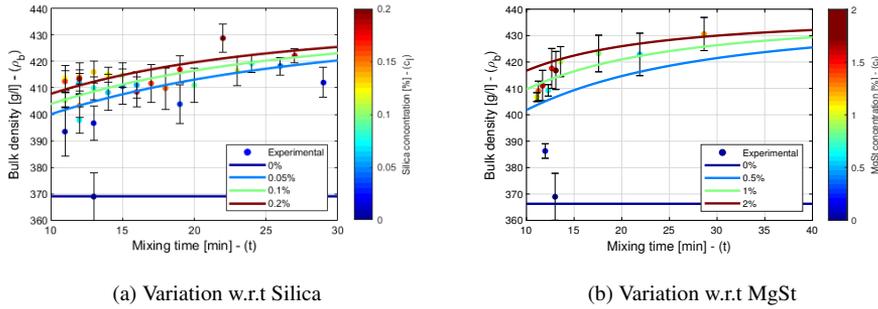


Figure 1: Bulk density increases with increase in concentrations for both Silica and MgSt.

proposed to capture this trend

$$\rho_b = \rho_{b,\infty} - \frac{\rho_{b,\infty} - \rho_{b,0}}{1 + C_\rho} \quad \text{with} \quad C_\rho = \frac{c_l^{r_1} (\gamma + \gamma_0)^{r_2}}{r_3} \quad (1)$$

where $\rho_{b,\infty}$ and $\rho_{b,0}$ represent the bulk densities when the shear imparted is infinite and zero respectively, C_ρ is a lumped parameter which defines the glidant or lubricant mixing conditions, where c_l is glidant or lubricant concentration, γ is the shear imparted to the powder during mixing, γ_0 is the initial shear imparted during pre-blending, and r_1, r_2, r_3 are fitting parameters. The total shear $\gamma + \gamma_0$ is considered proportional to mixing time. Figure 1 shows that increasing the concentration or mixing time of Silica or MgSt results in an increase in bulk density. The bulk density gradually increases and reaches an asymptotic value.

The tablet weight, W , is affected significantly by the process parameters such as the turret speed (n_T), feeder speed (n_F), dosing position (fill depth $-H^{\text{fill}}$) and diameter of the tablet (D). If turret speed is too high, then the die may be filled unevenly, resulting in undesired deviation in tablet weight, ultimately affecting the dosage of the active ingredient. The weight of a doubly-convex tablet formed by Natoli D-type tooling with cup-depth, h , is computed as follows:

$$\frac{W}{\rho_b V^{\text{fill}}} = -\xi_1 \frac{n_F}{n_T} + \xi_2 \frac{H^{\text{fill}}}{D} + \xi_3 \left(\frac{H^{\text{fill}}}{D} \right)^2 \quad (2)$$

with the volume of die-cavity, V^{fill} , given by

$$V^{\text{fill}} = \frac{\pi D^2 H^{\text{fill}}}{4} + \frac{\pi h}{6} \left(\frac{3D^2}{4} + h^2 \right) \quad (3)$$

where ξ_1, ξ_2, ξ_3 are fitting parameters. The same model describes the trends for both MgSt and Silica blends.

The main compaction force F_{punch} for the effect of Silica can be estimated using Kawakita equation (Kawakita and Lüdde, 1971):

$$\sigma_{\text{punch}} = \frac{4F_{\text{punch}}}{\pi D^2} = \frac{\rho^{\text{in-die}} - \rho_b / \rho_t}{[\rho^{\text{in-die}}(a-1) + \rho_b / \rho_t] b} \quad (4)$$

with the in-die relative density, $\rho^{\text{in-die}}$, given by

$$\rho^{\text{in-die}} = \frac{W}{\rho_t V^{\text{in-die}}} \quad (5)$$

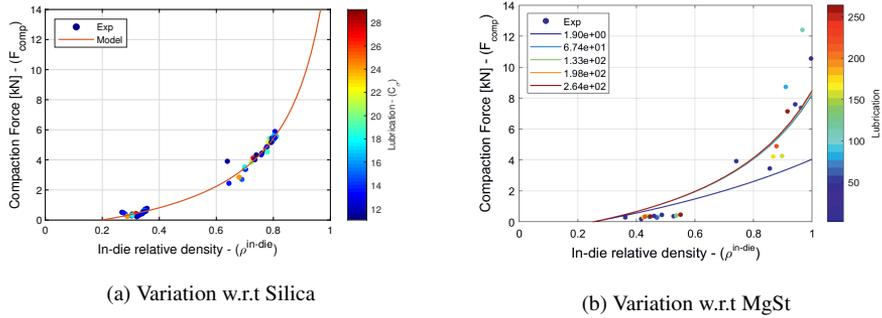


Figure 2: (a) shows that compaction force is independent of mixing conditions of Silica. (b) shows that compaction force decreases with increase in lubrication.

where σ_{punch} is the compaction pressure, ρ_t is the true density of the blend, a and b (MPa) are Kawakita parameters, $V^{\text{in-die}}$ is the volume of die-cavity with main compression thickness $H^{\text{in-die}}$, given by

$$V^{\text{in-die}} = \frac{\pi D^2 H^{\text{in-die}}}{4} + \frac{\pi h}{3} \left(\frac{3D^2}{4} + h^2 \right) \quad (6)$$

The compaction force for MgSt blends depends on lubrication conditions (Figure 2b). This effect is incorporated by modeling the parameter a as

$$a = \frac{a_0 - a_\infty}{1 + C_c} + a_{0,\infty} \quad \text{with} \quad C_c = \frac{c_l^{p_1} (\gamma + \gamma_0)^{p_2}}{p_3} \quad (7)$$

with C_c where $a_0, a_\infty, p_1, p_2, p_3$ are fitting parameters. The compaction force does not depend on the mixing conditions of Silica (Figure 2a), whereas it increases with increasing lubrication 2b.

Elastic recovery, ε_ρ , of a tablet is defined as

$$\rho^{\text{tablet}} = \rho^{\text{in-die}} (1 - \varepsilon_\rho) \quad (8)$$

with the out-of-die tablet relative density, ρ^{tablet} , given by

$$\rho^{\text{tablet}} = \frac{W}{\rho_t V^{\text{tablet}}} \quad (9)$$

and the out-of-die tablet volume after elastic recovery, V^{tablet} , with bellyband, H^{tablet} given by

$$V^{\text{tablet}} = \frac{\pi D^2 H^{\text{tablet}}}{4} + \frac{\pi h}{3} \left(\frac{3D^2}{4} + h^2 \right) \quad (10)$$

Elastic recovery is not sensitive to mixing conditions of Silica (Figure 3a) and is governed by (Gonzalez, 2019)

$$\varepsilon_\rho = \varepsilon_0 \frac{\rho^{\text{in-die}} - \rho_{c,\varepsilon}}{1 - \rho_{c,\varepsilon}} \quad (11)$$

However, with increase in lubrication with MgSt, the elastic recovery increases (Figure 3b). This trend is captured by modifying the ε_0 as follows:

$$\varepsilon_0 = \varepsilon_\infty + \frac{\varepsilon_\phi - \varepsilon_\infty}{1 + C_\varepsilon} \quad \text{with} \quad C_\varepsilon = \frac{c_l^{q_1} (\gamma + \gamma_0)^{q_2}}{q_3} \quad (12)$$

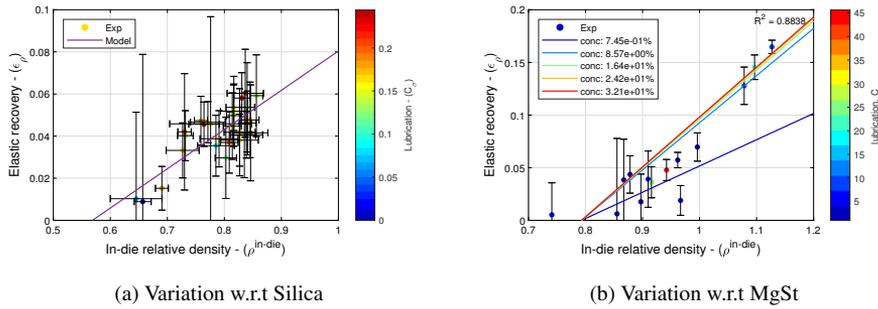


Figure 3: (a) shows that elastic recovery is independent of mixing conditions of Silica, (b) shows that elastic recovery increases with increase in lubrication.

where ϵ_∞ , ϵ_ϕ , q_1 , q_2 , q_3 are fitting parameters.

Tensile strength of a tablet is an important CQA since it is correlated with tablet dissolution. The tensile strength σ_t of a tablet depends on lubricant or glidant concentration and mixing time as follows (Kuentz and Leuenberger, 2000)

$$\sigma_t = \sigma_0 \left[1 - \left(\frac{1 - \rho^{\text{tablet}}}{1 - \rho_{c,\sigma_t}} \right) e^{(\rho^{\text{tablet}} - \rho_{c,\sigma_t})} \right] \quad (13)$$

where, ρ_{c,σ_t} is the critical relative density at which the tablet starts forming and σ_t goes to zero. The tensile strength at zero-porosity, σ_0 , is given by

$$\sigma_0 = \frac{\sigma_{0,\phi}}{1 + C_\sigma} \quad \text{with} \quad C_\sigma = \frac{c_l^{b_1} (\gamma + \gamma_0)^{b_2}}{b_3} \quad (14)$$

and

$$\rho_{c,\sigma_t} = \frac{\rho_{c,\sigma_t,\phi} - \rho_{c,\sigma_t,\infty}}{1 + C_p} + \rho_{c,\sigma_t,\infty} \quad (15)$$

where $\rho_{c,\sigma_t,\phi}$, $\rho_{c,\sigma_t,\infty}$, b_1 , b_2 , b_3 are the fitting parameters (Razavi et al., 2018). $\sigma_{0,\phi}$ and $\rho_{c,\sigma_t,\phi}$ represents the tensile strength and critical relative density corresponding to no lubrication, $C_\sigma = 0$. As the concentration or mixing time of Silica or MgSt increases in the formulation, softer tablets with lower tensile strength are formed (Figure 4). The decrease in tensile strength of lubricated tablets would be due to a combination of changes in physical properties of the blend, as well as the increased elastic recovery of lubricated tablets. Whereas, the tensile strength of tablets formed with Silica blends decreases solely due to changes in physical properties of blended material since elastic recovery is independent of Silica mixing conditions.

4. Conclusion

The results of the experiments demonstrate that the glidant Silica affects the bulk density of the blends, and the tensile strength, and consequently, the dissolution profile, of tablets. In particular, bulk density increases with an increase in glidant concentration or mixing time, whereas tensile strength decreases. However, interestingly, compaction force and elastic recovery show no dependency on the mixing conditions of Silica. In contrast, the lubricant MgSt affects all the CQAs of a tablet. Specifically, bulk density increases with an increase in lubrication, and tensile strength decreases. The elastic recovery and compaction force increase with the increase in lubrication.

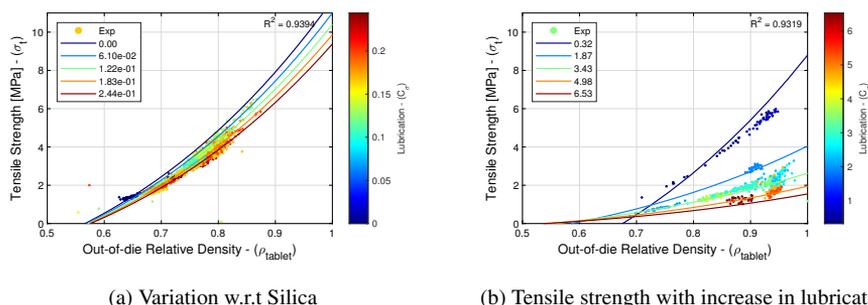


Figure 4: Tensile strength decreases with increase in concentration or mixing time, for both Silica and MgSt.

The effects of Silica and MgSt differ for compaction force and elastic recovery. The practical application of mechanistic models is to use them as steady-state models in MHE-NMPC control of rotary tablet press (Huang et al., 2021). Future work includes integrating the lubricant and glidant feeder with the tablet press to control tablet properties. Additionally, Residence time distribution models (RTDs) will be added to the MHE-NMPC framework to enhance real-time process control.

5. Acknowledgments

The authors gratefully acknowledge the support of the United States Food and Drug Administration under grant 1U01FD006487-01. The authors thank Natoli Engineering Company, and Carmelo Hernandez-Vega for his technical support. The authors thank Dominik Tomasz Nasilowski for his contributions to model development.

References

- Y. E. Apeji, A. K. Olowosulu, 2020. Quantifying the effect of glidant on the compaction and tableting properties of paracetamol granules. *Journal of Research in Pharmacy*.
- M. Gonzalez, 2019. Generalized loading-unloading contact laws for elasto-plastic spheres with bonding strength. *Journal of the Mechanics and Physics of Solids* 122, 633–656.
URL <https://www.sciencedirect.com/science/article/pii/S0022509618302308>
- Y.-S. Huang, M. Z. Sheriff, S. Bachawala, M. Gonzalez, Z. K. Nagy, G. V. Reklaitis, 2021. Evaluation of a combined mhe-nmpc approach to handle plant-model mismatch in a rotary tablet press. *Processes* 9 (9), 1612.
- K. Kawakita, K.-H. Lüdde, 1971. Some considerations on powder compression equations. *Powder technology* 4 (2), 61–68.
- M. Kuentz, H. Leuenberger, 2000. A new model for the hardness of a compacted particle system, applied to tablets of pharmaceutical polymers. *Powder Technology* 111 (1), 145–153.
URL <https://www.sciencedirect.com/science/article/pii/S0032591000002503>
- A. Mehrotra, M. Llusá, A. Faqih, M. Levin, F. J. Muzzio, 2007. Influence of shear intensity and total shear on properties of blends and tablets of lactose and cellulose lubricated with magnesium stearate. *International journal of pharmaceutics* 336 (2), 284–291.
- S. M. Razavi, M. Gonzalez, A. M. Cuitiño, 2018. Quantification of lubrication and particle size distribution effects on tensile strength and stiffness of tablets. *Powder Technology* 336, 360–374.
- Q. Su, S. Ganesh, M. Moreno, Y. Bommireddy, M. Gonzalez, G. V. Reklaitis, Z. K. Nagy, 2019. A perspective on quality-by-control (qbc) in pharmaceutical continuous manufacturing. *Computers & Chemical Engineering* 125, 216–231.
- B. Van Veen, G. Bolhuis, Y. Wu, K. Zuurman, H. Frijlink, 2005. Compaction mechanism and tablet strength of unlubricated and lubricated (silicified) microcrystalline cellulose. *European journal of Pharmaceutics and Biopharmaceutics* 59 (1), 133–138.
- F. A. Viana, 2013. Things you wanted to know about the latin hypercube design and were afraid to ask. In: *10th World Congress on Structural and Multidisciplinary Optimization*. Vol. 19. sn.

Health-aware control using hybrid models applied to a gas-lifted oil well network

José Matias^{a*}, Salmon Y. Ghebredngl^a and Johannes Jäschke^a

^a*Chemical Engineering Department, Norwegian University of Science and Technology, Sem Sælandsvei 4, Kjemiblokk 5, Trondheim, 7491, Norway*

Abstract

A health-aware controller (HAC) aims at finding the optimal trade-off between system reliability and productivity. By considering the system health, the controller can avoid unexpected equipment breakdown while improving the system economic performance. In subsea oil extraction industry, new and more stringent environmental, safety, and regulatory requirements are imposed every year. Thus, this industry can benefit from the implementation of this type of controllers for tasks such as equipment monitoring, or flow assurance. One challenge when implementing a health-aware controller is to derive models that represent the failure mechanism of subsea equipment using physics-based relationships. An alternative is to use data-based models. Instead of relying on physical knowledge, this type of models determine the input-output relationships using only data observed experimentally. In this paper, we propose the use of a hybrid HAC, where first-principles relations describe the mass and energy balances, while the system degradation evolution is represented by a data-driven model. We test two different structures, a linear regression and a neural network model. We investigate their performance in open-loop, carrying out an uncertainty analysis on the predicted degradation variability, as well as in closed-loop. The simulation results show that HAC's performance is dependent on the type of data-driven model used for predicting degradation, and the more complex model does not necessarily give a better overall result.

Keywords: Model predictive control, Hybrid Modeling, Diagnostics and prognostics

1. Introduction

In subsea oil and gas extraction systems, unexpected breakdowns and maintenance interventions are very costly. In such systems, choke valves are critical for the process safety because they are responsible for regulating the system pressure, and controlling the oil and gas flows from the reservoir to the system. The useful life of the choke valves is severely affected by sand particles, which are extracted from the reservoir together with the products of interest. The sand managing strategy is outlined early in the field development to ensure appropriate selection of equipment as well as instrumentation for monitoring, controlling and handling sand production. However, even with all these precautions, a very conservative operational strategy is often adopted. Typically, the production strategy is defined based on worst-case valve degradation scenarios, leading to sub-optimal operation and potential profit loss (Verheyleweghen and Jäschke, 2018).

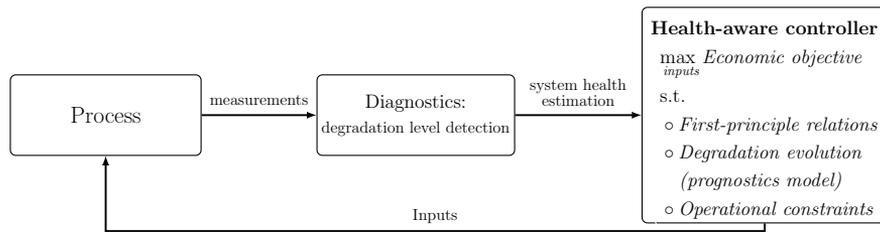


Figure 1: Health-aware controller block diagram, adapted from Matias et al. (2020)

Hence, there is a trade-off between maximizing production and minimizing equipment degradation in subsea production systems that needs to be addressed for optimal operation. Our paper uses a health-aware controller (HAC) (Escobet et al., 2012) strategy to deal with this problem. In HAC, equipment health monitoring tools are integrated to the control structure as illustrated in Fig. 1. The goal is to avoid conservative operation by actively steering plant degradation and preventing unexpected breakdowns.

A HAC was already implemented in a simulated subsea oil well network in Matias et al. (2020). The authors used the phenomenological model of DNV (2015) for erosive wear as the “real” degradation mechanism. They assumed perfect knowledge of the choke valve degradation evolution as well as of the current degradation state, which is unrealistic. In turn, we propose the use a data-driven models of the system, which were obtained by Jahren (2021), for inferring the current valve degradation (*diagnostics*) based on pressure, flow, and temperature measurements along the wells. Then, we use the same model inside the health-aware controller to predict the degradation evolution (*prognostics*).

2. Case Study: Gas-lifted oil well network

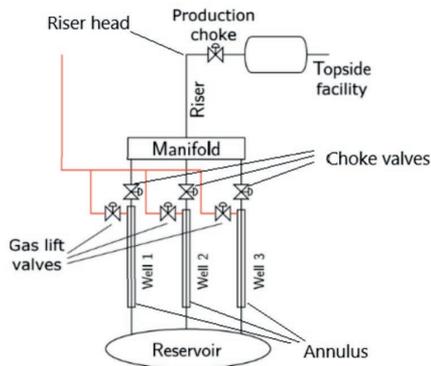


Figure 2: Gas lifted oil production system (Verheyleweghen and Jäschke, 2018)

In subsea systems, wells connect the oil and gas reservoirs to manifolds on the seabed. Then, long pipelines called risers direct the fluids to the topside facilities. The main driving force of this process is the reservoir natural pressure. However, in some cases, this pressure is not high enough to lift the fluids. One alternative to solve this problem is to use gas lift.

In gas lifted wells, gas is injected through an annulus, which is a void between the well and its external casing. By injecting gas, the fluid mixture density decreases. Consequently, the hydrostatic pressure on the reservoir outlet also decreases leading to larger oil outflows. However, if too much gas is injected, the pressure loss due

to friction overcomes the effects of the hydrostatic pressure difference, decreasing the well oil production.

Therefore, for optimizing the production, we want to find the gas lift injection flowrate that maximizes the oil flowrates. On the other hand, larger oil flowrates imply in a larger sand outflow from the reservoir, which decreases the remaining useful life of the choke valves. The degradation indicator here is the length of crack on the internal parts of the valve caused by erosive wear of the sand particles. Hence, by applying a health-aware controller to this system, we want to find the gas lift injection that optimizes this trade-off between optimizing production and extending the valve's remaining useful life.

3. Health-critical constraints and objective function

A challenge in the implementation of the health-aware controller is how to represent the trade-off between production and degradation, as well as defining health-critical constraints. In Matias et al. (2020), the authors represented the degradation threshold as a soft constraint, and then added slack terms in the objective function to penalize violations. This alternative presents two shortcomings. First, the performance of the controller is highly dependent on the weights of the economic and slack terms of the objective function, and tuning them is challenging. Second, from an operational point of view, it is difficult to define a representative degradation threshold for the equipment health.

For handling these issues, we use an exponential utility objective function Φ instead. It represents the trade-off between the risk of the choke valves breaking down and the economic gains. In the subsea oil well network of Fig. 2, the economic objective is to maximize the oil production of the three wells (i.e. $\phi = \sum_{i=1}^3 Q_{o,i}$). Assuming that we know an acceptable range for the degradation of the choke valves d , represented by $[d_L, d_U]$, the resulting exponential utility objective function at time k can then be written as:

$$\Phi(k) = \sum_{i=1}^3 \frac{(1 - e^{-\delta_i(k)Q_{o,i}(k)})}{(\delta_i(k) + \varepsilon)}, \quad \text{where} \quad \delta_i(k) = \frac{d_i(k) - d_L}{d_U - d_L} \quad (1)$$

where, ε is a small value added to avoid division by zero. The intuition behind this choice is the following: let us assume that the gas lift injection in well i , $Q_{gl,i}$, is chosen such that the oil production $Q_{o,i}$ in the well increases. However, the contribution of this gain to Φ depends on the normalized degradation δ_i , as illustrated by Fig. 3. Larger degradation levels lead to a risk-averse operation, in which the controller would accept lower levels of production rather than possibly breaking down the system. On the other hand, for degradation levels equal or lower than d_L , the controller would have a risk-taking approach and always seek the largest production level as possible.

4. Choosing the data-driven model

Next, we explore the combination of two of the data-driven modeling strategies presented in Jahren (2021) (linear regression (LR) and neural network regression (NNR)) with the

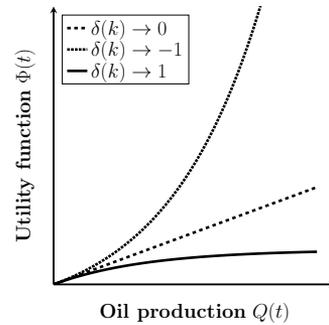


Figure 3: Exponential utility objective function.

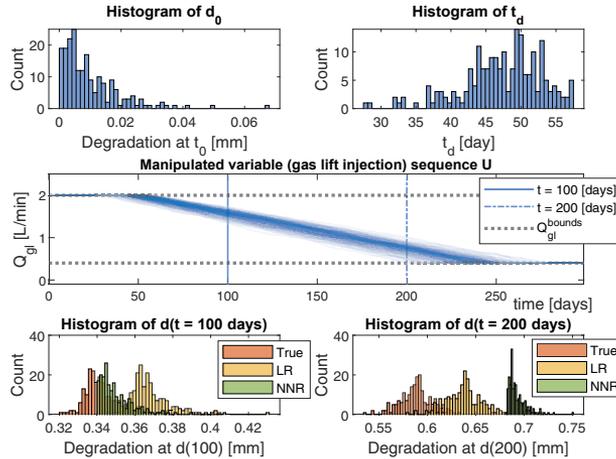


Figure 4: Open-loop Analysis. The bottom plot shows the variability of the predicted system degradation at $k = 100$ and $k = 200$ days.

health-aware controller. For training the models, the authors performed synthetic experiments using the valve degradation model of DNV (2015) as the real system. Both LR and NNR yielded very accurate predictions of the erosion rates on normalised unseen test data. For defining which one should be combined with the controller, we propose two different analysis, one in open-loop and another in closed-loop. In both cases, we also consider the valve degradation model of DNV (2015) as the real system. The structure of the LN and NNR and details about their training and validation can be found in Jahren (2021). The codes of all the simulations are available at <https://github.com/Process-Optimization-and-Control/Health-Aware-Controller>.

4.1. Open-loop analysis: uncertainty propagation through the models

In the open-loop tests, we run an uncertainty analysis that aims at quantifying how the variability of the model inputs affect the outputs of the data-driven models. The model output of interest is the system degradation at time k , $d(k)$, whereas the two model inputs subject to uncertainty are the choke valve degradation at the beginning of a given interval d_0 and the manipulated variable sequence \underline{U} :

$$d(k) = f(d_0, \underline{U}), \quad \text{where} \quad \underline{U} = [u_0, u_1, \dots, u_k] \quad (2)$$

Ideally, we would use historical data from operation under the influence of a health-aware controller to determine the input uncertainty. Since these data may not be available before implementing the controller, we use different probability density functions for representing our *a priori* knowledge/intuition about the operation. According to the results of Matias et al. (2020), we assume that, at the beginning of the simulation, the manipulated variables will be at the upper bound (i.e. trying to produce as much as possible). At time t_d , the controller infers that the degradation will be out of the acceptable degradation range if the production level does not decrease. Then, it starts changing u accordingly.

The described situation can be represented by:

$$\underline{\mathbf{U}} = [u_0, u_1, \dots, u_{t_d}, u_{t_d} - \Delta u_1, \dots, u_{t_d} - \sum_{t=1}^{k-t_d} \Delta u_t], \quad \text{with} \quad u_0 = u_1 = \dots = u_{t_d} = u_{\max} \quad (3)$$

In addition, we assume that the initial degradation d_0 can be higher than zero. For representing the described uncertainty, we use the following probability distributions:

$$d_0 \sim 0.01\text{Exp}(1), \quad \Delta u_t \sim \Delta u_{\max}\text{Beta}(1, 5), \quad t_d \sim 50\text{Weibull}(1, 10) \quad (4)$$

The histograms of d_0 and t_d , as well as the resulting input trajectories $\underline{\mathbf{U}}$ can be seen in Fig. 4. For the open-loop analysis, we generate 200 samples of the distributions, feed them to the model in Eq. 2, and integrate the system forward in time. The output uncertainty is analyzed at arbitrary time instants $k = 100$ and $k = 200$ days (bottom plots). For comparison, we also plot the “real” degradation uncertainty obtained with the model of DNV (2015). Although both data-driven models represented the test data set accurately in Jahren (2021), they were trained and validated in synthetic data generated randomly, which did not correspond to the operation of a health-aware controller. When using data that mimic the operation of a HAC, the neural network represents the output mean better at the beginning; however, the prediction drifts with time. Regarding variability, the linear regression shows a better approximation of the standard deviation at both time instants. Since HAC is implemented in a receding horizon fashion, one may choose to combine the neural network model to the HAC due to the smallest deviation of the mean degradation estimation in the short term. Moreover, due to the fact that its standard deviation is smaller than the true process, the worst-case open-loop predictions of the neural network is likely to be more conservative than the real one, which may be preferable.

4.2. Health-aware controller: Closed-loop results

Next, we run the hybrid health-aware controller in closed loop. The HAC has to manipulate the gas injection of three wells in order to maximize oil production, while prolonging the remaining useful life of the valves as much as possible. The simulation setup is similar to Matias et al. (2020). However, instead of defining a threshold for the degradation, we define an acceptable range, with $d_L = 0$ and $d_U = 0.6$ mm. We also impose bounds on the gas injection $[0.5, 2]$ kg/s. We use a prediction horizon of 50 days and a control horizon of 20 days. We consider that the wells have a constant sand production. However, well 1 produces less sand than the other wells, and wells 2 and 3 have exactly the same sand production. We also assume that well 2 has a larger reservoir outflow. Therefore, the choke valve in well 2 should have the smallest useful life. We also train the data-based models in data only from well 1, leading to plant-model mismatch in wells 2 and 3.

The results are shown in Fig. 5. We assume that the valves break after the crack length reaches the 0.65 mm threshold (which is unknown to the controller). Regarding the system diagnostics (top plot in Fig. 5), the strategies show a significant underestimation of the current valve degradation. We only show the value of well 2, which is the health-critical well. By analyzing the manipulated variables (middle plots in Fig. 5), we see that the NNR-based controller is not able to detect that the system will break until around 130 days. In turn, the LR-based starts to react around 80 days. We hypothesize that this unexpected behavior comes from the fact that the NNR-based controller prediction is worse

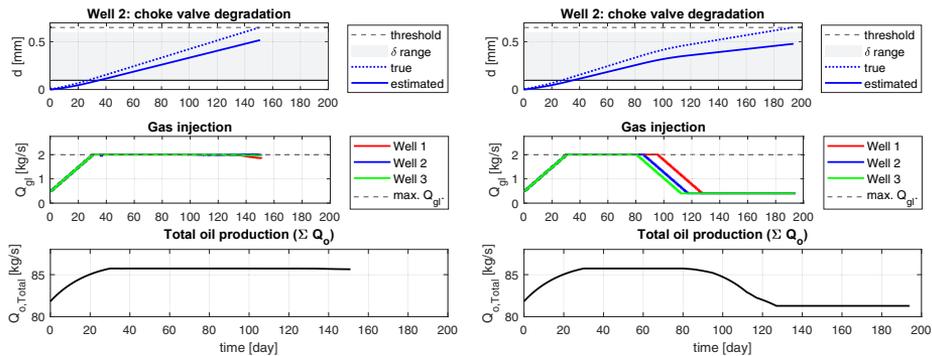


Figure 5: Closed-loop analysis. Plots on the left are related to the NNR-based HAC and plots on the right to LR-based HAC. The LR-based hybrid controller has a total production 25 % higher than the NNR-based. The breakdown time is 194 days for the LR-based versus 151 for the NNR-based.

when the degradation levels are higher, as indicated in the open-loop analysis. The solver then cannot properly extrapolate the system future behavior and fails to decrease the gas injection. However, a rigorous analysis of the neural network prediction behavior is much more challenging than the linear regression due to its black box nature.

5. Conclusion

The hybrid health-aware controller is applied to a synthetic case study of an oil and gas well network with artificial gas-lifting. We test two different data-driven model structures, a simple linear regression and a neural network regression model. The proposed controller is a possible alternative to find the optimal trade-off between increasing production and prolonging equipment useful life. However, plant-model mismatch due to the data-driven models for the diagnostics and prognostics steps had a detrimental effect on HAC's performance. The main conclusion is that HAC's performance is dependent on the type of data-driven model and the quality of its extrapolations capacity. Moreover, control strategies that rely on data-driven models should be tested in both *open* and *closed* loop simulations before implementation. Finally, we showed that using an exponential utility objective function in the health aware controller yields a good representation of the trade-off between system reliability and productivity.

References

- DNV, 2015. Managing sand production and erosion. Recommended Practice DNVGL-RP-0501. DNV GL Company, Oslo, Norway.
- T. Escobet, V. Puig, F. Nejari, 2012. Health aware control and model-based prognosis. In: 2012 20th Mediterranean Conference on Control & Automation (MED). IEEE, pp. 691–696.
- J. H. Jahren, 2021. Data-driven modelling of subsea equipment degradation using simulated and experimental case studies. Master's thesis, NTNU.
- J. Matias, J. Agotnes, J. Jäschke, 2020. Health-aware advanced control applied to a gas-lifted oil well network. IFAC-PapersOnLine 53 (3), 301–306.
- A. Verheyleweghen, J. Jäschke, 2018. Oil production optimization of several wells subject to choke degradation. IFAC-PapersOnLine 51 (8), 1–6.

Applying Ecological Interface Design for Modular Plants: Safety-Demonstrator Case Study

Nazanin Hamedi,^a Leon Urbas,^b Raimund Dachselt^c

^a *DFG Research Training Group 2323 CD-CPPS, Technische Universität Dresden, Georg-Schumann-Straße 7a, 01069 Dresden, Germany*

^b *Faculty of Electrical and Computer Engineering, Chair of Process Control Systems & Process Systems Engineering Group, Technische Universität Dresden, 01069 Dresden, Germany*

^c *Interactive Media Lab Dresden, Technische Universität Dresden, Dresden, Germany*
leon.urbas@tu-dresden.de

Abstract

A promising approach to reduce the time to market of chemical products is the application of modular plants, which are composed of standardized, previously developed Process Equipment Assemblies (PEAs). As this approach goes along with process intensification as well as continuous and highly autonomous operation, human operator performance might benefit from a more structured guidance to fully understand the system and interact with it whenever it is necessary. With the major goal of reducing human error, Vicente and Rasmussen (1990) introduced the Ecological Interface Design (EID), suggesting that an interface should provide a virtual ecology, connecting the work domain to the human operator. EID is based on Work Domain Analysis (WDA), which is a 2-dimensional Abstraction-Decomposition Space. The decomposition hierarchy in the horizontal direction gives information about the superficial structure of the plant; while, the abstraction hierarchy in vertical direction provides multi-level knowledge representation of how each component in the process is working.

In this study, the effects of applying EID for modular plants is investigated. To this end, a pilot modular plant, namely Safety-Demonstrator (Pelzer *et al.*, 2021), has been considered as the use case. It consists of a total of two PEAs: one for feeding the reactants and the other one for the reaction. At first, WDA has been done for this modular plant, and afterwards an EID has been developed. This approach not only helps operator understanding the system and making better decisions in challenging situations, but also shows potential to facilitate exploiting the changeability feature of MPs and selecting the best PEAs for a specific function of the process.

Keywords: Modular Plants; Ecological Interface Design; Safety-Demonstrator Case Study

1. Introduction

The availability of many different alternative products in food, pharmaceutical, and chemical industries shortens the products' lifecycles and results in volatile global market (Lier and Grünewald, 2011). To stay competitive in such a global market, companies should not only be able to produce the new demanded products in a shorter time, but also manage to change their production capacity based on the market. Moreover, to address the different global demands for a product, it is essential to develop mobile production technologies, allowing for the transport of a chemical plant from one location to another.

Designing Modular Plants (MPs) is a promising approach to fulfil these requirements (Holm *et al.*, 2015).

To facilitate flexible production, MPs are designed in a hierarchical structure. The lowest level of the hierarchy denotes to components, which are the non-separable, smallest units of a plant. A combination of components providing a specific process engineering function (e.g. heating) is called Functional Equipment Assemblies (FEA). Then one or more FEAs together build a Process Equipment Assembly (PEA) providing a procedural step (e.g. distillation). Each PEA has its own process automation allowing for decentralized operation. Exchanging information from one PEA to another and to the higher level control layer necessitates a standard data exchange format describing the PEAs. This standardized descriptive document is called Module Type Package (MTP).

MPs are at the top layer of the hierarchy, including at least one PEA. To control the connections between PEAs, a superior automation layer, called Process Orchestration Layer (POL), is also considered.

Considering the abovementioned features of MPs, they offer lots of important benefits. Apart from reducing the plant engineering time by a factor of 50% (Holm, 2016), they also increase the flexibility of production. In fact, not only the PEA change (inter-modular level) results in a new process, but also by changing FEAs (intra-modular level) the operating range can be modified (in the allowable range considered by the PEA manufacturer) (VDI 2776-1, 2020). Nevertheless, to help operators to deal with this complex system and exploit its outstanding features, the right information, with the best format should be provided to them.

In the quest to develop better human-machine interactions, the concept of Ecological Interface Design (EID) was introduced by Vicente and Rasmussen in 1990 (Vicente and Rasmussen, 1990). EID is based on Work Domain Analysis (WDA), which is a two dimensional hierarchical space (abstraction hierarchy (AH) in vertical direction and decomposition hierarchy in horizontal direction), aiming at providing a formalism to represent the work domain. The categorized information in WDA is then visualized in an operator understandable way. This information flow from the work domain to the operator and vice versa is called EID.

WDA and EID as a strong cognitive tool for visualization, has been used in many different fields, such as air traffic control (Ahlstrom, 2005), computer network management (Burns, Kuo and Ng, 2003), road transport structure (Salmon *et al.*, 2019), and medicine (Hajdukiewicz *et al.*, 2001). As a proof of concept, Jamieson and Vicente (Jamieson and Vicente, 2001) applied the EID to petrochemical process engineering. They applied the EID for the process of fluid catalytic cracking. Their results indicated that their developed EID can lead to higher operator adaptation, facilitate continuous learning, and assist distributed, collaborative work.

Considering the complexity of MPs; and at the same time, the promising results from previous applications of EID for chemical processes has motivated us to investigate the application of EID for MPs. To this end, safety demonstrator (Pelzer *et al.*, 2021), an experimental MP in Process to Order (P2O) lab in TU-Dresden has been considered as a use case. This MP consists of two PEAs for feeding and reaction, and each of them includes several FEAs. In the following subsections, at first, a WDA for this use case will be developed, and then, an interface will be suggested.

2. Development of Work Domain Analysis

Figure 1 illustrates the developed WDA for the safety demonstrator. As can be seen, the horizontal direction shows the part-whole relation (decomposition hierarchy) and for the

case of a MP, it can be considered as: $MP \rightarrow PEA \rightarrow FEA \rightarrow \text{Components}$. While this hierarchy deals with the superficial feature of the MP, the AH in the vertical line has a more in-depth viewpoint and focuses on how the plant works. Indeed, as we go down the AH, we are actually answering the question how the goals of previous layer are achievable. As a result, each cell of this two dimensional space provides a model with different amounts of details and yet complete for this work domain. It should be mentioned that describing the WD in the upper layers of the AH is more useful in the top layers of the part-whole hierarchy, and as we go down the AH, lower levels of part-whole decomposition hierarchy tend to be more informative (Son *et al.*, 2019). Each layer of the AH is described in the following:

Goal (also known as functional purpose): this layer illustrates the purpose of the process in the most abstract and therefore less technical way. As can be seen in Figure 1, in this study, three major goals, i.e. 1) production of a chemical, 2) safety, and 3) flexibility of MPs have been considered.

Function (also known as abstract function): this layer answers the question what are the functions that we need to reach the goals of the previous layer. As Figure 1 reflects, this layer of the AH involves the PEA, FEA, and components sections of the decomposition hierarchy. In this study, the control function has been also considered and their relationship between different sections of the process is identified. It should be mentioned that the heat transferred between the process and the ambient has been ignored. As can be seen, this layer does not contain any explanation of the technical terms or phenomena taking place in each equipment. It is noteworthy to mention that in Figure 1 only PEA2 is shown.

Behavior (also known as generalized function): with the help of a causal graph, this layer describes the phenomena taking place in the process. As can be seen in Figure 1, the control functions of previous layer are shown as conditional statements and their influence on each phenomena is clarified. To avoid a too complex graph, Figure 1 only shows the causal graph of PEA 2.

Structure (also known as physical function): the phenomena considered in previous layer are realized in this layer. Indeed, this layer relates the phenomena to the symbols of the equipment in Process and Instrumentation Diagrams (P&IDs). This layer is not drawn in Figure 1.

Physical form: this layer is the closest to the reality and it can be a photo or a video of the real plant, helping the operator making sure about the status of an equipment (e.g. looking at the color of the flames of a flare). In this study, this layer has not been considered.

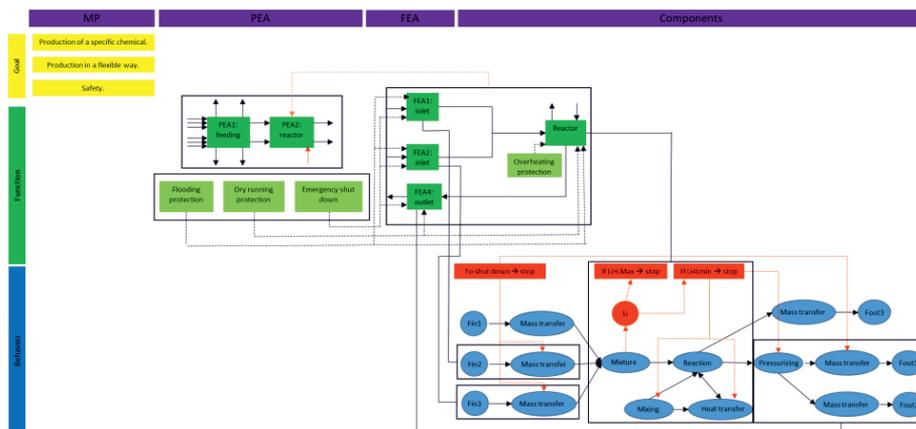


Figure 1: WDA for the safety demonstrator. (The lines between layers shows how to relationship. Function layer: dark green boxes show the PEAs and FEAs, light green boxes are the control functions of the process, solid arrows are mass (black) and energy (red) input and outputs, black dashed arrows show the composed of relation, and orange dashed arrows show aggregation relation. Behavior layer: arrows have causal meaning (black: process related, red: control function related)

3. Mock-up Interface Description

Based on the previously mentioned WDA and to illustrate its specific features, a Human Machine Interface (HMI) has been designed. It is noteworthy to mention that *Axure RP 10* has been used to develop this mock-up interface. In the following subsections the specific features of the EID will be discussed.

3.1. Different Views with Different Amounts of Details

One of the most obvious and yet important characteristics of WDA is its ability to provide different layers of details. Figure 2 (A) demonstrates a general view of the safety demonstrator. In the case of this figure, the green checkmark at the corner of PEA 1 indicates that it is working well; however, the red cross mark in the corner of PEA 2 suggests that something is wrong or maybe some action is necessary in this PEA. Therefore, the operator can further investigate the issue by looking at more details of PEA 2 (Figure 2 (B)). Moreover, different FEAs in a specific PEA are visualized in a categorized way (Figure 2 (B and C)). By using the home and upwards arrow buttons in the bottom of the page, one can respectively go to the overall view and one-layer upper view of the MP. These different views of the process help the operator to find the root cause of an alarm much easier.

3.2. Explanation of the automatic control system

While there has been a great improvement in automatic processes and cyber-physical production systems, it is impossible to omit the operator. In fact, the control systems are getting more and more complex and so does the interaction with them. Hence, if the control system is able to explain how it works and why it is active, it would be a significant help for the operator to integrate with the system (Gil *et al.*, 2019). In the case of this study, the consideration of the control functions (e.g. flooding protection) in the WDA offers an ability of the interface to explain the activated control function. For example, if the flooding protection control function of PEA 2 starts working, the interface should show a message that the liquid level in the reactor is higher than normal and therefore the valves in FEA1 and FEA2 are blocked (see Figure 2 (B)).

3.3. Phenomena representation of the process

Another very interesting feature of WDA is the behavior layer, in which the phenomena taking place in the process is explained in a machine-readable format (in this study with the aim of a causal graph (Figure 2 (D))). This representation of the process results in an ability of the interface to explain the effect of operator decisions on the performance of the process. For instance, there are three major phenomena occurring in the reactor of PEA 2: mixing, heat transfer, and reaction (see the row of behavior in Figure 1). If the operator decides to increase the mixing speed, the interface should be able to popup a message and tell him this action would result in more reaction and heat transfer. From a psychological viewpoint, these types of analysis involve the knowledge based behavior of the human operator (Vicente and Rasmussen, 1992; Jamieson and Vicente, 2001).

In this study, we have only focused on the phenomena of one equipment. However, there is a possibility to use this approach and estimate the effect of changing one parameter on

the next steps of the process. This has a specific importance for MPs, as they are highly changeable and it would be very beneficial if the system is able to qualitatively comment on the influence of changing a FEA on the sequential steps of the process. It is noteworthy to mention that the MPs vendors do not produce a module (PEA, or FEA) only for one exclusive chemical media, but the aim is to build the modules for as many chemical systems as possible. Thus, it would be really beneficial to add the phenomena hierarchical structure (which is not specific to a chemical system) to the MTP of PEAs.

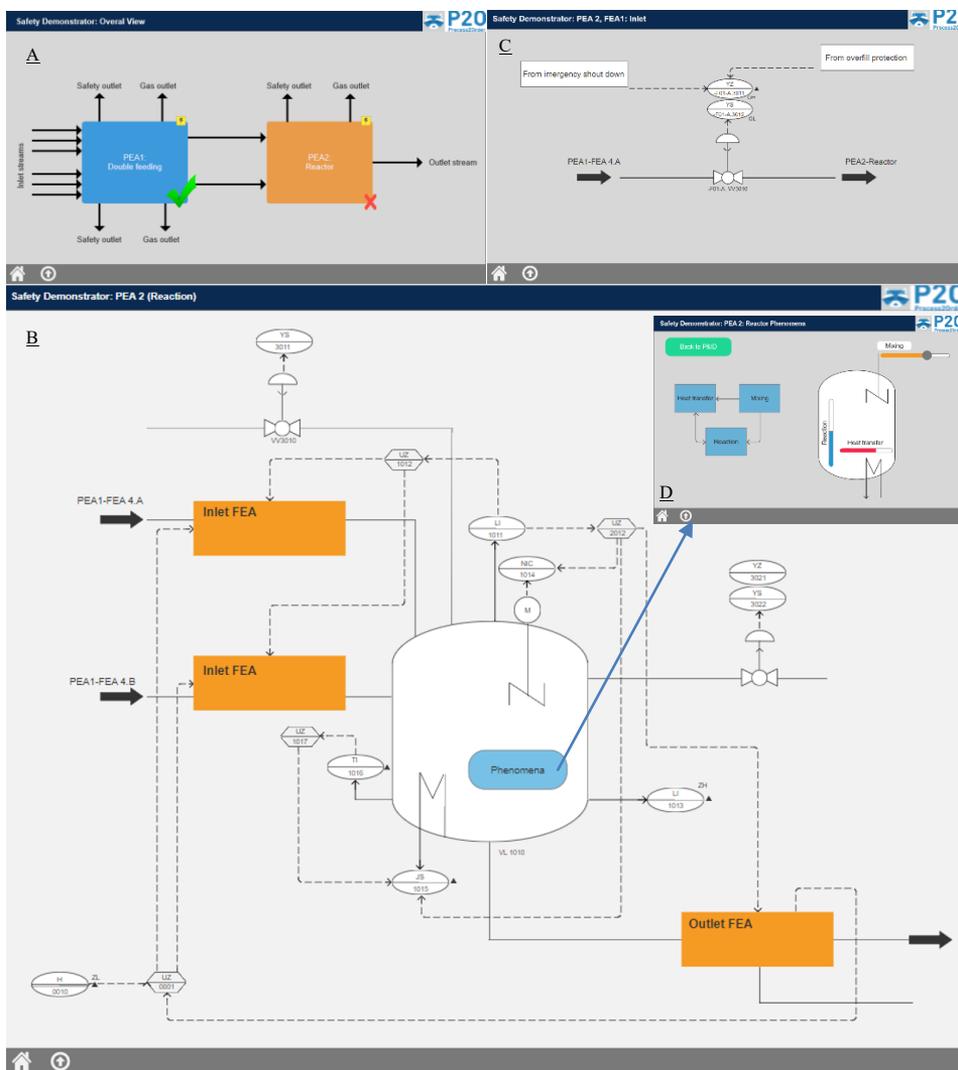


Figure 2: EID for safety-demonstrator, A) overall view to the process, B) view of the PEA2 (reaction), C) PEA2, FE1, D) Phenomena representation of the CSTR in PEA2

4. Conclusion

In this study, the concept of EID has been investigated for MPs. For this purpose, safety demonstrator, an experimental MP built in P2O lab, has been considered as the use case.

At first, WDA for the use case has been done to represent not only the superficial features of the plant (decomposition hierarchy), but also its more in-depth features of how it works (abstraction hierarchy). Based on the developed WDA, an EID has been suggested for the use case. Notable among important characteristics of this interface are its ability to show different views of the process (with different amounts of details), explanation of the control system, and representation of the phenomena taking place in the process. A crucial future step would be to evaluate the effectiveness of the developed interface by interviewing some experts. This study also shows some promising start point for using phenomena representation in the MTP of MPs to distinguish the effect of changing one FEA on the subsequent process steps.

Acknowledgments: This work was Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 319919706 /RTG2323.

References

- U. Ahlstrom, 2005, 'Work domain analysis for air traffic controller weather displays.', *Journal of safety research*, 36(2), pp. 159–169.
- C. M. Burns, J. Kuo, and S. Ng, 2003, 'Ecological interface design: a new approach for visualizing network management', *Computer Networks*, 43(3), pp. 369–388.
- M. Gil, et al., 2019, 'Designing human-in-the-loop autonomous Cyber-Physical Systems', *International Journal of Human-Computer Studies*, 130, pp. 21–39.
- J. R. Hajdukiewicz, et al., 2001, 'Modeling a medical environment: an ontology for integrated medical informatics design', *International Journal of Medical Informatics*, 62(1), pp. 79–99.
- T. Holm, et al., 2015, 'Engineering method for the integration of modules into fast evolving production systems in the process industry', in *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1042–1047.
- T. Holm, 2016, *Aufwandsbewertung im Engineering modularer Prozessanlagen*. Helmut-Schmidt-Universität.
- G. A. Jamieson, and K. J. Vicente, 2001, 'Ecological interface design for petrochemical applications: supporting operator adaptation, continuous learning, and distributed, collaborative work', *Computers & Chemical Engineering*, 25(7), pp. 1055–1074.
- S. Lier, and M. Grünewald, 2011, 'Net Present Value Analysis of Modular Chemical Production Plants', *Chemical Engineering & Technology*, 34(5), pp. 809–816.
- F. Pelzer, et al., 2021, 'Safety in modular process plants: demonstration of safety concepts', *e & i Elektrotechnik und Informationstechnik*, 138(7), pp. 462–468.
- P. M. Salmon, et al., 2019, 'Using the abstraction hierarchy to identify how the purpose and structure of road transport systems contributes to road trauma', *Transportation Research Interdisciplinary Perspectives*, 3, p. 100067.
- C. Son, et al., 2019, 'Reflecting Abstraction Hierarchy of a Chemical Processing System on Standard Operating Procedures', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), pp. 1806–1810.
- VDI 2776-1 (2020) *Process engineering plants - Modular plants - Fundamentals and planning modular plants*.
- K. J. Vicente, and J. Rasmussen, 1990, 'The Ecology of Human-Machine Systems II: Mediating "Direct Perception" in Complex Work Domains', *Ecological Psychology*, 2(3), pp. 207–249.
- K. J. Vicente, and J. Rasmussen, 1992, 'Ecological interface design: theoretical foundations', *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4), pp. 589–606.

Path-following for parametric MPCC: a flash tank case study

Caroline S. M. Nakama^a, Peter Maxwell^a and Johannes Jäschke^a

^a*Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

Abstract

Mathematical programs with complementarity constraints (MPCC) can arise in process models that contain discrete decisions such as switches, phase changes, and flow reversal. Path-following methods are an important part of advanced-step nonlinear model predictive control (NMPC) due to the ability to deal with changes in the active-set of constraints. In this work, we introduce a path-following algorithm for parametric MPCC demonstrated on a flash tank case study. We show that this algorithm can successfully track the solution without the need for fine discretization or identifying the exact points where active-set changes occur, which are important properties for NMPC implementation.

Keywords: Mathematical programming with complementarity constraints, parametric sensitivity, path-following

1. Introduction

Nonlinear model predictive control (NMPC) is a process control method that formulates and repeatedly solves an optimization problem using a nonlinear dynamic model representation of the process as constraints. When configuring an NMPC problem, it is important to have a model that can describe the process as accurately as possible within a defined range of the process variables, and that calculations can be performed during the time between two measurements, i.e., the optimization problem needs to be rapidly solved. Systems with switches, phase changes, or flow reversal, for example, result in models with nonsmooth decisions, which make optimization problems with dynamic models challenging to solve, especially with a limited time frame. For representing such processes, complementarity constraints can be used: they specify the relationship between two variables, enforcing that at least one of them must be at its bound. Optimization models with this type of constraint are called mathematical programs with complementarity constraints (MPCC). These models are inherently non-convex and fail to satisfy the Mangasarian–Fromovitz constraint qualification due to the complementarities, requiring reformulation strategies to handle these constraints with standard NLP solvers (Baumrucker et al., 2008).

Advanced-step NMPC (asNMPC) is a real-time control technique that uses a prediction of the next state variables based on the current control action as the initial values (here they can be seen as parameters) to solve the optimal control problem in advance between the sampling times. When the new sample is available, the solution is updated based on the sensitivity at the optimal solution with respect to the initial state (Diehl et al., 2005; Zavala and Biegler, 2009). Hence, computational delay between sampling and implementing the control action is reduced. A limitation of the original asNMPC is the assumption that the active-set of constraints does not change from the

optimal to the updated solutions. Path-following algorithms can be employed to handle this issue, since change in the active-set can be detected by discretizing the difference between the predicted and sampled states or by using an active-set identification method (Kungurtsev and Jaschke, 2017; Jäschke et al., 2014).

Literature on path-following of parametric MPCC (PMPCC) is scarce; to the best of the authors' knowledge, the only investigation on the topic was conducted by Kungurtsev and Jäschke (2019). They propose two algorithms: one is based on a penalty-term reformulation for MPCC (Baumrucker et al., 2008), while the other traces active-set bifurcations that stem from bi-active complementarity constraints, assuming that a reliable active-set method is available.

In general, PMPCCs pose a number of formidable challenges for numerical solution, arising from the inherent combinatorial nature of the problem and the interaction with the parameter dependence. Herein, we vastly simplify the problem by means of two key assumptions:

- i due to the nature of NMPC, the exact locations of active-set changes are not required but only determination of the solution at successively given values of a scalar parameter; and,
- ii that we are considering MPCC such that each complementarity constraint can only be non-simple (bi-active) at a small number of discrete points.

These two assumptions permit a much more straightforward handling of the problem yet still apply to a selection of practical problems. For example, the second assumption will be valid for most physical systems, whereby active-set changes occur only at a few discrete points (e.g. when phase transitions occur). We focus on the demonstration of this method with a flash tank case study, in which we obtain solutions for relevant points along the optimal path.

2. Background

In this section we present relevant definitions and concepts necessary for the algorithm described in the next section. We begin with the definition of complementarity constraint, roughly following the exposition in (Scheel and Scholtes, 2000). Consider a matrix-valued function $F : \mathbb{R}^n \rightarrow \mathbb{R}^{l \times q}$,

$$F(w) := \begin{bmatrix} F_{11}(w) & \dots & F_{1q}(w) \\ \vdots & \ddots & \vdots \\ F_{l1}(w) & \dots & F_{lq}(w) \end{bmatrix} \quad (1)$$

with $w \in \mathbb{R}^n$. A general complementarity constraint can be expressed as

$$F_{1k} \perp F_{2k} \perp \dots \perp F_{lk}, \quad F_{ik} \geq 0, \quad \text{for } i = 1, \dots, l \quad \text{and } k = 1, \dots, q \quad (2)$$

i.e., at least one entry of each column in F is zero with the remaining assuming nonnegative values. In practice, l is often equal to 2 and F_{ik} mostly represent variable bounds.

A parametric MPCC (PMPCC) is an extension of traditional parametric NLP optimization models in the sense that it contains at least one complementarity constraint and can be solved as a function of one or multiple parameters. For a parameter vector $p : \mathbb{R} \rightarrow \mathbb{R}^r$, we can define PMPCC($p(t)$) as

$$\min_{w(p(t))} \quad \varphi(w(p(t)), p(t)) \quad (3a)$$

$$\text{s.t. } h(w(p(t)), p(t)) = 0 \quad (3b)$$

$$g(w(p(t)), p(t)) \geq 0 \quad (3c)$$

$$F_{1k}(w(p(t)), p(t)) \perp \dots \perp F_{lk}(w(p(t)), p(t)) \quad \text{for } k = 1, \dots, q \quad (3d)$$

$$F_{ik} \geq 0, \quad \text{for } i = 1, \dots, l \quad \text{and } k = 1, \dots, q \quad (3e)$$

where $\varphi : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^s$, $g : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^q$, and $F : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^{l \times q}$ are smooth functions. Dropping explicit notation for dependence on t , the Lagrangian for Eqs. (3) is

$$\mathcal{L}(w(p), p, \lambda) := \varphi(w(p), p) - \mu(p)^T h(w(p), p) - \nu(p)^T g(w(p), p) - \Gamma(p)F(w(p), p) \quad (4)$$

where $\lambda = (\mu, \nu, \Gamma)$ is arranged in the natural manner by reshaping the matrix Γ columnwise into a vector. $\Gamma(p)^T F(w(p), p)$ is the inner product of the corresponding Γ with the F matrix.

The reason we define parametric optimization models is that we wish to map some interval $I_t = [t_a, t_b]$ to the solution curve $(w^*(p(t)), \lambda^*(p(t)))$, $t \in I_t$ of PMPCC($p(t)$) by calculating a piecewise homotopy along t . Locally, this requires calculating the sensitivity of w^* along t .

For the inequality and complementarity constraints, we define the corresponding g and F active-sets at some point $w(p)$ as $A_g(w(p)) := \{i : g_i(w(p)) = 0\}$ and $A_F(w(p)) := \{(i, j) : F_{ij} = 0\}$.

Mirroring terminology used in eigenvalue analysis, a complementarity constraint $F_{.j}$ is considered *simple* at point w if only one constraint in that column $F_{.j}$ is active. It is *nonsimple* if more than one constraint in $F_{.j}$ is active. This latter situation is often termed ‘bi-active’ in the case of two active constraints.

A point $w(p)$ feasible to PMPCC($p(t)$) is termed *weakly stationary* if there exists multipliers λ such that, where \circ is the Hadamard (elementwise) product,

$$\nabla_w \mathcal{L}(w(p), p, \lambda) = 0 \quad (5a)$$

$$\nu(p) \geq 0 \quad (5b)$$

$$\nu(p)^T g(w(p), p) = 0 \quad (5c)$$

$$\Gamma(p) \circ F(w(p), p) = 0. \quad (5d)$$

A point $w(p)$ which satisfies Eqs. (5) and further satisfies that $\Gamma_{ik} \geq 0$ if there exists some $j \neq i$ such that $F_{ik}(w(p), p) = F_{jk}(w(p), p) = 0$ is termed *strongly stationary*. In plain language, this condition is stipulating that if $F_{.k}$ is simple then there is no restriction on the associated Lagrange multiplier whereas if it is nonsimple then the associated Lagrange multipliers must be nonnegative. Thus, we see that for strongly stationary points, a simple complementarity constraint behaves akin to an equality constraint. The strongly active-set for g is defined as $A_g^+(w(p)) := \{i \in A_g(w(p)) : \exists \nu_i > 0 \text{ satisfying Eqs. (5)}\}$.

We now describe how we obtain the sensitivity of PMPCC($p(t)$) with respect to t by reduction locally to a parametric nonlinear program (PNLP). For $t \in I_t$, we assume that the F constraints are nonsimple or strong complementarity of the inequality constraints fails only at a (small) finite number of discrete points so that $I_t = \{t_a\} \cup I_1 \cup I_2 \cup \dots \cup \{t_b\}$, say, where each $I_i \subset I_t$ is an open interval. Within each I_i , the complementarity constraints can then be considered as equality constraints and strong complementarity of g holds, so the PMPCC reduces to a PNL, which is more amenable to known solution methods. Therefore, within each I_i , $A_F^+(w(p)) = A_F(w(p))$ and $A_g^+(w(p(t)))$ is invariant. For PNL($p(t)$), $t \in I_i$, assume that $\varphi(\cdot, \cdot)$, $h(\cdot, \cdot)$, $g(\cdot, \cdot)$, and $F(\cdot, \cdot)$ are twice continuously differentiable in a neighborhood of $w^*(p(t))$ satisfying the first-order optimality conditions, and that the linear independence constraint qualification (LICQ) and strong second order sufficient condition (SSOSC) hold. For PNL($p(t)$) $|_{t=t_0}$, LICQ implies that the dual space is a singleton (Kyparisis, 1985), i.e., the multipliers are unique. By taking the total derivative with respect to t of the stationarity conditions and active constraints of PNL, we obtain the following linear system

$$\underbrace{\begin{bmatrix} \nabla_{ww}^2 \mathcal{L} & \nabla_w h & \nabla_w gA & \nabla_w F_A \\ (\nabla_w h)^T & 0 & 0 & 0 \\ (\nabla_w gA)^T & 0 & 0 & 0 \\ (\nabla_w F_A)^T & 0 & 0 & 0 \end{bmatrix}}_{M(w(p(t)), p(t))} \underbrace{\begin{bmatrix} \dot{w} \\ \dot{\lambda} \end{bmatrix}}_{\dot{v}(p(t))} = - \underbrace{\begin{bmatrix} (\nabla_{wp}^2 \mathcal{L}) \dot{p} \\ (\nabla_p h) \dot{p} \\ (\nabla_p gA) \dot{p} \\ (\nabla_p F_A) \dot{p} \end{bmatrix}}_{b(w(p(t)), p(t))} \quad (6)$$

where $\dot{p} = \frac{dp}{dt}$, $\dot{w} = (\nabla_p w)\dot{p}$, $\dot{\lambda} = (\nabla_p \lambda)\dot{p}$, g_A includes only strongly active inequalities, F_a includes only (simple) active F constraints, and parameter dependence has been omitted for brevity.

3. Methodology

For the algorithm, we require access to a robust active-set identification method; we used an adaptation of a method in (Oberlin and Wright, 2006) along with some custom heuristics. The essential feature of the algorithm is to calculate a piecewise approximant $v(p(t)) = [w(p(t)) \lambda(p(t))]^T$ to the solution $v^*(p(t))$ by numerically integrating along the solution curve between active-set changes. In the canonical formulation of a first-order initial value problem, this can be expressed as $\dot{v}(p(t)) = f(t, v(p(t)))$ for $v(p(t_0)) = v_0$ where v_0 is a known initial point. From Eq. (6) and $M(p(t))$ invertible, $f(t, v(p(t))) = \dot{v}(p(t)) = M(w(p(t)), p(t))^{-1} b(w(p(t)), p(t))$; in practice, a linear solver is used to calculate $f(t, \cdot)$. In our implementation, we used an adaptive stepsize Runge–Kutta integrator, which maintains the truncation error within a predetermined bound by adjusting the stepsize. Event detection is used periodically to check whether an active-set change has occurred within the last integration step; this should not happen too often since the active-set identification is a comparatively expensive calculation. If an active-set change has occurred then, for the current $t_{(k)}$ held fixed, equality-constrained Newton iterations are performed with the new active-set until the error in the approximant $v(p(t))$ is sufficiently small; the number of elements in λ , and hence v , may also change if the cardinality of the active-set changes. The integration step can loosely be considered a ‘predictor’ step whereas the Newton iterations are a ‘corrector’ step. However, the error is nominally controlled by the adaptive stepsize numerical integration. Note that if the active-set identification encounters a nonsimple complementarity constraint, by assumption we can perturb $t_{(k)}$ by some small $\varepsilon > 0$ to again obtain simple complementarity constraints. This ensures the algorithm can always calculate the sensitivity.

Algorithm 1 Path-following for PMPCC(t) for $t \in I_t = [t_a, t_b]$

τ is set to maximum active-set recalculation interval
 $t_{(0)} \leftarrow t_a$, $\tilde{v}_{(0)} = [\tilde{w}_{(0)} \tilde{\lambda}_{(0)}]^T \leftarrow v_a$ where v_a is in a neighbourhood of the initial solution
 Calculate $A_{g,(0)}^+ = A_g^+(\tilde{w}_{(0)})$ and $A_{F,(0)}^+ = A_F^+(\tilde{w}_{(0)})$
 $v_{(0)} \leftarrow \text{Newton}(\tilde{v}_{(0)}, A_{g,(0)}^+, A_{F,(0)}^+)$
 $k \leftarrow 0$, $t_{AS} \leftarrow t_{(0)}$
repeat
 $k \leftarrow k + 1$
 Perform predictor integration step to obtain approximate $(t_{(k)}, \tilde{v}_{(k)})$
 if $t_{(k)} > (t_{AS} + \tau)$ **then**
 repeat
 Calculate $A_{g,(k)}^+$ and $A_{F,(k)}^+$
 if $A_{F,(k)}^+$ nonsimple **then**
 $t_{(k)} \leftarrow t_{(k)} + \varepsilon_t$
 end if
 until $A_{F,(k)}^+$ is simple
 $v_{(k)} \leftarrow \text{Newton}(\tilde{v}_{(k)}, A_{g,(k)}^+, A_{F,(k)}^+)$, $t_{AS} \leftarrow t_{(k)}$
 else
 $v_{(k)} \leftarrow \tilde{v}_{(k)}$
 end if
until $t \geq t_b$

It is immediately apparent that the relationship between integration stepsize $t_{(k)} - t_{(k-1)}$ and the active-set identification interval τ is important because it determines whether there will be any

transient ‘overshoot’ of an incorrect solution.

4. Case Study

We use the same case study presented in Kungurtsev and Jäschke (2019), in which a flash tank with a 3-component feed flow $Q \in \mathbb{R}$ and composition $z \in \mathbb{R}^3$ is simulated. We seek to analyze how the split between vapor $V \in \mathbb{R}$ and liquid $L \in \mathbb{R}$ products with composition $y \in \mathbb{R}^3$ and $x \in \mathbb{R}^3$ respectively vary with temperature $T \in \mathbb{R}$ for a fixed pressure $P \in \mathbb{R}$. For that, we set up the following optimization model

$$\min \frac{1}{2}(aQ - V)^2 dt \quad (7a)$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{C}} \frac{z_i(K_i - 1)}{1 + a_i(K_i - 1)} = 0 \quad (7b)$$

$$K_i = \frac{P_i^{\text{sat}}}{P} = \frac{y_i}{x_i} \quad \text{for } i \in \mathcal{C} \quad (7c)$$

$$\log_{10}(p_i^{\text{sat}}) = A_i - \frac{B_i}{T + C_i} \quad \text{for } i \in \mathcal{C} \quad (7d)$$

$$L + V = Q \quad (7e)$$

$$Lx_i + Vy_i = Qz_i \quad \text{for } i \in \mathcal{C} \quad (7f)$$

$$a - s_V + s_L - a_t = 0 \quad (7g)$$

$$0 \leq s_V \perp V \geq 0 \quad (7h)$$

$$0 \leq s_L \perp L \geq 0 \quad (7i)$$

$$0 \leq a, x, y \leq 1 \quad (7j)$$

$$K, p^{\text{sat}} \geq 0. \quad (7k)$$

Constraint (7b) is the Rachford–Rice equation, which calculates the fraction of the feed that goes to the vapor phase, V/Q , represented by $a_t \in \mathbb{R}$. $K \in \mathbb{R}^3$ is determined by Raoult’s law, given by constraint (7c). $P_i^{\text{sat}}(T) \in \mathbb{R}$ is the vapor pressure of the pure component $i \in \mathcal{C} = \{1, 2, 3\}$ at temperature T calculated using Antoine’s equation (7d), where $A_i \in \mathbb{R}$, $B_i \in \mathbb{R}$ and $C_i \in \mathbb{R}$ are constants for each compound i . Constraints (7e) and (7f) correspond to the total and component-wise mass balances respectively. Constraints (7g)–(7j) are necessary to ensure that $V/Q \in [0, 1]$. The Rachford–Rice equation results in negative values and values greater than one for a_t if T is lower than the mixture’s bubble point or larger than its dew point respectively, which would be physically impossible. Therefore, complementarity constraints (7h) and (7i) are considered; in the form of Eq. (2), $F_{11} = s_V$, $F_{21} = V$, $F_{12} = s_L$ and $F_{22} = L$. $s_L \in \mathbb{R}$ and $s_V \in \mathbb{R}$ are slack variables that represent how much a_t is lower than 0 and larger than 1 respectively. Variable $a \in \mathbb{R}$ represents the actual ratio V/Q , which is enforced by constraints (7g) and (7j). Note that $a = V/Q$ is not enforced as a hard constraint and, instead, is used as the objective function to be minimized. In this problem, temperature is the only parameter; we use $T(t_a) = 380$ and $T(t_b) = 400$.

5. Results

For the results presented here, the following values were used: $Q = 1$ kmol/s, $z = [0.5, 0.3, 0.2]^T$, $P = 5$ bar, $A = [3.98; 4.00; 3.93]^T$, $B = [1065; 1171; 1183]^T$, and $C = [-41.14; -48.83; -52.53]^T$. In Fig. 1a, the solution paths of primal variables L and V are shown. The maximum integration stepsize and τ in this run was set small, at 0.1, so the active-set changes are detected almost immediately after they happen. The corresponding complementarity slack variables s_l and s_v are shown in Fig. 1b. For a larger maximum stepsize of 5.0, there is less certainty where the actual active-set change happened and a jump is clearly visible, see Fig. 1d.

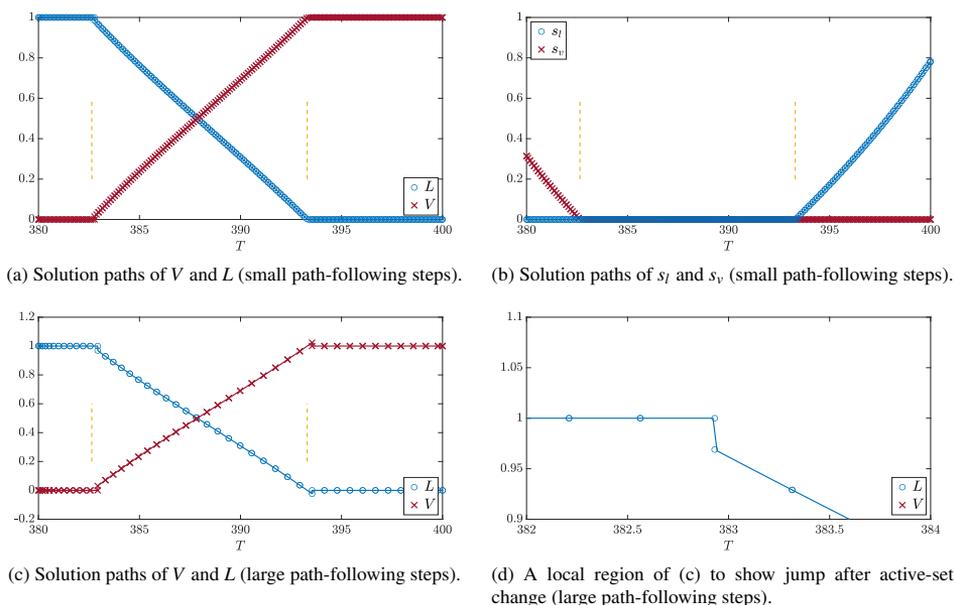


Figure 1: Top row, stepsize of 0.1: (left) L (circle) and V (cross) solution paths; (right) s_L and s_V paths. Bottom, stepsize of 5.0: (left) as above; (right) zoomed-in to region showing jump. Active-set changes indicated in vertical dashed lines (yellow).

6. Conclusion

We have demonstrated on the flash tank case study that the presented algorithm is a suitable method for path-following PMPCC. Since it does not require the identification of the exact location of active-set changes and that relatively coarse discretization can be used, this algorithm is a promising candidate for use in advanced-step NMPC of models with complementarity.

Acknowledgement: The authors acknowledge financial support from NRC (Norwegian Research Council) through FRIPRO Project SensPATH.

References

- B. Baumrucker, J. G. Renfro, L. T. Biegler, 2008. MPEC problem formulations and solution strategies with chemical engineering applications. *Comput. Chem. Eng.* 32 (12), 2903–2913.
- M. Diehl, H. G. Bock, J. P. Schlöder, 2005. A real-time iteration scheme for nonlinear optimization in optimal feedback control. *SIAM J Control Optim* 43 (5), 1714–1736.
- J. Jäschke, X. Yang, L. T. Biegler, 2014. Fast economic model predictive control based on NLP-sensitivities. *J. Process Control* 24 (8), 1260–1272.
- V. Kungurtsev, J. Jäschke, 2017. A predictor-corrector path-following algorithm for dual-degenerate parametric optimization problems. *SIAM J. Optim.* 27 (1), 538–564.
- V. Kungurtsev, J. Jäschke, 2019. Pathfollowing for parametric mathematical programs with complementarity constraints, preprint at www.optimization-online.org/DB_FILE/2019/02/7084.pdf.
- J. Kyparisis, 1985. On uniqueness of Kuhn–Tucker multipliers in nonlinear programming. *Math. Program.* 32 (2), 242–246.
- C. Oberlin, S. J. Wright, 2006. Active set identification in nonlinear programming. *SIAM J. Optim.* 17 (2), 577–605.
- H. Scheel, S. Scholtes, 2000. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Math. Oper. Res.* 25 (1), 1–22.
- V. M. Zavala, L. T. Biegler, 2009. The advanced-step NMPC controller: Optimality, stability and robustness. *Autom.* 45 (1), 86–93.

Feature engineering for neural network-based oscillation detection in process industries

Aswin Krishna M^a, Ayush Kedawat^a, Abhishek Bansal^a and Resmi Suresh^{a,*}

^a*Department of Chemical Engineering, IIT Guwahati-781039, Assam, India
resmis@iitg.ac.in*

Abstract

Process industries frequently encounter oscillations in control loops. Oscillations present in control loops could be an indication of one or more faults in the process such as valve stiction and controller tuning issues. These faults would negatively impact the performance of any process by degrading the quality of the product/output, reducing average throughput, increasing energy consumption etc. In this article, we propose a neural-network-based approach for oscillation detection. Various feature engineering strategies based on domain knowledge are adopted to improve the accuracy, precision and recall of oscillation detection while reducing the computational efforts. Fast Fourier Transform (FFT) and FFT of ACF (Autocorrelation function) of the dynamic process data are used as input features to the neural network. Feature selection based on peaks in the frequency domain data (for both FFT and FFT of ACF) is used for reducing the number of features. A sensitivity study on the variation of accuracy, precision and recall on the number of input features is also part of this work. An 80% reduction in number of input features is obtained compared to methods available in literature without compromising performance, and thus can be easily embedded in chips for online implementation. An accuracy of 96% and a recall of 0.95 for oscillatory data are obtained for the proposed algorithm.

Keywords: Oscillation detection, Feature selection, Feature engineering, Machine learning, Neural network, ACF, FFT

1. Introduction

The presence of oscillations in process data is a very common problem in industries. Oscillations are one of the major causes that deteriorate the performance of any control loop in process industries Srinivasan et al. (2011); Jiang et al. (2007). The presence of oscillations in process variables would result in poor product quality and production losses, and are not desirable. Oscillations can be due to various reasons like stiction in valves, controller tuning issues or due to some external disturbances. Early knowledge of oscillations present in the data and possible cause for the oscillations can help in mitigating its effect on the plant performance.

Although various techniques have been proposed in literature for oscillation detection (Thornhill et al., 2003; Jelali and Huang, 2010; Dambros et al., 2019b; Venkatasubramanian et al., 2003), their performance is poor for actual plant data. With the advent of machine learning, a few machine learning (ML)-based techniques have also been proposed recently for oscillation detection and diagnosis Dambros et al. (2019c,a). Although these techniques show high accuracy in detecting oscillations, they rely on a large number of input features in the frequency domain (obtained from FFT or Fast Fourier Transform of the time-domain signals). Dependency on a large number of

input features may result in erroneous detection of oscillations especially in presence of noise and may limit its industrial application. Moreover, the period and amplitude identification is poor.

This article proposes feature engineering and feature selection strategies based on domain knowledge for improving neural network-based oscillation detection. The proposed approach helps to reduce the number of features required for accurate detection of oscillations in process data.

2. Dataset available

The proposed approach was trained and validated using the synthetic dataset provided by Dambros et al. (2019c). The dataset contains dynamic process data for 110,000 process variables that contain non-oscillatory (43705 process variables), oscillatory (33174 process variables), and irregular oscillatory variables (33121 process variables). Class 0 refers to non-oscillatory, class 1 refers to oscillatory and class 2 refers to irregular oscillatory variables. Since this is a synthetic data, we have the information on which variables are non-oscillatory/ oscillatory/ irregular oscillatory. This is a multi-class classification problem. To avoid bias in oscillation detection due to varying magnitudes, the data is normalized before sending them for oscillation detection.

3. FFT-based Neural Network Approach for Oscillation Detection

Dambros et al. (2019c) presented a deep feedforward neural network (NN) for automatic oscillation detection using frequency domain data as input features. As different variables would be sampled at different rates, the dynamic data would be of different lengths and cannot be used as input features for NN directly. Hence, the dynamic process data was transformed to frequency domain through Fourier transform such that each variable had a length of 4097 samples. Dambros et al. (2019c) created a large synthetic dataset and the algorithm (after training) when tested on a subset of this synthetic data (test set) showed a high accuracy of 97% in detecting oscillations. However, the NN model uses 4097 input features and thus have a very large set of weights making it a computationally expensive method. The need to process large number of features makes it difficult to embed the algorithm in chip for online implementation. The proposed oscillation detection algorithm in chips and Optimization of hyperparameters is also difficult due to large size of the neural network. In this section, we discuss a Neural network model for oscillation detection that uses a subset of frequency domain features selected based on the peaks in FFT spectrum.

Similar to Dambros et al. (2019c), the time-domain data is first converted to frequency domain and the amplitudes corresponding to non-negative frequencies is chosen as FFT is symmetric. This results in 4097 amplitude values corresponding to 4097 frequencies for each variable. However, instead of using the entire frequency domain data for NN model, we select a few significant features based on domain knowledge. A peak in the FFT spectrum indicates presence of oscillations in the data. Practically, we can only find the maximum value of the FFT spectrum. To see whether the identified maxima is a peak, one has to rely on some heuristics based on adjacent amplitude values and thus, are not very accurate. This is why we needed an alternate data-based method for oscillation detection. Here, we use this idea that the maximum amplitude in FFT of the data and its adjacent amplitude values (corresponding to adjacent frequencies) are the most significant features in determining whether that variable data is oscillatory or not. A subset of significant features is chosen by identifying the maximum amplitude and a small neighborhood around it. Considering m amplitude values on both sides of the maximum value, the total number of input features will be $2m + 1$ including the maximum FFT value. We can alter the complexity of the network by varying the value of m . Sensitivity analysis was performed to find the optimal value of m . Various values of m are chosen and a neural network model (of same structure) is built for each m . This method would be referred as Method 1 hereafter. The structure of the neural network model used and the results obtained are described next.

Table 1: Oscillation detection results and sensitivity studies when a subset of FFT of data is used as input features

No. of Features ($2m + 1$)	Confusion Matrix	Accuracy	Precision			Recall			Computational time
			Class 0	Class 1	Class 2	Class 0	Class 1	Class 2	
11 ($m = 5$)	$\begin{bmatrix} 3450 & 100 & 435 \\ 1158 & 1828 & 41 \\ 2035 & 58 & 895 \end{bmatrix}$	61.73%	0.519	0.92	0.653	0.866	0.604	0.299	14.1s
21 ($m = 10$)	$\begin{bmatrix} 3653 & 40 & 292 \\ 965 & 1992 & 70 \\ 1840 & 29 & 1119 \end{bmatrix}$	67.64%	0.566	0.966	0.755	0.917	0.658	0.374	14.5s
31 ($m = 15$)	$\begin{bmatrix} 3710 & 38 & 237 \\ 905 & 2086 & 36 \\ 1788 & 40 & 1160 \end{bmatrix}$	69.56%	0.579	0.964	0.810	0.931	0.689	0.388	15.8s
101 ($m = 50$)	$\begin{bmatrix} 3782 & 5 & 198 \\ 644 & 2314 & 69 \\ 1107 & 30 & 1851 \end{bmatrix}$	79.47%	0.684	0.985	0.874	0.949	0.764	0.620	17.1s
201 ($m = 100$)	$\begin{bmatrix} 3856 & 3 & 126 \\ 345 & 2616 & 66 \\ 596 & 51 & 2341 \end{bmatrix}$	88.13%	0.804	0.980	0.924	0.968	0.864	0.783	18.9s
401 ($m = 200$)	$\begin{bmatrix} 3945 & 6 & 34 \\ 147 & 2837 & 43 \\ 227 & 32 & 2729 \end{bmatrix}$	95.11%	0.913	0.987	0.973	0.990	0.938	0.913	20.6s
4097	$\begin{bmatrix} 3968 & 2 & 15 \\ 152 & 2810 & 65 \\ 201 & 52 & 2735 \end{bmatrix}$	95.13%	0.918	0.981	0.972	0.996	0.928	0.915	64.7s

3.1. Neural-network classifier

A three layer feed forward neural-network with neurons 400, 100 and 20 respectively in each layer is used for detecting oscillations in data. This neural-network was implemented in Python software by setting the hyperparameters 'batch size' as 10000, 'activation function' as 'Hard Sigmoid', and 'Optimization algorithm' as 'adamax'. The network is trained using a training dataset of 100,000 variables' data. The remaining 10,000 variables' data have been used for testing which contains 3985 non-oscillatory variables, 3027 oscillatory and 2988 irregularly oscillatory variables.

The results of oscillation detection for the test set using the trained model for various number of input features are provided in Table 1. The confusion matrix provided is such that the rows represent the true classes and the columns represent the predicted classes. The variation of accuracy, precision and recall with number of features is depicted in Figure 1. As the number of features increase from 11 ($m = 5$) to 4097 (full data), accuracy monotonically increases from 61.73% to 95.13% on the test set. Note that there isn't any significant improvement in accuracy by increasing the number of features from 401 to 4097. At the same time, the model precision for classes 1 and 2, and recall for class 1 have reduced when number of features were increased from 401 to 4097. Further, model training time for 401 features was around 21sec whereas it took more than a minute for training the model with 4097 features was around 65 sec for the training set. Note that these computational times are reported without including the time required for hyper-parameter optimization. This is because we have not performed hyper-parameter optimization since we are using the same network structure for all m values. Hyper-parameter optimization would improve the performance of each model.

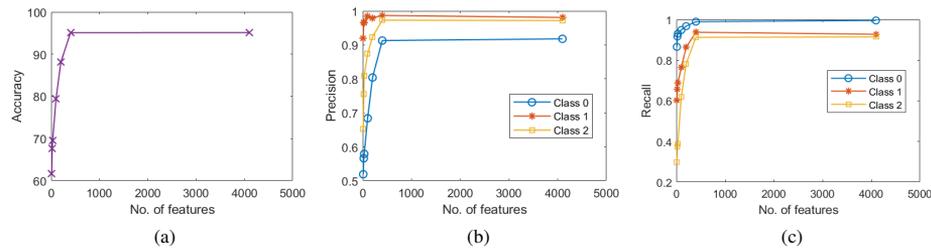


Figure 1: Variation of accuracy, precision and recall for the neural network classifier based on FFT features

4. FFT of ACF as input features

Although we were able to obtain almost similar accuracy of the full model using reduced number of features, another feature engineering strategy was attempted in search of an approach which can provide better performance with reduced number of features.

Autocorrelation Function (ACF) provides the correlation of a signal to itself at different time lags. It helps in finding repeating patterns in a signal. An important property of ACF is that it is oscillatory for an oscillatory signal with the same period of oscillation as the signal. This domain knowledge is used to extract useful features from process data. If the original variable data is oscillatory, ACF of the data will also be oscillatory. Hence, the presence of peaks in FFT of ACF of the data is an indication of oscillations in the original data. ACF shows noise free oscillations even if the original data is noisy and thus can be used as a better tool to identify oscillations.

As a first attempt, instead of using FFT of data as features, FFT of ACF of the data was used as input features for the neural network. The same approach as described in Section 3 is used to select a subset of features from FFT of ACF. Maxima of FFT of ACF is identified for each variable and $2m + 1$ amplitude values in the neighborhood of maxima are used as features. For each m , the same neural network structure as described in Section 3.1 is used. The network needs to be trained for each value of m as the input features are different. The method of using a subset of features from FFT of ACF of data is referred to as Method 2 hereafter. Depending on the value of m chosen, we obtain different performance in detecting oscillations as shown in Table 2. Although we obtained an accuracy greater than 90% for 401 features, we see that the performance is poorer than the earlier case with FFT of data as input features (Method 1). However, the predictions showed that we were able to correctly predict the class for some of the cases where Method 1 failed. Hence, the algorithm was further improved by incorporating features from both FFT of data and FFT of ACF of data.

The final method (Method 3) adopted for oscillation detection is a neural network that uses combination of subsets of FFT of data and FFT of ACF of data as input features. Again, maxima of FFT of data was identified and $2m + 1$ features in the neighborhood were picked. Similarly, $2m + 1$ features in the neighborhood of maxima of FFT of ACF of data were picked resulting in a total of $4m + 2$ input features. The results obtained for various values of m are provided in Table 3. Note that the accuracy has been improved to 95.96% for 802 features which is better than the accuracy obtained for Method 1 using 4097 features. Thus, a better accuracy is obtained with 80% reduction in the number of features. Computational time has also reduced. As the workstation used is of high configuration, the reduction in computational time is not much. However, for a normal desktop, computational time would be improved considerably with 80% reduction in number of input features. Moreover, since the purpose of the algorithm is to detect oscillations, recall of

Table 2: Oscillation detection results and sensitivity studies when a subset of FFT of ACF of data is used as input features

No. of Features ($2m + 1$)	Confusion Matrix	Accuracy	Precision			Recall			Computational time
			Class 0	Class 1	Class 2	Class 0	Class 1	Class 2	
11 ($m = 5$)	$\begin{bmatrix} 3046 & 481 & 458 \\ 699 & 2224 & 104 \\ 1492 & 271 & 1225 \end{bmatrix}$	64.95%	0.582	0.747	0.686	0.764	0.735	0.410	14.8s
21 ($m = 10$)	$\begin{bmatrix} 3582 & 194 & 209 \\ 800 & 2164 & 63 \\ 1687 & 124 & 1177 \end{bmatrix}$	69.23%	0.590	0.872	0.812	0.899	0.715	0.394	14.8s
31 ($m = 15$)	$\begin{bmatrix} 3781 & 67 & 137 \\ 818 & 2121 & 88 \\ 1708 & 63 & 1217 \end{bmatrix}$	71.19%	0.599	0.942	0.844	0.949	0.701	0.407	15.1s
101 ($m = 50$)	$\begin{bmatrix} 3842 & 4 & 139 \\ 332 & 2275 & 420 \\ 583 & 67 & 2338 \end{bmatrix}$	84.55%	0.808	0.970	0.807	0.964	0.752	0.782	17s
201 ($m = 100$)	$\begin{bmatrix} 3897 & 10 & 78 \\ 181 & 2575 & 271 \\ 397 & 117 & 2474 \end{bmatrix}$	89.46%	0.871	0.953	0.876	0.978	0.851	0.828	19.6s
401 ($m = 200$)	$\begin{bmatrix} 3899 & 12 & 74 \\ 121 & 2800 & 106 \\ 340 & 118 & 2530 \end{bmatrix}$	92.29%	0.894	0.956	0.934	0.978	0.925	0.847	21.2s
4097	$\begin{bmatrix} 3933 & 5 & 47 \\ 81 & 2737 & 209 \\ 194 & 90 & 2704 \end{bmatrix}$	93.74%	0.935	0.966	0.914	0.987	0.904	0.905	65.2s

classes 1 and 2 are more significant. Using methods 1 and 2, maximum recall obtained was 0.938 for class 1 and 0.913 for class 2. Using this approach, we are able to increase the recall to 0.95 for class 1 and 0.94 for class 2 using just 802 features. Also, with 8194 features, we are able to achieve 96.61% accuracy. Note that these results are obtained using a neural network for which hyper-parameter optimization have not been performed. Performance of the algorithm is expected to increase with optimized hyper-parameters. On the other hand, with respect to Dambros et al. (2019c), the work had been carried out using 4097 features of the FFT of the signal alone as the input and had achieved an accuracy of 97% in their work. The proposed method, albeit having 80% less number of features, provides an accuracy of 96%.

Hence, from various feature engineering and selection strategies, it was found that a neural network that uses 802 input features combined from FFT of data and FFT of ACF of data is the best choice in terms of accuracy as well as computational time. It was also found that there is a trade-off between accuracy and computational time. Based on application, if a lower computational time is required, a lower number of features can be used at the cost of lesser accuracy.

5. Conclusions

Neural network-based algorithms for accurate and fast detection of oscillations in process data are discussed in this article. Various feature engineering strategies based on domain knowledge have been implemented to improve the accuracy and computational effort of the method. FFT of data and FFT of ACF of data have been used as input features. To reduce computational complexity, significant features are selected based on the peaks or maximas in FFT and FFT of ACF. It was observed that neural network classifier performed better when a combination of FFT and FFT of

Table 3: Oscillation detection results and sensitivity studies when a combinations of subsets of FFT of data and FFT of ACF of data is used as input features

No. of Features ($4m+2$)	Confusion Matrix	Accuracy	Precision			Recall			Computational time
			Class 0	Class 1	Class 2	Class 0	Class 1	Class 2	
22 ($m=5$)	$\begin{bmatrix} 3420 & 77 & 488 \\ 857 & 2109 & 61 \\ 1443 & 98 & 1447 \end{bmatrix}$	69.76%	0.598	0.923	0.725	0.859	0.697	0.484	28.9s
42 ($m=10$)	$\begin{bmatrix} 3505 & 62 & 418 \\ 810 & 2139 & 78 \\ 1387 & 65 & 1536 \end{bmatrix}$	71.80%	0.615	0.944	0.756	0.880	0.707	0.514	31.7s
62 ($m=15$)	$\begin{bmatrix} 3538 & 62 & 385 \\ 851 & 2145 & 131 \\ 1201 & 63 & 1724 \end{bmatrix}$	74.07%	0.633	0.945	0.770	0.888	0.686	0.577	30.1s
202 ($m=50$)	$\begin{bmatrix} 3765 & 15 & 205 \\ 447 & 2347 & 233 \\ 650 & 53 & 2285 \end{bmatrix}$	83.97%	0.774	0.972	0.840	0.945	0.775	0.765	44.2s
402 ($m=100$)	$\begin{bmatrix} 3845 & 18 & 122 \\ 201 & 2709 & 117 \\ 329 & 111 & 2548 \end{bmatrix}$	91.02%	0.879	0.954	0.914	0.965	0.895	0.853	41.1s
802 ($m=200$)	$\begin{bmatrix} 3914 & 14 & 57 \\ 75 & 2876 & 76 \\ 115 & 67 & 2806 \end{bmatrix}$	95.96%	0.954	0.973	0.955	0.982	0.950	0.940	47.2s
8194	$\begin{bmatrix} 3961 & 5 & 19 \\ 82 & 2887 & 58 \\ 128 & 47 & 2813 \end{bmatrix}$	96.61%	0.950	0.982	0.973	0.994	0.954	0.941	174.7s

ACF features were used as input features rather than using FFT features alone or FFT of ACF features alone as the input to the neural network. The performance saturates while computational time increases with increasing number of features. The neural network classifier based on 802 input features (401 FFT and 401 FFT of ACF) was found to be a best compromise in terms of accuracy/precision/recall and computational time. Efficacy of the proposed method on real plant data needs to be tested. Prediction of amplitude and period of oscillation using the proposed feature set would be attempted in future.

References

- J. W. Dambros, M. Farenzena, J. O. Trierweiler, 2019a. Oscillation detection and diagnosis in process industries by pattern recognition technique. *IFAC-PapersOnLine* 52 (1), 299–304.
- J. W. Dambros, J. O. Trierweiler, M. Farenzena, 2019b. Oscillation detection in process industries–part i: Review of the detection methods. *Journal of Process Control* 78, 108–123.
- J. W. Dambros, J. O. Trierweiler, M. Farenzena, M. Kloft, 2019c. Oscillation detection in process industries by a machine learning-based approach. *Industrial & Engineering Chemistry Research* 58 (31), 14180–14192.
- M. Jelali, B. Huang, 2010. *Detection and diagnosis of stiction in control loops: state of the art and advanced methods*. Springer.
- H. Jiang, M. S. Choudhury, S. L. Shah, 2007. Detection and diagnosis of plant-wide oscillations from industrial data using the spectral envelope method. *Journal of Process Control* 17 (2), 143–155.
- B. Srinivasan, U. Nallasivam, R. Rengaswamy, 2011. Diagnosis of root cause for oscillations in closed-loop chemical process systems. *IFAC Proceedings Volumes* 44 (1), 13145–13150.
- N. F. Thornhill, B. Huang, H. Zhang, 2003. Detection of multiple oscillations in control loops. *Journal of Process control* 13 (1), 91–100.
- V. Venkatasubramanian, R. Rengaswamy, K. Yin, S. N. Kavuri, 2003. A review of process fault detection and diagnosis: Part i: Quantitative model-based methods. *Computers & chemical engineering* 27 (3), 293–311.

Multi-parametric Model Predictive Control Strategies for Evaporation Processes in Pharmaceutical Industries

Ioana Naşcu^{a,b}, Nikolaos A. Diangelakis^c, and Efstratios N. Pistikopoulos^c,

^a *Department of Automation Technical University of Cluj Napoca, Romania*

^b *Key Laboratory of Advanced Control and Optimization for Chemical Processes,
Ministry of Education, East China University of Science and Technology, Shanghai
200237, China*

^c *Texas A&M Energy Institute and Artie McFerrin Department of Chemical
Engineering, Texas A&M, College Station, TX, USA*

Ioana.Nascu@aut.utcluj.ro

Abstract

In this paper we develop an advanced multi-parametric model predictive control approach for the control of an evaporation process in the pharmaceutical industry. The proposed strategies set the foundation for the development of controllers that aim to work with different molecules and different thermodynamic scenarios without repeating the process design and process control design steps. First, a comprehensive mathematical model of the process for one molecule type is developed and implemented within gPROMS. The model, along with its experimental set-up validation, is then used for the development of the advanced control strategies. Finally, the performances of the control strategies are validated against the original high-fidelity model, thus closing the loop. The simulations show good performances and satisfactory behavior.

Keywords: multi-parametric/explicit model predictive control, pharmaceutical processes, evaporation process.

1. Introduction

Due to their complex nature, pharmaceutical plants are required to operate near operational constraints with very strict product quality specifications and deal with complex and highly integrated processes, varying production targets, raw material variability and process/model uncertainty (Ierapetritou et al. 2016, Su et al. 2019, Seborg et al. 2017, Rantanen and Khinast 2015). Using model based control approaches greatly affects the time and resource utilization for the development process; certain pharmaceutical process design problems, under assumptions, can be transformed into process control problems (Nascu et al. 2016, Politis et al. 2017, Wang et al. 2017). If we consider a pre-existing pharmaceutical process in which a new molecule comes in, i.e. in the form of a newly developed active pharmaceutical ingredient (API), the standard approach requires a series of Design of Experiments (DoE) to be performed for the new process to be designed. However, only information regarding the material properties of the new molecule are required. A model based, material properties aware controller can make the necessary adjustment to adapt to the new molecule thus alleviating the need of computationally expensive and time consuming DoE's.

In this work we set the foundation for advanced multi-parametric model predictive control (mp-MPC) systems for a continuous evaporation process that are designed to work with different molecules without repeating the process design and process control design steps. The first step is to design a model of the process for one molecule type. This model will be then used to design advanced mp-MPCs for different molecules. Explicit/multi-parametric model predictive control (mp-MPC), solves offline the optimization problem using multi-parametric programming and derives the control inputs as a set of explicit linear functions of the system states, disturbances, set-points, etc (Pistikopoulos et al. 2002, Diangelakis et al. 2017). The performances and limits of the designed control schemes are tested on the model developed within the gPROMS platform for varying operating targets and process disturbances. The designed methodologies show good performances: fast settling time and no significant overshoot or undershoot. Moreover, this work represents the first step towards the development of advanced MPCs that are designed to work with different molecules and different thermodynamics scenarios without redoing the process design and process control steps.

2. Theoretical background

2.1. Process Description

Evaporation is a unit operation that separates liquids from solids by means of heat transfer via vaporization or boiling. The purpose of evaporation is to concentrate a solution containing a non-volatile solute (e.g., solids) and a solvent (e.g., liquid). Evaporating a portion of the solvent concentrates the solute into a more viscous liquid product. (Bryan W. Hackett 2018, Govatsmark and Skogestad 2001, Bloore and O-Callaghan 2009). Evaporation may be carried out as a batch or continuous process. This work focuses on evaporation as a semi continuous process as presented in Figure 1, in which the feed streams are continuous and the product is pushed out every 10 minutes.

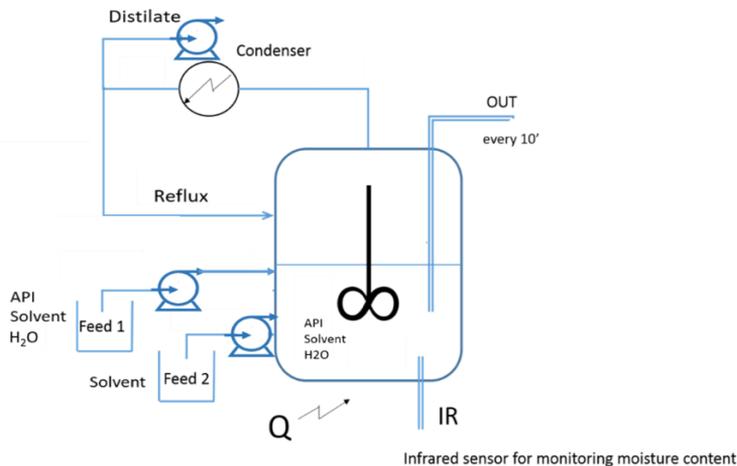


Figure 1 Evaporator scheme

The crude solutions (Feed 1 and Feed 2) are continuously pumped into the evaporator. A calculated mass flow rate for Feed 1 solution is set to achieve a target mass feed rate of API. A calculated mass flow rate of solvent is also continuously pumped into the evaporator (Feed 2). The vapour phase inside the evaporator consists of water and solvent only. The vapour condenses in the condenser and the excess distillate continuously flows

back into the evaporator as reflux. The calculated mass flow rate of solvent, and the calculated distillate mass flow rate, are set to achieve a given target composition of API and water in the solution flowing from the evaporator into the crystallization process. The solution from the evaporator is transferred intermittently about once every 10 minutes. The two feed streams into the evaporator and the distillate exiting the evaporator have mass flow control with a Coriolis mass flow meter in the loop. The mass flow rate of concentrated API solution flowing out of the evaporator is indirectly controlled because it is a consequence of the three mass flow inputs. Heat input to the evaporator is set at a constant amount which is higher than needed for the distillate mass flow. A Fourier Transform Infrared Spectrometer (FTIR) is installed into the bottom of the evaporator, measuring in real time the water and API content in the evaporator.

2.2. Process Model

The process model for the Evaporation process, presented in Figure 1 was adapted from the out-of-the-box evaporation model of gPROMS for the needs of this work. The model uses gSAFT thermodynamic package to model the thermodynamic properties. This model will be further used for the design of the control strategies as well as to test the performances of the designed controllers. It can operate regardless the choice of API and solvents (given thermophysical properties libraries). Moreover, the user can change the parameters of the two feeds, the rate of heat flow to the evaporator (Q - the amount of heat energy transferred to the evaporator per unit of time), the reflux constant as well as the configurations and initial conditions for the drum. The reflux constant is a value between 0 and 1 specifying the ratio in which reflux goes back in the evaporator.

The inputs of the model are Feed 1 flow rate (API, Solvent, H₂O), Feed 2 flow rate (Solvent), the reflux flow rate (Solvent, H₂O) as well as the heat flow rate to the evaporator, (Q). The controlled variables are the concentrations of API and H₂O. The measured variables for this process are the feed rates from Feed 1, Feed 2, reflux, temperature and the concentration of API, Solvent and H₂O inside the evaporator.

The model was tested for different APIs and different Solvents where open loop simulations were run for different initial conditions. The initial conditions are the same as the ones used in the real experiment. The flowrates variable are measured in [kg/sec], the concentrations for the feeds, the concentrations in the evaporator as well as the concentrations at the output of the evaporator are given in wt% and the input energy is given in [J/sec].

2.3. Multi-parametric Model Predictive Control

The following mp-QP optimization problem is solved to obtain the control laws using the POP toolbox (Oberdieck et al. 2016, Pistikopoulos et al. 2015, Pistikopoulos et al. 2007a, Pistikopoulos et al. 1999) and determine the controller:

$$\begin{aligned}
 \min_u J &= \hat{x}'_N P \hat{x}_N + \sum_{k=1}^{N-1} x'_k Q_k x_k + \sum_{k=1}^{N-1} (y_k - y_k^R)' Q R_k (y_k - y_k^R) + \\
 &\quad + \sum_{k=0}^{N_u-1} (u_k - u^R)' R_k (u_k - u^R) + \sum_{k=0}^{N_u-1} \Delta u'_k R 1_k \Delta u_k \\
 \text{s.t.} \quad &x_{t+1} = A x_t + B u_t + G w_t \\
 &y_t = C x_t + v_t \\
 &y_{\min} \leq y \leq y_{\max} \\
 &u_{\min} \leq u \leq u_{\max} \\
 &x_t \in X \subseteq \mathfrak{R}^p, u_t \in U \subseteq \mathfrak{R}^s
 \end{aligned} \tag{1}$$

where \hat{x} are the estimated states given by the state estimator, y outputs and u controls, w are the process disturbances and v the measurement noise, all (discrete) time dependent vectors. The subsets of output variables that get tracked have time-dependent set points y^R . Finally, Δu are changes in control variables, $\Delta u(k) = u(k) - u(k-1)$. The prediction horizon is denoted by N and control horizon by N_u . X , U are the sets of the state and input constraints that contain the origin in their interior. Both $Q > 0$, the objective coefficient for the states and $P > 0$, the terminal weight matrix for the states, are symmetric semi-positive definite matrices. The quadratic matrix for manipulated variables $R > 0$ is a symmetric positive matrix, QR is the quadratic matrix for tracked outputs and RI is a weight matrix for the control action changes (Δu).

3. Results

For the design of the mp-MPC controller the methodology presented in the previous section is used. The following tuning parameters are used: the objective coefficients for states (x), $Q=0$ when we have no state estimation and $Q=1$ in the case with state estimation, the control horizon $N_u=1$ and the prediction horizon $N=20$. The control actions for this process are in different ranges, for the first control action, Q , the nominal value is 45 while for the second control action the nominal value is $8.95e-6$. Moreover, it can be observed from Figure 2 that the control action on Feed 2 is penalized more than the control action on Q , H_2O concentration changes will be slower and API concentration changes will be faster. The higher the value of the weight factor, the more the command will be penalized.

First the optimization problem (1) is solved offline using the POP toolbox which will result in a map of all possible solutions. Once the look-up table is obtained, the controller will only have to perform simple function evaluations to derive the optimal control action that will be given to the process to take it to the desired setpoint value.

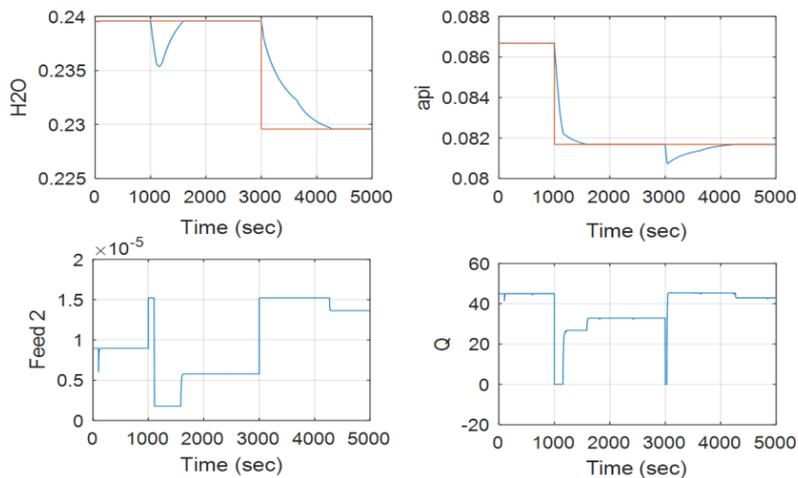


Figure 2. Closed loop response (mp-MPC). Step change on API and H_2O concentration. -4.16% step change on H_2O concentration and -5.7% API concentration. The H_2O and API concentration is given in wt%, U control action Feed 2 is the flowrate in [kg/sec] and U control action Q is in [J/sec].

A step change of -5.7% is given to the API concentration set point and a step of -4.16% is given to the H_2O concentration set point and the results are presented in Figure 2. The

designed controller due to its tuning parameters has a less aggressive behaviour, a fast enough settling time, and does not result in any undershoot/overshoot.

It can be observed that the step on the API reference will also affect the H₂O concentration and the H₂O reference will affect the API concentration but the controller is able to deal with the step changes and bring both concentrations back to the setpoint values.

To further test the performances of the designed controllers, a more realistic disturbance signal is designed as presented in Figure 3. The three different disturbance pattern signals are applied to the three concentrations of Feed 1: H₂O, Solvent and API once the process reaches steady state.

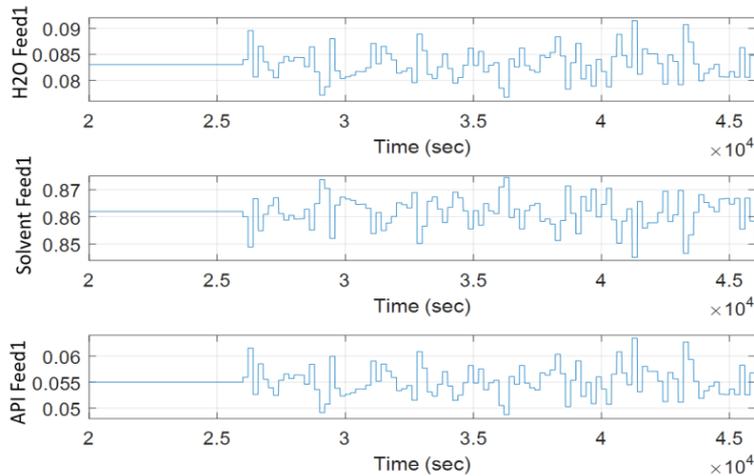


Figure 3. Realistic disturbance signal. The concentrations of API, Solvent and H₂O are represented in wt%

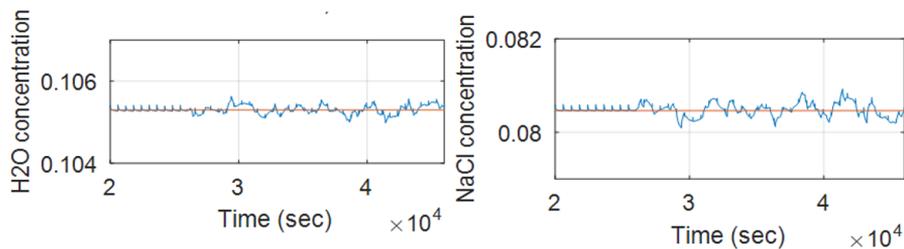


Figure 4. Closed loop response using a realistic disturbance signal

From Figure 4 it can be observed that the mpMPC controller is able to compensate for the disturbances introduced by Feed 1 (H₂O, Solvent, API). Since the mpMPC is multivariable, each control action will be responsible for both outputs.

4. Conclusions

In this paper we develop a multi-parametric model predictive control strategy for the control of the concentration inside of a continuous evaporation process. We start by developing a comprehensive mathematical model of the process for one molecule type

and implement it within the gPROMS platform. To analyze the performances the designed controllers are tested on the developed gPROMS model for: reference tracking, step changes in references as well as disturbances on the input concentration. The developed strategies show good performances without having significant overshoot or undershoot as well as a fast settling time. Moreover, the controllers are capable of maintaining the desired setpoint values while rejecting disturbances.

Acknowledgments

The authors would like to gratefully acknowledge Dr. Salvador Garcia Munoz and Dr. Martin Johnson for their knowledge and help as well as financial support from Eli Lilly and Company, Texas A&M University and Texas A&M Energy Institute is gratefully acknowledged.

References

- Bloore, C. G. and O-Callaghan, D. (2009) 'Process Control in Evaporation and Drying', *Dairy Powders and Concentrated Products*, 332-350.
- Bryan W. Hackett, P. e. (2018) 'The Essentials of Continuous Evaporation', *CEP, AIChE*.
- Diangelakis, N. A., Burnak, B., Katz, J. and Pistikopoulos, E. N. (2017) 'Process design and control optimization: A simultaneous approach by multi-parametric programming', *AIChE Journal*, 63(11), 4827-4846.
- Govatsmark, M. and Skogestad, S. (2001) 'Control structure selection for an evaporation process', *En Proceedings of ESCAPE*, 9, 657-662.
- Ierapetritou, M., Muzzio, F. and Reklaitis, G. (2016) 'Perspectives on the continuous manufacturing of powder-based pharmaceutical processes', *AIChE Journal*, 62(6), 1846-1862.
- Nascu, I., Diangelakis, N. A., Oberdieck, R., Papathanasiou, M. M., Pistikopoulos, E. N. and Ieee (2016) 'Explicit MPC in real-world applications: the PAROC framework' in *2016 AMERICAN CONTROL CONFERENCE (ACC)*, 913-918.
- Oberdieck, R., Diangelakis, N. A., Nascu, I., Papathanasiou, M. M., Sun, M. X., Auraamidou, S. and Pistikopoulos, E. N. (2016) 'On multi-parametric programming and its applications in process systems engineering', *CHEMICAL ENGINEERING RESEARCH & DESIGN*, 116, 61-82.
- Pistikopoulos, E. N., Bozinis, N. A. and Dua, V. (1999) 'POP: A Matlab implementation of multi-parametric quadratic programming algorithm', *London, UK: Centre for Process System Engineering, Imperial College*.
- Pistikopoulos, E. N., Diangelakis, N. A., Oberdieck, R., Papathanasiou, M. M., Nascu, I. and Sun, M. (2015) 'PAROC - An integrated framework and software platform for the optimisation and advanced model-based control of process systems', *Chemical Engineering Science*, 136, 115-138.
- Pistikopoulos, E. N., Dua, V., Bozinis, N. A., Bemporad, A. and Morari, M. (2002) 'On-line optimization via off-line parametric optimization tools', *Computers and Chemical Engineering*, 26(2), 175-185.
- Pistikopoulos, E. N., Georgiadis, M. and Dua, V. (2007a) *Multi-parametric Programming: Theory, Algorithms and Applications*, Wiley-VCH.
- Politis, S., Colombo, P., Colombo, G. and Rekkas, D. (2017) 'Design of experiments (DoE) in pharmaceutical development', *Drug Development and Industrial Pharmacy*, 43, 1-36.
- Rantanen, J. and Khinast, J. (2015) 'The Future of Pharmaceutical Manufacturing Sciences', *Journal of pharmaceutical sciences*, 104(11), 3612-3638.
- Seborg, D. E., Edgar, T. F., Mellichamp, D. A. and Doyle, F. J. (2017) *Process dynamics and control*.
- Su, Q., Ganesh, S., Moreno, M., Bommireddy, Y., Gonzalez, M., Reklaitis, G. V. and Nagy, Z. K. (2019) 'A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing', *Computers and Chemical Engineering*, 125, 216-231.
- Wang, Z., Escotet-Espinoza, M. S. and Ierapetritou, M. (2017) 'Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models', *Computers and Chemical Engineering*, 107, 77-91.

Process as a battery: Electricity price based optimal operation of zeolite crystallization in a COBR

Robin Semrau^a, Jiadi Yang^a and Sebastian Engell^a

^a*TU Dortmund University, August-Schmidt-Straße 1, 44227 Dortmund, Germany
Robin.Semrau@tu-dortmund.de*

Abstract

This paper investigates the flexible operation of a Continuous Oscillatory Baffled Reactor for the hydrothermal synthesis of zeolites to find the economically optimal operating trajectory in the presence of varying electric energy prices. The process and the rigorous dynamic model are introduced. The performance of the dynamic optimization scheme is evaluated in simulation studies. A clear benefit of the dynamic adaptation to the variations of the energy price is demonstrated.

Keywords: dynamic optimization, Continuous Oscillatory Baffled Reactor, energy price based operation, demand side management

1. Introduction

The electric energy supply is in a transition towards renewable power sources. Due to their high volatility, the inclusion of electric energy renewable sources is a major challenge for the process industry. The fluctuating power supply and the largely steady demand of continuous processes create a supply-demand mismatch. This leads to high variations in the electric energy prices, which are expected to be amplified in the future as demonstrated by Göransson et al. (2019).

The ability to react to energy price variations, therefore, has two benefits, first an economic benefit of using electric energy at low-price levels and second to increase the stability of the energy grid by adapting the demand. The approach presented here dynamically adapts the throughput and the heating duties to optimize the productivity of the process, while ensuring a mean production rate and the specified product quality.

Similar approaches have already been reported in the literature for scheduling problems under the headline of demand-side management (DSM) as presented by e.g. Leo et al. (2021) or Brée et al. (2018). In these, the plant dynamics are neglected and steady-state models used instead. However, with fast-changing prices of electric energy, the steady-state assumptions do not hold. In those cases, calculated schedules using steady-state models can lead to suboptimal performance or constraint violation as shown by Caspari et al. (2019). Therefore, the non-linear dynamic plant model is used to calculate an optimal dynamic schedule.

The approach of an optimized time-varying continuous operation is investigated in this paper for a hydrothermal production of zeolites in a continuous plant as presented in Ramirez Mendoza et al. (2020). The production of zeolites is an energy-intensive process which is typically performed in batch reactors. The shift to continuous operation has several advantages such as heat integration and lower variations of product quality. The sedimentation problem of the suspended solid is solved by using the Continuous Oscillatory Baffled Reactor (COBR) concept. However, the adaptive flexible operation of the distributed continuous plant is a main challenge, which motivates this paper.

2. Process overview

A schematic representation of the process is shown in Figure 1. The reactants are previously mixed and fed as a suspension of amorphous solids. The feed is preheated with an integrated heat exchanger. The suspension is fed into the Continuous Oscillatory Baffled Reactor, which is equipped with three direct electrical heaters. The three heaters have different lengths as shown in the schematic. The heating duties of the three heaters can be set individually. Within the COBR reactor, the hydrothermal formation of zeolites is performed at elevated temperatures, the residence time is approximately 80 min. The suspension of crystalline zeolites leaves the reactor and is used to heat the feed stream. The operational limits of the reactor are a minimal flow rate of 60 L/h and a maximum temperature of 413.15 K of the reaction medium and of the reactor wall. The product requirements are a minimum crystallinity of 98% and a fixed average throughput.

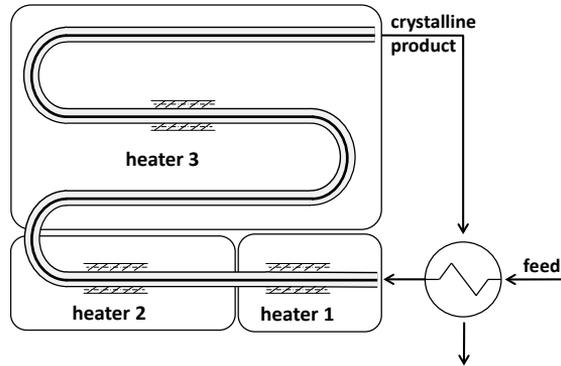


Figure 1: Schematic of the process

The suspension of crystalline zeolites leaves the reactor and is used to heat the feed stream. The operational limits of the reactor are a minimal flow rate of 60 L/h and a maximum temperature of 413.15 K of the reaction medium and of the reactor wall. The product requirements are a minimum crystallinity of 98% and a fixed average throughput.

2.1. Model

The reactor dynamics are described with the partial differential equations (1).

$$\rho c_p \partial_t T_w = \dot{Q}_E + \dot{Q}_H - \dot{Q}_{Rw} \quad (1a)$$

$$\rho c_p \partial_t T = -u \rho c_p \partial_z T + \rho c_p \partial_z [\lambda \partial_z T] + \dot{Q}_{Rw} / \rho c_p \quad (1b)$$

$$\partial_t c_i = -u \partial_z c_i + \partial_z [D \partial_z c_i] + r_i \quad (1c)$$

$$\partial_t n = -u \partial_z n + \partial_z [D \partial_z n] + G \partial_L n + B \quad (1d)$$

$$B = k_n(T)(c_I - c_{I,eq})c_{am} \quad (1e)$$

$$G = k_g(T)(c_I - c_{I,eq}) \quad (1f)$$

The energy balance of the wall temperature is shown in (1a) in which the effects of the heat transfer to the reaction medium, the loss to the environment and heat input of the direct electrical heaters are incorporated. The axial dispersion model for the energy and concentration balances of the reaction suspension is shown in (1b),(1c). The reaction term r_i describes the conversion from the amorphous state via dissolved ions in the liquid phase to the crystalline product. The population balance (1d) describes the particle density distribution (n) in size (L) with a heterogeneous nucleation term (1e) and the size-independent growth model (1f). It is solved using the method of moments. The spatial discretization is approximated using the finite differencing method on 30 grid points. The resulting non-linear ordinary differential equation model has 240 dynamic states.

2.2. Process economics

The profit of the process includes the influence of the product revenue, the reactant and energy costs and the fixed costs as shown in (2).

$$\text{Profit} = \underbrace{P_p \cdot \dot{m}}_{\text{product revenue}} - \underbrace{P_{Rs} \cdot \dot{m}}_{\text{reactant cost}} - \underbrace{P_E \cdot \dot{Q}_H}_{\text{energy cost}} - \underbrace{C_{Fix}}_{\text{fixed cost}} \quad (2)$$

The fixed costs can be neglected when optimizing the operating conditions since the additive term in the objective does not influence the solution of an optimization problem. The other terms are scaled with a reference operating point and divided by the (negative) net revenue. The scaled operational cost is shown in equation (3), in which the parameter γ described the scaled price of electric energy.

$$\Phi(u, \gamma) = -\dot{m}/\dot{m}_{ref} + \gamma \dot{Q}_H/\dot{Q}_{H,ref} \quad (3)$$

3. Optimization problem

3.1. Dynamic optimization problem

The formulation of the dynamic optimization problem of the flexible operation is shown in equation (4). The operation over one and a half day is minimized assuming perfect knowledge of the future price of electric power.

$$\min_{u,x} \int_0^{t_f} \Phi(u, \gamma(t)) dt \quad (4a)$$

$$\text{s.t.: } \dot{x} = f(x, u), \quad (4b)$$

$$0 \geq g(x, u), \quad (4c)$$

$$\bar{X}_j \geq 0.98 \quad \forall j \in [1, K] \quad (4d)$$

$$\bar{V}_j = \dot{V}_{fix} \quad \forall j \in [1, K] \quad (4e)$$

$$\int_0^L T(t_f, \cdot) dz = \int_0^L T(0, \cdot) dz \quad (4f)$$

$$\int_0^L m_3(t_f, \cdot) dz = \int_0^L m_3(0, \cdot) dz \quad (4g)$$

The objective (4a) is the scaled cost defined in (3). The dynamic process model is represented by (4b). The process bounds defined in section 2 are represented by (4c). The restrictions on the mean throughput and the mean product quality are included as averaging constraints (4d)-(4e) in the dynamic optimization problem. These constraints must be met on three ($K = 3$) subsequent averaging horizons which have the length of 12 h each. The specified crystallinity and the specified throughput should be reached on average over 12, 24 and over 36 hours. This reduces the needed storage capacity without losses of productivity. The optimization will be repeated after 24 h and therefore the solution for the horizon between 24 h and 36 h will not be applied but in combination with the endpoint hold-up constraints (4f),(4g), terminal sell-off effects can be reduced in this manner which is further discussed in 4.1.

The resulting non-linear optimization problem was implemented using the software CasADi by Andersson et al. (2018) and solved with the IPOPT solver. The ode model (4b) is solved using orthogonal collocation on two finite elements using second order Radau polynomials. The optimization problem has 103.968 variables and is solved within 3-4 h.

3.2. Static optimization problem

Additionally, a static optimization problem is defined in (5) which represents a scenario in which the plant dynamics is infinitely fast. The objective function and process bounds are the same as in (4). Constraint (5b) enforces a steady-state operation, while the mean crystallinity and throughput are enforced by (5d) and (5e).

$$\min_{u_i, x_i} \sum_{i=0}^N \Phi(u_i, \gamma) \quad (5a)$$

$$\text{s.t.: } 0 = f(x_i, u_i), \quad (5b)$$

$$0 \geq g(x_i, u_i), \quad (5c)$$

$$\bar{X}_i \geq 0.98 \quad \forall i \in [0, N] \quad (5d)$$

$$\bar{V}_i = \dot{V}_{fix} \quad \forall i \in [0, N] \quad (5e)$$

4. Results

4.1. Terminal sell-off

The terminal sell-off commonly observed in economic MPC literature e.g. by Ellis et al. (2014), is caused by the approximation of the infinite horizon optimization problem by a finite horizon optimization problem which results in suboptimal performance for a continuous operation. The results of the optimization problem (4) with a constant price of energy are shown in Figure 2. In the second and third plot, the results with the end-point constraints (4f),(4g) are shown, while in the first plot the results without end-point constraints are displayed. In the first and the second plot, one averaging horizon was used ($K = 1$), whereas in the third plot, two averaging horizons were used ($K = 2$). The black dotted line represents the steady-state optimum. The

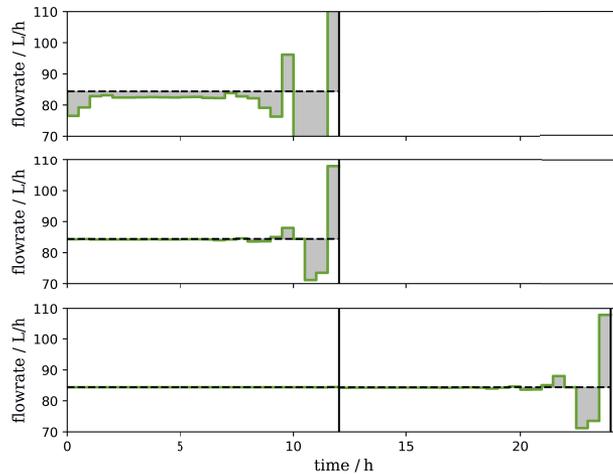


Figure 2: Optimal calculated flowrates for different number of averaging horizons and different end-point constraints

influence of the final horizon can clearly be observed by the deviation of the dynamic optimization results from the steady-state result. However, the effect is far more prominent in the first case without terminal constraints. The terminal sell-off leads to a deviation even at the beginning of the optimization horizon. Using the end-point constraints (4f),(4g) this effect is reduced. However, it still leads to a deviation close to the end of the horizon. Using the end-point constraints and an additional averaging horizon the deviation from the steady-state optimum is shifted to the end of the second averaging horizon, whereas no deviation in the first horizon is observed. The solution proposed in this work is similar to the approach in MPC, i.e. discarding the results for the final horizon and re-solving the optimization problem with a shifted time frame.

4.2. Dynamic optimization

The results of the solution of the dynamic optimization problem (4) with a time-varying energy price are shown in Figure 3. The variation of the energy price is assumed to be a square wave with a relative amplitude of 0.56 and a frequency of 0.167 1/h which is displayed in the bottom plot. The initial state $x(t = 0)$ is the optimal steady-state for the specified production rate. The temperature and the crystallinity at the reactor outlet are displayed in the upper two plots. The corresponding inputs, the flow rate and the heating powers are shown below. The green solid line in the plots represents the calculated optimal trajectory. In the fourth plot, the three areas represent

the heating powers of the first, the second and the third heater. The coloured dotted line shows the solution of the static optimization problem (5) for the given price variation while the black dotted line is the optimal steady-state for the average production rate. The average values of the flow rate and of the crystallinity are constrained. The deviation to these average values is shown with the grey shaded areas. The vertical black lines show the end of the averaging horizons after 12 h, 24 h and 36 h.

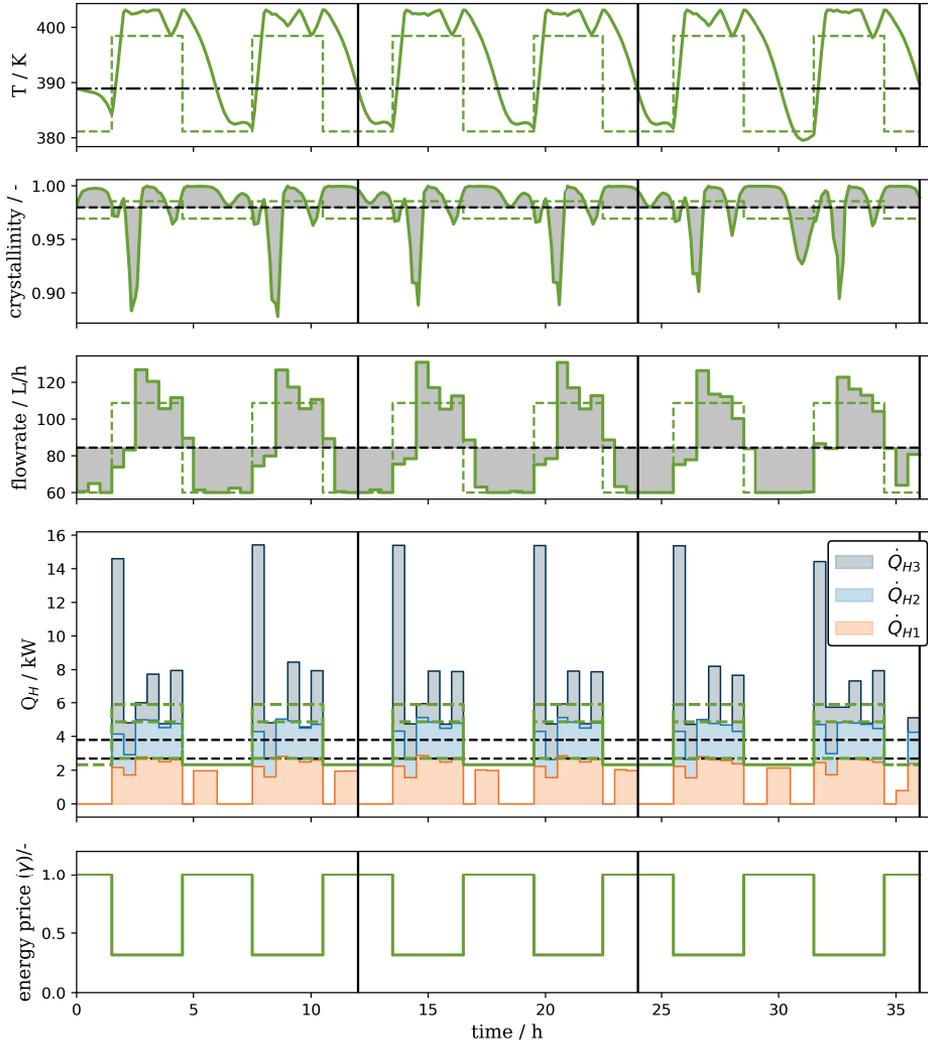


Figure 3: Result of the dynamic optimization with the a varying energy price

The periodic energy price variation leads to a periodic pattern of the operation of the COBR. At a high price of energy, the flow rate is close to its lower bound and the heating power is small. When the price of energy switches to a low value (e.g. after 13.5 h) the heating power of all three heaters reaches a maximum, while the flow rate increases to the mean value. The extensive heating reduces the drop of the product quality which is caused by the low energy input in the high

energy price period. The usage of the third heater is detrimental from an energetic point of view but is compensated by the savings due to the lower price of energy. The temperature is increased and stays close to the steady-state level. One hour after the price change, the flow rate reaches a maximum and slowly drops afterwards. Within that period, the first two heaters have a constant value. The heat input of the third heater oscillates around the static value. The oscillation ends with a maximum in order to store energy in the system before the increase of the price. After the transition to high values of the price of energy, the flow rate is rapidly decreased to the lower bound and the heaters are turned off. In this dynamic transition period, the residual heat is used. Later, to ensure the desired crystallinity, the first heater is turned on again and operated close to the steady-state optimum around the middle of the high energy price period. The variations of the different periodic patterns from each other in the beginning and to the end are caused by the initial and terminal constraints. The benefit of the shown operation in comparison to a static optimal operation can be calculated as an energy cost saving of 32.4% for this scenario. The proposed method is capable to calculate an optimal dynamic schedule assuming perfect knowledge of the variation of electric energy price. The complexity of the calculated purely dynamic operation of the COBR underlines the necessity for the inclusion of the full dynamic plant model in the optimization. Using this method, the optimal operation schedule dependent on the variation of the price of electric power can be investigated.

5. Conclusion

In this work, a flexible operation of a continuous zeolite crystallization process in a COBR is demonstrated. The plant and the model are introduced and the resulting dynamic optimization problem is described. The influence of the finite horizon approximation is further analysed and its influence on the optimized schedule is minimized using end-point constraints and an extended optimization horizon. The dynamic optimization results are presented which show a periodic pattern of operation, following the variations of the price of electric power. The results show a large potential of the adaptive dynamic operation. Further extensions of this work concern including alternative electrical energy inputs, such as sonication or vapour compression technologies and the handling of the uncertainty in the predictions of the variations of the price of energy similar to the approach presented by Leo and Engell (2018) for scheduling problems.

References

- Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., Diehl, M., 2018. CasADi: a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11 (1), 1–36.
- Brée, L. C., Perrey, K., Bulan, A., Mitsos, A., 2018. Demand side management and operational mode switching in chlorine production. *AIChE Journal* 65 (7), e16352.
- Caspari, A., Offermanns, C., Schäfer, P., Mhamdi, A., Mitsos, A., 2019. A flexible air separation process: 2. optimal operation using economic model predictive control. *AIChE Journal* 65 (11).
- Ellis, M., Durand, H., Christofides, P. D., 2014. A tutorial review of economic model predictive control methods. *Journal of Process Control* 24 (8), 1156–1178.
- Göransson, L., Lehtveer, M., Nyholm, E., Taljegard, M., Walter, V., 2019. The benefit of collaboration in the north european electricity system transition—system and sector perspectives. *Energies* 12 (24), 4648.
- Leo, E., Ave, G. D., Harjunkoski, I., Engell, S., 2021. Stochastic short-term integrated electricity procurement and production scheduling for a large consumer. *Computers & Chemical Engineering* 145 (1), 107191.
- Leo, E., Engell, S., 2018. Integrated day-ahead energy procurement and production scheduling. *at - Automatisierungstechnik* 66 (11), 950–963.
- Ramirez Mendoza, H., Valdez, M. L. P., Van Gerven, T., Lutz, C., 2020. Continuous flow synthesis of zeolite FAU in an oscillatory baffled reactor. *Journal of Advanced Manufacturing and Processing* 2 (2).

Acknowledgement

The project leading to this publication has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820716.(project SIMPLIFY)

Leveraging Deep Learning for Efficient Explicit MPC of High-Dimensional and Non-linear Chemical Processes

Ahmed Shokry ^{a*}, Mehdi Abou El Qassime ^{a,b}, Eric Moulines ^{a,b}

^a*École polytechnique, Route de Saclay 91128 Palaiseau Cedex, France*

^b*Emines School of Industrial Management, LOT 660 43150, Ben Guerir, Morocco*

**ahmed.shokry@polytechnique.edu*

Abstract

This paper presents an efficient Deep Learning (DL) based method for explicit model predictive control (E-MPC) of high-dimensional and/or nonlinear chemical processes for which mathematical E-MPC approaches are difficult to apply. The method uses DL models for off-line development of control laws that accurately approximate the relationship between the optimal values of the control variables to be applied in the next sampling period (SP) as a function of the values of the state variables in the current SP. The training data are generated by solving the MPC problem considering different initial values of the state variables selected using design of computer experiments (DOCE) techniques. The obtained DL-based control laws are then integrated into a closed-loop for online control of the process. A numerical validation procedure is used to evaluate the performance of the developed control laws in terms of their accuracy and computational cost. The method is applied to case studies for which a “direct” application of mathematical E-MPC techniques is difficult due to their high dimensionality and nonlinearity. The results show a high performance of the proposed method and a reduction in the complexity of the solution procedure compared to mathematical E-MPC.

Keywords: Control; Explicit MPC; Chemical Processes; Deep Learning; ANNs.

1. Introduction

MPC is one of the most widely used control methods in engineering because it can efficiently deal with multivariate systems, hard constraints on control and/or state variables, and different types of control objectives, e.g., economic objectives. However, MPC faces the challenge that its expensive online computations require high computational resources: at each SP, MPC solves a dynamic optimization problem (open-loop optimal control) that requires repeated evaluation of a dynamic process model (Karg and Lucia, 2020). Therefore, MPC applications may be infeasible for processes characterized by high dimensionality, nonlinearity, and/or fast dynamics (Rivotti et al., 2012). To overcome this challenge, E-MPC methods - also called MultiParametric MPC - have been developed to solve the MPC problem off-line, providing simple mathematical relationships that describe the future values of the optimal control variables as a function of the current values of the state variables, where each relationship applies only to a specific subregion of the state variables space, called critical region (Chen et al., 2018). These mathematical laws are then used online to drive the process, requiring very little computational time. Although E-MPC methods have been successfully applied in many cases in various engineering fields, they are limited to linear, discrete-time state-space models with moderate dimensionality (Rivotti et al., 2012). Recently, the use of machine

learning has been proposed to extend these limitations by exploiting their universal approximation and dimensionality reduction capabilities to develop accurate and computationally cheap data-driven control laws that approximate the relationship between the optimal future control values and the current state values. For example, Shokry et al. (2016) used multilayer artificial neural networks (ANNs) to develop explicit control laws and applied them to a simple continuous stirred tank reactor (CSTR). Chen et al. (2018) used constrained ANNs to develop control laws that explicitly learn the mathematical constraints of the MPC problem, and applied them to simple linear, discrete-time state-space models. Karg et al. (2020) used ANNs to approximate control laws and applied them to linear, discrete-time state-space models of mass-spring systems. However, most of these methods are tailored to the existence of simple linear discrete-time state-space models to which mathematical E-MPC methods can be easily applied, and also require large amounts of training data. Therefore, in this paper, an efficient DL-based method for E-MPC (DL-E-MPC) of high-dimensional and nonlinear chemical processes is presented, which uses DL models for off-line development of control laws that accurately approximate the relationship between the optimal values of the control variables to be applied at the next SP as a function of the values of the state variables at the current SP. The training data are generated by solving the MPC problem considering different initial values of the state variables selected by DOCE techniques. The obtained DL-based control laws are then integrated into a closed-loop for online control of the process. A validation procedure is used to evaluate the performance of the control laws in terms of i) open-loop response accuracy (i.e., over a SP, without feedback), ii) closed-loop response accuracy (i.e., over the entire control trajectory, with feedback), iii) state accuracy, and vi) online computational cost. The proposed method is applied to two case studies from the literature (Rivotti et al., 2012), to which a "direct" application of E-MPC methods is not possible due to their high dimensionality and/or nonlinearity.

2. Problem statement

In a generic MPC problem (Eqs. (1:3)), a process is controlled over a finite time horizon composed of a number N^{fnl} of equal SP, relying on an accurate process model (Eq. (2)). At each SP $k, k = 1, 2, \dots, N^{fnl}$, the optimal trajectory of the manipulated inputs $[u_{t+1}^*, \dots, u_{t+N_p}^*]$ over a specific prediction horizon N_p is calculated by solving a dynamic optimization problem that considers a desired performance index J (Eq. (1)) and subjected to a set of constraints $g_l, l = 1, \dots, L$ (Karg and Lucia, 2020). The measured values of the state variables at the previous SP, x_{k-1} , are used as the initial conditions for this optimization problem. Then, only the values of the calculated optimal control profile corresponding to the first SP, u_{t+1}^* , are implemented in the real process, and so on. Notice that in Eq. (1), $Q \in R^{m \times m}$ and $\mathfrak{R} \in R^{v \times v}$ represent weight metrics of the objective function J , while \check{r} is the setpoint.

$$\min_{u_{t+1}, \dots, u_{t+N_p}} J = \sum_{k=1}^{N_p-1} (x_{t+k} - \check{r})' Q (x_{t+k} - \check{r}) + \Delta u_{t+k}' \mathfrak{R} \Delta u_{t+k} \quad (1)$$

$$S.T. \quad x_{t+1} = F(x_t, u_t), \quad x \in R^m, u \in R^v \quad (2)$$

$$g_l(x_t, u_t) \leq 0, \quad l = 1, 2, \dots, L \quad (3)$$

$$u_{min} \leq u_k \leq u_{max}, \quad \Delta u_{t+k} = u_{t+k} - u_{t+k-1}, \quad k = 1, \dots, N_p.$$

3. Proposed methodology

Motivated by situations where F (Eq.(2)) is a highly nonlinear and/or high-dimensional model that hinders the application of mathematical E-MPC methods, this paper proposes an efficient DL-E-MPC for such processes, consisting of the following steps:

3.1. Sampling of the initial values of state variables

In the first step, different n^{tr} values of the initial state variables $[x_{0,i}]$, $i = 1, \dots, n^{tr}$, $x \in R^m$ are selected within their expected/feasible variability domain. The selection is done using a DOCE technique that covers the entire variability space of the state variables to obtain information about the control behavior over the whole state domain. In this work, we use a hybrid DOCE technique that combines Hammersley and fractional designs, which generates uniform sampling schemes with very low computational requirements.

3.2. Training and testing data generation by solving the MPC problem

In the second step, the MPC problem is solved n^{tr} times as formulated in Eqs. (1:3) but only over a finite duration $N^{trn} \leq N^{fnl}$, where in each time, one of the previously generated initial state values, $x_{0,i}$, is considered as the initial conditions for the $i - th$ MPC problem. This allows to obtain the optimal closed-loop trajectories of the control inputs $[u_{1,i}^*, u_{2,i}^* \dots, u_{N^{trn},i}^*]$ and the associated trajectories of the state $[x_{0,i}, x_{1,i}, \dots, x_{N^{trn}-1,i}]$, $i = 1, \dots, n^{tr}$. Then, the n^{tr} pairs of control-state trajectories are folded into an input-output dataset in the form of $[x_{t,i}, u_{t+1,i}^*]$, $i = 1, \dots, n^{tr} \times N^{trn}$. Another testing dataset is generated in the same way, as described in Sections 3.1 and 3.2, but in this time, the MPC problem is solved over the entire time horizon N^{fnl} to obtain n^{ts} pairs of optimal closed-loop control-state trajectories, i.e., $[u_{1,i}^{*ts}, u_{2,i}^{*ts}, \dots, u_{N^{fnl},i}^{*ts}] - [x_{0,i}^{ts}, x_{1,i}^{ts}, \dots, x_{N^{fnl}-1,i}^{ts}]$, $i = 1, \dots, n^{ts}$. These test trajectories are also folded into input-output dataset in the form of $[x_{t,i}^{ts}, u_{t+1,i}^{*ts}]$, $i = 1, \dots, n^{ts} \times N^{fnl}$. The folded test dataset will be used to assess the open-loop accuracy of the control laws, while the n^{ts} test pairs of control-state trajectories will be used to assess their closed-loop accuracy.

3.3. Development of DL-based control laws

In this step, the generated training data $[x_{t,i}, u_{t+1,i}^*]$, $i = 1, \dots, n^{tr} \times N^{trn}$ are used to develop a number of v DL-based control laws $\hat{u}_{i,t+1}^* = \mathcal{F}_{DANNi}(x_t)$ $x \in R^m$, $i = 1, 2, \dots, v$, where \mathcal{F}_{DANNi} is a feedforward deep ANN (DANN) for regression. The selection of the DANNs' structure (number of hidden layers and their respective sizes, type of activation functions, training loss, etc.) and the training algorithm is done by a cut-and-try procedure that seeks for the best compromise between the prediction accuracy and the simplicity of the structure (Shokry et al., 2016). After training the control laws, their open-loop response accuracy is evaluated using the folded test dataset $[x_{t,i}^{ts}, u_{t+1,i}^{*ts}]$, $i = 1, \dots, n^{ts} \times N^{fnl}$ generated in Section 3.2, and an accuracy measure, such as the Normalized Room Mean Square Error (NRMSE) (Eq. (4)), can be calculated for each of the v control laws.

$$NRMSE = 100 \times \frac{\sqrt{\frac{1}{n^{ts} \times N^{fnl}} \sum_{n^{ts} \times N^{fnl}} (u_{t+1,i}^{*ts} - \hat{u}_{t+1,i}^{*ts})^2}}{(u_{max}^{*ts} - u_{min}^{*ts})} \quad (4)$$

where u_{max}^{*ts} and u_{min}^{*ts} are the maximum and minimum values of the control variables in the testing dataset, respectively.

3.4. Online (closed-loop) deployment of the control laws

Finally, the control laws based on DL are integrated into a closed-loop to predict the entire control trajectory starting from arbitrary initial conditions of the state variables: at each SP, the real outputs measured from the process are sent to the control laws to calculate the optimal values of the control variables, which are sent to the actuators to implement them in the process, and so on. The closed-loop accuracy (in terms of NRMSE) of the control laws is evaluated using the n^{ts} pairs of optimal control-state trajectories generated in Section 3.2. This closed-loop NRMSE is calculated for each predicted trajectory individually and then averaged over the n^{ts} trajectories. A third NRMSE is calculated to evaluate the accuracy of the process state variables driven by the application of the predicted control trajectories compared to their exact values calculated by mathematical MPC.

4. Applications

4.1. Case 1: MPC of two connected CSTRs

A series of two CSTRs is considered in which an irreversible reaction $A \Rightarrow B$ takes place (Rivotti et al., 2012), where reactant A is fed into the first reactor and the resulting mixture is then fed into the second reactor. A nonlinear model describing the dynamic relationships between six state variables and two manipulated variables is used to control the process. The state variables include the reactant concentration, temperature, and volume in each reactor, $CA1, V1, T1, CA2, V2, T2$, and the manipulated variables are the heat supplied to the first reactor, η , and the outlet flow from the second reactor, ω . The objective is to bring the controlled state variables, volume and temperature of the second reactor, to their setpoints $V2 = 100 L$, $T2 = 463 c^o$ through manipulating the control variables within their allowable limits: $0.75 \leq \eta \leq 1.1$ and $-0.75 \leq \omega \leq 1.1$, assuming that all state variables are measured and there are no external disturbances (Rivotti et al., 2012). The problem is formulated as in Section 2 with $Q = [1 \ 0; 0 \ 1]$, $R = [0.05 \ 0; 0 \ 0.05]$, $N_p = 10$ and $N^{fnl}=200$ SPs each of 0.1 min.

The proposed method is applied as described in Section 3. A sampling plan of $n^{tr} = 400$ value combinations of the initial state variables is generated, $[CA1_{o,i}, V1_{o,i}, T1_{o,i}, CA2_{o,i}, V2_{o,i}, T2_{o,i}]_{400}$, using the hybrid DOCE technique. Then the MPC problem is solved 400 times over the horizon $N^{trn} = 10$ to obtain the control trajectories $[\eta_{1,i}^*, \eta_{2,i}^* \dots, \eta_{10,i}^*]_{400}$, $[\omega_{1,i}^*, \omega_{2,i}^* \dots, \omega_{10,i}^*]_{400}$ and the associated state trajectories, which are then folded into the input-output training matrix $[CA1_{t,i}, V1_{t,i}, T1_{t,i}, CA2_{t,i}, V2_{t,i}, T2_{t,i}, \eta_{t+1,i}^*, \omega_{t+1,i}^*]_{4000}$. Figure 1-(a) shows the input training data projected onto the space of two state variables, and Figure 1-(b, c) shows the distributions of the output training data, much of which lies at the extremes of the distributions. This is because the MPC problem is solved over a short period of time, N^{trn} , that represents the early period immediately after the initial state where the controller tends to apply allowable extreme control values to quickly stabilize the process. The training set is used to build two DANNs $\hat{\eta}_{t+1}^* = f_{DANN1}^{\eta}(CA1_t, V1_t, T1_t, CA2_t, V2_t, T2_t)$ and $\hat{\omega}_{t+1}^* = f_{DANN2}^{\omega}(CA1_t, V1_t, T1_t, CA2_t, V2_t, T2_t)$. A number of $n^{ts} = 30$ testing optimal closed-loop control profiles are generated, as described in Section 3.2, to evaluate the developed control laws. Table 1 shows the promising performance of the developed control laws in terms of accuracy, which is well below 1% of NRMSE in all cases (open-loop, closed-loop, and states), and the very low online computational overhead required to predict the

control actions in each SP (0.015 s) compared to that of MPC (2.55 s). Figure 2-(a, b) shows the three predicted control trajectories with the highest closed-loop NRMSE (dashed lines) compared to their exact behavior (solid lines), while Figure 2-(c) shows the corresponding state trajectories. The figure again highlights the promising performance of the methodology.

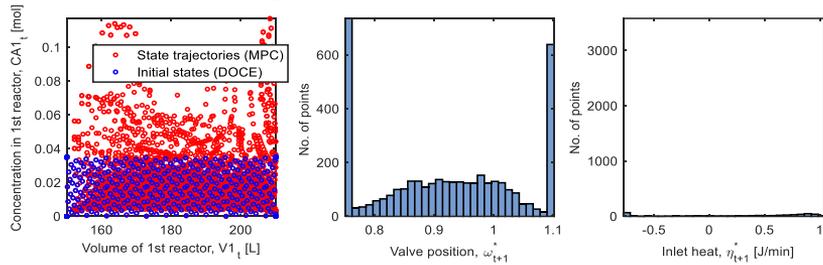


Figure 1. Folded training dataset of the 2-CSTRs case study.

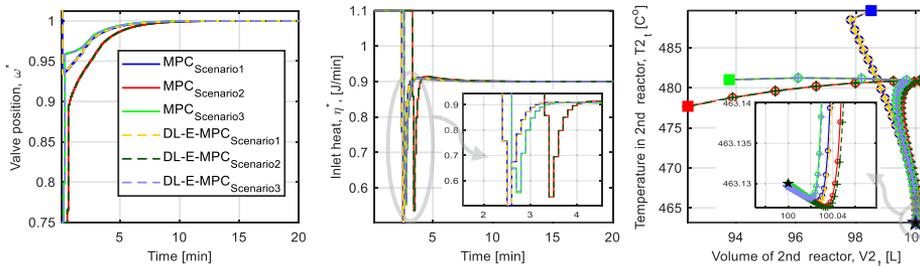


Figure 2. (a,b) Three predicted control trajectories compared to their exact values, and (c) corresponding states trajectories, where squares are initial conditions and star is the setpoint.

4.2. Case 2: MPC of a distillation column

In this case, a distillation column consisting of 30 trays is considered, which is fed with liquid mixtures at the 17th tray (Rivotti et al., 2012). The system is represented by a differential model describing the dynamic relationships between 32 state variables, which include the liquid composition at each tray, y_2, y_3, \dots, y_{31} , distillate composition, y_1 , bottom-product composition, y_{32} , and a control variable represented by the reflux ratio, γ . The goal is to maintain the purity of the distillate at the setpoint, $y_{32} = 0.935$, by manipulating the reflux ratio within its limits $0 \leq \gamma \leq 5$. The parameters of the MPC problem are $Q = 20$, $R = 1 \times 10^{-6}$, $N_p = 15$, and $N^{fnl} = 200$ SPs each of 0.1 min. The method is applied as described in Sections 3 and 4.1 starting by the generation of a sampling plan involving $n^{tr} = 400$ different initial state values $[y_{1,o,i}, \dots, y_{32,o,i}]_{400}$. The MPC problem is, then, solved with $N^{trn} = 40$ to collect the control $[\gamma_{1,i}^*, \dots, \gamma_{40,i}^*]_{400}$ and the related state trajectories. The folded input-output training matrix $[y_{1,t,i}, \dots, y_{32,t,i}, \gamma_{t+1,i}^*]_{16000}$ is used to develop a DANNs-based control law $\hat{\gamma}_{t+1}^* = f_{DANN1}^{\gamma}(y_{1,t,i}, \dots, y_{32,t,i})$, which is assessed using a number of $n^{ts} = 30$ testing control profiles. Table 1 and Figure 3 show the promising assessment results.

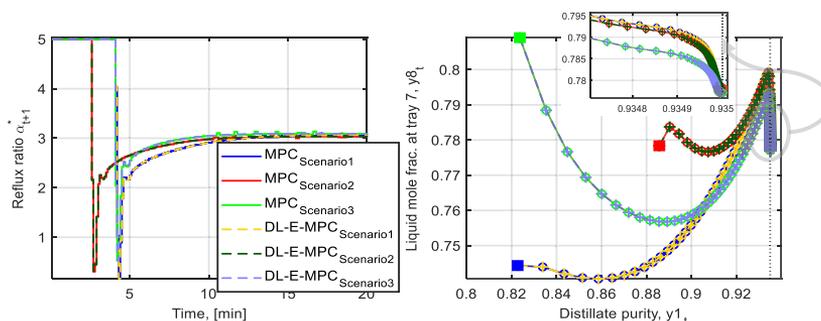


Figure 3. (a) Three predicted control trajectories -with the highest closed-loop NRMSE- compared to their exact values, and (b) corresponding states trajectories, where squares are initial conditions and dotted line is the setpoint.

Table 1. Computational cost and accuracy of the developed DL-based control laws.

		Off-line CPU time (s)*			NRMSE (%)			Online CPU time per SP	
		Training data gen.	Testing data gen.	DANNs training	Open-loop	Closed-loop	Final states	DL-E-MPC	MPC
Case1	f_{DANN1}^{η}	5174	15135	65	0.023	0.066	0.051	0.015	2.55
	f_{DANN2}^{ω}				0.038	0.106			
Case2	f_{DANN1}^{γ}	65306	24124	2922	0.974	0.400	0.003	0.010	4.02

*Intel core (TM) i7-8565U CPU@ 1.80GHz, 12GB RAM.

5. Conclusion

In this paper, we present an efficient DL-E-MPC method for highly nonlinear and/or high-dimensional chemical processes, consisting of four steps and incorporating various tools and techniques, such as DOCE, state-of-the-art dynamic optimization methods, and machine learning techniques. The core of the method is the off-line training of DL models to serve as accurate and computationally cheap control laws, which are then integrated into a closed-loop control scheme to predict the optimal control of the process online. The method is applied to two case studies for which direct application of mathematical E-MPC is difficult. The results are very promising in terms of prediction accuracy (NRMSE less than 1% in all cases) and online computational cost (saving at least 99.4% of computational effort in all cases). Future work will investigate extending the capabilities of the method to handle state constraints, unknown disturbances, and process-model mismatches.

References

- A. Shokry, C. Dombayci, A. Espuña, 2016, Multiparametric Metamodels for Model Predictive Control of Chemical Processes, Computer Aided Chemical Engineering, 2016, 937-942.
- B.Karg, S.Lucia, 2020, Efficient representation and approximation of model predictive control laws via deep learning. IEEE Trans Cybern, 50, 3866-3878.
- P. Rivotti, R. Lambert, E. Pistikopoulos, 2012, Combined model approximation techniques and multiparametric programming for explicit nonlinear MPC, Comp Chem Eng, 277-287.
- S. Chen et al., 2018, Approximating Explicit Model Predictive Control Using Constrained Neural Networks, American Control Conference, 1520-1527.

Incorporation of error propagation into an elemental balancing based soft-sensor for improved online monitoring of microbial fed-batch processes

Don Fabian Müller^a, Konrad Lagoda^a, Daniel Wibbing^b, Christoph Herwig^a and Julian Kager^{c*}

^a*Institute of Chemical, Environmental and Biological Engineering, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria*

^b*Festo SE & Co. KG, 73770 Denkendorf, Germany*

^c*Competence Center CHASE GmbH, 4040 Linz, Austria*

julian.kager@chasecenter.at

Abstract

Monitoring and control of biological processes is still mainly limited to only a few physical and chemical properties that can be measured online with relatively low efforts. Important parameters and variables of a bioprocess are predominantly determined by offline sampling as they are difficult to accurately measure online. Therefore closed-loop control of those critical process variables poses great difficulties to the bioprocessing industry. To overcome this challenge state estimation techniques can be used. For their proper function they need observable process models, which are still not available for every bioprocess. In this work a first principle based soft-sensor is presented, which allows for real-time estimation of biomass and specific reaction rates for cell growth, substrate uptake, O₂ consumption and CO₂ formation in biotechnological production processes. The proposed algorithm utilizes elemental balancing of the carbon flux alongside the degree of reduction balance to incorporate gross error detection and data reconciliation between the two balances. The functionality of gross error detection and data reconciliation strongly rely on the errors of the used measurements. Most of the recent works employed elemental balancing for state estimation considering only constant measurement errors over time. The high dynamic range during fed-batch operation as well as the carbon evolution and oxygen uptake rates being calculated from several measurements necessitates a proper and dynamic error propagation procedure. We successfully applied an adaptive propagation of measurement uncertainties to substrate limited fed-batch cultivations of recombinant *Escherichia coli*. The novel soft-sensor algorithm led to a reduction of the biomass estimation error from formerly 10.8% to 5.3% NRMSE. Especially the consideration of additional uncertainty derived from supplemented pure oxygen increased data reconciliation effectiveness. In addition gross error detection is more sensitive to better indicate sensor faults or mismatches between carbon and degree of reduction balances. Elemental balancing including true measurement errors and error propagation is a valuable tool to estimate biomass and biomass specific reaction rates of microbial fed-batch processes. The basis on the law of mass conservation rather than on very specific process models makes it a generically applicable soft-sensor with reduced efforts for implementation and measurement needs that can also be used for real-time control purposes.

Keywords: bioprocessing, soft sensor, process analytical technology, elemental balancing, error propagation to measurement uncertainties

1. Introduction

The Process Analytical Technology (PAT) initiative, which was introduced by the FDA in 2002, should ensure the quality of bioprocesses not only for the end product but also throughout the whole pharmaceutical manufacturing process. It facilitates modeling approaches as well as new sensor technologies to monitor and control critical process parameters (CPPs) and critical quality attributes (CQAs) during the process [Hausmann et al. (2017)]. Model-based methods are gaining importance as they are able to increase process understanding and automation as well as guide the decision making [Narayanan et al. (2020)]. A software sensor or short soft sensor is the combination of one or several hardware sensors and a software model to derive new estimated variables from those measurements that are hard or impossible to measure directly or measured at too low frequency [Luttmann et al. (2012), Hausmann et al. (2017)]. Soft sensors can be used to effectively monitor and control fermentation processes, estimate important variables that are hard to measure and deal with the complex and strong nonlinear nature of a bioprocess [Zhu et al. (2020)]. Recent works presented material balance based soft-sensor algorithms for growth rate and biomass estimation [Wechselberger et al. (2013), Sagmeister et al. (2013)]. In simulation studies it was already shown that propagation of uncertainties from the measurements to the conversion rates is advantageous compared to an arbitrarily chosen static error assumption [Steinwandter et al. (2017)]. In this contribution we incorporated a first order uncertainty propagation procedure into an elemental balancing soft sensor and compared reaction rate and biomass estimation accuracies to the static error consideration based on real experimental data of a typical fed-batch bioprocess.

2. Methods

Fermentations with a recombinant *Escherichia coli* BL21 (DE3) were carried out in 3.3 L Labfors 5 bioreactors (Infors, Bottmingen, Switzerland). A minimal medium according to DeLisa et al. (1999) was used. The process design was composed of a batch phase with 10 g L^{-1} glucose as the main carbon source, an uninduced fed batch phase with a predefined exponential glucose feed profile to keep the specific growth rate q_X at a constant level and a production phase with constant feed rate, induction with 1 mM isopropyl β -d-1-thiogalactopyranoside (IPTG) and lowering of the temperature from $37 \text{ }^\circ\text{C}$ to $30 \text{ }^\circ\text{C}$. The pH was controlled with ammonia and H_3PO_4 to 7.0. For offline biomass quantification the dry cell weight (DCW) and the optical density at 600 nm (OD_{600}) were determined. Glucose concentration was measured with the Cedex Bio HT and the Glucose Bio HT photometric assay using hexokinase. Real-Time process control was realized through the Lucullus Process Information Management System PIMS (Securecell, Urdorf, Switzerland) together with MATLAB[®]. The soft sensor algorithm shown in chap. 3 was implemented in a Simulink[®] model connected to the Real-Time environment. MATLAB[®] was also used for data preprocessing for the soft sensor inputs and to compute important process variables. The substrate feed rate F_R was calculated by numeric integration of the feed weight signal combined with a Savitzky-Golay-Filter approach [Savitzky and Golay (1964)]. The Jacobian matrix was also computed in MATLAB[®] using the Symbolic Math Toolbox.

3. Soft sensor

3.1. Calculation of C-molar reaction rates from prime variables

The measured rate vector r_m , which is composed of the rates for substrate uptake r_S , O_2 uptake r_{O_2} and CO_2 evolution r_{CO_2} , can be calculated according to the formulas below under the assumption that the reaction rate is equal to the transfer rate measured. The formulas contain the aeration rates for air $F_{n\text{AIR}}$ and oxygen $F_{n\text{O}_2}$, the substrate feed rate F_R and the measured offgas mole fractions of oxygen $x_{\text{O}_2\text{Gout}}$ and CO_2 $x_{\text{CO}_2\text{Gout}}$. These are called the prime variables which are the inputs to the soft sensor. $x_{\text{O}_2\text{Gin}}$ and $x_{\text{CO}_2\text{Gin}}$ are the mole fractions of oxygen and CO_2 respectively in the gas

intake, which can be calculated from gas flow balancing. Parameters are V_{nM} which is the molar gas volume at standard temperature and pressure, c_S , which is the substrate concentration of the feeding solution and M_S , which is the C-molar mass of the substrate used.

$$r_{CO_2} = \frac{F_{nAIR} + F_{nO_2}}{V_{nM}} \cdot (x_{CGout} - x_{CGin}) \quad (1)$$

$$r_{O_2} = \frac{F_{nAIR} + F_{nO_2}}{V_{nM}} \cdot (x_{OGout} - x_{OGin}) \quad (2)$$

$$r_S = -\frac{F_R \cdot c_S}{M_S} \quad (3)$$

3.2. Unknown rate calculation

Essentially the material balance system can be compactly described in matrix form according to eq. 4. E is the Elemental Composition Matrix where the rows represent the element contents (e.g. carbon, hydrogen, oxygen, nitrogen) and the columns represent the metabolic compounds in the process (e.g. biomass, substrate, metabolic products). r is the corresponding vector of reaction rates. For the further analysis E and r are divided into measured rates r_m with the corresponding subset E_m and non-measured unknown rates r_c with E_c .

$$Er = E_m r_m + E_c r_c = 0 \quad (4)$$

The addition of the generalized degree of reduction balance (DoR) described in Roels (1983) to the equation system leads to an overdetermined system. This enables besides the calculation of the unknown rates also a data reconciliation approach where the accuracy of the calculations can be increased in a least squares manner. For the use case of growth rate estimation in *Escherichia coli* fermentations eq. 5 was derived as a special form of eq. 4 including the C-molar carbon balance as well as the DoR balance yielding in an overdetermined system for growth rate estimation.

$$E \cdot r = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 4 & -4 & 4.113 & 0 \end{bmatrix} \cdot \begin{bmatrix} r_S \\ r_{O_2} \\ r_X \\ r_{CO_2} \end{bmatrix} = 0 \quad (5)$$

Using the Moore-Penrose pseudoinverse of E_c , which can be calculated by $E_c^+ = (E_c^T E_c)^{-1} E_c^T$, the least-squares fit of the unknown reaction rate vector is given by:

$$r_c = E_c^+ E_m r_m \quad (6)$$

3.3. Data reconciliation

Before calculating the unknown rate vector r_c from r_m with eq. 7 a reconciliation procedure can be applied to check the data consistency and detect gross errors. Inserting eq. 7 into eq. 4 leads to eq. 8 with R being the redundancy matrix $R = E_m - E_c(E_c^T E_c)^{-1} E_c^T E_m$ and R_r the reduced redundancy matrix containing only the independent rows of R .

$$Rr_m = 0 \quad (7)$$

The optimal estimate for measured rates \hat{r}_m under rate uncertainty is then formulated with F being the error variance covariance matrix and P being the variance covariance matrix of the residuals

described by $P = R_r F R_r^T$. I is the identity matrix. The detailed derivation of eq. 8 is described in Stephanopoulos et al. (1998).

$$\hat{r}_m = (I - F R_r^T P^{-1} R_r) r_m \quad (8)$$

The test function h computed in eq. 9 provides information about the statistical significance of the result, where h represents the sum of weighted squares of the residual vector $\varepsilon = R_r r_m$. The h -value is computed before the actual data reconciliation and can be evaluated against the χ^2 distribution at the degrees of freedom of R to detect systematic errors in the data to a certain level of confidence as described in Wang and Stephanopoulos (1983).

$$h = \varepsilon^T P^{-1} \varepsilon \quad (9)$$

3.4. Dynamic propagation of uncertainties vs. static error assumption

The definition of the true error variance covariance matrix F reflecting the errors accurately is not trivial. Often, the matrix is assumed to be diagonal with static relative error assumptions meaning that the measurement errors are uncorrelated. However, in many cases the uncertainties are actually correlated to each other resulting in a non-diagonal covariance matrix. A more accurate description of F can be achieved by propagation of uncertainties from the prime variables, which are in this case the inputs substrate mass flow rate F_R , the aeration rates for air F_{nAIR} and oxygen F_{nO2} and the offgas measurements x_{OGout} and x_{CGout} . After computing the Jacobian matrix J of the vector valued function, which describes the mapping from prime variables to the measured rate vector r_m , the first order propagation of uncertainty can be evaluated as shown in Madron et al. (1977):

$$F = J F_{in} J^T \quad (10)$$

with F_{in} being a diagonal error covariance matrix of the input prime variables mentioned above. The uncertainties of the prime variables were derived from the specification data of the equipment used, in particular the balances with an error of 0.1 g, the mass flow controllers with an error of $0.9 \text{ L h}^{-1} + 0.005x$ and offgas analysers with errors for O_2 of $0.1\% + 0.03x$ and CO_2 of $0.025\% + 0.03x$. x denotes the measurement value. In general, Jacobian calculation strongly depends on model correctness and is sensitive to model uncertainties. Here, the model structure is solely based on first principles with the parameters V_{nM} and M_S being well defined and the only source of parameter uncertainty is the substrate concentration of the feed c_S (chap. 3.1). c_S should be accurately measured before fermentation to ensure proper Jacobian calculation.

4. Results & Discussion

Both soft sensor approaches, that were described in chapter 3.4, have been compared on the data of a typical fed batch fermentation with batch phase, uninduced exponential feed and constant feed rate after induction with IPTG at 1.0 d. Fig. 1 shows the reconciled rate estimates for static errors (left) and dynamic error propagation (right). It can be observed, that the growth rate r_X is shifting to negative values utilizing error propagation especially in the late production phase after induction, although the noise level is considerably larger than under static error consideration. In fig. 2 the process results are shown including the biomass measurements and estimates obtained by integration of the growth rate r_X for both approaches. While the biomass estimates in the early growth phase do not differ significantly from each other, they start to deviate later in the production phase where the static error assumption leads to a major discrepancy of estimation and measured biomass with a NRMSE of 10.8% with respect to the DCW as the reference data. However, calculating the biomass under dynamic error propagation leads to more accurate estimates especially in the late production phase and a decrease of NRMSE to 5.3% corresponding to an increase of

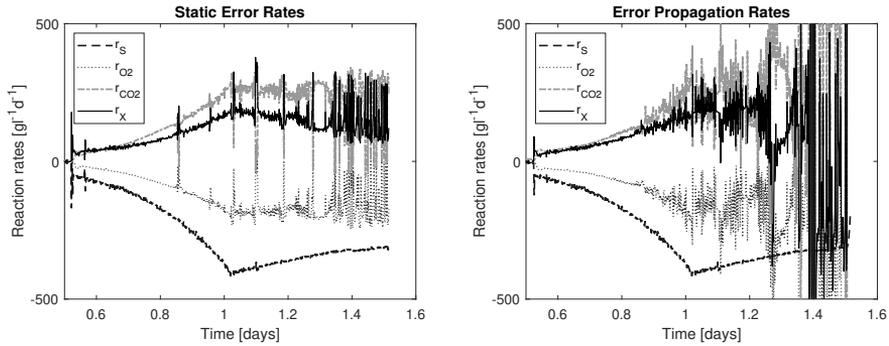


Figure 1: Reconciled rates are shown for both approaches from the start of the fed-batch phase.

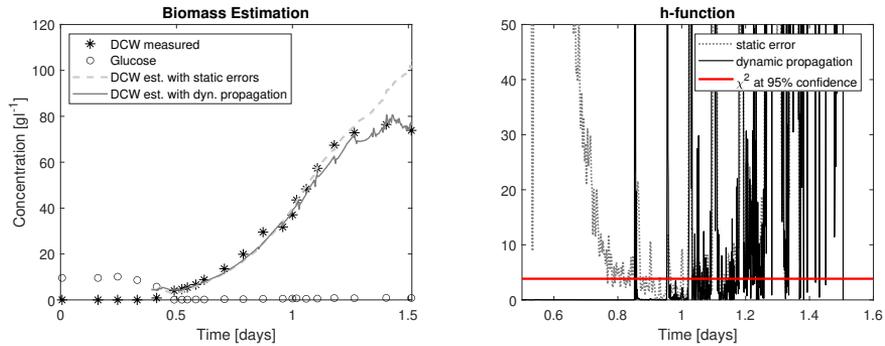


Figure 2: The left plot shows the measured concentrations of DCW and glucose as well as the biomass estimates with static error assumption (NRMSE = 10.8%) and dynamic error propagation (NRMSE = 5.3%). Biomass estimates are obtained by numerical integration of r_x . The right plot shows the value of the h-function for both approaches and the critical χ^2 distribution of 3.84 at a 95% confidence level for a degree of freedom of 1.

Table 1: The error variance covariance matrices F [mol d^{-1}] are shown for both approaches at a process time of 1.5 d.

Static Errors	Dynamic Error Propagation
$F = \begin{bmatrix} 0.485 & 0 & 0 \\ 0 & 0.301 & 0 \\ 0 & 0 & 0.273 \end{bmatrix}$	$F = \begin{bmatrix} 2.6 \times 10^{-5} & 0 & 0 \\ 0 & 3.686 & -3.623 \\ 0 & -3.623 & 3.728 \end{bmatrix}$

accuracy over 50%, which is similar to the improvements observed in Steinwandter et al. (2017). Fig. 2 also shows the h-function test described in chap. 3.3. Before induction the high static error h-values above the critical χ^2 indicate inconsistencies in between the balances. By incorporating dynamic error propagation the h-value is well below the critical χ^2 indicating a consistent closing of the balances. This is probably because the variance covariance matrix F reflects the true errors much better. A snapshot of both F matrices at the end of the process after 1.5 d is shown in tab. 4. Obviously, assuming static errors, F is just a diagonal matrix with scaled versions of the values itself, whereas covariances are present with error propagation. Here, the variance of r_s is very low

($2.6 \times 10^{-5} \text{ mold}^{-1}$) compared to r_{O_2} and r_{CO_2} , such that r_{O_2} and r_{CO_2} are much more influenced by the data reconciliation than r_S .

5. Conclusion & Outlook

The goal of this contribution was to investigate, whether the dynamic propagation of measurement errors has a positive impact on the accuracy of data reconciliation in a microbial fed-batch process. We propose the usage of dynamic error propagation from prime variables in order to estimate the reaction rates. The effectiveness of the data reconciliation procedure is increased as the error variance covariance matrix reflects the true variances more accurately leading to better estimates of the rates and biomass concentration. The prerequisite for that is, that the uncertainties of the prime variables are known, which could be derived from the sensor and equipment specification data. However, the soft sensor presented here has some limitations, that should be addressed. One major limitation is the assumption of steady state substrate concentration, meaning that the whole substrate feed is assumed to be taken up by the cells instantaneously neglecting potential substrate accumulation. This steady state assumption holds well enough, when the specific substrate uptake rate q_S is significantly below the maximum substrate uptake capacity q_{Smax} and the cell growth is substrate limited. In the case of substrate excess, where cells take up substrate near at q_{Smax} , the C-balance does not close properly anymore and the rate estimates become inaccurate. Therefore, the soft sensor is not applicable to the batch phase, to high feed pulses or whenever overfeeding occurs. Furthermore, the soft-sensor is not considering cell death leading to overestimation of cell mass especially in late production phase. Despite those limitations, we believe that elemental balancing is a solid foundation for soft sensors in bioprocessing. The generic first principle nature of elemental balancing improves transferability and generalizability of the algorithms. Limitations as mentioned above should be tackled by targeted and reasonable extensions of the elemental balance system with reaction kinetics and model based state estimation techniques.

References

- M. P. DeLisa, J. Li, G. Rao, W. A. Weigand, W. E. Bentley, Oct. 1999. Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. *Biotechnology and Bioengineering* 65 (1), 54–64.
- R. Hausmann, M. Henkel, F. Hecker, B. Hitzmann, 2017. Present Status of Automation for Industrial Bioprocesses. In: *Current Developments in Biotechnology and Bioengineering*. Elsevier, pp. 725–757, imp, DOI: 10.1016/B978-0-444-63663-8.00025-2.
- R. Luttmann, D. G. Bracewell, G. Cornelissen, K. V. Gernaey, J. Glassey, V. C. Hass, C. Kaiser, C. Preusse, G. Striedner, C.-F. Mandenius, Aug. 2012. Soft sensors in bioprocessing: A status report and recommendations. *Biotechnology Journal* 7 (8), 1040–1048, roam.
- F. Madron, V. Veverka, V. Vaněček, Jul. 1977. Statistical analysis of material balance of a chemical reactor. *AIChE Journal* 23 (4), 482–486.
- H. Narayanan, M. F. Luna, M. Stosch, M. N. Cruz Bournazou, G. Polotti, M. Morbidelli, A. Butté, M. Sokolov, Jan. 2020. Bioprocessing in the Digital Age: The Role of Process Models. *Biotechnology Journal* 15 (1), 1900172.
- J. A. Roels, 1983. *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press.
- P. Sagmeister, P. Wechselberger, M. Jazini, A. Meitz, T. Langemann, C. Herwig, Jun. 2013. Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. *Chemical Engineering Science* 96, 190–198.
- A. Savitzky, M. J. E. Golay, Jul. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36 (8), 1627–1639, publisher: American Chemical Society.
- V. Steinwandter, T. Zahel, P. Sagmeister, C. Herwig, Jan. 2017. Propagation of measurement accuracy to biomass soft-sensor estimation and control quality. *Analytical and Bioanalytical Chemistry* 409 (3), 693–706, imp.
- G. N. Stephanopoulos, A. A. Aristidou, J. Nielsen, 1998. *Metabolic Engineering*. Elsevier.
- N. S. Wang, G. Stephanopoulos, Sep. 1983. Application of macroscopic balances to the identification of gross measurement errors. *Biotechnology and Bioengineering* 25 (9), 2177–2208.
- P. Wechselberger, P. Sagmeister, C. Herwig, Sep. 2013. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess and Biosystems Engineering* 36 (9), 1205–1218.
- X. Zhu, K. U. Rehman, B. Wang, M. Shahzad, Mar. 2020. Modern Soft-Sensing Modeling Methods for Fermentation Processes. *Sensors* 20 (6), 1771.

In-silico Formulation of Iterative Learning Control for Chromatographic Purification of Biopharmaceuticals

Daniel Espinoza^a, Niklas Andersson^a and Bernt Nilsson^{a,*}

^a*Lund University Department of Chemical Engineering, Naturvetarvägen 14, Lund 221 00, Sweden*
bernt.nilsson@chemeng.lth.se

Abstract

This study presents a novel application of multivariate iterative learning control (ILC) to the continuous, chromatographic ion-exchange purification of biopharmaceuticals for consistent and reliable production. The ILC algorithm used was based on a linearized model of the retention times of two compounds as functions of the starting and ending values of a linear salt gradient. Simulations of an ion-exchange purification process were used to identify the non-linear model behavior and thus suggest an appropriate linearization of the model for use in the ILC. Two control configurations were compared: one using direct inversion of the resulting linear model, and another using a least-squares, quadratic-criterion objective function for optimal control in conjunction with the model. The result was an ILC configuration based on a simple model with parameters that only required 3 experiments to compute, that was capable of controlling the retention times of two compounds simultaneously. This leads to more reliable and flexible operation of continuous and integrated biopharmaceutical purification in the future, and serves as a foundation for further development of other ILC-based control strategies within biopharmaceutical purification.

Keywords: preparative chromatography, ion-exchange, simulation, iterative learning control, model-based control

1. Introduction

Production of biopharmaceuticals has traditionally been performed in batch mode. However, as societal pressures for cheaper pharmaceuticals increase along with the need for faster development of new pharmaceuticals, research within the field has shifted towards development of flexible, integrated and continuous manufacturing processes that allow for consistent product quality (Zydney, 2015; Gronemeyer et al., 2014; Jungbauer, 2013). The purification, or downstream processing, of biopharmaceutical manufacture is of particular importance due to the high requirements for purity in the finished pharmaceutical, as well as the desire to recover as much of the desired product as possible. This has given rise to a need for monitoring and control during continuous downstream processing, so that a desired yield and purity of the biopharmaceutical can be maintained.

The core of many downstream processes for biopharmaceutical production is chromatography. While it is inherently a batch process, chromatography has been adapted for continuous and integrated processing in several ways, often operating in parallel with other chromatography columns in a cyclic fashion (Jungbauer, 2013). This cyclic batch-wise operation makes iterative learning control (ILC) a prime candidate for control of chromatography steps in continuous downstream

processes. Hailing from the field of robotics, ILC is a control technique that is used to improve the operation of processes that repeat over time based on the information gathered from previous process runs (Moore, 1993). It has been successfully applied to control the temperature trajectory of batch reactors (Lee and Lee, 2007), and more recently to single-input-single-output (SISO) control of the retention time of a protein in an ion-exchange (IEX) chromatography column (Holmqvist and Sellberg, 2016). In this latter study, an approach based on direct model inversion was used in which the retention time was modelled as a function of the time duration of the elution step, where the separation of the protein from other molecules takes place. Finally, another direct-inversion SISO ILC application has been done on the loading of columns in periodic counter-current (PCC) chromatography (Löfgren et al., 2020).

When separating a target protein from a known impurity by means of ion-exchange, it would be useful to control the retention times of both the target compound and the impurity. This increases the magnitude of the problem to a multiple-output system with increased complexity, which requires a careful choice of model to obtain the desired control. The purpose of this study was to use in-silico simulation of IEX chromatography to develop a control approach based on a linear model of the retention times of two proteins as a function of a set of process inputs. The result was a multiple-input-multiple-output (MIMO) ILC that made use of a quadratic-criterion objective function in conjunction with a linear map from beginning and ending values of a linear salt gradient, to the retention times of the two proteins. The choice of ILC is motivated by the repetitious nature of chromatographic purification steps, and by that ILC hinges on the process returning to its initial conditions between each control action (Longman, 2000), which in this study was achieved by the regeneration and equilibration steps following the elution. The efficacy of the ILC was demonstrated by running a simulated sequence of chromatography processes and letting the ILC correct for a simulated disturbance to the process inputs. To the authors' knowledge, application of MIMO ILC to chromatographic purification of biopharmaceuticals has not previously been presented.

In the following section, a brief background of the IEX chromatography process is provided along with the strategy for formulation of a linear process model by means of in-silico simulation. This section is followed by a description of the ILC algorithm and the least-squares, quadratic-criterion objective function used in conjunction with the model, along with the results of the simulated and controlled sequence of chromatography processes. Finally, some concluding remarks and suggestions for future studies will be made.

2. Simulations and model formulation

In liquid chromatography, substances in a liquid, mobile phase, are separated based on interactions with a solid, stationary phase. In ion-exchange chromatography, these interactions take place between ionically charged substances in the mobile phase and ligands on the stationary phase. Ion-exchange liquid chromatography with linear gradient elution is commonly performed in a five step process, as illustrated by Figure 1. In the first step, the load step, a certain volume of a mixture of a product and its impurities is loaded onto a chromatography column. The product binds to the stationary phase along with any other ionically charged impurities. Any substance that does not bind to the column is washed out in the wash step by a low salt concentration buffer solution, hereafter called buffer A. In the elution step, the column is loaded with a mixture of buffer A and buffer B, which is a solution with a high salt concentration. The percentage of buffer B in the mobile phase is increased linearly over the duration of the elution step, and as the concentration of salt in the column increases, the molecules bound to the column release and are eluted at different rates. The result is that the substances in the mixture have different retention times. Following the elution step, the column is regenerated by loading it with 100% buffer B and then equilibrated with 100% buffer A, to ensure that the column is restored to its initial conditions before the next sample is loaded and the process can be repeated.

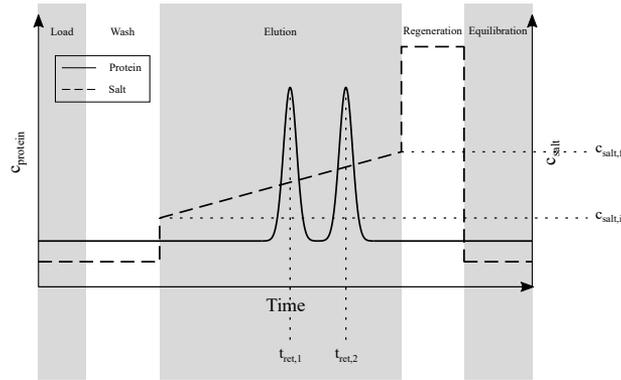


Figure 1: The ion-exchange purification process and its steps.

In an elution step with fixed time duration, the initial and final salt concentrations of the gradient are what determine the retention times of the substances. In a real chromatography process, the salt concentration during the elution step is controlled by two pumps, one each for buffer A and B, and the percentage of the flowrate that is driven by each pump. The initial and final percentages of buffer B, $x_{B,i}$ and $x_{B,f}$, are thus directly manipulable values by the process and were chosen as the control signals for this study. To control the retention times of the two compounds, $t_{ret,1}$ and $t_{ret,2}$, using model-based ILC, a model using $x_{B,i}$ and $x_{B,f}$ as inputs was needed.

To evaluate the relation between the retention times and $x_{B,i}$ and $x_{B,f}$, response surfaces were generated in-silico by modelling and simulating the mass transfer in a chromatography column. The mass transfer along the column in the z -direction can be described with a convection-dispersion model along with an adsorption term:

$$\frac{\partial c_i}{\partial t} = -\frac{F}{A\varepsilon} \frac{\partial c_i}{\partial z} + D_{app} \frac{\partial^2 c_i}{\partial z^2} - \frac{(1-\varepsilon_c)}{\varepsilon} \frac{\partial q_i}{\partial t} \quad (1)$$

where c_i is the concentration of component i in the mobile phase, F is the volumetric flowrate of the mobile phase, A is the cross-section area of the column and D_{app} is the apparent dispersion coefficient. ε and ε_c are the total porosity and column void fraction, respectively, and q_i is the concentration of component i in the stationary phase. The change in q_i can, in turn, be modelled by means of the Langmuir mobile phase modulator:

$$\frac{\partial q_i}{\partial t} = k_{kin,i} \left(B_i c_i c_s^{-\beta_i} \left(1 - \sum_{j=1}^N \frac{q_j}{q_{max,j}} \right) - q_i \right) \quad (2)$$

where β_i is a parameter that describes the ion-exchange characteristics, c_s is the salt concentration in the mobile phase, q_j is the concentration of compound j in the stationary phase and $q_{max,j}$ is its maximum possible concentration in the stationary phase. $k_{kin,i}$ and B_i are lumped parameters that describe the adsorption kinetics and equilibria. (Karlsson et al., 2004)

The response surfaces were generated by creating lists of $x_{B,i}$ and $x_{B,f}$ ranging from 0 to 100 and recalculating them to salt concentrations c_s . The above model was then simulated for each combination of initial and final salt concentrations. The simulations were carried out in Python by discretizing the space dimension (the length of the column) into 100 finite volumes, thus reformulating the partial-differential model into a system of ordinary differential equations (ODEs) using the finite-volume method. The second-order derivative in space, i.e., the dispersion term, was discretized using a 3-point central derivative approximation, while the first-order derivative in

space, i.e., the convection term, was discretized using a 2-step backwards derivative approximation (Shyy, 1994). The system of ODEs was then solved using the backward-differential formula (BDF) method in the Python SciPy library. The retention times were found by taking the solution at the column outlet and finding the maximum values of the peaks using the first and second order derivatives of the concentration data (Felinger, 1998).

As the response surfaces in Figure 2 show, the retention volumes for both peaks decreased as both $x_{B,i}$ and $x_{B,f}$ increased. In particular, the response surfaces showed that the model could be linearized in the zone $0 \leq x_{B,i} \leq 40$ and $20 \leq x_{B,f} \leq 100$ and yield marginal errors. Thus, the model was linearized around a set of inputs $\mathbf{x}_B^d = [x_{B,i}^d, x_{B,f}^d]^T$ that yielded a desired separation between peaks, by means of a finite-difference derivative Jacobian matrix:

$$\mathbf{G} = \frac{\partial \mathbf{F}(\mathbf{u})}{\partial \mathbf{u}} = \begin{bmatrix} \frac{t_{ret,1}(\mathbf{x}_B^{pert,1}) - t_{ret1}(\mathbf{x}_B^d)}{\varepsilon} & \frac{t_{ret,1}(\mathbf{x}_B^{pert,2}) - t_{ret1}(\mathbf{x}_B^d)}{\varepsilon} \\ \frac{t_{ret,2}(\mathbf{x}_B^{pert,1}) - t_{ret2}(\mathbf{x}_B^d)}{\varepsilon} & \frac{t_{ret,2}(\mathbf{x}_B^{pert,2}) - t_{ret2}(\mathbf{x}_B^d)}{\varepsilon} \end{bmatrix} \quad (3)$$

where \mathbf{F} denotes the non-linear model and \mathbf{G} denotes the linearized model as a 2×2 matrix. $\mathbf{x}_B^{pert,1}$ and $\mathbf{x}_B^{pert,2}$ each denote a set of inputs \mathbf{x}_B where either $x_{B,i}$ or $x_{B,f}$ are perturbed by a small value ε . Smaller values of ε result in a locally more accurate \mathbf{G} , but depending on the variations in slope of the function \mathbf{F} this may not be desirable. In this case, as seen in Figure 2, the slope is almost linear in all directions in the zone of interest and thus the choice of ε was assumed to be less important. For the controller as detailed in the next section, \mathbf{G} was computed around the point $x_{B,i} = 10$, $x_{B,f} = 30$ for $\varepsilon = 3$, which was deemed small enough since $x_{B,i}$ and $x_{B,f}$ range from 0 to 100. The retention times at this point were 19.7 and 25.8 minutes for the left and right peak, respectively, and were used as set points for the controller.

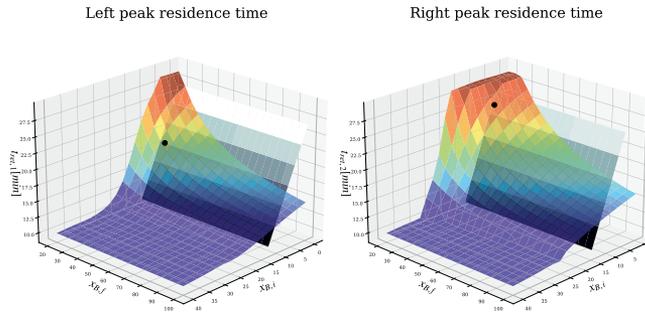


Figure 2: Response surfaces (jet) from the simulations of the chromatography process. The retention times of the first (left) and second (right) peak are plotted as functions of the initial ($x_{B,i}$) and final ($x_{B,f}$) proportions of buffer B in the linear gradient, along with the linearization (grey) around the point $x_{B,i} = 10$, $x_{B,f} = 30$ (black).

3. Control algorithm

In order to maintain a desired process output \mathbf{y}_d in a repeating process, the following ILC algorithm can be used (Lee and Lee, 2007; Holmqvist and Sellberg, 2016):

$$\mathbf{u}_{k+1} - \mathbf{u}_k = \mathbf{K} \mathbf{e}_k \quad (4)$$

where the control inputs \mathbf{u} on process cycle k are updated for the following cycle $k + 1$ based on the error $\mathbf{e}_k = \mathbf{y}_d - \mathbf{y}_k$, i.e., how the process output \mathbf{y} deviates from the desired output. \mathbf{K} , in turn,

is a learning filter that can be designed in different ways depending on the system to be controlled. In this study, the learning filter was designed using the model-based approach detailed by Lee and Lee (2007), in which the following quadratic-criterion objective function is used:

$$\min_{\mathbf{u}_{k+1}} \|\mathbf{e}_{k+1}\|_{\mathbf{Q}}^2 + \|\Delta\mathbf{u}_{k+1}\|_{\mathbf{R}}^2 \tag{5}$$

The objective function has a penalty term for changes in process input on cycle $k + 1$, and for process errors \mathbf{e}_k , weighted by the positive-definite matrices \mathbf{Q} and \mathbf{R} . The solution to the objective function becomes:

$$\mathbf{K} = [\mathbf{G}^T \mathbf{Q} \mathbf{G} + \mathbf{R}]^{-1} \mathbf{G}^T \mathbf{Q} \tag{6}$$

The benefit of using quadratic-criterion-based ILC, or Q-ILC, is that it becomes possible to apply damping to the control action by penalizing changes to the process input, which in turn allows for robust control of processes with less precise model estimations or with disturbances to the process dynamics and outputs. For the purposes of this study, \mathbf{Q} was selected to a 2x2 identity matrix \mathbf{I} , and \mathbf{R} to $r\mathbf{I}$, where r was a scalar value. With this selection of weighting matrices, the damping of the controller could be increased by increasing the value of r , while the special case $r = 0$ resulted in a direct inversion controller, $\mathbf{K} = \mathbf{G}^{-1}$. This latter choice would result in a complete elimination of the process error on the following cycle for perfect process models (Lee and Lee, 2007).

To test the proposed process control configuration, a sequence of 12 chromatography cycles were simulated. On cycle 3, a disturbance to the process inputs was introduced by increasing the salt concentration in buffer B by 10%. The sequence was run with both direct inversion ($r = 0$) and Q-ILC ($r = 0.5$). The results, displayed in Figure 3, show that the direct inversion approach failed to correct for the error on cycle 4, instead overshooting the set points. It failed to correct for the overshoot perfectly on cycle 5 as well and undershot them, only landing stably on the set points on cycle 6 and onwards. Even in an in-silico study with perfect repeatability, the direct inversion controller proved to be slightly unstable due to the imperfect process model, which caused this "bouncing" behavior.

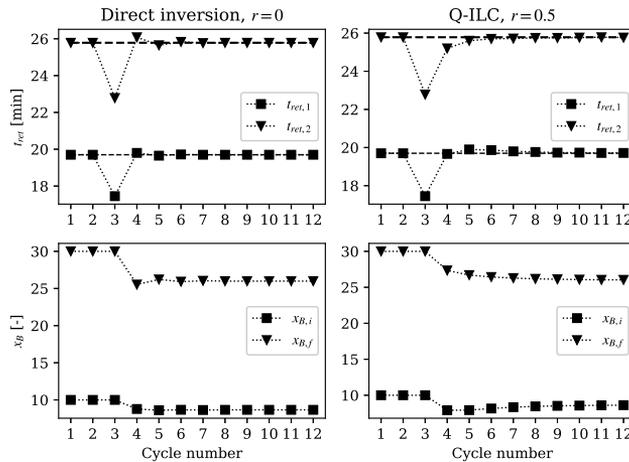


Figure 3: The 12-cycle controlled sequence. The direct inversion (left) sequence showed an overshoot when correcting for the disturbance on cycle 3, stemming from an imperfect process model estimation. The Q-ILC approach (right) displayed no overshoot of the first peak, but instead showed coupling behavior between the two peaks.

Q-ILC with $r = 0.5$, on the other hand, proved to avoid the overshoot of the first peak, but took 2 cycles to restore to the setpoint. Additionally, it overshoot the second peak on cycles 5 and 6. This is likely due to coupling between the two retention times, and could be remedied by some form of decoupling, i.e., by reparametrizing the control inputs into decoupled control inputs by means of a decoupling matrix, \mathbf{D} (Waller et al., 2003). If \mathbf{D} is chosen carefully, this would make it possible to adjust each peak individually without shifting the other. In a real case, Q-ILC would be useful due to variations in the process dynamic caused by ambient conditions or human error, or to compensate for stochasticity in the outputs caused by the UV detectors used to measure the presence of biopharmaceuticals and impurities at the outlet of the column. Thus, decoupling would also be crucial in a real case.

4. Concluding remarks

The successful application of multivariate, model-based ILC using a quadratic-criterion objective function to a simulated chromatographic separation contributes to the field of biopharmaceutical purification by providing a method of maintaining the desired separation when disturbances are applied to the process. Of particular note is the simple model estimation, which resulted in a functioning controller while requiring only 3 experiments to compute. For the purposes of this study, a single model estimation was performed with three simulated experiments. This single model estimation may be particularly suitable for repetitious chromatography steps in continuous downstream processes, where the goal is to improve the consistency of product quality and thus any disturbances to the system can be expected to happen slowly and in the long term. However, the model-based approach can be modified to fit any desired process and its dynamics, for example, non-linear dynamics or additional process inputs/outputs. Thus, the proposed control approach provides a solid starting point for control of other cyclically repeating process operations in biopharmaceutical purification. To further examine the efficacy of the Q-ILC, practical studies are required. Thus, the application of this controller to a real process is a natural next step.

References

- A. Felinger, 1998. Data Analysis and Signal Processing in Chromatography. Data Handling in Science and Technology. Elsevier.
- P. Gronemeyer, R. Ditz, J. Strube, 2014. Trends in upstream and downstream process development for antibody manufacturing. *Bioengineering* 1, 188–212.
- A. Holmqvist, A. Sellberg, 2016. A generic pat software interface for on-line monitoring and control of chromatographic separation systems. *Computer Aided Chemical Engineering* 38, 811–816.
- A. Jungbauer, 2013. Continuous downstream processing of biopharmaceuticals. *Trends in Biotechnology* 31 (8), 479–492.
- D. Karlsson, N. Jakobsson, A. Axelsson, B. Nilsson, 2004. Model-based optimization of a preparative ion-exchange step for antibody purification. *Journal of Chromatography A* 1005 (1-2), 29–39.
- J. H. Lee, K. S. Lee, 2007. Iterative learning control applied to batch processes: An overview. *Control Engineering Practice* 15 (10), 1306–1318.
- R. W. Longman, 2000. Iterative learning control and repetitive control for engineering practice. *International Journal of Control* 73 (10), 930–954.
- A. Löfgren, J. Gomis-Fons, N. Andersson, B. Nilsson, L. Berghard, C. L. Häggglund, 2020. An integrated continuous downstream process with real-time control: A case study with periodic countercurrent chromatography and continuous virus inactivation. *Biotechnology and Bioengineering* 118 (4), 1645–1657.
- K. L. Moore, 1993. *Iterative Learning Control for Deterministic Systems*. Advances in Industrial Control. Springer-Verlag London Limited.
- W. Shyy, 1994. *Computational Modeling for Fluid Flow and Interfacial Transport*, 1st Edition. Transport Processes in Engineering. Elsevier.
- M. Waller, J. B. Waller, K. V. Waller, 2003. Decoupling revisited. *Industrial & Engineering Chemistry Research* 42 (20), 4575–4577.
- A. L. Zydny, 2015. Perspectives on integrated continuous bioprocessing - opportunities and challenges. *Current Opinion in Chemical Engineering* 10, 8–13.

Application of Real-Time Optimization with Modifier Adaptation to the Reactive Extrusion of Hydrophobically Modified Ethoxylated Urethanes

Maximilian Cegla^{a,*} and Sebastian Engell^a

^a*TU Dortmund University, Biochemical and Chemical Engineering, Process Dynamics and Operations Group, Emil-Figge-Straße 70, 44227 Dortmund, Germany*
E-mail: maximilian.cegla@tu-dortmund.de

Abstract

This work presents the application of modifier adaptation with quadratic approximation (MAWQA) to the production process of hydrophobically modified ethoxylated urethanes (HEUR) by reactive extrusion in a simulation study. A plant-model mismatch in the conveying rate in the twin-screw extruder model as well as a mismatch in the reaction kinetics are investigated. These uncertainties result from a limited availability of experiments for the chemical system and the extrusion system during the up-scaling to production size. It is shown that the application of MAWQA can drive the plant to the true economic optimum after about 25 set-points and satisfy the constraints for both cases even under the presence of noise in the inputs and measurements. These 25 set-point changes take about 2 hours which makes this method applicable for frequent product changes.

Keywords: RTO, MAWQA, Reactive Extrusion, Model-Based Control, Plant-Model Mismatch

1. Introduction

Hydrophobically modified ethoxylated urethanes (HEUR) are a group of polymers mainly used as rheology modifiers in paints and coatings as reported by Reuvers (1999). These rheology modifiers are conventionally produced in batches of several cubic meters and directly formulated with water within the same vessel. This production process takes multiple hours per batch. Within the SIMPLIFY project, the transition to a continuous production of these paint thickeners by reactive extrusion on a twin-screw extruder is investigated, which offer numerous advantages. This transition significantly reduces the time and costs associated with cleans and allows a fully electrified production using renewable energy sources. To operate the process at the economic and ecological optimum, an advanced process control strategy is required. Applying classical model-based control methods such as conventional real-time optimization would lead to suboptimal performance as the optimal set-point of the model does not coincide with the optimal point of the real process due to the plant-model mismatch. This plant-model mismatch is caused due to the limited availability of experiments to determine the chemical and rheological system. Especially relevant in the case of the production of HEUR is the high sensitivity to traces of moisture in the reaction system, decreasing the reproducibility of the experiments. Considering the model of the extrusion system, errors can occur during scale-up to a bigger extruder. To overcome this mismatch, we suggest the application of modifier adaptation with quadratic approximation (MAWQA) for the reactive extrusion process. This method extends the method of iterative optimization using plant and model gradients Gao and Engell (2005) and Marchetti et al. (2009) by a quadratic approximation of the cost and constraint functions (Gao et al. (2016)). This extension increases the robustness for noisy

process data and simplifies the estimation of real plant gradients in practice. An example for a successful implementation of MAWQA in an industrial environment was reported by Gottu Mukkula et al. (2020). This method is suitable for the reactive extrusion process as the main requirements are met: The process can be safely probed and exited, reaches steady state within five minutes and the cost and constraint function can be measured. For the production of paint-thickeners, the off-spec product produced during the probing phase can be blended in other products minimizing waste. In the following the model for the twin-screw extruder, the chemical system and the applied control method are presented. Afterwards the application of this method to the case of a plant-model mismatch in the reaction model and a mismatch in the extruder model are investigated.

2. Models and Method

2.1. Twin-Screw Extruder Model

The extruder model used in this work is based on a mechanistic twin-screw extruder model developed by Eitzlmayr et al. (2014). The approach has been extended by a singular perturbation approach for the pressure description by Cegla and Engell (2021) so that it can be used in efficient optimization algorithms. This model accounts for different flow patterns within the extruder caused by different screw geometries. It discretizes the extruder into finite volumes in which the polymerization reaction is occurring, in this case 29. The finite volumes are connected by internal flows caused by a pressure difference or by the rotation and geometry of the screw. The conveying capacity for a given screw element must either be determined experimentally or approximated from the knowledge of their geometry and therefore is prone to errors. With the progressing polymerization reaction in the extruder, the viscosity of the processed material increases, amplifying the energy dissipation and affecting the internal flows. Due to this strong coupling of the states, the model is highly non-linear. Furthermore, due to the high complexity of the switching of the model when a section is entirely filled, it is not possible to transform the model into a steady state model. Therefore, the steady state is computed by simulation of the model over a long time horizon. By the nature of the model, input multiplicities are not possible. The extruder assumed in this work is a Leistritz 18 mm ZSE MAXX extruder with a length to diameter ratio of 60. The screw concept is consisting of double flighted conveying elements with a pitch of 20mm and a cylindrical die with a diameter of 6 mm and a length of 40mm.

2.2. Kinetic Model

For the production polyurethane paint thickeners, initially a long chained polyethylene glycol (PEG) is reacted with an aliphatic diisocyanate (ISO). Further along the extruder, a short chained mono alcohol (A) is fed to stop the reaction and give the product the hydrophilic properties by end-capping of the formed urethanes. The reaction kinetics are approximated by a differential equation for the weight average molecular weight M_W similar to Verhoeven et al. (2004) and extended by an inhibition term by the reacted mono alcohol A_R . The parameter $M_{w,rep}$ is the molecular weight of a polymer repetition unit and is used to describe the stoichiometry of the end-capping. The maximum achievable molecular weight for a given ratio between ISO and PEG is given by $M_{w,max}$ and can be calculated from the reaction stoichiometry. The parameter k_{inhib} quantifies the effect of reacted mono alcohol on the molecular weight. Transurethanisation is not observed in this temperature range. The reaction system is described by:

$$\frac{dM_w}{dt} = k_0 \cdot \exp\left(\frac{-E_A}{RT}\right) \cdot \left| \frac{M_{w,max} - M_w - [A_R] \cdot k_{inhib}}{M_{w,max} - [A_R] \cdot k_{inhib}} \right| \quad (1)$$

$$\frac{d[A]}{dt} = -\frac{d[A_R]}{dt} = [A] \cdot \left(\frac{2M_{w,rep}}{M_w} - [A_R] \right) \cdot k_{0,alc} \cdot \exp\left(\frac{-E_{A,alc}}{RT}\right) \quad (2)$$

The melt behaves quasi-Newtonian. The temperature influence of the viscosity is described by an Arrhenius approach, the influence of the molecular weight is described by a power law. The exponent 3.4 is chosen as exponent as reported in literature being applicable for linear polymers such as polyurethanes (Verhoeven et al. (2004)). Summarizing, the viscosity is described by:

$$\eta(T, M_w) = \eta_0 \cdot M_w^{3.4} \cdot \exp\left(\frac{-E_A}{RT}\right) \quad (3)$$

2.3. Modifier Adaptation with Quadratic Approximation

The basis for the MAWQA method is the extension of the steady state optimization problem by contributions for the differences of the plant and model gradients multiplied by the change in the input u for the objective function and the constraint function as proposed by Gao and Engell (2005) and Marchetti et al. (2009). F_m denotes the plant model, J the cost function, G the constraint function, y the model output, the indices m and p denote the model and plant respectively:

$$\min_u J_m^{ad,k}(\hat{y}, u) \quad (4)$$

$$s.t. \hat{y} = F_m(u) \quad (5)$$

$$G_m^{ad,k}(\hat{y}, u) \leq 0. \quad (6)$$

The adapted cost function for the model J_m^{ad} at the k -th time point is therefore:

$$J_m^{ad,k} = J_m(\hat{y}, u) + (\nabla J_p(y_p^k, u^k) - \nabla J_m(\hat{y}^k, u^k))^T (u - u^k) \quad (7)$$

and adapted constraint function for the model G_m^{ad} :

$$G_m^{ad,k} = G_m(\hat{y}, u) + G_p(y_p^k, u^k) - G_m(\hat{y}^k, u^k) + (\nabla G_p(y_p^k, u^k) - \nabla G_m(\hat{y}^k, u^k))^T (u - u^k). \quad (8)$$

To robustly determine the plant gradients and to improve the optimization, Gao et al. (2016) proposed the approximation of the plant cost and constraint functions J_m and G_m with the set of inputs \mathbf{u} by quadratic functions with the set of parameters P :

$$\min_P \sum_{i=1}^n (J_p(u^i) - J_Q(u^i, P))^2 \quad J_Q(u, P) = \sum_{i=1}^{n_u} \sum_{j=1}^{n_u} a_{ij} u_i u_j + \sum_{i=1}^{n_u} b_i u_i + c. \quad (9)$$

This two quadratic function can then be analytically derived with respect to the inputs to calculate the plant cost gradient ∇J_p and plant constraint gradient ∇G_p for equation (7) and (8). The prediction quality is described by ρ_m^k for the adjusted optimization problem (4) and by ρ_Q^k of the quadratic approximation for both the cost and constraint function for a single constraint:

$$\rho_m^k = \max \left\{ \left| 1 - \frac{J_m^k - J_M^{k-1}}{J_p^k - J_p^{k-1}} \right|, \left| 1 - \frac{G_m^k - G_m^{k-1}}{G_p^k - G_p^{k-1}} \right| \right\} \quad \rho_Q^k = \max \left\{ \left| 1 - \frac{J_Q^k - J_Q^{k-1}}{J_p^k - J_p^{k-1}} \right|, \left| 1 - \frac{G_Q^k - G_Q^{k-1}}{G_p^k - G_p^{k-1}} \right| \right\} \quad (10)$$

For the case that $\rho_Q^k < \rho_m^k$, the quadratic model is used to calculate the next input:

$$\min_u J_Q(u, P) \quad (11)$$

$$s.t. G_Q(u, P) \leq 0 \quad (12)$$

$$(u - u^k)' \cdot \text{cov}(\mathbf{u}^k) \cdot (u - u^k) \leq \gamma^2 \quad (13)$$

Equation (13) constrains the inputs to the validity range of the quadratic approximation to prevent extrapolation and can be tuned by the distance parameter γ . For the case that the new input is an exploration step and not improving the cost function, u^k will not be updated.

3. Results

3.1. Initialization of the method

The method is initialized with 12 set-points which are distributed using Latin hyper cube sampling and 3 set-points for the computation of the finite differences to get a good coverage of the search space and to have $\frac{(n_u+1)(n_u+2)}{2}$ sample points as required for the quadratic approximation. The nominal plant optimum is at a throughput of 20 kg/h, a rotation speed of 200 RPM and a barrel temperature of 132 °C with a total specific energy demand of 163.36 kJ/kg. Gaussian noise with a standard deviation of 0.25 kg/h is added to the throughput and with a standard deviation of 1.5 kg/mol on the measurement of the product molecular weight. The rotational speed and the individual barrel temperatures are assumed noise free as these are precisely controlled by feedback control. The throughput is bounded between 5 and 20 kg/h, the rotation speed between 200 and 600 RPM, and the barrel temperature between 120 and 160 °C. A combined economic and ecological cost function is used aiming at the minimization of the energy input required to produce a given quantity of product. The contributions to the energy inputs are the mechanical energy input that is provided by the motor, the heat input provided by the heating blocks and the required cooling energy provided by the cooling system of the block. The process constraint is a minimum molecular weight of the product of 55 kg/mol to fulfil the desired rheological properties of the HEUR. The validity range of the quadratic model is constrained with $\gamma = 1$ and the model is trained on the 12 inputs closest to the last-set point.

3.2. Mismatch in the extruder model

In this section, an mismatch in the extruder model is assumed additional to the noise described in the last subsection. The conveying capacity of all extruder elements is assumed to be 20% higher in the model than in the plant. A higher conveying capacity causes a shorter residence time as less material is present in the screw section and it is transported faster out of the extruder. This causes the model optimum to be very conservative in the satisfaction of the molecular weight constraint.

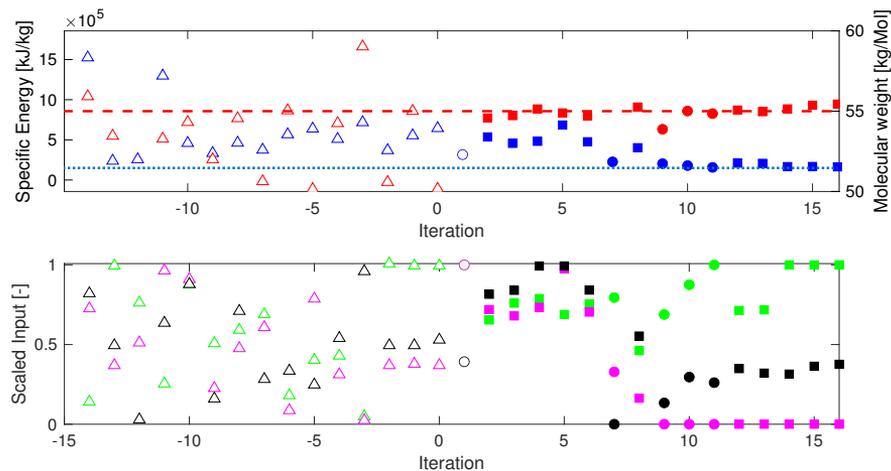


Figure 1: Results for the application of the method with an plant-model mismatch in the conveying capacity of the screw. \triangle denotes the initial probing steps, \circ denotes improvement steps and \square indicates exploration steps. The inputs are scaled within the bounds of the optimization problem. Filled symbols mark steps computed from the quadratic approximation, empty symbols mark steps based on the extended model. Red indicates the molecular weight, blue the specific energy consumption, green the throughput, magenta the rotation speed and black the barrel temperature.

The results of the application of the MAWQA method are shown in Figure 1. After the initial probing, the method uses the gradients determined by finite differences in the extended optimization problem to determine the first improving set-point. Both rotation speed and throughput are set to the upper limit while the barrel temperature is set close to the optimum. Although the cost function is improved, due to the high noise in the measurement the prediction of the finite differences for the constraint function, the resulting set-point does not satisfy the constraints. Therefore from the next iteration on, the quadratic approximation is used for optimization and significantly reduces the constraint violation. Between iterations five and ten, the method is driving the plant to its true optimum by decreasing the rotation speed from the upper limit to the lower limit. These big changes can be explained by the low sensitivity of the molecular weight and the specific energy use to the rotation speed. This low sensitivity can be explained physically as with the screw design used here, most of the residence for the reaction is generated by the back-pressure zone of the die and not in the upstream screw elements. Furthermore in a 18mm extruder, the dominant introduction of energy is via the electrical heating, especially for polymers with a high melt flow index like the investigated HEUR. At iteration 10 the true plant optimum is reached with satisfaction of the constraints. As the presented method depends on the random initial probing, the method was repeated 10 times showing similar or better performance.

3.3. Mismatch in the kinetic model

In this section, the method is applied for the case of a plant-model mismatch in the reaction kinetics. The same noise as in the first case is present. The pre-exponential factor k_0 for the polymerization reaction is assumed to be 10% lower in the model. With this mismatch, the result of a conventional steady state optimization would be very conservative as more residence time would be provided and a more viscous product would be produced, increasing the specific energy consumption. The results obtained using MAWQA for this case are shown in Figure 2.

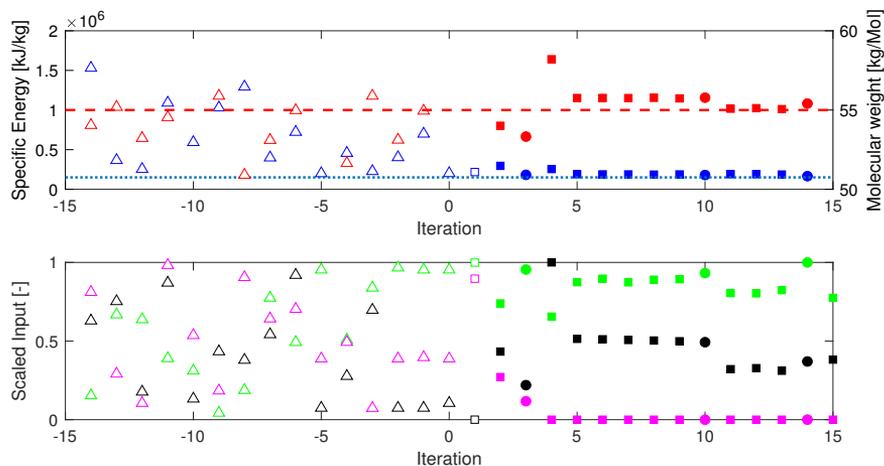


Figure 2: Results for the application of the method with an uncertainty in the reaction kinetics. Δ denotes the initial probing steps, \circ denotes improvement steps and \square indicates exploration steps. The inputs are scaled within the bounds of the optimization problem. Filled symbols mark steps computed from the quadratic approximation, empty symbols are based on the extended model. Red indicates the molecular weight, blue the specific energy consumption, green the throughput, magenta the rotation speed and black the barrel temperature.

After probing, the first evaluation is performed by solving the extended optimization problem with the gradients computed by finite differences. Similar to the other presented case, the resulting set-point results in a low specific energy use, but a very low molecular weight caused by moving the temperature to the lower bound. As in the first case, after the first iteration the model quality of the quadratic approximation was at every point superior to the extended model. In the following steps the rotation speed is decreased and the barrel temperature increased. Worth mentioning here is the fact that after the fifth iteration the constraints are fulfilled at a cost function value close to the optimum, the next improvement is the 10th iteration. Therefore, it might be worthwhile to base the evaluation process of a solution not only on the cost function but also on the constraint violation. This change would influence the range of validity of the quadratic approximation in equation (13).

4. Conclusion

In this work, the benefits of the application of MAWQA to the production of HEUR in reactive extrusion are shown. For both investigated cases of a mismatch in the structure of the model as well as a mismatch in the reaction kinetic the controller drives the plant to the true optimum after a maximum of 10 iterations. These 10 iterations and the 15 initial probed set-points in total correspond to a total duration of about two hours operation time. This makes this method suitable for a flexible production with frequent product changes. In the situation of few product changes but a discontinuous production, which is the case for a day shift production, this method can be initialized with the quadratic models of past shifts with a similar product. Using this information, the required iterations can be reduced to a minimum and no further time consuming initialization is required in the regular operation. Furthermore on a long term, the information of the quadratic models can be used to detect and to quantify effects such as wear in the production process. In our future work, we will experimentally apply the method proposed here to the real reactive extrusion system. Furthermore, different approaches for the selection of the points for the quadratic approximation and the use of more optimization variables will be investigated.

5. Acknowledgements

The work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820716 (SIMPLIFY). (www.spire2030.eu/simplify)

References

- M. Cegla, S. Engell, 2021. Application of Model Predictive Control to the reactive extrusion of ϵ -Caprolactone in a twin-screw extruder. *IFAC-PapersOnLine* 54 (3), 225–230.
- A. Eitzlmayr, G. Koscher, G. Reynolds, Z. Huang, J. Booth, P. Shering, J. Khinast, 2014. Mechanistic modeling of modular co-rotating twin-screw extruders. *International Journal of Pharmaceutics* 474 (1-2), 157–176.
- W. Gao, S. Engell, 2005. Iterative set-point optimization of batch chromatography. *Computers and Chemical Engineering* 29 (6 SPEC. ISS.), 1401–1409.
- W. Gao, S. Wenzel, S. Engell, 2016. A reliable modifier-adaptation strategy for real-time optimization. *Computers and Chemical Engineering* 91, 318–328.
- A. R. Gottu Mukkula, S. Kern, M. Salge, M. Holtkamp, S. Guhl, C. Fleicher, K. Meyer, M. P. Remelhe, M. Maiwald, S. Engell, 2020. An application of modifier adaptation with quadratic approximation on a pilot scale plant in industrial environment. *IFAC-PapersOnLine* 53 (2), 11773–11779.
- A. Marchetti, B. Chachuat, D. Bonvin, 2009. Modifier-adaptation methodology for real-time optimization. *Industrial and Engineering Chemistry Research* 48 (13), 6022–6033.
- A. J. Reuvers, 1999. Control of rheology of water-borne paints using associative thickeners. *Progress in Organic Coatings* 35 (1-4), 171–181.
- V. W. Verhoeven, M. P. Van Vondel, K. J. Ganzeveld, L. P. Janssen, 2004. Rheo-kinetic measurement of thermoplastic polyurethane polymerization in a measurement kneader. *Polymer Engineering and Science* 44 (9), 1648–1655.

32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING

VOLUME 3

COMPUTER-AIDED CHEMICAL ENGINEERING, 51

**32nd EUROPEAN SYMPOSIUM ON
COMPUTER AIDED PROCESS
ENGINEERING**

VOLUME 3

Edited by

Ludovic Montastruc

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

ludovic.montastruc@ensiacet.fr

Stephane Negny

Professor, Laboratoire de Genie Chimique

Universite de Toulouse, France

stephane.negny@ensiacet.fr



ELSEVIER

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

Elsevier

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

Copyright © 2022 Elsevier B.V. All rights reserved.

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

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

Notices

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

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

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

British Library Cataloguing in Publication Data

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

Library of Congress Cataloging-in-Publication Data

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

ISBN (Volume 3): 978-0-443-18633-2

ISBN (Set) : 978-0-323-95879-0

ISSN: 1570-7946

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



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Candice Janco

Acquisition Editor: Anita Koch

Editorial Project Manager: Lena Sparks

Production Project Manager: Paul Prasad Chandramohan

Designer: Mark Rogers

Typeset by STRAIVE

Contents

T5: Concepts, Methods and Tools

- 200. Development of a virtual platform for the metaheuristic optimization of heat exchangers**
Oscar D. Lara-Montaño, Fernando I. Gómez-Castro, Claudia Gutierrez-Antonio **1195**
- 201. Combining optimization and life cycle assessment: Design of low-carbon multi-energy systems in the SecMOD framework**
Christiane Reinert, Lars Schellhas, Julia Frohmann, Niklas Nolzen, Dominik Tillmanns, Nils Baumgärtner, Sarah Deutz, André Bardow **1201**
- 202. Robust Design of Distributed Energy Systems Within Unbalanced Power Networks**
Ishanki De Mel, Oleksiy V. Klymenko, Michael Short **1207**
- 203. SCR: A novel surrogate-based global optimization algorithm for constrained black-box problems**
S. A. Zaryab, A. Manno, E. Martelli **1213**
- 204. Comparative Life Cycle Assessment of Demand-Side Management via Operational Optimization**
Benedikt Nilges, Célia Burghardt, Kosan Roh, Christiane Reinert, Niklas von der Aßen **1219**
- 205. A rigorous synthesis and optimal design methodology for chemical and biochemical processes**
Jorge Chavez-Salas, Fernando Israel Gómez-Castro, Ricardo Morales-Rodriguez **1225**
- 206. Renewables-Based Multigeneration System for District Energy Supply**
Houd Al-Obaidli, Rajesh Govindan, Tareq Al-Ansari **1231**
- 207. Topology-Based Construction of Business-Integrated Material Modelling Workflows**
Heinz A. Preisig, Peter Klein, Natalia Konchakova, Thomas F. Hagelien, Jesper Friis and Martin T. Horsch **1237**
- 208. Investigation of optimal blending of livestock manures to produce biocrude via hydrothermal liquefaction**
Mohammad Alherbawi, Prakash Parthasarathy, Gordon Mckay, Hamish R. Mackey, Tareq Al-Ansari **1243**

209. Systematic Pairing Selection for Economic-oriented Constraint Control <i>Risvan Dirza, and Sigurd Skogestad</i>	1249
210. Quality assessment for dynamic, hybrid semi-parametric state observers <i>Isabell Viedt, Jonathan Mädler, Leon Urbas</i>	1255
211. ANN-assisted optimization-based design of energy-integrated distillation columns <i>Kai Fabian Kruber, Anna Miroschnitschenko and Mirko Skiborowski</i>	1261
212. Application of multiplicative homomorphic encryption in process industries <i>Zarina Chokparova, Leon Urbas</i>	1267
213. PPOPT - Multiparametric Solver for Explicit MPC <i>Dustin Kenefake and Efstratios N. Pistikopoulos</i>	1273
214. Improved Sequential Least Squares Programming–Driven Feasible Path Algorithm for Process Optimisation <i>Yingjie Ma, Nan Zhang, Jie Li</i>	1279
215. Parallel Simulated Annealing approach for optimal process plants instrumentation <i>José Hernández, Gabriela Minetti, Carolina Salto, Mercedes Carnero and Mabel Sánchez</i>	1285
216. A sparse polynomial surrogate model for the shrinking core model in phosphate ore digestion <i>Sanae Elmisaoui, Saad Benjelloun, Abdellah Chkifa and Abderrazak M. Latifi</i>	1291
217. On the use of data engineering and machine learning in global optimization applications with cutting plane approximations <i>Vasiliki Deligianni, Asimina Marousi, Apostolos Chalkis and Antonis Kokossis</i>	1297
218. A systematic approach for the processing of experimental data from anaerobic syngas fermentations <i>Eduardo Almeida Benalcázar, Henk Noorman, Rubens Maciel Filho, John Posada</i>	1303
219. Process Systems Engineering prospects in Circular Economy implementation in industry <i>Konstantina G. Stylianopoulou, Emilia M. Kondili, John K. Kaldellis</i>	1309

220. Incremental financial analysis of black liquor upgraded gasification in integrated kraft pulp and ammonia production plants under uncertainty of feedstock costs and carbon taxes <i>Meire Ellen Gorete Ribeiro Domingos, Daniel Flórez-Orrego, Moisés Teles dos Santos, Silvio de Oliveira Júnior, François Maréchal</i>	1315
221. A Capex Opex Simultaneous Robust Optimizer: Process Simulation-based Generalized Framework for Reliable Economic Estimations <i>Kristiano Prifti, Andrea Galeazzi, Massimo Barbieri and Flavio Manenti</i>	1321
222. A soybean supply chain model to analyze the greenhouse gas emissions of the transport sector <i>Milagros Verrengia, Aldo Vecchiatti</i>	1327
223. Ontology for Enhanced Industrial Process Control <i>Renata Samara Rodrigues de Sousa, Song Won Park</i>	1333
224. Complementing Natural Gas Driven Syngas with Optimum Blends of Gasified Biomass Waste <i>Ahmed AlNouss, Gordon Mckay, Tareq Al-Ansari</i>	1339
225. How tools and technologies can help drive sustainability <i>Julien de Beer, Mihaela Hahne</i>	1345
226. The use of game theory in the analysis of marine lubricant markets in ports <i>Dimitra Tsakona and Antonis Kokossis</i>	1351
T6: Digitalization and Artificial Intelligence	
227. Application of Outlier Treatment Towards Improved Property Prediction Models <i>Adem R.N. Aouichaoui, Seyed Soheil Mansouri, Jens Abildskov, Gürkan Sin</i>	1357
228. Data-driven, Image-Based Flow Regime Classification for Stirred Aerated Tanks <i>Corinna Kröger, Valentin Khaydarov, Leon Urbas</i>	1363
229. Data-driven modeling for physical property prediction of polypropylene composites using artificial neural network and principal component analysis <i>Chonghyo J, Hyundo P, Seokyoung H, Jongkoo L, Hyungtae C, Il M, Junghwan K</i>	1369

230. Optimization of the Wastewater Treatment Plant Aeration Using Artificial Neural Networks Models <i>Norbert-Botond Mihály and Vasile Mircea Cristea</i>	1375
231. The truncated Q statistic for Statistical Process Monitoring of High-Dimensional Systems <i>Marco S. Reis, Ricardo Rendall, Tiago J.Rato, Cristina Martins, Pedro Delgado</i>	1381
232. Development of a predictive emission measurement system using hybrid models with industrial data <i>Ataide S. Andrade Neto, Argimiro R. Secchi, Bruno D. O. Capron, Antônio Rocha, Luciana N. Loureiro and Patrícia R. Ventura</i>	1387
233. On The Interpretability of Graph Neural Networks in QSPR modeling <i>Fan Fan, Adem R.N Aouichaoui, Gürkan Sin</i>	1393
234. Predictive Maintenance in the Digital Era <i>Aaron S. Yeardley, Jude O. Ejeh, Louis Allen, Solomon F. Brown and Joan Cordiner</i>	1399
235. An evolutionary approach for techno-economic assessment <i>Martin C. De Meio Reggiani, Luciana B. Villar, Hernán P. Vigier, Nélide B. Brignole</i>	1405
236. Optimization of an artificial neural network structure for modelling carbon capture in spray columns <i>Ulderico Di Caprio, Emine Kayahan, Min Wu, Siegfried Mercelis, Peter Hellinckx, Tom Van Gerven, Steffen Waldherr and M. Enis Leblebici</i>	1411
237. Hierarchical Statistical Process Monitoring based on a Functional Decomposition of the Causal Network <i>Rodrigo Paredes, Tiago J. Rato, Lino O. Santos, Marco S. Reis</i>	1417
238. A Recurrent Neural Networks-Based Approach for Modeling and Control of a Crystallization Process <i>Fernando Arrais R. D. Lima, Gabriel F. M. de Miranda, Marcellus G. F. de Moraes, Bruno D. O. Capron, Maurício B. de Souza Jr.</i>	1423
239. Lithium-air battery electrocatalyst identification using Machine Learning and SciBERT word embeddings <i>Arun Muthukkumaran, Arjun Ravichandran, Sai Shanbhag, Ramprasad Arjun, Raghunathan Rengaswamy</i>	1429

- 240. A digital twin-concept for smart process equipment assemblies supporting process validation in modular plants**
Jonathan Mädler, Julian Rahm, Isabell Viedt, Leon Urbas 1435
- 241. Bayesian Optimization for techno-economic analysis of pressure swing adsorption processes**
Leif Erik Andersson, Johannes Schilling, Luca Riboldi, André Bardow and Rahul Anantharaman 1441
- 242. Online state of charge estimation of lithium-ion battery using surrogate model based on electrochemical model**
Seunghyeon Oh, Kyojin Jang, Jiyong Kim, Il Moon 1447
- 243. Machine-learning based prediction of infinite-dilution activity coefficients of ionic liquids using physicochemical properties**
Seongju Lee, Changsu Kim, Thai Ngan Do, and Jiyong Kim 1453
- 244. The Impact of Reward Shaping in Reinforcement Learning for Agent-based Microgrid Control**
Valentin Père, Fabien Baillon, Mathieu Milhé and Jean-Louis Dirion 1459
- 245. Knowledge mining from scientific literature for acute aquatic toxicity: classification for hybrid predictive modelling**
Gulnara Shavaliyeva, Stavros Papadokostantakis, Greg Peters 1465
- 246. Machine Learning for the prediction of the thermochemical properties (enthalpy and entropy of formation) of a molecule from its molecular descriptors**
Cindy Trinh, Dimitrios Meimaroglou, Silvia Lasala and Olivier Herbinet 1471
- 247. A Novel Machine Learning-Based Optimization Approach for the Molecular Design of Solvents**
Zihao Wang, Teng Zhou, Kai Sundmacher 1477
- 248. A methodology for gray-box modeling of nonlinear ODE systems**
Joschka Winz and Sebastian Engell 1483
- 249. Optimal aeration of wastewater treatment plants based on recurrent neural networks and real plant data**
Christian Hoffmann, Joris Weigert, Gerardo Brand Rihm, Erik Esche, and Jens-Uwe Repke 1489
- 250. Soft Sensor of Key Components in Recirculating Aquaculture Systems, using Feedforward Neural Networks**
Allyne M. dos Santos, Espen Karlsen, Sigurd Skogestad and Kari J.K. Attramadal 1495

251. Interpretability of neural networks predictions using Accumulated Local Effects as a model-agnostic method <i>Tina Danesh, Rachid Ouaret, Pascal Floquet and Stéphane Negny</i>	1501
252. Hybrid AI Models in Chemical Engineering – A Purpose-driven Perspective <i>Arijit Chakraborty, Sven Serneels, Heiko Claussen and Venkat Venkatasubramanian</i>	1507
253. Determination of Adsorption Energies from DFT databases using Machine Learning techniques <i>José I. Arsuaga and Ana I. Torres</i>	1513
254. Probabilistic graphical models for the identification and analysis of reaction pathways in water treatment <i>Rachid Ouaret, Ali Badara Minta, Claire Albasi, Jean-Marc Choubert and Antonin Azaïs</i>	1519
255. Integrating autoencoder and heteroscedastic noise neural networks for industrial data analysis and soft-sensor design <i>Sam Kay, Harry Kay, Max Mowbray, Amanda Lane, Cesar Mendoza, Philip Martin, Dongda Zhang</i>	1525
256. Economic risk analysis of emergent green hydrogen facilities – A Monte Carlo approach <i>Bruno Gerard, Eduardo Carrera, Denis Lun, Lu Zhao</i>	1531
257. Machine Learning-Based Surrogate Models and Transfer Learning for Derivative Free Optimization of HTPeM Fuel Cells <i>Luis A. Briceno-Mena, Christopher G. Arges, Jose A. Romagnoli</i>	1537
258. A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning <i>Andrea Galeazzi, Kristiano Prifti, Francesco Gallo and Flavio Manenti</i>	1543
259. Bayesian Neural Network-Based Calibration for Urban Air Quality Sensors <i>Gustavo R. Taira, Adriano G. Leal, Alessandro S. Santos, Song W. Park</i>	1549
260. Hyperconic Machine Learning to Predict Microbial Growth <i>J. Cortez-González, J.P. Serrano-Rubio, R. Murrieta-Dueñas; I. Segovia-Dominguez, V. López-Ramírez</i>	1555

T7: CAPE Applications Addressing Societal Challenges

- 261. Optimal Contract Selection for Contract Manufacturing Organisations in the Pharmaceutical Industry Under Uncertainty**
Apostolos P. Elekidis, Michael C. Georgiadis **1561**
- 262. A Framework for Economic Optimization of Carbon Capture and Sequestration from Italian Industrial Sources Under Seismic Risk Constraints**
Federico d'Amore, Gianmarco Marcato, Paolo Mocellin, Fabrizio Bezzo **1567**
- 263. Energy-water Scheduling Decisions for Agricultural Scenario Planning**
Marcello Di Martino, Sarah Namany, Styliani Avraamidou, Tareq Al-Ansari, Patrick Linke and Efstratios N. Pistikopoulos **1573**
- 264. Optimizing the allocation of resources for the security of the water-energy-food nexus**
Brenda Cansino-Loeza, Aurora del Carmen Munguía-López, José María Ponce-Ortega **1579**
- 265. Application of CAPE Tools into Prospective Life Cycle Assessment: A Case Study in Recycling Systems Design of Lithium-Ion Battery**
Yasunori Kikuchi, Aya Heiho, Yi Dou, Izuru Suwa, Chiharu Tokoro **1585**
- 266. A Fair-Sustainable Approach for the Optimization of an Integrated Fuel Production System**
Aurora del Carmen Munguía-López, Aurora de Fátima Sánchez-Bautista, Mahmoud M. El-Halwagi, José María Ponce-Ortega **1591**
- 267. The advancement of zero-emission natural gas power plants and their role in future energy supply**
Aibo Zhang, Haoshui Yu, Emre Gençer, Adem A.R. Nielsen, Gürkan Sin, Songjie Shi **1597**
- 268. A Process Integration-Based Optimal Decarbonisation Policymaking Software Framework**
Purusothmn Nair S Bhasker Nair, Michael Short, Dominic C. Y. Foo, Raymond R. Tan **1603**
- 269. Interplaying of industry 4.0 and circular economy in cyber-physical systems towards the mines of the future**
Mohammed Yaqot, Brenno C. Menezes, Robert E. Franzoi **1609**

270. Environmental Impacts of Rice Husk-Derived Silica under Uncertainty: Is “Bio” better?	
<i>Ethan Errington, Miao Guo and Jerry Y.Y. Heng</i>	1615
271. Economic Analysis of Novel Pathways for Recovery of Lithium Battery Waste	
<i>Khalid Amao, Bogdan Dorneanu, Harvey Arellano-Garcia</i>	1621
272. Sensitivity Analysis and Risk Assessment for the In-Silico Design and Use of Optimized Cell Factories in a Xylitol Biorefinery	
<i>Nikolaus I. Vollmer, Krist V. Gernaey, Gürkan Sin</i>	1627
273. Optimal design and planning of supply chains for viral vectors and RNA vaccines	
<i>Dauda Ibrahim, Zoltán Kis, Kyungjae Tak, Maria Papathanasiou, Cleo Kontoravdi, Benoît Chachuat, and Nilay Shah</i>	1633
274. Different approaches to epidemic modeling - The Covid-19 case study	
<i>Davide Manca</i>	1639
275. An integrated and sustainable Ethanol-Hydrogen-Ammonia-Urea plant	
<i>Artur S. Bispo, Leonardo O. S. Santana, Fernando L. P. Pessoa, Ana L. B. de Souza, Ewerton E. S. Calixto</i>	1645
276. A methodology based on social life cycle assessment for social hotspots identification	
<i>Andreia Santos, Madalena Veloso, Ana Carvalho, Ana Barbosa-Póvoa</i>	1651
277. A game-theoretical approach for the analysis of waste treatment and circular economy networks	
<i>Antonis Kokossis and Evripidis Melampianakis</i>	1657
T8: Education in CAPE and Knowledge Transfer	
278. Agile Process Systems Engineering (PSE) education – 1. What should be taught to achieve desired outcomes mastery by graduates?	
<i>Emilia M. Kondili, Ian T. Cameron, Grégoire Léonard, Daniel R. Lewin, Seyed Soheil Mansouri, Fernando G. Martins, Luis Ricardez-Sandoval, Hirokazu Sugiyama, Edwin Zondervan</i>	1663
279. You have nothing to lose: Worry-free flipping for PSE	
<i>Daniel R. Lewin</i>	1669

280. Agile Process Systems Engineering (PSE) education – 2. How to teach to achieve desired outcomes mastery by graduates <i>Daniel R. Lewin, Ian T. Cameron, Emilia M. Kondili, Grégoire Léonard, Seyed Soheil Mansouri, Fernando G. Martins, Luis Ricardez-Sandoval, Hirokazu Sugiyama, Edwin Zondervan</i>	1675
281. Decision-making framework for improved educational resilience under pandemic events <i>Robert E. Franzoi, Noof AlQashouti, Brenno C. Menezes</i>	1681
282. Teaching courses heavily dependent on computational resources to STEM students during Pandemics <i>Nelson Chibeles-Martins, Lourdes B. Afonso</i>	1687
Author Index	1693

Development of a virtual platform for the metaheuristic optimization of heat exchangers

Oscar D. Lara-Montaña,^a Fernando I. Gómez-Castro,^{a,*} Claudia Gutierrez-Antonio^b

^a*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato, Guanajuato, 36050, México.*

^b*Facultad de Ingeniería, Universidad Autónoma de Querétaro, Campus Amazcala, Carretera a Chichimequillas s/n km. 1, Amazcala, El Marqués, Querétaro, 76465, México.*

Abstract

Independently of the implemented heat exchange technology for industrial applications, it is necessary to apply an optimization routine to obtain the best heat exchanger design according to an objective function. This work presents a design and optimization environment for heat exchangers where the total annual cost is used as an objective function. Eight metaheuristic algorithms are employed to carry out the optimization procedure independently: Cuckoo Search, Differential Evolution, Grey Wolf Optimizer, Jaya Algorithm, Particle Swarm Optimization, Univariate Marginal Optimization Algorithm, and Whale Optimization Algorithm. Three different types of heat exchangers are contemplated: shell-and-tube heat exchanger, plate fin heat exchanger and plate heat exchanger. The performance of each optimization algorithm is assessed for each kind of exchanger.

Keywords: heat exchanger optimization, optimization environment, metaheuristic optimization.

1. Introduction

Heat exchanger devices are applied to exchange thermal energy between two or more fluid streams or to supply heating and cooling services in industry. Different technologies are available, some examples of heat exchangers configurations include double pipe, shell-and-tube, finned-tube, plate-fin, plate or spiral plate (Thulukkanam, 2013). The heat exchanger technology is selected according to different criteria, such as, operational conditions, available space, physical properties of the fluids, among others.

Three types of heat exchangers are considered: shell-and-tube heat exchanger (STHE), plate-fin heat exchanger (PFHE), and plate heat exchanger (PHE). STHEs are the most versatile, being able to operate at high temperatures and pressures. Also, they have a good heat transfer area-to-volume ratio. PFHEs are categorized as compact heat exchangers because of their high heat transfer area-to-volume ratio. They are conformed by a series of flat plates and layers of corrugated fins and its main use is in gas-to-gas systems. PHE are flexible devices; its heat transfer area can be modified by adding or removing plates. Regardless of the heat exchanger technology, it is necessary to produce an optimal design according to an objective function. Some common optimization objectives are the minimization of the heat transfer area, total annual cost, the entropy generation, or the maximization of the effectiveness. The optimization process is not a trivial task. All design methods available, independently of the heat exchange technology, contain non-

linear, non-continuous and non-differentiable equations which results in complex optimization problems (Onishi et al., 2013). The resulting optimization problems depends on continue and discrete variables that increase the complexity. There is plenty of different gradient-based optimization algorithms. Gradient-based algorithms tend to fast converge, but most of times converge in a local optimum in non-convex optimization problems. On the other hand, metaheuristic algorithms employ two search mechanisms known as exploration and exploitation, increasing the probability to converge in the global optima (Yang, 2014).

Metaheuristic optimization algorithms are employed to produce the best heat exchanger design. The user can choose between eight optimizer options; these are Cuckoo Search (CS) (Yang and Suash Deb, 2009), Differential Evolution (DE) (Storn and Price, 1997), Grey Wolf Optimizer (GWO) (Mirjalili et al., 2014), Jaya Algorithm (JA) (Venkata Rao, 2016), Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1995), Univariate Marginal Optimization Algorithm (UMDA) (Larrañaga and Lozano, 2001), and Whale Optimization Algorithm (WOA) (Mirjalili and Lewis, 2016). The design and optimization environment is designed in a modular way, all the codification was done in Python. The total annual cost (TAC) is used as objective function. Specific constraints are employed for each type of heat exchanger.

2. Heat exchanger design and optimization environment

This section explains how the design and optimization environment works, which is designed in a modular way. Different modules that perform specific task interacts with each other to obtain the best heat exchanger design. Once the optimization procedure is finished, the most relevant variables are shown. Six modules are involved, these are: the *Main* module, the *Optimizers* module, the *OptProblem* module, the *DesignMethods* module and the *Costing* module. Figure 1 shows how the modules are interlinked.

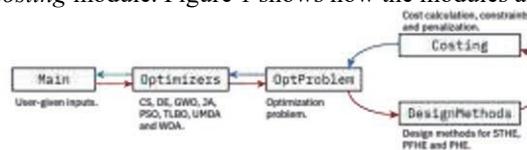


Figure 1. Modular configuration of the design and optimization environment.

2.1. Main module

The *Main* module is the only one the user can directly interact with. The user provides the different parameters for the design and cost calculation of the heat exchanger, and for optimization procedure. The required parameters regarding the design of the heat exchanger are the type of heat exchanger (STHE, PFHE or PHE), the inlet and outlet temperatures of the cold and hot fluids, the physical properties of the fluids. For the cost calculation it is necessary to introduce the interest rate, the number of operating hours per year, the electricity cost per 1 kWh, the value of the Chemical Engineering Plant Cost Index (CEPCI), and the value of the factors for the construction and operating conditions. Regarding to the metaheuristic optimization routine, it is necessary to select the optimization algorithm to be used, indicate the upper and lower bounds for each decision variable, specify the number of iterations and individuals (solutions) produced per iteration, and the number of experiments (only if the user requires to obtain statistical information such as the standard deviation, median, mean, and minimum value). Once all this information is given, the design and optimization procedure can be carried out.

All the information that the user provides is distributed to the other modules. The *Main* module directly interacts with the *Optimizers* module, which receives all the information regarding the optimization procedure.

2.2. Optimizers module

After the user introduce all the necessary information and runs the environment, the *Main* module calls the *Optimizers* module. The *Optimizers* module contains the codes of the eight metaheuristic optimization algorithms that can be employed to optimize the selected type of heat exchanger. The included metaheuristic algorithms are CS, DE, GWO, JA, PSO, TLBO, UMDA and WOA. The value of parameters used by Lara-Montaño et al. (2021) were employed in this work.

This module generates all solutions according to the chosen optimization algorithm. It takes as inputs the parameters involved in the optimization procedure and is linked with the *OptProblem* module which in turn is linked to the *DesignMethods* and *Costing* modules. The outputs of the *Optimizers* module is the best solution i.e. the best heat exchanger design and its total annual cost (value of the objective function).

2.3. Modules that define the optimization problem

The module *Optimizers* contains the optimization problem that depend on the selected heat exchanger. This module takes as inputs the decision variables on which each optimization problem depends. The value for the objective function for each design is the output. The module is divided in two parts. In the first part a heat exchanger design is calculated from the decision variables provided by the *Optimizers* module, this occurs in the *DesignMethods* module. The second part consist of obtaining the value of the evaluation of the objective function for each heat exchanger design produced.

The *DesignMethods* module contains the equations, variables and parameters involved in the design methods of STHEs, PFHEs and PHEs. The Bell-Delaware method is implemented to design STHEs, the details of this design methodology can be found in Shah and Sekulic (2003). The procedure to design PFHEs is taken from Mishra et al. (2009) and Rao and Patel (2010). Finally, the methodology to design PHEs is obtained from Kakaç et al. (2002).

The *Costing* module receives information from the *DesignMethods* module and calculates the operating cost (C_{op}) given by the cost of pumping, and the fixed cost (C_f) that mainly depends on the value of the heat transfer area. Then, the total annual cost, which is employed as objective function, is calculated by the addition of the (C_{op}) and (C_f). The (C_{op}) is computed according to Towler and Sinnott (2012). The (C_f) for STHEs and PHEs is calculated according to Smith (2005), whereas for PFHE it is computed as shown in Towler and Sinnott (2012). Also, the *Costing* module implements the constraints for each optimization problem and applies a penalization if any constraint is not satisfied. The value of the total annual cost is sent back to the *Optimizers* module that produce potentially new heat exchanger designs in each iteration. Once the optimization procedure finishes, the results are printed and exported. The results are the relevant variables and parameters of the heat exchanger design, the operation cost, fixed cost, total annual cost, the statistical results, and the best value of the objective function found.

3. Case studies and optimization problems

Two case studies are considered. The first case study, for STHE and PHE, involves methanol with a flow rate of 27.8 kg/s and inlet and outlet temperatures of 95°C and 40°C, respectively. The other fluid is sea water with a flow rate of 68.9 kg/s and inlet and outlet temperatures of 25°C and 40°C, respectively. In the second case study, for PFHE, both

fluids are air. The hot fluid has an input and output temperature of 240°C and 171°C, respectively and a flow rate of 0.8962 kg/s.

The cold fluid has an input and output temperature of 4°C and 79.5°C, respectively and a flow rate of 0.8296 kg/s. A project lifetime of 20 years, an interest rate of 5%, 8,000 operating hours per year, an electricity cost of 0.1 kWh and a CEPCI equal to 686.7 (Jenkins, 2021) are used.

Eleven decision variables are involved in the design of a STHE, seven in the PFHE and eight in the PHE. Upper and lower boundaries for continuous decision variables are shown in Table 1. The allowed values for discrete variables are shown in Table 2.

For STHE the pressure drop in shell side, ΔP_s and tube side, ΔP_t must be lower or equal to 70,000 Pa. The fluid velocity in tubes must be found between 1 m/s and 3 m/s (Caputo et al., 2008). And the ratio between the inner diameter of the shell and the length of tubes (D_s/L) must be lower or equal to 15. For PFHE the maximum allowed pressure drop in hot side, ΔP_h , and cold side, ΔP_c , are 300 Pa and 2,000 Pa, respectively (Rao and Patel, 2010). For PHE the maximum allowed pressure drop is 70,000 Pa in both sides.

The three optimization problems were solved using all the optimization algorithms. In Lara-Montaño et al. (2021) it is determined through a sensitivity analysis that 50 individuals are required, these are values for the number of individuals and iterations used. To obtain statistical information 30 experiments were run per optimization problem and optimization algorithm.

Table 1. Bounds for continuous design variables.

Heat exchanger	Design variable		Lower bound	Upper bound
STHE	Diameter of shell	D_s	300 mm	1,000 mm
	Baffle spacing at center	L_{bc}	$0.2D_s$	$0.55D_s$
	Baffle spacing at the inlet and outlet	L_{bo}, L_{bi}	L_{bc}	$1.6L_{bc}$
	Tube-to-baffle diametrical clearance	δ_{tb}	$0.01d_o$	$0.1d_o$
	Diametrical clearance of shell-to-baffle	δ_{sb}	$0.01D_s$	$0.1D_s$
	Outer diameter of tube bundle	D_{otl}	$0.8(D_s - \delta_{sb})$	$0.95(D_s - \delta_{sb})$
PFHE	Heat exchanger length in cold size	L_{cold}	0.1 m	1 m
	Heat exchanger length in hot size	L_{hot}	0.1 m	1 m
	Height of fin	H	2 mm	20 mm
	Fin thickness	F_t	0.1 mm	2 mm
	Lance length of the fin	l	1 mm	10 mm
PHE	Horizontal distance between ports	L_h	0.3 m	0.7 m
	Vertical distance between ports	L_p	1.1 m	2 m
	Port diameter	D_p	0.1 m	0.4 m
	Plate thickness	P_t	0.001 m	0.0003 m
	Plate pitch	P	0.0015 m	0.005 m
	Enlargement factor	E_f	1.15 m	1.25 m

Table 2. Allowed values for discrete variables

Heat exchanger	Design variable		Allowed values
STHE	Outer and inner tube diameter	d_o, d_i	Taken from Flynn et al. (2019)
	Tube pitch	P_t	[1.25 d_o , 1.5 d_o]
	Tube layout angle	TL	[30°, 45°, 90°]
	Baffle cut	B_c	[25%, 30%, 40%, 45%]
	Number of tube passes	s	[1, 2, 4]
PFHE	Fin frequency	n	[200, 201, ..., 1,000]
	Number of fin layers in hot side	N_h	[2, 3, ..., 40]
PHE	Chevron angle	D	[30°, 45°, 50°, 60°, 65°]
	Number of plates	N_p	[2, 3, ..., 700]

4. Results

Table 3 shows the statistical results after 30 experiments for each optimization problem and optimization algorithm. DE and GWO have the best performance in the design optimization of a STHE. These optimization algorithms produce the smaller SD, they can obtain the best design at least in one experiment, and according to the median in most of the experiment converge about the best design. WOA, PSO and UMDA present the worse performance according to the produced SD. PSO do not has a good performance, but it was able to find the STHE design with the minimum cost in at least one experiment.

Table 3. Statistical results

Heat exchanger	Statistical variable	CS	DE	GWO	JA	PSO	TLBO	UMDA	WOA
STHE	Mean (USD/year)	12,899.59	12,862.28	12,873.76	13,365.17	13,142.55	13,266.87	13,237.23	13,890.98
	Median (USD/year)	12,892.48	12,860.52	12,862.92	13,178.23	12,931.32	13,174.19	13,120.50	13,294.01
	SD (USD/year)	27.81	6.44	13.15	311.91	534.86	271.60	448.81	1,067.88
	Minimum (USD/year)	12,866.71	12,860.34	12,860.64	12,871.41	12,860.31	12,951.26	12,865.27	12,892.19
	Mean (USD/year)	2,266.49	2,221.56	2,227.36	2,290.75	2,232.46	2,231.20	2,276.12	2,249.30
PHE	Median (USD/year)	2,271.82	2,222.59	2,226.54	2,289.67	2,304.05	2,227.85	2,299.17	2,402.80
	SD (USD/year)	36.41	2.12	3.57	50.74	119.98	5.24	66.44	243.65
	Minimum (USD/year)	2,225.11	2,221.19	2,221.19	2,231.03	2,221.19	2,221.19	2,237.31	2,221.19
	Mean (USD/year)	3,969.98	3,969.97	3,969.97	3,969.97	3,995.01	3,969.97	3,997.30	3,969.97
	Median (USD/year)	3,973.55	3,969.97	3,969.97	3,975.38	4,080.23	3,969.97	4,012.79	3,969.97
PFHE	SD (USD/year)	7.76	0.00	0.00	9.51	447.08	0.00	35.96	0.00
	Minimum (USD/year)	3,969.97	3,969.97	3,969.97	3,969.97	3,969.97	3,969.97	3,976.86	3,969.97

Regarding to the optimization problem where the best design for a PHE must be obtained, PSO and WOA have the worst performance, but both algorithms are able to converge in the best solution in at least one experiment. DE, GWO and TLBO have the best performance, these algorithms produce the smaller SD, and their median is close to the minimum value for TAC found.

Even though PSO and UMDA have the worst performance to find the best design of a PFHE, due their large SD, they can obtain the best design in at least one experiment. That is, in at least one experiment these algorithms converge in the neighborhood of the best solution. In this optimization problem all the algorithms can obtain the best design, and most of them have a produce a median close to the minimum value found for the TAC.

The more complex optimization problem is the one that involves the design of a STHE. This optimization problem produces relatively large SD while using multiple optimization algorithms. The best overall optimization algorithms are DE and GWO. The values found of the decision variables for the best designs of the heat exchangers are presented in Table 4.

5. Conclusion

The presented design and optimization environment for heat exchangers can find the best design, using the TAC as objective function, for STHE, PFHE, and PHE. Currently, it is limited to heat exchangers with no phase change. Although multiple metaheuristic optimization algorithms are implemented, DE has the overall best performance independently of the heat exchanger that requires to be optimized. Thus, according to the results, DE is the most recommended optimization method for the non-linear mixed-integer models representing the heat exchangers. The environment also supports the

execution and statistical analysis of multiple experiments as a tool of analysis for the obtained results.

6. Acknowledgements

The authors acknowledge the support of Universidad de Guanajuato and Universidad Autónoma de Querétaro, and the scholarship provided by CONACYT to support the graduate studies of O.D. Lara-Montaño.

References

- A. C. Caputo, P. M. Pelagagge, P. Salini, jul 2008. Heat exchanger design based on economic optimisation. *Applied Thermal Engineering* 28 (10), 1151–1159.
- A. M. Flynn, T. Akashige, L. Theodore, 2019. *Kern's Process Heat Transfer*, 2nd Edition. John Wiley & Sons, New Jersey.
- S. Jenkins, Sep. 2021. 2021 cepci updates: July (prelim.) and june (final).
URL <https://www.chemengonline.com/2021-cepci-updates-july-prelim-and-june-final/>
- S. Kakaç, H. Liu, A. Pramuanjaroenkij, 2002. *Heat Exchangers: Selection, Rating, and Thermal Design*, 2nd Edition. CRC Press.
- J. Kennedy, R. Eberhart, 1995. Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. Vol. 4. IEEE, pp. 1942–1948.
- O. D. Lara-Montaño, F. I. Gómez-Castro, C. Gutiérrez-Antonio, 2021. Comparison of the performance of different meta-heuristic methods for the optimization of shell-and-tube heat exchangers. *Computers and Chemical Engineering* 152, 107403.
- P. Larrañaga, J. A. Lozano, 2001. *Estimation of distribution algorithms: A new tool for evolutionary computation*. Vol. 2. Springer Science & Business Media.
- S. Mirjalili, A. Lewis, 2016. The Whale Optimization Algorithm. *Advances in Engineering Software* 95, 51–67.
- S. Mirjalili, S. M. Mirjalili, A. Lewis, 2014. Grey Wolf Optimizer. *Advances in Engineering Software* 69, 46–61.
- M. Mishra, P. K. Das, S. Sarangi, 2009. Second law based optimisation of crossflow plate-fin heat exchanger design using genetic algorithm. *Applied Thermal Engineering* 29, 2983–2989.
- V. C. Onishi, M. A. Ravagnani, J. A. Caballero, oct 2013. Mathematical programming model for heat exchanger design through optimization of partial objectives. *Energy Conversion and Management* 74, 60–69.
- R. V. Rao, V. Patel, 2010. Thermodynamic optimization of cross flow plate-fin heat exchanger using a particle swarm optimization algorithm. *International Journal of Thermal Sciences* 49 (9), 1712–1721.
- R. K. Shah, D. P. Sekulic, 2003. *Fundamentals of Heat Exchanger Design*, 1st Edition. John Wiley & Sons, New Jersey.
- R. Smith, 2005. *Chemical Process Design and Integration*, 2nd Edition. John Wiley & Sons, Ltd, Chichester.
- R. Storn, K. Price, 1997. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 11 (4), 341–359.
- K. Thulukkanam, 2013. *Heat Exchanger Design Handbook*, 2nd Edition. CRC Press.
- G. Towler, R. Sinnott, 2012. *Chemical engineering design: principles, practice, and economics of plant and process design*, 2nd Edition. Butterworth-Heinemann, Oxford Waltham, Mass.
- R. Venkata Rao, 2016. Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *International Journal of Industrial Engineering Computations* 7, 19–34.
- X. Yang, Suash Deb, 2009. Cuckoo Search via Lévy flights. In: *2009 World Congress on Nature Biologically Inspired Computing (NaBIC)*. IEEE, pp. 210–214.
- X. S. Yang, 2014. *Nature-Inspired Optimization Algorithms*, 1st Edition. Elsevier.

Combining optimization and life cycle assessment: Design of low-carbon multi-energy systems in the SecMOD framework

Christiane Reinert^a, Lars Schellhas^a, Julia Frohmann^a, Niklas Nolzen^{a,c},
Dominik Tillmanns^a, Nils Baumgärtner^a, Sarah Deutz^a, André Bardow^{a,b,c,CA}

^a *Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstraße 8,
52062 Aachen, Germany*

^b *Institute of Energy and Climate Research, Energy Systems Engineering (IEK-10),
Forschungszentrum Jülich GmbH, Jülich 52425, Germany*

^c *ETH Zürich, Department of Mechanical and Process Engineering, Energy & Process
Systems Engineering, Tannenstr. 3, Zürich 8092, Switzerland, CA: abardow@ethz.ch*

Abstract

The decarbonization of industrial utility systems is an important step to reduce greenhouse gas emissions. Decarbonization can be enabled by sector coupling, which can also enhance system flexibility of low-carbon utility systems. However, exploiting sector coupling efficiently for the design of low-carbon utility systems is complex and, therefore, best addressed by mathematical optimization. Recently, the open-source framework SecMOD was introduced for the linear optimization of multi-energy models with adjustable spatial and temporal resolution. SecMOD considers environmental impacts by fully integrating life cycle assessment. In this work, we optimize a sector-coupled utility system supplying electricity, heating, and cooling. For this purpose, we extend SecMOD to allow mixed-integer decisions. Further, we investigate the benefits of a pumped thermal energy storage system consisting of a high-temperature heat pump, heat storage, and an organic Rankine cycle. We identify trade-offs in system design by comparing a least-cost design to a design with minimal greenhouse gas emissions. Combining heat pumps and heat storage is economically and environmentally viable and leads to synergies between sectors. However, the reconversion of stored heat to electricity cannot compete with alternatives, such as battery storage.

Keywords: utility systems; decarbonization; MILP; pumped thermal energy storage; sector-coupling

1. Introduction: Low-carbon industrial utility systems

Industry significantly contributes to overall economic value added but is also subject to high greenhouse gas emissions, leading to an increase in climate change. Hence, beyond the traditional search for cost reductions, improvements now also aim increasingly at low-carbon energy supply (Ringkjøb et al. 2018). Integrating low-carbon electricity from renewable sources is vital to reduce greenhouse gas emissions. While renewable energy is increasingly integrated in the electricity sector, its contribution to supply heat has stagnated. Sector-coupling may help to bridge this gap and lead to synergy effects between sectors (Guelpa et al. 2019).

Pumped thermal energy storage (PTES) systems can strengthen sector-coupling by addressing volatility in electricity supply to flexibly provide electricity and heat (Dumont

et al. 2020): During high feed-in of renewable energy converters, PTES systems convert electricity into heat and store heat that can not be used immediately. The stored heat can then either fulfill a heat demand or be re-converted to electricity at a later time. In this way, PTES systems can contribute to a secure and low-carbon energy supply.

The need to consider volatile electricity supply of renewables, sectoral interaction, and environmental aspects leads to highly complex systems. To cope with this complexity, effective energy systems are usually designed based on mathematical modeling and optimization. Hence, the above-mentioned challenges need to be addressed when designing optimization methods and software frameworks. Although many models have been developed for the transition to low-carbon energy systems, reusability is limited when these models are not openly available (Pfenninger et al. 2018). Generalized and modular open-source software frameworks ensure reusability and, hence, contribute to accelerating transparent research. For this purpose, the object-oriented framework SecMOD was recently introduced that considers energy conversion, transport, and storage (Reinert et al. 2022).

In this work, we apply SecMOD to a case study of an industrial utility system providing electricity, heating, and cooling. First, we introduce the general formulation of SecMOD and discuss how we integrate life cycle indicators as an objective function. To model the utility system as a mixed-integer linear program (MILP), we modify SecMOD to account for integer decisions, such as the modelling of part-load behaviour. We then identify trade-offs between the least-cost design and the design with the lowest overall greenhouse gas emissions. Two energy supply scenarios are considered: with and without grid supply. In these scenarios, we examine to what extent sector-coupling by a PTES system can contribute to the optimal design of the utility system.

2. Method: Utility system optimization considering life cycle indicators

In this Section, we present the general problem statement of SecMOD (Reinert et al. 2022). We then discuss the integration of life cycle assessment (LCA) into the energy system optimization and further specify how SecMOD is implemented.

Generally, the problem statement of SecMOD is as follows: Given

- a potentially spatially and temporally resolved exogenous energy demand,
- a set of components to convert and store energy,
- and additional constraints (such as emission limits),

find the energy system that minimizes the objective function (e.g., total annualized cost or overall annual global warming impact) and cover the energy demands in each time step. As decision variables, we consider the capacity expansion of energy converters and storage components and respective component operation.

In this work, we extend the framework formulation to a mixed-integer linear program (MILP). The MILP optimization considers discrete component expansion and integer decisions, e.g., to model part-load behavior. We optimize and evaluate the system using the LCA methodology, standardized in ISO 14040:2006 and ISO 14044:2006. LCA quantifies the impacts of a component or system on the environment over the whole life cycle. In contrast to classical optimization of utility systems, we extend all component models by life cycle inventories (LCIs). Here, we employ the LCI database ecoinvent 3.5 (Wernet et al. 2016). Each LCI quantifies the material and energy flows needed to build and operate the component. We model the direct inputs and outputs of component operation using piecewise-affine functions to reflect part-load efficiencies.

For the optimization, the impacts of the infrastructure and operation are calculated by multiplying the LCIs with the respective environmental impact of each energy or mass

flow. The environmental impacts are determined using the Environmental Footprint 2.0 methodology recommended by the European Commission (Joint Research Center 2010). Due to space limitations, this work only reports on the impact category “climate change”, quantifying the global warming impact (GWI). The optimization problem is implemented in Python using Pyomo (Bynum et al. 2021; Hart et al. 2011) and is solved using Gurobi (Gurobi Optimization 2021), employing the branch-and-cut algorithm. The time-series aggregation module by Hoffmann et al. (2020) is integrated to decrease the number of time steps for time-dependent variables. The SecMOD code and documentation are publicly accessible (Reinert et al. 2022).

3. Case Study: Integrating pumped thermal energy storage in a utility system

The utility system assessed in this study (Figure 1) was first published by Voll et al. (2013) and extended by Baumgärtner et al. (2019). The system satisfies temporally resolved demands for electricity, heating, and cooling. Import is possible for grid electricity and natural gas, each associated with specific costs and environmental impacts. Overproduction of electricity may occur at no costs or revenues to account for curtailment. However, overproduction of heat or cooling is not allowed.

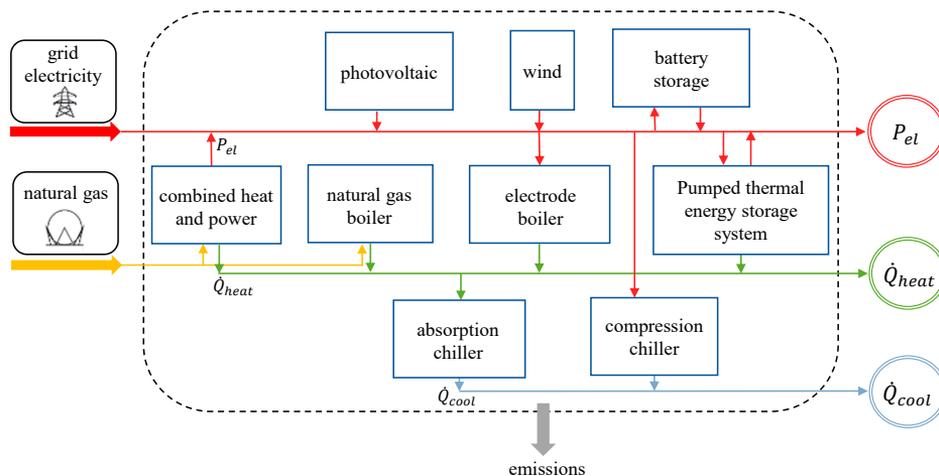


Figure 1- Simplified structure of the utility system providing electricity, heating, and cooling supply (extending on Voll et al. (2013)). For chillers, boilers, and the combined heat and power engine, multiple units can be employed, accounting for scaling effects.

The utility system consists of the components shown in Figure 1. Combined heat and power engines (CHP) and the electricity grid provide electricity. We further add wind turbines, battery storage and photovoltaic systems, as modeled in Baumgärtner et al. (2021). Heat can be provided by the CHPs, an electrode boiler, or a natural gas boiler. Cooling power is provided by absorption chillers or compression chillers, respectively. In addition to the existing components, we further model a PTES system as an emerging technology. The PTES system consists of a heat pump (based on Baumgärtner et al. (2021)), sensible heat storage with water as storage medium (Baumgärtner et al. 2019), and an organic Rankine cycle (ORC) (based on Stoppato and Benato (2020), Tartièrre and Astolfi (2017), and Tillmanns et al. (2022)). The heat storage unit can supply heat directly or feed the ORC for the reconversion to electricity, thereby adding additional flexibility.

Scaling effects during sizing are taken into account by allowing small, medium, and large components for the chillers, boilers, and CHP. The components have different specific costs and part-load behavior. For all units, sizing can be chosen continuously between the minimum and maximum component size.

We consider one year of operation, aggregated to six typical days with hourly resolution. As clustering technique for temporal aggregation, we use k-medoids. We optimize the system for two scenarios: Firstly, we consider the system given above (current system) and optimize it by an economic and an environmental cost function. The economic cost function is the total annualized system cost. As environmental objective, we minimize the overall annual GWI. The annual GWI considers the whole life cycle of the system components. For this purpose, the impacts from construction are annualized over the unit life time without discounting. Secondly, we model the utility system as a stand-alone system where electricity cannot be imported from the grid to evaluate sector-coupling in volatile energy systems on a small scale.

4. Results: Economically and environmentally optimal energy supply structure

Figure 2 shows the annual supply of electricity, heating, and cooling for the current and stand-alone utility system for all scenarios. The optimality gap is less than one percent.

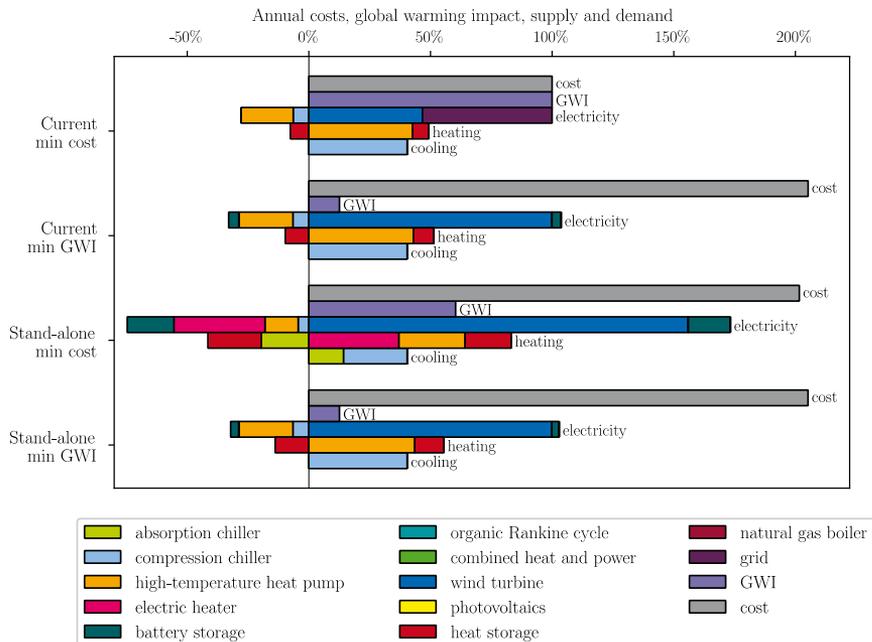


Figure 2 Annual electricity, heating, and cooling supply and endogenous demand in the current and stand-alone utility system for the cost-optimal and environmentally optimal cases. All energy flows are normalized by the electricity flow in the case “current min cost”. Further, the annualized cost and global warming impacts (GWI) are shown normalized by the case “current min cost”.

The current cost-optimal system provides electricity partly by the grid and partly by wind turbines. However, the minimal-GWI system fully relies on wind electricity. Except the current cost-optimal system, all systems store both electricity and heat to decouple supply

and demand. The PTES system as a whole is not used in any case: While some components of the PTES system - high-temperature heat pump and heat storage - are used in all systems to couple the electricity and heating sectors, the organic Rankine cycle is never used to provide electricity. Cooling is provided mainly by a compression chiller in all cases.

The environmentally optimized designs are similar in the current and stand-alone systems, as the supply of grid electricity is associated with relatively high environmental burdens and therefore not used even when available. A comparison of the stand-alone systems shows that the environmentally optimal system is near-optimal regarding cost, as the cost are less than 2 % higher than in the cost-optimal case. However, their resulting utility supply is rather different: While heat provision is slightly cheaper by combining a heat pump and an electric boiler, the environmentally optimal system employs a heat pump only. Similarly, the cost-minimal system operates the absorption chiller, whereas the environmental optimization uses the compression chiller only. In the stand-alone cost minimization, stored heat is used to provide cooling by the absorption chiller, as storing heat and using the heat directly for cooling is cost-efficient. Alternative routes that exploit the low costs of heat storage are inferior, for example, to provide cooling by combining the PTES system and the compression chiller. Overall, the studied PTES system is outperformed by other available flexibility options and therefore not employed as a whole to flexibly provide electricity.

5. Conclusions

In this work, we optimize an industrial utility system using the SecMOD framework. To account for part-load behavior in utility systems, we extend SecMOD to a MILP formulation. We show trade-offs between economically and environmentally optimal designs for a grid-integrated and a stand-alone utility system. While a significant trade-off is found between cost and GWI for the grid-integrated scenario, the trend is different in the stand-alone system: At comparably low cost, the environmentally optimal system has a five times lower GWI than the cost-optimal system.

The results further confirm that sector-coupling and energy storage are essential for industrial decarbonization. Especially in the stand-alone scenario, we observe high synergies between sectors. Regarding pumped thermal energy storage as a system of components to enhance flexibility, we find that some components (heat pump and heat storage) of PTES system are economically and environmentally viable in all cases. However, reconversion to electricity is neither economically nor environmentally competitive for the studied system.

We hope that our results can contribute to understand trade-offs in utility systems and hence help realizing environmental improvements that can be achieved cost-efficiently.

Acknowledgments

CR and AB gratefully acknowledge the financial support of SCI4climate.NRW (grant number: EFO 0001G) and by the Swiss Federal Office of Energy's SWEET programme as part of the project PATHFINDER, respectively.

References

N. Baumgärtner, R. Delorme, M. Hennen, A. Bardow, 2019, Design of low-carbon utility systems: Exploiting time-dependent grid emissions for climate-friendly demand-side management, *Appl. Energy*, 247, pp. 755–765.

- N. Baumgärtner, S. Deutz, C. Reinert, N. Nolzen, E. Küpper, M. Hennen, D. Hollermann, A. Bardow, 2021, Life-Cycle Assessment of Sector-Coupled National Energy Systems: Environmental Impacts of Electricity, Heat, and Transportation in Germany Till 2050, *Front. Energy Res.*, 9, p. 27.
- M. L. Bynum, G. A. Hackebeil, W. E. Hart, C. D. Laird, B. L. Nicholson, J. D. Sirola, J.-P. Watson, D. L. Woodruff, 2021, *Pyomo — Optimization Modeling in Python*, Third edition, Springer Science & Business Media, 67.
- O. Dumont, G. F. Frate, A. Pillai, S. Lecompte, M. de paepe, V. Lemort, 2020, Carnot battery technology: A state-of-the-art review, *Journal of Energy Storage*, 32, August, p. 101756.
- ISO 14040:2006, 2006, Environmental management - Life Cycle Assessment - Principles and framework. EN ISO 14040:2006 + AMD 1:2020.
- ISO 14044:2006, 2006, Environmental management - Life Cycle Assessment - Requirements and guidelines. EN ISO 14044:2006+AMD 1:2017+AMD 2:2020.
- E. Guelpa, A. Bischi, V. Verda, M. Chertkov, H. Lund, 2019, Towards future infrastructures for sustainable multi-energy systems: A review, *Energy*, 184, pp. 2–21.
- L. L.C. Gurobi Optimization, 2021, *Gurobi Optimizer Reference Manual*, Available online at <https://www.gurobi.com>.
- W. E. Hart, J.-P. Watson, D. L. Woodruff, 2011, *Pyomo: modeling and solving mathematical programs in Python*, *Math. Prog. Comp.*, 3, 3, pp. 219–260.
- M. Hoffmann, L. Kotzur, D. Stolten, M. Robinius, 2020, A Review on Time Series Aggregation Methods for Energy System Models, *Energies*, 13, 3, p. 641.
- Joint Research Center, 2010, *International Reference Life Cycle Data System. General guide for Life Cycle Assessment - Detailed guidance*.
- S. Pfenninger, L. Hirth, I. Schlecht, E. Schmid, F. Wiese, T. Brown, C. Davis, M. Gidden, H. Heinrichs, C. Heuberger, S. Hilpert, U. Krien, C. Matke, A. Nebel, R. Morrison, B. Müller, G. Pleßmann, M. Reeg, J. C. Richstein, A. Shivakumar, I. Staffell, T. Tröndle, C. Wingenbach, 2018, Opening the black box of energy modelling: Strategies and lessons learned, *Energy Strategy Rev.*, 19, pp. 63–71.
- C. Reinert, L. Schellhas, J. Mannhardt, D. Shu, A. Kämper, N. Baumgärtner, S. Deutz, A. Bardow, 2022, *SecMOD: An open-source modular framework combining multi-sector system optimization and life-cycle assessment*, In preparation.
- H.-K. Ringkjøb, P. M. Haugan, I. M. Solbrekke, 2018, A review of modelling tools for energy and electricity systems with large shares of variable renewables, *Renewable Sustainable Energy Rev.*, 96, pp. 440–459.
- A. Stoppato, A. Benato, 2020, Life Cycle Assessment of a Commercially Available Organic Rankine Cycle Unit Coupled with a Biomass Boiler, *Energies*, 13, 7, p. 1835.
- T. Tartière, M. Astolfi, 2017, A World Overview of the Organic Rankine Cycle Market, *Energy Procedia*, 129, 10, pp. 2–9.
- D. Tillmanns, D. Pell, J. Schilling, A. Bardow, 2022, The thermo-economic potential of Pumped-Thermal Electricity Storage: Insights from the integrated design of processes and working fluids, In preparation.
- P. Voll, C. Klaffke, M. Hennen, A. Bardow, 2013, Automated superstructure-based synthesis and optimization of distributed energy supply systems, *Energy*, 50, 11, pp. 374–388.
- G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, B. and Weidema, 2016, The ecoinvent database version 3 (part I): overview and methodology., *Int J Life Cycle Assess*, 21, 9, pp. 1218-1230.

Robust Design of Distributed Energy Systems Within Unbalanced Power Networks

Ishanki De Mel^a, Oleksiy V. Klymenko^a, Michael Short^{a*}

*^aDepartment of Chemical and Process Engineering, University of Surrey, Guildford,
United Kingdom GU2 7XH
m.short@surrey.ac.uk*

Abstract

The design and operation of distributed energy systems (DES) have often been modelled as linear optimisation problems. Although DES are increasingly connected to existing alternating current (AC) distribution networks, state-of-the-art DES modelling frameworks use oversimplified approximations which either exclude network constraints or overlook the inherent three-phase unbalance present in distribution networks. This can lead to poor designs which amplify network operational issues and result in greater costs to both the network and consumers/producers. This study presents a new modelling framework for DES design, which incorporates unbalanced optimal power flow within DES models for the first time. Furthermore, Robust Optimisation is included in this detailed modelling framework to ensure design feasibility under worst-case scenarios. Results show that previous frameworks tend to either overestimate or underestimate objectives when compared with the DES model combined with unbalanced power flow. Robust scenarios demonstrate that the new combined model is capable of closing the gap between objectives when compared with a linear DES-only model, albeit with different designs that do not violate grid constraints during baseline operation. These results suggest that this detailed framework can be utilised for DES design and network planning, as it produces more robust designs which can potentially help avert operational issues.

Keywords: Distributed Energy, Distribution Network, Unbalanced Power Flow, Robust Optimisation, Nonlinear Programming.

1. Introduction

With persisting efforts to integrate more renewable energy resources and improve network resilience, a global growth in community DES and microgrids has been predicted. Designing DES requires the consideration of many constraints, and these are best assessed using optimisation models. Mixed-Integer Linear Programming (MILP) models have been commonly used to design DES. However, with the increasing integration of DES into existing AC distribution networks, the linear DC power flow approximations used in these models do not accurately represent complex network constraints. Several studies have attempted to bridge this gap by incorporating AC optimal power flow (OPF) models for balanced networks with DES design, which are labelled DES-OPF models (De Mel et al., 2021; Mashayekh et al., 2017; Morvaj et al., 2016). Comparisons have shown that solutions from DES-only frameworks can be infeasible with respect to the network. Despite these efforts, all existing studies assume that the overall network is balanced, which is a characteristic of transmission networks. In reality, most DES are connected to unbalanced distribution networks, particularly in European radial distribution networks (Ma et al., 2020). The impacts of considering phase

unbalance in distribution networks and DES design simultaneously have not been investigated before, despite the potential implications on network longevity, integration of renewable resources, and costs to the prosumers. A few studies have incorporated network unbalance within DES operational models. However, the significant gap observed in DES design should also be addressed, as a robust DES design can help alleviate operational issues in advance. Furthermore, despite the growing number of studies on designing and operating DES under uncertainty using techniques such as Robust Optimisation (RO) and Stochastic Programming (Mavromatidis et al., 2018a), detailed DES-OPF models are notably absent from such frameworks. This could be due to the increased complexity of these models, both to formulate and solve. This study aims to bridge these gaps by advancing the modelling framework previously presented by the same authors (De Mel et al., 2021) to incorporate unbalanced power flows within DES design for the first time, and the uncertainty of key inputs using RO (Li et al., 2011). RO was chosen as it can assess worst-case scenarios without assuming probability distributions. Overall, the study contributes to more realistic DES design and operation by improving system feasibility and robustness.

2. Methodology

The bi-level model used in this study first determines the binary topology of the DES design using an MILP, followed by a nonlinear programming (NLP) model which re-optimises the design under unbalanced network constraints (De Mel et al., 2021). Such a strategy is used as it avoids the need to solve a large, nonconvex Mixed-Integer and Nonlinear Programming (MINLP) model, which is difficult to converge and computationally costly. The combined model is labelled DES-UPF. The overall objective, which minimises total annualised cost for a DES with houses $i \in \mathbf{I}$ operating in time $t \in \mathbf{T}$, is transformed to its robust counterpart (Z) to consider uncertainty in electricity and gas prices. A polyhedral uncertainty set (Bertsimas and Sim, 2004) has been chosen for its balanced conservativeness and linearity in the MILP, where an auxiliary variable v_0 and an adjustable sizing parameter Γ_0 are introduced to the reformulation:

$$\min Z$$

$$Z - \left(\left(\sum_{i \in \mathbf{I}, t \in \mathbf{T}} a \bar{P}_t^{\text{grid}} E_{i,t} \right) + \left(\sum_{i \in \mathbf{I}, t \in \mathbf{T}} b \bar{C}^{\text{gas}} H_{i,t} \right) + C^{\text{IN,OM}} - I \right) - v_0 \Gamma_0 \geq 0 \quad (1)$$

$$v_0 \geq a \hat{P}_t^{\text{grid}} E_{i,t} \quad \forall i \in \mathbf{I}, t \in \mathbf{T} \quad (2)$$

$$v_0 \geq b \hat{C}^{\text{gas}} H_{i,t} \quad \forall i \in \mathbf{I}, t \in \mathbf{T} \quad (3)$$

where a and b are scalars, \bar{P}_t^{grid} and \bar{C}^{gas} are nominal values of electricity and gas prices, \hat{P}_t^{grid} and \hat{C}^{gas} are their respective perturbations, $E_{i,t}$ and $H_{i,t}$ are variables for electricity and heat produced and/or bought, $C^{\text{IN,OM}}$ represents all other investment and operational costs, and I is the income generated. Uncertainty associated with demands and irradiance exist in the right-hand side coefficient of the associated constraints, and therefore can be represented by the following generalised counterpart:

$$x_{i,t} - \varepsilon \hat{c}_{i,t} \leq \bar{c}_{i,t} \quad \forall i \in \mathbf{I}, t \in \mathbf{T} \quad (4)$$

where $x_{i,t}$ is a variable, $\bar{c}_{i,t}$ is the nominal value of the uncertain parameter, $\hat{c}_{i,t}$ is the perturbation, and ε represents the parameter controlling the size of the uncertainty, which is $\varepsilon \in [0,1]$ for a polyhedral set with one uncertain parameter in the constraint. Note that the sizing parameters can be used to control the level of conservativeness in the decision, i.e., the higher the value taken by the sizing parameter, the more conservative the decision will be. As the upper bounds of the two sizing parameters, Γ_0 and ε , are different, Γ_0 is

fixed to its upper bound and multiplied by ε to ensure that the conservativeness is varied by the same amount across the constraints.

The NLP includes nonconvex constraints to calculate the active P_n^α and reactive Q_n^α power at each phase $\alpha = \{A, B, C\}$ and node $n \in \mathbf{N}$, connected to branch $(n, m) \in \mathbf{L}$:

$$P_{n,t}^\alpha = V_{n,t}^\alpha \sum_{m \in \mathbf{N}} \sum_{\varphi \in \{a,b,c\}} V_{m,t}^\varphi (g_{mn,t}^{\alpha\varphi} \cos(\theta_{n,t}^\alpha - \theta_{m,t}^\varphi) + b_{mn,t}^{\alpha\varphi} \sin(\theta_{n,t}^\alpha - \theta_{m,t}^\varphi)) \quad \forall n \in \mathbf{N}, t \in \mathbf{T} \quad (5)$$

$$Q_{n,t}^\alpha = V_{n,t}^\alpha \sum_{m \in \mathbf{N}} \sum_{\varphi \in \{a,b,c\}} V_{m,t}^\varphi (g_{mn,t}^{\alpha\varphi} \sin(\theta_{n,t}^\alpha - \theta_{m,t}^\varphi) - b_{mn,t}^{\alpha\varphi} \cos(\theta_{n,t}^\alpha - \theta_{m,t}^\varphi)) \quad \forall n \in \mathbf{N}, t \in \mathbf{T} \quad (6)$$

where $g_{mn}^{\alpha\varphi}$ and $b_{mn}^{\alpha\varphi}$ are conductance and susceptance derived from the real and imaginary parts of the complex phase admittance matrix Y , as shown below:

$$Y = \begin{bmatrix} Y_{11}^{ABC} & \dots & Y_{1n}^{ABC} \\ \vdots & \ddots & \vdots \\ Y_{n1}^{ABC} & \dots & Y_{nn}^{ABC} \end{bmatrix} \quad (7)$$

Each element of Y represents a 3×3 matrix, obtained using line parameters and the Approximate Line Model (Grigsby, 2018). The Y submatrices at the transformer primary and secondary connections have been obtained from Chen et al. (1991). The DES and UPF are linked at load (house) nodes, where active and reactive power injections occur at the respective phase, while injections at the other phases remain zero.

3. Case study

A modified version of the IEEE European Low Voltage Test Feeder (IEEE, 2020) has been considered as a realistic unbalanced three-phase test case. A reduced version of the 906-node test feeder is used to demonstrate the proposed framework, as presented in Figure 1, with 22 of the 55 loads and associated key network connections. A Wye-Wye transformer is considered at the slack node, i.e., the point at which the local network is connected to the wider grid. It is assumed that all the consumers in this network are opting to install distributed energy resources, which include solar PVs, lithium-ion batteries, and boilers. All loads are assumed to have a constant power factor of 0.95.

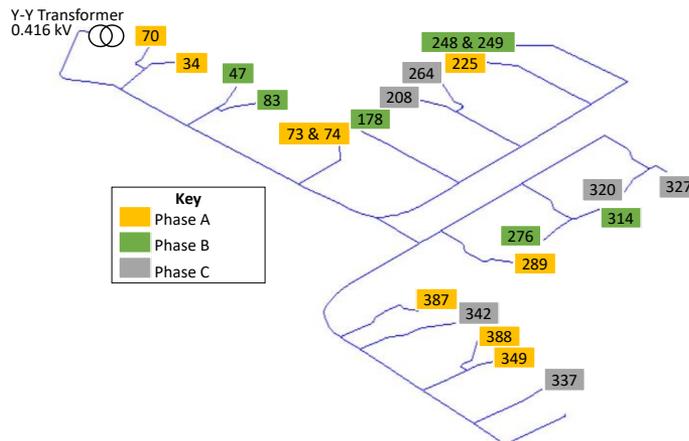


Figure 1. The subset of loads and branches from the IEEE European LV Test Feeder which are used in this study. Note the unequal number of loads connected to each phase.

Averaged demand and irradiance data for four representative seasonal days have been discretised into hourly timestamps. Other inputs include renewable energy generation and export tariffs, day and night electricity purchasing prices, available roof area (35 m²) for PV installation, available volume (0.5 m³) for battery installation, and technology performance parameters and capital costs. Solar irradiance, heat and electricity demand, and pricing are considered as the subset of uncertain inputs. These have been identified by Mavromatidis et al. (2018b) to have the most influence on model outputs, via Global Sensitivity Analysis. As the DES lifespan is considered to be 20 years, harsh daily and seasonal variations are likely to be dampened when averaged over several years. Therefore, overall uncertainty intervals of 20%, 10% and 25% have been assumed for demands, irradiance, and pricing, respectively.

4. Results and Discussion

The MILP and NLP models are solved on Pyomo (Hart et al., 2011) using CPLEX (IBM, 2019) and CONOPT (Drud, 1985), respectively. The overall deterministic NLP contains 386,404 continuous variables and 627,169 constraints, with 4,312 fixed binary variables determined by the MILP.

4.1. Deterministic scenario

Table 1 presents the results for the deterministic scenario, where four models are evaluated: 1) DES-only which represents the MILP, 2) DES-DPF containing the DC power flow approximation, 3) DES-OPF containing balanced AC OPF, and 4) DES-UPF with unbalanced optimal power flow.

Table 1. Comparison of deterministic results for four models.

Breakdown	DES only (MILP)	DES-DPF	DES-OPF	DES-UPF
Objective (£)	21,297	21,297	23,692	21,840
% Difference (Objective)	-	0	11	3
Grid electricity (£)	4,962	4,962	4,876	4,917
PV investment (£)	19,422	19,422	17,789	18,915
PV operation (£)	1,375	1,375	1,259	1,339
Boiler investment (£)	3,359	3,359	3,359	3,359
Boiler operation (£)	13,804	13,804	13,804	13,804
Battery investment (£)	64	64	148	95
Battery operation (£)	27	27	62	39
Export income (£)	6,368	6,368	3,548	5,681
Generation income (£)	15,348	15,348	14,058	14,947
<i>Time taken (s)</i>	<i>23</i>	<i>91</i>	<i>353</i>	<i>1,968</i>

Voltage violations are detected in both DES-OPF and DES-UPF as a consequence of higher power injections from the prosumer due to high generation. The DES-DPF fails to detect this due to the assumption that voltages always remain at nominal values. Significant differences in objective values and designs are observed in DES-OPF when compared with DES-UPF, suggesting that the former overestimates both cost and design for this type of network. This overestimation is a result of a higher number of violations detected in DES-OPF across the network, as all loads are assumed to be connected to a balanced network represented by a single phase. On the other hand, the unbalanced model is capable of distinguishing voltage violations at specific phases and remedying them, rather than reducing PV capacity across most of the network. While the DES-UPF

objective has a low percentage difference when compared with that of DES-only, more prominent design differences can be observed, especially with respect to battery investment as larger batteries reduce network violations. Note that the high computational time recorded in Table 1 for DES-UPF is mostly used for model building on Pyomo, while only approximately 20% of the total time is used by the solver.

4.2. Robust scenario

Figure 2 summarises the percentage differences observed between the DES-only (MILP) model and DES-UPF at different levels of conservativeness. As conservativeness increases, the objective values achieved become more similar, as seen in Figure 2a. Despite these small differences in total annualised costs, Figure 2b shows that differences in design remain prominent at each level. Further tests show that fixing the design as predicted by the MILP would result in greater costs when unbalanced network constraints are included, due to solar power curtailment resulting in loss of export income. The consistently higher battery capacities chosen by the DES-UPF once again help minimise costs while reducing network violations. It is evident that the differences in objectives between the MILP and fixed DES-UPF design reduce as the levels of conservativeness increase. This suggests that MILP designs, despite the lack of detailed network constraints, may be practically feasible with respect to the network if uncertainty is considered with high levels of conservativeness in these formulations, while being less computationally expensive.

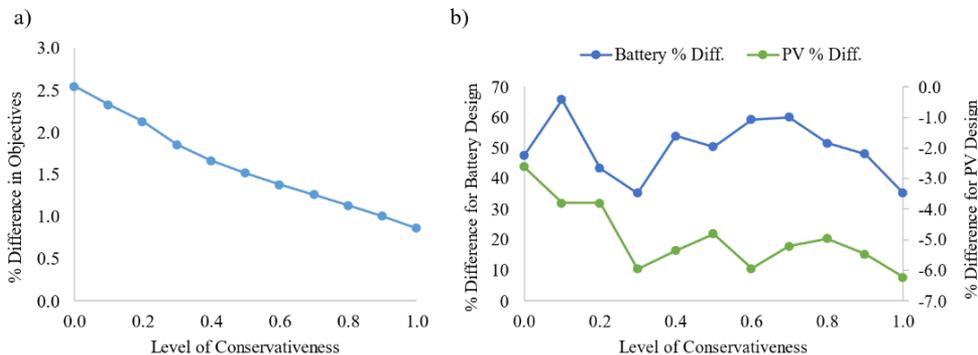


Figure 2. Percentage differences observed with respect to different levels of conservativeness (ϵ) in a) objective values, and b) installed battery and PV capacities.

5. Conclusion

The modelling framework proposed in this work incorporates more realistic power flow constraints associated with unbalanced distributed networks for the first time, to which DES have been increasingly connected. It also uses Robust Optimisation techniques to optimise under uncertainty, which has not been included in detailed and combined models before. The framework is tested using an unbalanced feeder and compared with state-of-the-art models, including DES combined with balanced AC power flow (DES-OPF). Results demonstrate that assuming the network is balanced can lead to oversized DES, much higher capital costs, and less income. They also confirm the unsuitability of DC approximations, as this model fails to detect any network violations. It is evident that the DES-UPF framework proposed in this paper achieves more robust designs at a slightly higher cost when compared to the globally optimal DES-only model. Results also suggest that DES-only and DES-UPF solutions could converge at higher levels of

conservativeness, emphasising the need to consider uncertainty in the design phase. The significantly higher computational expense of this model, compared with state-of-art DES-OPF, poses a major limitation to evaluating larger and more realistic test cases with higher nodes and branches. Therefore, decomposition techniques would be essential to improve the efficiency of this framework and further increase its utility. Other improvements include the addition of more transformer connections and network infrastructure to study potential mitigation strategies at the operational stage. Finally, the study emphasises the need for more investigations using high-fidelity models to obtain locally optimal but more robust DES designs while considering existing infrastructure and networks. This would give both network operators and prosumers greater confidence in implementing DES, and jointly address the need for more renewable energy resources.

References

- D. Bertsimas, M. Sim, 2004, The Price of Robustness. *Oper. Res.* 52, 35–53
- T. H. Chen, M. S. Chen, T. Inoue, P. Kotas, and E. A. Chebli, 1991, Three-phase cogenerator and transformer models for distribution system analysis, *IEEE Trans. Power Deliv.*, vol. 6, no. 4, pp. 1671–1681.
- I. A. De Mel, O. V. Klymenko, and M. Short, 2021, Levels of Approximation for the Optimal Design of Distributed Energy Systems, *Comput. Aided Chem. Eng.*, vol. 50, pp. 1403–1408.
- A. Drud, 1985, CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems, *Math. Program.*, vol. 31, no. 2, pp. 153–191.
- L. L. Grigsby, 2018, *Electric Power Generation, Transmission, and Distribution: The Electric Power Engineering Handbook*, CRC Press.
- W. E. Hart, J. P. Watson, and D. L. Woodruff, 2011, Pyomo: Modeling and solving mathematical programs in Python, *Math. Program. Comput.*, vol. 3, no. 3, pp. 219–260.
- IBM, 2019, IBM ILOG CPLEX 12.9 User’s Manual, https://www.ibm.com/support/knowledgecenter/SSSA5P_12.9.0/ilog.odms.cplex.help/CPLEX/homepages/usrmanplex.html.
- IEEE, 2020, IEEE PES Test Feeder, <https://cmte.ieee.org/pes-testfeeders/resources/>
- Z. Li, R. Ding, and C. A. Floudas, 2011, A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: I. Robust Linear Optimization and Robust Mixed Integer Linear Optimization, *Ind. Eng. Chem. Res.*, vol. 50, no. 18, pp. 10567–10603.
- K. Ma, L. Fang, and W. Kong, 2020, Review of distribution network phase unbalance: Scale, causes, consequences, solutions, and future research directions, *CSEE J. Power Energy Syst.*, vol. 6, no. 3, pp. 479–488.
- S. Mashayekh, M. Stadler, G. Cardoso, and M. Heleno, 2017, A mixed integer linear programming approach for optimal DER portfolio, sizing, and placement in multi-energy microgrids, *Appl. Energy*, vol. 187, pp. 154–168.
- G. Mavromatidis, K. Orehounig, and J. Carmeliet, 2018a, A review of uncertainty characterisation approaches for the optimal design of distributed energy systems, *Renew. Sustain. Energy Rev.*, vol. 88, pp. 258–277.
- G. Mavromatidis, K. Orehounig, and J. Carmeliet, 2018b, Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems, *Appl. Energy*, vol. 214, pp. 219–238.
- B. Morvaj, R. Evins, and J. Carmeliet, 2016, Optimization framework for distributed energy systems with integrated electrical grid constraints, *Appl. Energy*, vol. 171, pp. 296–313.

SCR: A novel surrogate-based global optimization algorithm for constrained black-box problems

S. A. Zaryab^a, A. Manno^b, E. Martelli^{a*}

^a, *Politecnico di Milano, Department of Energy, Via Lambruschini 4, Milan 20156, Italy*

^b, *Università degli Studi dell'Aquila, Department of Information Engineering, Computer Science and Mathematics, Via Vetoio, 67100, L'Aquila, Italy*

^{*}, emanuele.martelli@polimi.it

Abstract

This paper proposes a novel optimization algorithm for constrained black-box problems, where the objective function and some constraints are computed by a simulation code. The basic idea of the optimization algorithm, referred to as SCR (Surrogate-CMAES-RQLIF), is to (i) build separate Kriging surrogates for the objective function and black-box constraints, (ii) use the global search algorithm CMAES to find the global optimum region of the surrogate, (iii) use the recent algorithm RQLIF to refine the search locally, (iv) use all the points sampled by RQLIF and additional points within the optimal region located by CMAES to update the surrogate model. Tests on 46 constrained and unconstrained test problems show that SCR outperforms the benchmark algorithms in terms of fraction of problems solved, specially at low function evaluations (< 300).

Keywords: process optimization, surrogate-based optimization, Black-box optimization, Global optimization.

1. Introduction

A vast range of process engineering and economic problems [1] can be tackled as black box optimization problems where the objective function and the constraints are computed by a simulation code (e.g., flow sheeting software, CFD code, etc). The simulation code works as a noisy and computationally expensive black-box function called by the optimization algorithm. In such optimization problems gradient information are not available so gradient-based algorithms are not suitable, and it is necessary to rely on derivative-free methods [2]. Moreover, simulations codes are computational expensive making it necessary to reduce the number of black-box function evaluations. Thus, an efficient derivative-free algorithm is required to optimize these simulation codes with a limited number of function evaluations, this can be done by creating a surrogate model of the simulation code.

The goal of this paper is to present a novel derivative-free, surrogate-based global optimization algorithm called SCR which is well suited for black-box problems with general nonlinear constraints. The problem that we consider is of the form

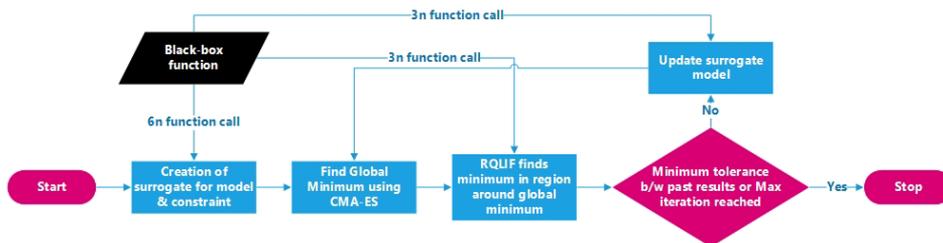
$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s. t. : } & lb \leq x \leq ub \\ & h(x) = 0 \\ & g(x) \leq 0 \end{aligned}$$

where lb , ub are the bounds, g is the vector of inequality constraints, h is the vector of equality constraints and f is the objective function. One or all the functions may be nonlinear, nonconvex, computationally expensive, non-smooth and noisy black-box functions (e.g., computed by a process simulation code).

2. SCR algorithm

SCR (Surrogate-CMAES-RQLIF) is a surrogate based derivative-free optimization algorithm which is suitable for problems with expensive and possibly noisy black-box functions. The flowchart of SCR is provided in Figure 1 and summarized below in four steps.

- 0) Initially $6n$ (where n is the number of optimization variables) random points are generated using latin hypercube.
- 1) Separate Kriging surrogate models of the black box problem and its constraints are created using the Matlab based Kriging toolbox DACE [3].
- 2) The surrogates of the objective function and constraints are then used by the evolutionary algorithm CMAES [4] to find the global minimum of the constrained surrogate model. The constraints and objective surrogates are combined in CMAES using the quadratic penalty approach.
- 3) If the optimum found by CMAES is relatively close to the best solution found in the previous iterations (i.e., there is the need of performing a local search in that region), the local search algorithm RQLIF [5] is called with a limit of $3n$ black-box function evaluations. RQLIF is a hybrid implicit-filtering-model based (using a regularized quadratic model) algorithm specifically developed for expensive black-box functions and tested on a wide range of test and engineering problems.
- 4) At the end of the RQLIF run, the surrogate models are updated by adding the points tested by RQLIF along with $3n$ new points found using latin hypercube. The algorithm stops if RQLIF reaches the convergence tolerance on the search step length, or the maximum number of black-box evaluations is reached. If none of these stopping criteria is met, the algorithm begins a new iteration from Step 2.



3. Assessment of SCR performance on test functions

The Matlab implementation of SCR was tested in comparison with two well-known derivative free algorithms: NOMAD [6] and CMAES [4]. These algorithms were selected from many other alternatives (see [7] for a review) as they are widely used and very effective, as reported in literature [5] [8]. Furthermore, both these algorithms use different strategies, NOMAD samples points on an adaptive mesh grid using direct search methods

while CMAES uses an evolutionary optimization strategy with covariance matrix adaptation.

The tests were carried out on 25 unconstrained nonlinear test problems taken from [9] and 21 constrained nonlinear test problems taken from GLOBALlib and Floudas' Collection [10]. These test problems feature 2-13 real variables and up to 17 constraints. Each run of each algorithm was repeated ten times starting from initialization points was found using Latin Hyper Cube Method. The average results of the ten runs are used to plot the performance profiles reported in Figure 2. The performance parameter shows the fraction of problems solved within a certain tolerance “ τ ”, as defined by [11], with the maximum allowed budget of black-box function evaluations.

Figure 2 shows that SCR outperforms the benchmark algorithms in terms of fraction of problems solved for both constrained and unconstrained problems. The figure also shows that SCR performs quite well at low function evaluations (< 300) in terms of fraction of problems solved and solution quality.

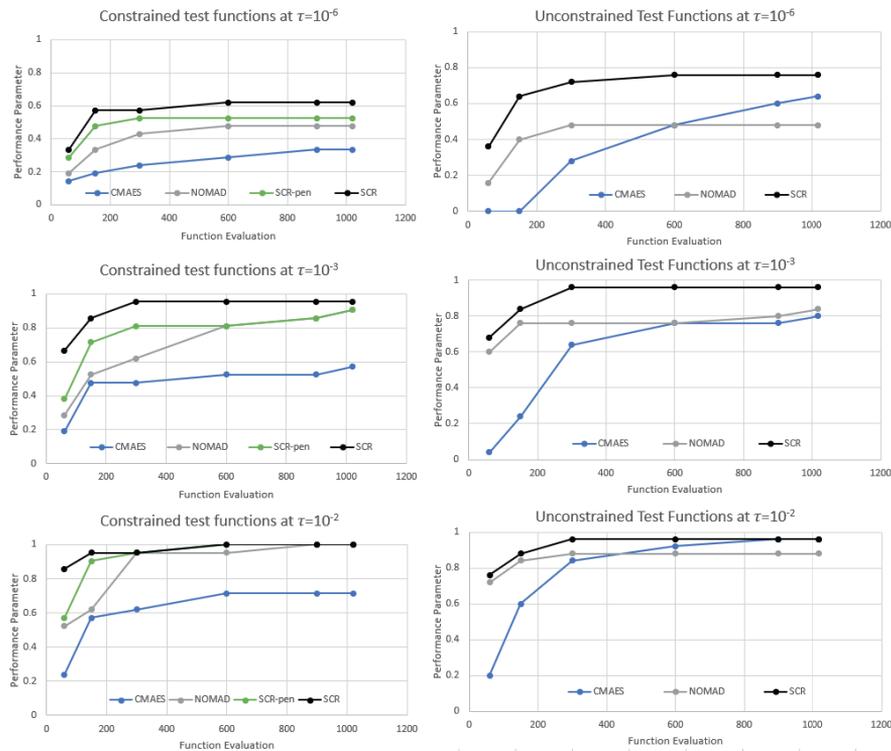


Figure 2: Comparison of SCR with benchmark algorithms

Figure 2 reports the performance of a modified version of SCR (called SCR-pen) which generates directly the surrogate of the penalized function (obj. function + penalty term proportional to the quadratic violation of constraints). Compared to directly generating the surrogate of the penalized objective function (SCR-pen), the strategy of SCR of generating separate surrogates for the black-box function and each of its constraints (g and h) leads to the following main advantages: (i) the surrogates are more accurate since the constraint functions do not have steep valleys/curvatures like those of the quadratic penalty term, (ii) the surrogate generation algorithm and the optimization algorithm do

not incur in ill-conditioning problems caused by the quadratic penalty term outside the feasible region.

4. Optimization of a CO₂ Purification Unit

SCR was applied to the techno-economic optimization of a CO₂ Purification unit (CPU) developed for the purification of CO₂ captured from a Cement Plant [12]. The CPU scheme is reported in Figure 3 and it is based on a patent by Air Products [13], in the version published by [14]. The CPU model was developed in Aspen Plus® and its convergence needs 10-20 seconds depending on the input conditions. The optimization problem is to minimize the total annual cost (sum of annualized capital cost and operating costs) subject to four main nonlinear inequality constraints: (i) the recovery of CO₂ should be greater than 95%, (ii) the purity of CO₂ should be greater than 98%, (iii) the oxygen in the outlet stream should be less than 75 ppm, and (iv) the mole fraction of nitrogen in the outlet stream should be less than 3%.

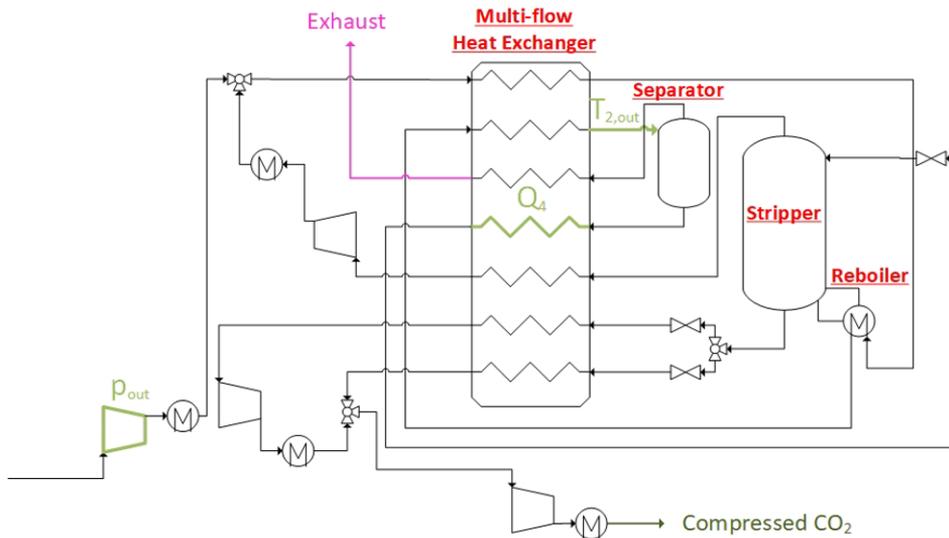


Figure 3: Scheme of CO₂ purification unit (CPU)

The process variables that are optimized (shown in green in Figure 3) are (1) p_{out} , the compressor outlet pressure, (2) Q_4 , the heat absorbed by the liquid stream coming from the separator and going to the stripper, and (3) $T_{2,out}$, the temperature of the inlet stream of the separator.

Figure 4 shows the convergence plot (objective function value of the best-found solution as a function of the number of simulation runs) of SCR and NOMAD. The figure indicates that SCR is able to find a good solution with only 50 simulation calls, and further improving it with the successive 150 calls. However, NOMAD was able to find a similar solution after 400 calls. The comparison of the yearly capital cost of each component and the total yearly operational cost of the CPU for the optimized case for SCR and NOMAD is provided in Table 1. The values of the main process parameters affected by the optimization process are shown in Table 1.

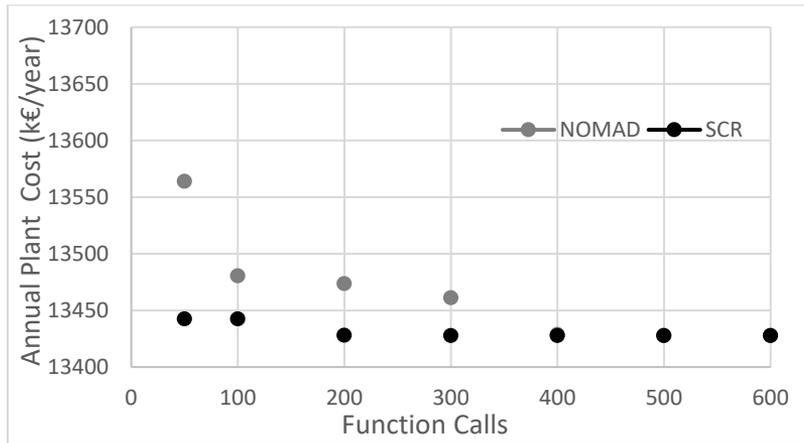


Figure 4: Comparison between SCR and NOMAD for the considered case study

Table 1: Main process parameters and the yearly cost of CPU components optimized with SCR and NOMAD

	SCR	NOMAD
p_{out} (bar)	38.3	37.7
$T_{2,out}$ ($^{\circ}C$)	-44.3	-44.6
Q_4 (kW)	564.8	562.3
CO ₂ Recovery (captured CO ₂ /inlet CO ₂)	95.00%	95.00%
Purity (mol. conc. of CO ₂ in captured stream)	99.99%	99.99%
Oxygen conc. in captured stream (ppm)	7.26	7.13
N ₂ +AR conc. in captured stream	2.95E-06	2.88E-06
Total Area of main HX (m ²)	892.6	948.2
Compressor power (kW)	11206.2	11171.2
Costs		
Annualized compressors cost (k€/year)	5678	5664
Annualized Intercoolers cost (k€/year)	305	305
Annualized Heat Exchanger cost (k€/year)	865	900
Other Equipment cost (k€/year)	126	125
Yearly Operational Cost (k€/year)	6453	6433
Total yearly cost	13428	13428

5. Conclusion

SCR is a surrogate based derivative free algorithm which is well suited for black-box optimization problems with expensive function evaluations. SCR creates surrogate model of the black-box problem and its constraints to accelerate convergence and handle the noise. Performance profiles of SCR compared with two well-known optimization algorithms (CMAES & NOMAD) indicate that SCR provides good quality solutions in limited number of functional evaluations on a large set of benchmarking test problems.

When applied to a real engineering black-box optimization problem, the optimization of a CO₂ purification process, SCR needs about half simulation runs (about 200) compared to NOMAD to find the same optimum solution.

References

- [1] C. Audet and M. Kokkolaras, "Blackbox and derivative-free optimization: theory, algorithms and applications," *Optim Eng*, no. 177, pp. 1-2, 2016.
- [2] A. Conn, K. Scheinberg and L. Vicente, Introduction to derivative-free optimization, vol. 8, Philadelphia: SIAM, 2009.
- [3] S. N. Lophaven, H. B. Nielsen and J. Sondergaard, "Dace: a MATLAB Kriging toolbox.," Technical University of Denmark: Technical Report No. IMM-TR-2002-12., Kongens Lyngby, 2002.
- [4] N. Hansen and S. Kern, "Evaluating the CMA Evolution Strategy on Multimodal Test Functions.," *Eighth International Conference on Parallel Problem Solving from Nature PPSN*, vol. VIII, no. Berlin: Springer, pp. 282-291, 2004.
- [5] A. Manno, E. Amaldi, F. Casella and E. Martelli., "A local search method for costly black-box problems and its application to CSP plant start-up optimization refinement," *Optimization and Engineering*, 2020.
- [6] S. Le Digabel, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software (TOMS)*, vol. 37(4), pp. 1-15, 2011.
- [7] L. Rios and N. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," *J Global Optim*, vol. 56(3), pp. 1247-1293, 2013.
- [8] E. Martelli and E. Amaldi, "PGS-COM: A hybrid method for constrained non-smooth black-box optimization Problems: brief review, novel algorithm and comparative evaluation," *Computers & Chemical Engineering*, no. 63, pp. 108-139, 2014.
- [9] M. Jamil and X.-S. Yang, "A Literature Survey of Benchmark Functions For Global Optimization Problems," *Int. Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, no. 2, pp. 150-194, 2013.
- [10] C. C. -. U. t. Technology, "GLOBALLib and Floudas' Collection," [Online]. Available: https://www.mat.univie.ac.at/~neum/glopt/coconut/Benchmark/Library1_new_v1.html. [Accessed 14 August 2020].
- [11] E. Dolan and J. Mor'e, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91(2), pp. 201-213, 2002.
- [12] M. Voldsund, R. Anantharaman, D. Berstad, E. D. Lena, C. Fu, O. S. Gardarsdottir, A. Jamali, J.-F. Pérez-Calvo, M. Romano, S. Roussanaly, J. Ruppert, O. Stallmann and D. Sutter, "D4.6 CEMCAP comparative techno-economic analysis of CO2 capture in cement plants," DOI: 10.5281/zenodo.2597090, 2019.
- [13] V. White and R. Allam, "Purification of carbon dioxide". Patent US 2008/0173584A1, 2008.
- [14] R. Strube and G. Manfrida, "CO2 capture in coal fired power plants — Impact on plant performance," *International Journal of Greenhouse Gas Control*, vol. 5, pp. 710-726, 2011.

Comparative Life Cycle Assessment of Demand-Side Management via Operational Optimization

Benedikt Nilges^a, Célia Burghardt^a, Kosan Roh^{b,c}, Christiane Reinert^a, Niklas von der Aßen^{a,*}

^a*Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstraße 8, 52062 Aachen, Germany*

^b*Process Systems Engineering (AVT.SVT), RWTH Aachen University, Forckenbeckstraße 51, 52074 Aachen, Germany*

^c*Department of Chemical Engineering and Applied Chemistry, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Republic of Korea*

*Corresponding author: niklas.vonderassen@ltt.rwth-aachen.de

Abstract

Due to the volatile nature of renewable energy sources, balancing electricity generation and consumption in electric power systems becomes more challenging with an increasing share of renewable generation technologies. A promising option to face fluctuations in the electricity grid is demand-side management (DSM), the adjustment of electricity consumption driven by varying electricity prices. To assess the environmental impacts of DSM, the holistic method of life cycle assessment can be applied. In literature, life cycle assessment of DSM receives increased attention. However, the goals of operating DSM are not consistently reflected in the definition of the functional unit. In our view, the goals of DSM include both the original products of a process as well as the flexible electricity consumption. So far, no approach exists that considers both goals in the functional unit.

This work aims to close this gap by developing an approach that jointly considers products and flexibility in the functional unit of DSM. We compare DSM with a steady-state operation of the process as a reference system, given that both systems provide the same products. Furthermore, we expand the reference system with energy storage technologies, which enable short- and long-term flexibility in electricity consumption. We demonstrate the proposed approach on the switchable chlor-alkali electrolysis, using hourly electricity prices and emission factors. Our findings show that comparing the switchable electrolysis to the reference system indicates only small differences in the environmental impacts. Thus, the results indicate that the flexible operation of the chlor-alkali electrolysis has comparable environmental impacts as integrating electricity storage.

Keywords: functional unit, flexibility, chlor-alkali electrolysis

1. Introduction

The integration of renewable energies into the electricity grid aims to reduce greenhouse gas (GHG) emissions. However, due to the volatile nature of renewable energies, the temporal mismatch between demand and supply of electricity is likely to increase. This mismatch can be reduced by demand-side management (DSM) - the load management of

processes or appliances to reduce the overall energy costs (Warren, 2014). To analyze if DSM is suitable to stabilize the electricity grid while avoiding significant environmental impacts, the environmental assessment of DSM is necessary. Currently, the environmental impacts of DSM receive increased attention in literature (e.g., Schäfer et al., 2020). For the environmental assessment of DSM, the method of life cycle assessment (LCA) can be applied. LCA is a holistic method, covering potential environmental impacts over the entire life cycle of a product system (International organization for standardisation, 2021). The LCA is conducted by comparing DSM with a reference operation mode. For this comparison, it is essential that both modes of operation provide the same functions, e.g., the production of a certain product. These functions are defined and quantified by the functional unit. However, concerning the environmental assessment of DSM, the functional unit differs in literature. Some studies focus on the products of the DSM and consider the flexible electricity consumption as a secondary benefit to reduce process costs (Walzberg et al., 2019). Other studies consider the grid service due to the flexible electricity consumption as the only function, especially when non-industrial DSM is considered (Milovanoff et al., 2018). In the environmental assessment of industrial DSM, the products should be integrated into the functional unit since they are essential for downstream processes. Furthermore, considering the importance of system services to the electricity grid, the flexible electricity consumption of DSM should also be included in the functional unit.

In this work, we present an approach to integrate both the products and the flexible electricity consumption in the functional unit of DSM. The approach is introduced and applied to the switchable chlor-alkali electrolysis, which is a potential DSM process in the chemical industry.

2. LCA of demand-side management processes

The proposed approach aims to assess environmental impacts of DSM via a comparative LCA, in which products and flexible electricity consumption are included in the functional unit. Since industrial processes usually supply downstream processes, we define a constant supply rate of the products in the functional unit. Therefore, the DSM system generally needs an additional product storage to enable a constant supply rate. The reference system for comparison comprises the conventional steady-state operation, providing a constant supply of products. However, since the steady-state operation cannot provide a flexible electricity consumption, the reference system is expanded with energy storage units. To consider different time periods of flexibility, short-term as well as long-term storage technologies are used for the expansion. The power and capacity of these storage units depend on the flexibility defined in the functional unit. Since DSM is assessed, we define the flexible electricity consumption by the DSM as the required flexibility. But quantifying this flexibility is challenging because it depends on the flexible power that the DSM can provide and the duration of provided flexibility. Thus, we propose the following four-step approach to determine the environmental impacts of the DSM and the reference system while guaranteeing the same flexibility (Figure 1):

1. To quantify the flexibility of the DSM, the operation profile of the process is required. For this, the operation of the DSM process is optimized while minimizing operating costs. The time horizon of this optimization covers one year so that seasonal variations in DSM are considered. The time resolution of the optimization should be defined in accordance with the ramping constraints of the DSM. From the results of this optimization, the electrical load profile of the DSM process is obtained. The derivative of this load profile, i.e., the

marginal electricity consumption, characterizes the flexibility of the DSM. In the functional unit, flexibility is defined as the potential to provide the same flexibility as the DSM, thus the potential to operate with the same marginal electricity consumption.

2. In the reference system, the steady-state operation and the storage capacities need to be designed so that they can provide the products and flexibility as stated in the functional unit. For this, an optimization minimizing design and operational costs of the reference system is conducted. Since the reference system should provide the same flexibility as the DSM system, the derivative of the electrical load profile is predefined for the reference system by the derivative of the DSM system as obtained in Step 1. The same derivative results in the same shape of the electric load profile. Although the shape of electricity consumption is predefined, it does not restrict the total amount of consumed electricity, i.e., there can be a constant offset in electricity consumption between the DSM and the reference system. The amount of the constant offset depends on the process and storage efficiencies. Furthermore, the constant offset enables the assessment of different DSM types such as load shifting, peak clipping, or valley filling.
3. Subsequent to the design optimization, the reference system can provide the same flexibility as the DSM. However, the restriction of the derivative of electricity consumption can lead to a suboptimal economic operation of the reference system. Economically, the reference system would probably be operated differently due to the diverse efficiencies and ramping constraints of the storage technologies. Thus, an optimization minimizing the operational costs of the reference system with the design results of step 2 is conducted while providing the same supply rate as defined in the functional unit.
4. As a last step, the life cycle impact assessment of both systems is conducted. For this, the respective operation determined in Steps 1 and 3 is applied to calculate the environmental impacts in each time step. These impacts are accumulated for one year of operation. Furthermore, the construction and disposal impacts of the production and storage facilities are included to consider the complete life cycle of both systems. The construction impacts of each component are proportionally

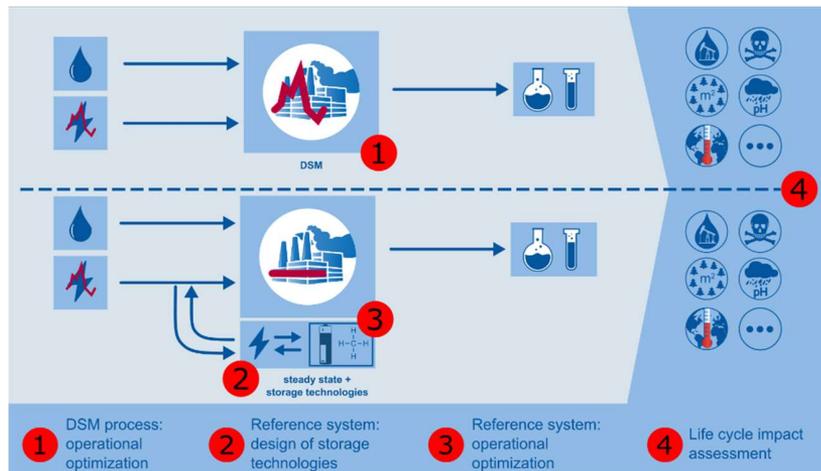


Figure 1: Life cycle assessment of demand-side management (DSM) with proposed four step approach.

allocated to one year, depending on the lifetime of the component. Finally, the impacts of operation and construction are added, and both systems are compared.

3. LCA of the switchable chlor-alkali electrolysis

The presented approach is applied for the environmental assessment of the switchable chlor-alkali electrolysis. Chlor-alkali electrolysis is an essential process in the chemical industry, producing the bulk chemicals chlorine and sodium hydroxide. The switchable chlor-alkali electrolysis can apply DSM by utilizing a bifunctional electrode which allows changing operation between two different modes (Brée et al., 2018). The standard mode converts a brine solution into chlorine, sodium hydroxide, and hydrogen (H₂-mode). The typical electricity consumption ranges between 2.1 and 3.0 kWh per kg of produced chlorine. The second mode requires oxygen as an additional input and suppresses the hydrogen production (O₂-mode). This mode consumes around 30% less electricity. When switching between these modes, cleaning the electrolysis cell is necessary to avoid oxyhydrogen reaction. The cleaning leads to downtimes in the production. Therefore, to enable a constant supply of products, the electrolysis is capable of overproduction, and a small intermediate storage for the products is integrated into the system.

The mathematical model of the switchable chlor-alkali electrolysis is adapted from Roh et al. (2019). The model is formulated as a mixed-integer linear program (MILP). Since the downtimes due to the cleaning of the electrolysis cell do not represent the controllable flexibility of the DSM, the downtimes and the overproduction are neglected for the determination of the flexibility in the functional unit. The supply of products defined in the functional unit is 1 kg chlorine and 1.128 kg sodium hydroxide per hour over the time horizon of one year. The produced hydrogen is assumed to be burned for supplying heat to other processes, which is often applied in the chemical industry (Jörissen et al., 2011). Therefore, a heat supply is also defined in the functional unit by multiplication of the produced hydrogen and its lower heating value.

In the reference system, the H₂-mode is chosen as the steady-state process since it is the conventional process of the chlor-alkali electrolysis. Because the H₂-mode consumes more electricity than the O₂-mode, the DSM application represented by the switchable chlor-alkali electrolysis is peak-clipping. The short-term storage technology considered for the system expansion of the reference system is a vanadium redox flow battery. This type of battery was chosen because it allows a separate design of storage power and capacity. For long-term storage, the power-to-methane-to-power technology is considered. This technology consists of water electrolysis, a methanation plant, a methane storage tank, and a natural gas turbine. Although this technology is immature, it is one of the most promising options for long-term storage of electricity without being regionally bounded (Jülch, 2016).

The exogenous electricity price is obtained from Bundesnetzagentur | SMARD.de (2020) and represents the German electricity prices on the wholesale market in 2019. The electricity price is increased by 1.5ct/kWh, representing levies and taxes for the energy intensive industry. For the environmental assessment, the hourly grid mix emissions are calculated with emission factors of the generation technologies provided by ecoinvent 3.6 (Wernet et al., 2016) and the hourly mix of generating technologies, which is also obtained from Bundesnetzagentur | SMARD.de (2020). The environmental impacts of the chlor-alkali electrolyzer construction are integrated by the work of Jung et al. (2014). For the calculation of the vanadium redox flow battery's construction impacts, the modelling of Weber et al. (2018) is applied. Construction impacts of the water electrolysis are taken from Baumgärtner et al. (2021). All other construction impacts are modeled with the life

cycle inventory database ecoinvent 3.6 (Wernet et al., 2016). The considered life cycle impact assessment method is Environmental Footprint 2.0, which is recommended by the European Commission's Joint Research Center (Sala et al., 2014).

4. Results of the switchable chlor-alkali electrolysis

The operational optimization of the switchable chlor-alkali electrolysis shows that the electrolysis operates more often in the H₂-mode during the year (from 22.7% to 73.6% per month, 56.4% annual average, see Figure 2). The O₂-mode is the preferred mode during January and February, since the electricity price in these months is higher than in the summer. The DSM results in a seasonal variation of electricity demand, which is represented in the functional unit and favors a system expansion with long-term storage in the reference system. Actually, the power-to-methane-to-power facility provides all the flexibility with a charging power of 0.47 kW, discharging power of 0.16 kW and a methane storage of 25.94 kg, representing an electrical storage capacity of 360.29 kWh.

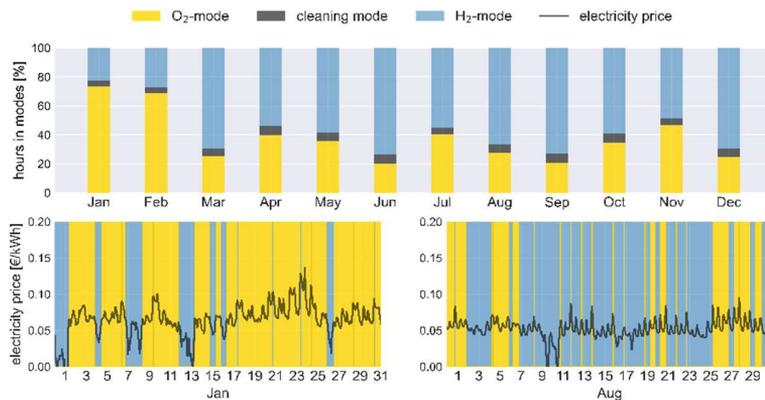


Figure 2: Operation overview of the switchable chlor-alkali electrolysis for the complete year (top plot) and detailed operation for January and August (bottom plots).

The results of the life cycle impact assessment show that DSM of the switchable chlor-alkali electrolysis leads to similar environmental impacts compared to the reference system in all impact categories except human toxicity (non-cancer). The reasons for the similar environmental impacts are the high impacts of sodium chloride and electricity in the operation of the electrolysis. Impacts due to construction are mostly negligible in comparison to the impacts of process operation. However, the impact in human toxicity (non-cancer) is remarkably higher for the reference system, because of the high impact of storage construction in this category. As the results are mostly in the range of impact assessment uncertainties, no recommendations for operating the chlor-alkali electrolysis can be given. However, the results indicate that applying DSM for the chlor-alkali electrolysis has comparable environmental impacts as integrating storage technologies into the grid to match the electricity supply with its demand.

5. Conclusions

In this paper we propose an approach to assess the environmental impacts of DSM via LCA by integrating flexibility in the functional unit. The approach enables a comparison between applying DSM and the utilization of storage technologies to provide flexibility. We demonstrate our approach on the environmental assessment of the switchable chlor-alkali electrolysis. The environmental impacts of the switchable chlor-alkali electrolysis

and the reference system are mostly in the range of impact assessment uncertainties. However, the results indicate that applying DSM for the chlor-alkali electrolysis has comparable environmental impacts as integrating storage technologies.

Acknowledgments

BN and KR gratefully acknowledge the financial support of the Kopernikus project SynErgie (funding number 03SFK3L1-2) by the Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich.

References

- Baumgärtner, Nils; Deutz, Sarah; Reinert, Christiane; Nolzen, Niklas; Kuepper, Lucas Elias; Hennen, Maike et al. (2021): Life-Cycle Assessment of Sector-Coupled National Energy Systems: Environmental Impacts of Electricity, Heat, and Transportation in Germany Till 2050. In: *Front. Energy Res.* 9, S. 886. DOI: 10.3389/fenrg.2021.621502.
- Brée, Luisa C.; Perrey, Karen; Bulan, Andreas; Mitsos, Alexander (2018): Demand Side Management and Operational Mode Switching in Chlorine Production. In: *AIChE J* (65). DOI: 10.1002/aic.16352.
- Bundesnetzagentur | SMARD.de (2020): Strommarktdaten, Stromhandel und Stromerzeugung in Deutschland. Bundesnetzagentur. Available online at <https://www.smard.de/home>, last accessed on 27.10.2021.
- International organization for standardisation ISO 14040, 2021: DIN ISO 14040.
- Jörissen, Jakob; Turek, Thomas; Weber, Rainer (2011): Chlorherstellung mit Sauerstoffverzehrkatoden. In: *Chemie in unserer Zeit* 45 (3), S. 172–183. DOI: 10.1002/ciuz.201100545.
- Jülich, Verena (2016): Comparison of electricity storage options using leveled cost of storage (LCOS) method. In: *Applied Energy* 183, S. 1594–1606. DOI: 10.1016/j.apenergy.2016.08.165.
- Jung, Johannes; Postels, Sarah; Bardow, André (2014): Cleaner chlorine production using oxygen depolarized cathodes? A life cycle assessment. In: *Journal of Cleaner Production* (80), S. 46–56. DOI: 10.1016/j.jclepro.2014.05.086.
- Milovanoff, Alexandre; Dandres, Thomas; Gaudreault, Caroline; Cheriet, Mohamed; Samson, Réjean (2018): Real-time environmental assessment of electricity use: a tool for sustainable demand-side management programs. In: *Int J Life Cycle Assess* 23 (10), S. 1981–1994. DOI: 10.1007/s11367-017-1428-2.
- Roh, Kosan; Brée, Luisa C.; Perrey, Karen; Bulan, Andreas; Mitsos, Alexander (2019): Flexible operation of switchable chlor-alkali electrolysis for demand side management. In: *Applied Energy* 255 (c), S. 113880. DOI: 10.1016/j.apenergy.2019.113880.
- Sala, Serenella; Pant, Rana; Hauschild, Michael Zwicky, Pennington, David (2014): Recommendations for Life Cycle Impact Assessment in the European context.
- Schäfer, Pascal; Daun, Torben M.; Mitsos, Alexander (2020): Do investments in flexibility enhance sustainability? A simulative study considering the German electricity sector. In: *AIChE J* 66 (11), S. 1. DOI: 10.1002/aic.17010.
- Walzberg, Julien; Dandres, Thomas; Merveille, Nicolas; Cheriet, Mohamed; Samson, Réjean (2019): Accounting for fluctuating demand in the life cycle assessments of residential electricity consumption and demand-side management strategies. In: *Journal of Cleaner Production* 240 (2), S. 118251. DOI: 10.1016/j.jclepro.2019.118251.
- Warren, Peter (2014): A review of demand-side management policy in the UK. In: *Renewable and Sustainable Energy Reviews* 29 (4), S. 941–951. DOI: 10.1016/j.rser.2013.09.009.
- Weber, Selina; Peters, Jens F.; Baumann, Manuel; Weil, Marcel (2018): Life Cycle Assessment of a Vanadium Redox Flow Battery. In: *Environmental science & technology* 52 (18), S. 10864–10873. DOI: 10.1021/acs.est.8b02073.
- Wernet, Gregor; Bauer, Christian; Steubing, Bernhard; Reinhard, Jürgen; Moreno-Ruiz, Emilia; Weidema, Bo (2016): The ecoinvent database version 3 (part I): overview and methodology. In: *Int J Life Cycle Assess* 21 (9), S. 1218–1230. DOI: 10.1007/s11367-016-1087-8.

A rigorous synthesis and optimal design methodology for chemical and biochemical processes

Jorge Chavez-Salas^a, Fernando Israel Gómez-Castro^a, Ricardo Morales-Rodriguez^{a*}

^a*Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato, Gto., 36050, México.*
ricardo.morales@ugto.mx

Abstract

This study presents a systematic methodology for process synthesis based on superstructure optimization, which identifies the optimal configuration of a biorefinery to produce high value chemical and biochemical products, by maximizing product profits and minimizing process energy costs and the investment costs. The case study employed agave bagasse as a feedstock. The results indicated that furfural production is the best option, followed by lactic acid production as the second best and, ethanol, ethylene glycol and diethylene glycol production was the third best option, relying on the objectives previously mentioned. The optimization was performed using a variant of genetic algorithms with multi-objective approach (Gamultiobj), through a COM® interface linking the computational tools such as, Aspen Plus and MATLAB, which allowed to perform a rigorous energy balance calculation employing the proper thermodynamic models.

Keywords: superstructure, optimization, agave bagasse, genetic algorithms.

1. Introduction

The production of biofuels and/or high value chemical products derived from lignocellulosic biomass is one of the alternatives that has been studied and developed in recent years (Ren et al., 2009). These chemical compounds can be produced from a variety of non-food crops, such as lignocellulosic residues from agriculture, forestry, and food industry, which has a direct impact on reducing dependence on fossil resources. Lignocellulosic biomass is mainly composed by three biopolymers: cellulose, hemicellulose, and lignin. The cellulose and hemicellulose fractions can be subjected to diverse chemical and/or biological processes and a wide variety of compounds can be obtained. There is great interest in proposing new optimal industrial processes that use lignocellulosic biomass as feedstock, with the purpose of producing high value chemical compounds and/or biofuels, through production platforms called biorefineries.

A superstructure represents all possible pathways by synthesizing distinctive alternatives at different stages of the processing network to optimize a series of possible processing approaches (Quaglia et al., 2015). This methodology has been widely used to optimize multi-criteria processing alternatives, for example, to select the optimal pathway for the conversion of microalgae to biodiesel (Rizwan et al., 2013), among others. Although the results were promising in terms of different parameters such as economic and environmental, the study did not cover other aspects, such as rigorous energy balances

analysis, which is a paramount aspect that must be included in the multidimensional analysis for finding optimal (bio)process routes.

The optimization task in process synthesis employing superstructures is another issue to solve. Some computer-aided tools have been employed to solve that kind of problems, but it is necessary to represent and develop the mass balance process model and most of the time, energy balance is rarely included (Rizwan et al., 2013). On the other hand, the use of ActiveX-OLE(COM)[®] technology allows the implementation of an interface between MATLAB R2017a[®] and Aspen Plus V.8.8[®] (Darkwah et al., 2018; Ponce-Rocha et al., 2021). The utilization of the NSGA-II algorithm is a well-established multi-objective evolutionary algorithm, which has been successfully implemented in a wide range of applications (Punnathanam and Kotecha, 2016). Matlab [®] offers a suite focused on multi-objective optimization (Gamultiobj) to create a set of points on the Pareto front. Gamultiobj uses a controlled and elitist genetic algorithm (a variant of NSGA-II). Moreover, Aspen Plus[®] provides a suitable environment for robust process simulation. Thereby, the objective of this research was to create and implement, a new methodology based on a genetic algorithm for the optimization of superstructures, integrating two computational programs through an interface: Aspen Plus (AP) and Matlab. This allowed the use of a rigorous thermodynamic approach by combining computer-aided tools, to find the best configuration of a biorefinery considering different processing routes to obtain high value biofuels and chemical products. Expanding how process synthesis has mostly been performed so far.

2. Methodology

The proposed methodology includes the collection of data such as, potential process routes to transform lignocellulosic biomass. The methodology was used to identify the optimal route for processing one or several chemical products on a commercial scale from agave bagasse, since it is a natural residue from the production of tequila and mezcal in some regions of Mexico. The processed agave bagasse used in this study was 1000 kg/day. The costs used in the economic evaluation were based on the current commercial price available (Alibaba, 2021). The thermodynamic model employed in this work was NRTL, while the parameters were obtained from Wooley and Putsche (1996). The compounds that were not found in the aspen plus database were generated using published experimental data and using group contribution approach.

2.1. Superstructure generation

Five major processing sections are included in the superstructure: feedstock, pretreatment, hydrolysis, fermentation, separation and/or purification (see Fig. 1).

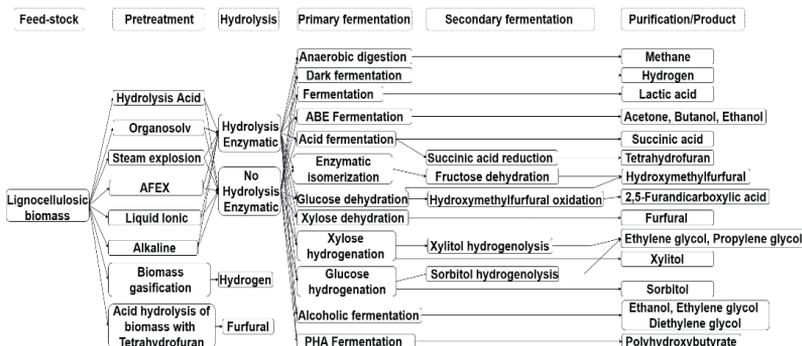


Fig. 1 Biorefinery superstructure to produce diverse chemical products.

The different production routes of biofuels and/or chemicals from lignocellulosic biomass were developed and simulated in Aspen Plus® through a superstructure. Aspen Plus was selected for its versatility and thoroughness for obtaining the results of mass balances and energy requirements, in addition to the availability of appropriate thermodynamic models. Aspen Plus and Matlab were connected to use the optimization suites available in Matlab. The optimization technique used in this work was based on stochastic multi-objective evolutionary methods.

2.2. Superstructure objective function

The multi-objective optimization problem was formulated to maximize the profit per day (GAN) and minimize the energy consumption per day (Q), given by:

$$\min(-GAN, Q) = f(x_{BM}, x_{sugars}, x_{sec-product}) \quad (1)$$

$$GAN = \sum_{i=1}^n (P_i * CV_i) - CB * BM - \sum_{j=1}^m MP_j \quad (2)$$

$$Q = \sum_{k=1}^p C_{vapor} * Q_{in_k} + C_{water} * Q_{out_k} \quad (3)$$

where P_i is the amount of product obtained in kg/day, CV_i are the selling prices of product obtained in USD/kg, CB is the cost of feedstock in USD/kg, BM is the processed biomass (1,000 kg/day) and MP_i is the cost of reactants and additives of each process in USD/kg. Q_{in} and Q_{out} are the amount of energy to be supplied and removed from the equipment respectively in kJ/day, C_{steam} and C_{water} are the cost of steam and cooling water, respectively, in USD/kg.

The multi-objective function is formulated in terms of decision variables as the treated mass fractions distribution in each section considered in the superstructure, that is, the mass fractions of fed biomass, mass fractions of the pretreatment, enzymatic hydrolysis, and fermentation stages. Optimization of the superstructure was performed using a stochastic multi-objective evolutionary method (Gamultiobj®), where population sizes of 100 individuals, 15 generations, crossover fraction 0.8, and a feasible adaptive mutation fraction were used. The implementation of the COM® interface used the computational tools Aspen Plus V.8.8® - Matlab R2017a®. The investment costs estimation was also included as a decision variable and calculated with the Guthrie method, to obtain an approximate investment cost for each of the scenarios to be considered. So, this parameter will be a decision criterion for an optimal biorefinery route.

3. Results

3.1 Optimal product and its processing route.

Table 1 presents a summary of the scenarios obtained by the optimizer, including GAN (product profit per day), Q (energy cost per day), IC (investment costs) and the result of the multi-objective function; in addition to the products obtained for each scenario, where the production of furfural is common for all the scenarios. Fig. 2a illustrates the optimal results in a 3D plot, Fig. 2b presents a ternary plot that illustrates the scenarios among the three objectives mentioned above. The optimal process flow diagram of the optimal processing route is shown in Fig. 2c. The results shown in Table 1 indicate that the processing route presented in Fig. 2c leads to the highest GAN (393.73 USD/day) of all the candidates, having a minimum cost of energy expenses (13 USD/day) and a minimum in the investment costs (3,230,000 USD). The optimal processing route for furfural

production consists of direct treatment of agave bagasse by acid hydrolysis with tetrahydrofuran.

Table 1 Results of superstructure optimization.

Scenario	Multi-objective function	GAN (USD/day)	Q (USD/day)	IC (USD)	Products
A	1,114.47	1,237.51	123.04	150,744,112	furfural, HMF, FDCA
B	380.73	393.73	13.00	3,230,000	furfural
C	408.53	521.92	113.39	249,740,863	furfural, HMF, FDCA
D	575.94	692.92	116.98	143,792,625	furfural, HMF, FDCA
E	1,092.27	1,215.31	123.04	150,694,121	furfural, HMF, FDCA
F	467.48	582.63	115.15	249,740,863	furfural, HMF, FDCA
G	611.67	731.53	119.86	149,839,974	furfural, HMF, FDCA
H	894.81	1,016.99	122.18	146,775,251	furfural, HMF, FDCA
I	861.94	983.44	121.50	148,226,897	furfural, HMF, FDCA

Nomenclature: HMF = hydroxymethylfurfural; FDCA = 2,5-Furandicarboxylic acid

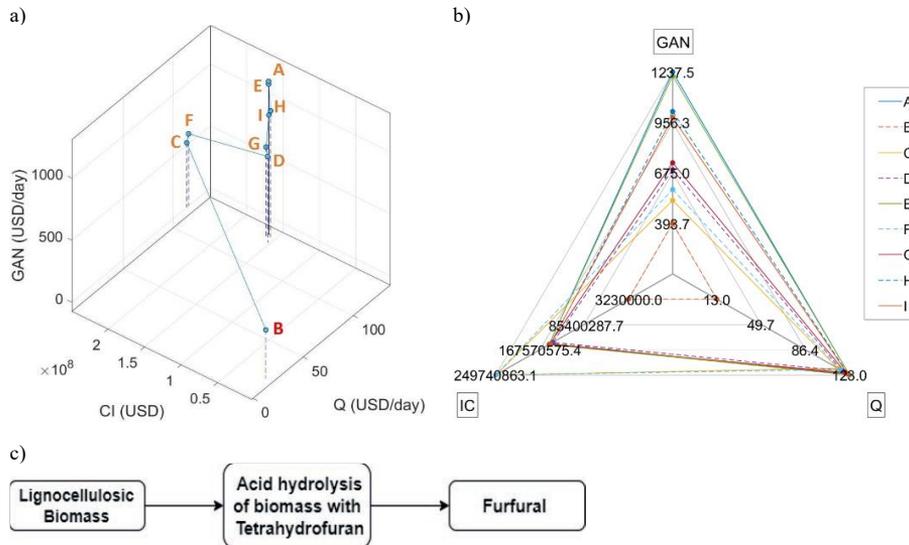


Fig. 2. a) 3D plot of the optimal processing scenarios, b) Ternary plot of optimal processing scenarios, c) Optimal processing route for agave bagasse.

3.2 Suboptimal products.

A subsequent optimization was performed not considering furfural production by acid hydrolysis with tetrahydrofuran in the superstructure, to find the second-best process configuration. Fig. 3a and 3b show the 3D plot and the ternary graph, respectively, where the optimal scenarios are illustrated. The process flow of the optimal processing route is shown in Fig. 3c that correspond to lactic acid production. The process layout includes a steam explosion, enzymatic hydrolysis, fermentation and separation and purification stage. The results indicated that is possible to have a profit of 939.70 USD/day, energy expenses (Q) of 26.13 USD/day and an IC of 6,187,109 USD.

In the same direction, a third optimal scenario was determined eliminating the furfural and lactic acid production processes of the superstructure. The results showed that the

third best option was the production of ethanol, ethylene glycol and diethylene glycol, including a pretreatment stage using steam explosion, followed by an enzymatic hydrolysis and fermentation stages. The profit (GAN) was 178.49 USD/day, with an energy cost (Q) of 13.75 USD/day and an IC of 8,339,205 USD. Fig. 4a-4b shows the obtained optimal scenarios in a 3D and ternary plots, respectively. Fig. 4c illustrates the process flow diagram of the optimal processing route for ethanol, ethylene glycol and diethylene glycol production.

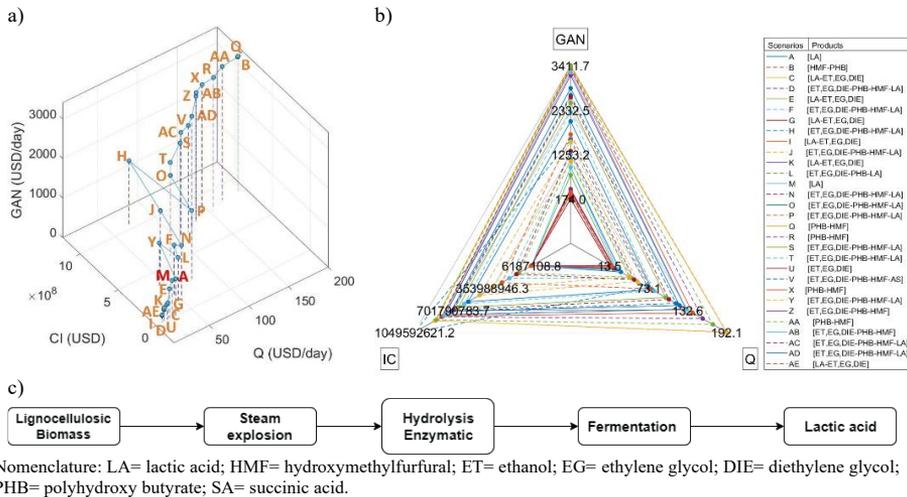


Fig. 3. a) 3D plot of the suboptimal processing scenarios by eliminating the furfural production process, b) Ternary plot of the suboptimal processing scenarios by eliminating the furfural production process, c) Second optimal processing route for agave bagasse: lactic acid production.

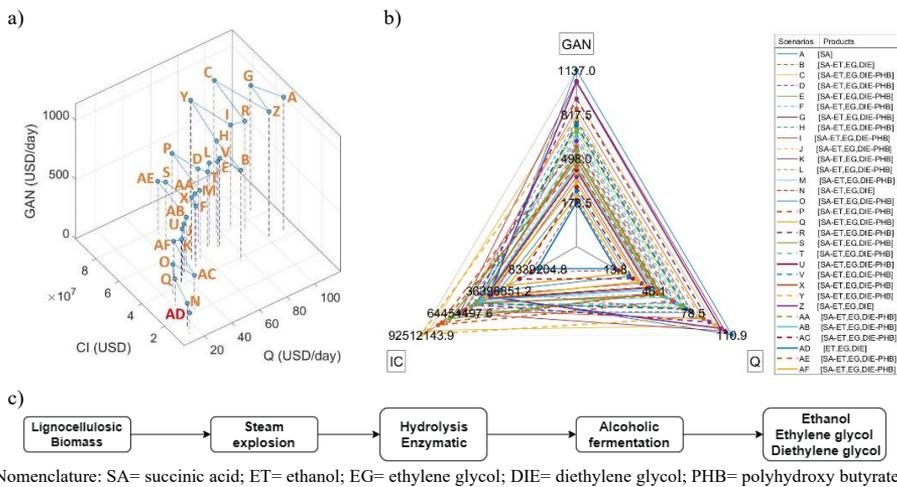


Fig. 4. a) 3D plot of the suboptimal processing scenarios by eliminating the furfural and lactic acid production process, b) Ternary plot of the suboptimal processing scenarios by eliminating the furfural and lactic acid production processes, b) Third optimal processing route for agave bagasse: ethanol, ethylene glycol and diethylene glycol production.

4. Conclusion

In this study, a new methodology for process synthesis with a superstructure approach was performed, in which different optimal process routes to produce high value chemicals were determined using agave bagasse lignocellulosic biomass as feedstock. The methodology involves a multi-objective optimization, maximizing profit, minimizing energy requirements, and including the lowest capital cost in the selection criteria. The integration of the Aspen Plus and Matlab computer-aided tools showed several advantages, one of them being that there was no need to explicitly introduce the equations representing each of the processes considered in the superstructure, such as mass and energy balances.

In addition, the simulation of the superstructure in Aspen Plus allowed to calculate the thermodynamic properties, through equations of state and solution models that consider the non-idealities of the mixtures involved in the phase equilibria and the interactions of components; in this way, a rigorous approach was used, avoiding disregarding important data in the synthesis of processes.

Different process scenarios were obtained in which the three decision criteria were satisfactorily fulfilled: maximization of total profit, minimization of energy expenses and the investment cost, permitting to obtain feasible process routes for the conversion of agave bagasse into a high value chemical product, such as the production of furfural, as first option, as second option the production of lactic acid and, as third option the production of ethanol, ethylene glycol and diethylene glycol.

References

- Alibaba, 2021. <https://www.alibaba.com>. Consultation date June 2021.
- K. Darkwah, B.L. Knutson, J.R. Seay, 2018, Multi-objective versus single-objective optimization of batch bioethanol production based on a time-dependent fermentation model, *Clean Technol. Environ. Policy*, 20, 6, 1271–1285.
- J.D. Ponce-Rocha, M. Picón-Núñez, A. Carvalho, A. Santos, F.I. Gómez-Castro, R. Morales-Rodríguez, 2021, A design and optimization framework for (bio-) chemical process based on exergo-economic and environmental aspects, *Computer Aided Chemical Engineering*, 50, 1897-1902.
- V. Punnathanam, P. Kotecha, 2016, Effective multi-objective optimization of Stirling engine systems, *Appl. Therm. Eng.*, 108, 261–276.
- A. Quaglia, C.L. Gargalo, S. Chairakwongsa, G. Sin, R. Gani, 2015, Systematic network synthesis and design: Problem formulation, superstructure generation, data management and solution. *Computers & Chemical Engineering*, 72, 68–86.
- N. Ren, A. Wang, G. Cao, J. Xu, L. Gao, 2009, Bioconversion of lignocellulosic biomass to hydrogen: Potential and challenges. *Biotechnology Advances*, 27, 1051-60.
- M. Rizwan, J.H. Lee, R. Gani, 2013, Optimal processing pathway for the production of biodiesel from microalgal biomass: A superstructure based approach. *Computers & Chemical Engineering*, 58, 305–314.
- R.J. Wooley, V. Putsche, 1993, Development of an ASPEN PLUS Physical Property Database for Biofuels Components. NREL/TP-425–20685.

Renewables-Based Multigeneration System for District Energy Supply

Houd Al-Obaidli, Rajesh Govindan, Tareq Al-Ansari*

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

Abstract

District energy systems (DES) provide electric and thermal energy from a central plant to a developed area through underground distribution networks. Fossil fuels are widely used as a primary energy source in DES. In 2021, energy-related CO₂ emissions from residential and commercial buildings reached 10% worldwide. In this paper, a renewables-based multigeneration system is proposed as a sustainable alternative for DES. The conventional system is based on a gas-fired combined cycle turbine and an absorption cooling system (ACS). The proposed renewables-based system utilises biomass and solar thermal energies and an ACS to provide district electricity, heating and cooling. A stochastic optimisation approach based on multi-objective generic algorithm (MOGA) is used to support the process modelling and identify the optimal pre-defined configurations based on their proximity to the Pareto curve. A techno-economic assessment is performed to compare the levelised cost of energy (LCOE) and evaluate the anticipated environmental footprint reduction. Results show that although the LCOE of the proposed system is higher than the conventional system, the reduction in CO₂ emissions converts directly into annual savings in districts covered by a carbon tax regime.

Keywords: Carbon tax, Multigeneration, Optimisation, Renewable energy, Techno-economic assessment.

1. Introduction

DES is one of the vital infrastructural elements of any modern urban development. It enables the efficient supply of critical utilities such as electricity, cooling and heating to contemporary residential and commercial districts. Conventional district supply entails fossil fuels as a primary energy source, such as coal, oil, and natural gas. Currently, residential and commercial buildings contribute 10% of the global carbon emissions. Therefore, renewable energy sources are now targeted for DES integration to enhance the environmental performance of the sector (Lake *et al.*, 2017; IEA, 2021). Recently, several studies featured renewable energy integration into DES. Al-Obaidli *et al.* (2020) designed a renewables-based multigeneration system to supply electricity, heating, cooling and freshwater. The study's main objective was to assess the energetic and exergetic efficiencies of the system compared to conventional supply methods. Hou *et al.* (2018) evaluated expansion strategies for existing conventional DES through demand-side management efficiencies and the use of renewable technologies. The study found that a potential reduction in CO₂ emissions between 6% and 8% was possible. Cheng *et al.* (2020) investigated the cost impact of CO₂ emissions and renewable energy targets in two real-world districts in China using a mixed-integer linear programming model (MILP). The study found that the average cost of reducing carbon emissions by 40% is 66 \$/tCO₂,

while the average cost of achieving 25% renewable energy penetration is 65 \$/MWh. Chen et al. (2021) proposed a 100% renewable energy DES for 30 different regions in China. The optimal portfolio for each district was determined based on economic performance and geographical features. The study identified the feasibility of the proposed system based on current market conditions with an average payback period of 6 years. Whilst earlier studies may have looked at various indicators (cost, emissions, etc...) individually, they have not attempted to find the optimal solutions based on these indicators combined.

This paper aims to develop a techno-economic assessment and optimisation tool for evaluating renewables-based DES compared to conventional supply systems. The MOGA process is used as a trade-off mechanism between LCOE and life-cycle emissions (LCE) and Net present project valuation (NPV) is performed subsequently for the identification of project profitability. The main objective is to couple the techno-economic evaluation and the trade-off analysis using MOGA in order to compare and evaluate the pre-defined configurations with the baseline for DES application.

2. Methodology

A case study was built for Al-Khor and Al-Thakira district, located in the North-East of Qatar. It is the closest residential district to the Ras Laffan industrial city (RLIC), which houses some of the world's most extensive natural gas processing and liquefaction facilities. Over the last 30 years, the district has witnessed cycles of growth and decline in both population and commercial activities due to peaking and diminishing project and construction work in the nearby RLIC. According to the 2020 census, the population total was 140,453, nearly 28% less than the population in 2010 (PSA, 2021). By dividing the total energy production over the total population, it is estimated that the average energy consumption per capita in Qatar is 17,307 kWh per annum. A DES is designed to supply electricity, heating and cooling for the district where commercial supply requirements are assumed to be equivalent to the residential energy requirements. Meanwhile, industrial requirements are excluded from this study. Table 1 provides the design specifications of the DES system.

Table 1. Design specifications of DES system (EIA, 2020).

Technology	CF (%)	Unit capacity (MW)	Number of units	Plant capacity (MW)
CCGT	56.6	430	3	1,290
PV	24.9	350	7	2,450
CSP	43	900	2	1,800
BIGCC	58.4	90	11	990
BECCS	49.8	90	13	1,170

Figure 1 depicts the DES system potential configurations. Renewables-based DES system using biomass and solar technologies is considered. Four technologies are envisaged: photovoltaics (PV) with battery storage, concentrated solar power (CSP) with thermal storage, a carbon-neutral biomass integrated gasification combined cycle (BIGCC) and a carbon-negative bioenergy with carbon capture and storage (BECCS) (Namany, 2019). A conventional system based on a natural gas-fired combined-cycle gas turbine (CCGT) is also included as a baseline configuration for comparison purposes. Since a PV-only

configuration cannot supply the required thermal energy for heating, a hybrid system will be considered and evaluated.

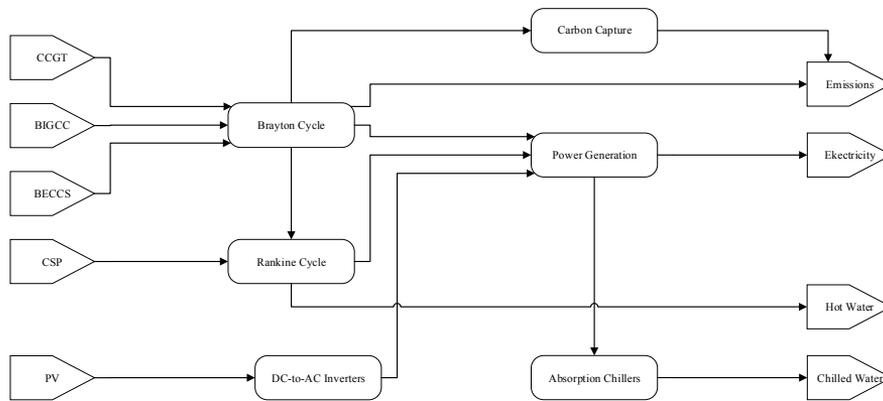


Fig. 1. DES block diagram.

A techno-economic assessment is performed based on the levelised cost of energy and life-cycle emissions. The LCOE is defined as lifetime costs over lifetime energy produced and is calculated using Eq. (1).

$$LCOE = \frac{CC + TOC \times DF}{E_{prod} \times DF} \quad (1)$$

The capital cost (CC) is the product of the plant's nameplate capacity by the capital cost rate, the total operating cost (TOC) is the sum of fixed (FOC) and variable operating costs (VOC), and E_{prod} is the annual energy production. The discount factor (DF) is determined using a facility lifetime of 24 years and a discount rate of 6.5%. Technology-specific cost and LCE parameters are listed in Table 2. Natural gas fuel cost is added to the VOC of CCGT only.

Table 2. Cost and LCE data (EIA, 2020; IPCC, 2014; Namany, 2019).

Technology	CC rate (\$/kW)	FOC rate (\$/kW)	VOC rate (\$/MWh)	LCE (gCO ₂ eq/kWh)
CCGT	1,084	14.1	2.55	490
PV	1,755	31.27	-	48
CSP	7,221	85.4	-	27
BIGCC	4,468	132.82	-	-
BECCS	5,865	146.32	-	(65)

A multi-objective genetics algorithm is used to determine the optimal Pareto frontier. The two competing objectives of the optimisation problem are:

$$\text{minimise } \sum_i c_i x_i$$

$$\text{minimise } \sum_i e_i x_i$$

where c_i is the unit LCOE per technology, e_i is the unit LCE per technology, and x_i is the technology fraction of the portfolio. The optimisation problem is subject to the following constraints:

$$\begin{aligned} \sum_i x_i &= 1 \\ 0 \leq x_i &\leq 1 \end{aligned}$$

The net present value is calculated using Eq. (2).

$$NPV = -CC + AP \times DF \quad (2)$$

The annual profit (AP) is determined by subtracting the TOC plus carbon taxes from the annual revenues.

3. Results and Discussion

Average US wholesale electricity price, Henry Hub natural gas spot price, and EU ETS CO₂ price data were used. Table 3 presents the LCOE for each of the configurations.

Table 3. LCOE results.

Configuration	Power generation mix	LCOE (\$/MWh)
A	CCGT	48
B	PV	89
C	CSP	254
D	BIGCC	103
E	BECCS	153
F	50% PV + 50% BIGCC	108
G	50% PV + 50% BECCS	129

Figure 2 presents the trade-off chart between LCOE and LCE. The different configurations create an artificial Pareto front. It is noted that the baseline configuration provided the lowest cost, and configuration E provided the lowest emissions. The points that lie on the Pareto front are considered "optimal", and the points that lie outside are considered "suboptimal" relative to the optimisation objectives at hand.

From Figure 2, configuration E was chosen from the set of non-dominating solutions (B, F, D, G, and E) to perform the comparison with the baseline case. MOGA was applied, and 700 solutions were identified. Most of the MOGA solutions lie on and between configurations B and E, creating a Pareto front starting from configuration B through configurations F, D, and G to configuration E. This means that all of these configurations are optimal, and configuration C is suboptimal. Configuration A was not evaluated as part of the optimisation run and is only presented for comparison purposes.

The wholesale electricity price follows the LCOE of the power mixture in the case of market equilibrium. This study assumed that the average wholesale electricity price would rise from 40 to 200 \$/MWh. Figure 3 provides the NPV for all the configurations.

It is noted that configuration A provided the highest NPV, and configuration C provided the lowest NPV. It is also noted that the optimal configurations are not necessarily the configuration with the highest NPV.

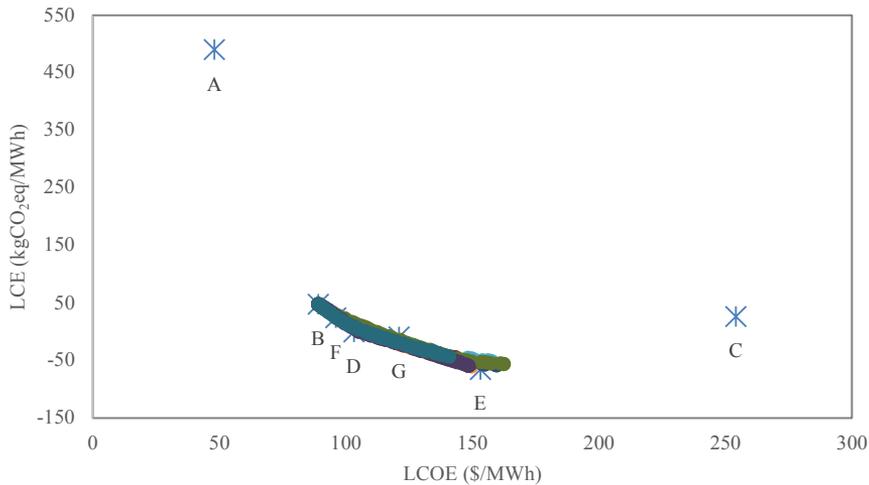


Fig. 2. LCOE and LCE trade-off chart.

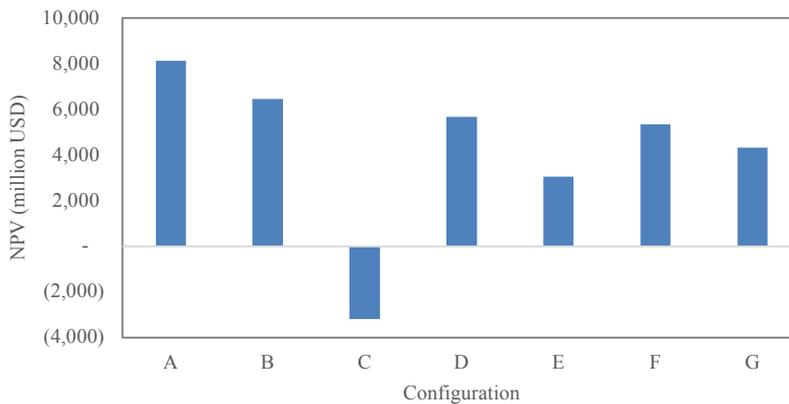


Fig. 3. Net present asset valuation results.

A sensitivity analysis for the CO₂ price is performed on the two profitable configurations, as shown in Figure 4. It is noted that when CO₂ price exceeds 240 \$/tCO₂, the selected non-dominating optimal configuration also becomes the most profitable one since carbon taxes limit the profitability of configuration A as CO₂ price increases. On the other hand, configuration E is positively influenced by rising CO₂ market prices.

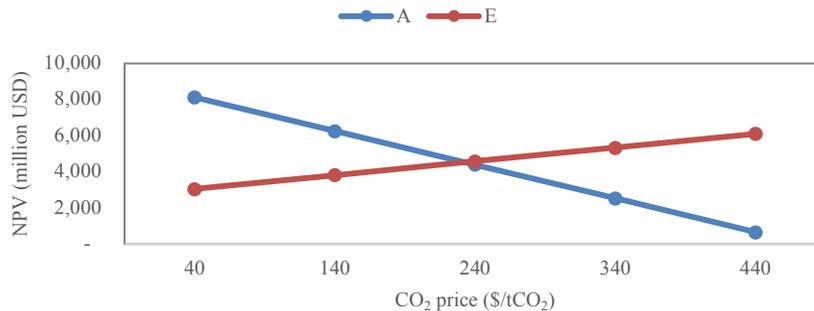


Fig. 4. Sensitivity analysis of CO₂ price impact on NPV.

4. Conclusions

A techno-economic assessment was conducted to evaluate an alternative DES for Al-Khor and Al-Thakira district in Qatar. The renewables-based configurations were compared to a conventional natural gas-fired combined cycle in terms of LCOE and LCE. The results were characterised by a Pareto front where CCGT had the lowest Levelised cost and BECCS had the lowest emissions. Several renewable energy configurations had the best trade-off between cost and emissions and only configuration C (CSP) was sub-optimal. Moreover, carbon-neutral and carbon-negative configurations provide a significant advantage in carbon taxes compared to the conventional system especially in a volatile carbon market. NPV provides a good indication of project profitability; however, when it comes to emissions, CO₂ prices would need to increase by more than six times the current levels for the renewable energy-based DES systems to present a more favourable NPV.

References

- H. Al-Obaidli, Y. Bicer, T. Al-Ansari, 2020, Performance comparison of a natural gas and renewable-based power and desalination system for polygeneration, *Greenhouse Gases Science and Technology*, 10, 678–702.
- X. Chen, J. Xiao, J. Yuan, Z. Xiao, W. Gang, 2021, Application and performance analysis of 100% renewable energy systems serving low-density communities, *Renewable Energy*, 176, 433–446.
- Y. Cheng, N. Zhang, D. S. Kirschen, W. Huang, C. Kang, 2020, Planning multiple energy systems for low-carbon districts with high penetration of renewable energy: An empirical study in China, *Applied Energy*, 261, 114390.
- EIA, 2020, Capital Cost and Performance Characteristic Estimates for Utility Scale Electric Power Generating Technologies.
- J. Hou, P. Xu, X. Lu, Z. Pang, Y. Chu, G. Huang, 2018, Implementation of expansion planning in existing district energy system: A case study in China, *Applied Energy*, 211, 269–281.
- IEA, 2021, Global energy-related CO₂ emissions by sector.
- IPCC, 2014, Annex III: Technology-specific cost and performance parameters.
- A. Lake, B. Rezaie, S. Beyerlein, 2017, Review of district heating and cooling systems for a sustainable future, *Renewable and Sustainable Energy Reviews*, 67, 417–425.
- S. Namany, T. Al-Ansari, R. Govindan, 2019, Optimisation of the energy, water, and food nexus for food security scenarios, *Computers and Chemical Engineering*, 129, 106513.
- PSA, 2021, Qatar Census 2020 Main Results.
- QGEWC, 2021, Annual Statistics Report 2020.

Topology-Based Construction of Business-Integrated Material Modelling Workflows

Heinz A. Preisig^a, Peter Klein^b, Natalia Konchakova^c, Thomas F. Hagelien^d, Jesper Friis^d and Martin T. Horsch^{e, f}

^a*Dept of Chemical Engineering, NTNU, 7491 Trondheim, Norway*

^b*Fraunhofer Institute for Industrial Mathematics, Kaiserslautern, Germany*

^c*Institute of Surface Science, Helmholtz-Zentrum Hereon, Geesthacht, Germany*

^d*SINTEF, NO-7465 Trondheim, Norway*

^e*School of Psychology and Computer Science, University of Central Lancashire, Preston, UK*

^f*STFC Daresbury Laboratory, UK Research and Innovation, Daresbury, UK*

Heinz.Preisig@chemeng.ntnu.no

Abstract

Designing a new product requires information from the business and physical domains, which implies integrating business decision tools with process and material simulation processes to form an overall workflow. The integration involves coupling the business workflow management systems with analysis tools, optimisation and decision support systems, which require process simulations and an integrated data transfer service. The process simulation, in turn, will in general model multiple layers of time scales and thus is also in need of data transfer between different solvers. Here we discuss the main components in the light of a coating-design project.

Keywords: Computational engineering, multi-scale modelling, applied ontology

1. Project Background

The Horizon 2020 project VIPCOAT constructs an innovation platform for new active protective coatings based on materials modelling and optimisation. The project's application is coatings for the aerospace industry, and its objective is to introduce novel approaches for corrosion protection by active inhibiting pigments as nano-additives. VIPCOAT is to support end-users in developing new and effective corrosion barriers for metal surfaces, deploying environmentally friendly technologies, and, in parallel, provide a decision support system for business integration. VIPCOAT implements a multi-layer digital structure to enable the open innovation process of the production processes and the design of value-adding product chains. VIPCOAT uses a conceptually new idea, namely a Pareto chain along a value-adding B2B2B (business-to-business-to-business) sequence enabling collaborative, transparent decision processes. The approach builds on multi-criteria optimisation (MCO) and implements a set of decision tools that allow for exploring the individual Pareto fronts interactively.

Usually, coatings contain several components. Primarily the coating matrix is the material that makes up the coating itself. In addition, several additives are mixed in, protecting from UV, giving colour, and inhibiting corrosion in coating defects. When producing a paint, one has several objectives: purpose-related performance, the toxicity of the involved materials, and costs of production and application. Typically, these objectives are not possible to satisfy simultaneously. Also, with

having different components, usually, several suppliers are part of the production process. Therefore, it is natural that each member of the suppliers and customers has another set of objectives. And each member is affected by all of those being close to them in the process. Consequently, if one has to define *optimality*, one is confronted with multiple, incompatible business-related objectives or KPIs (Key Performance Indicators). Hence, overall, one deals with two levels of multi-objective optimisation problems: the lower one is the product, the upper the multi-player business layer. The project will create a digital platform for coating formulation, development and optimization, which could serve as a computational coating marketplace.

The VIPCOAT platform will comprise a collection of physics and data-based materials models, a data space, and an associated ontology-driven service that enables a smooth and simple data transfer between Business Decision Support Systems (BDSS), and the MoDeNa software orchestrator. For this purpose, the team develops new semantic network-based technologies and approaches to ensure an automatic data exchange for multi-scale simulations and multi-level material and business software components. These semantic technology developments are aligned with the Elementary Multiperspective Material Ontology (EMMO), cf. Francisco Morgado et al. (2020), in line with a series of efforts coordinated through activities of the European Materials Modelling Council (EMMC ASBL) and the Innovation Centre for Process Data Technology (Inprodat e.V.).

2. Business-to-Business process level

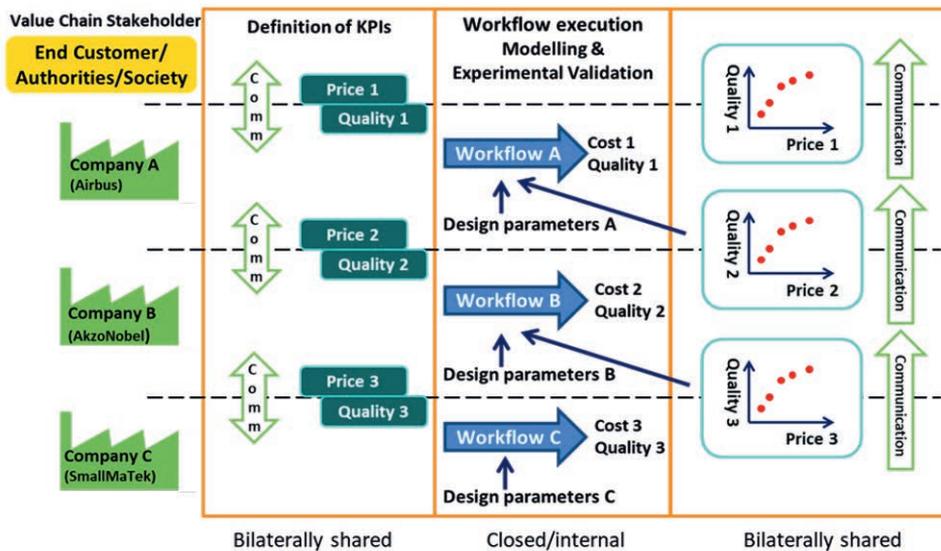


Figure 1: Business-to-Business-to-Business approach implemented in VIPCOAT. Bilateral communication starts the processes, followed by business internal developments and ending in an open access to Pareto-optimal product variants to be used in decision making upstream the production chain.

The VIPCOAT approach supports collaborative decision making along with production and/or value-added chains, using a conceptually new idea: the Pareto chain along Business-to-Business-to-Business (B2B2B) value-added chains as pictorially presented in Figure 1. The value-added chain goes from the bottom to the top. On the bottom, in VIPCOAT SmallMaTek (SMT) producing inhibiting pigments, Company C has some production capabilities associated with costs. As high quality and low cost are almost always contradicting goals, the capabilities to deliver products upstream the value chain are best described using a Pareto front of the best possible compromises.

For a company B downstream, in VIPCOAT represented by AkzoNobel, the best possible compromises form a sub-space of the Companies B opportunities and thus enter their design space for fabrication for company A, in VIPCOAT represented by Airbus. Therefore, we have a B2B2B environment in which collaborative decision making on top of transparent decision processes will make a lot of sense. In VIPCOAT, this is directly supported by the concept of a Pareto chain of interactive decision tools, designed as interactive explorers of the individual Pareto fronts. One Business executes a development and/or production process, supported by materials modelling to be described next, internally. Thus the internal cost structure is kept confidential and the transparent part of the B2B2B relation is the Pareto front of prices and associated product variants.

3. Business internal process level

One of the project's main objectives is to reduce the number of experiments, or in other words, replace experiments with predictions based on a physical model of the experimental setup. The first step is thus to design and discuss a model topology using the minimal graphical language we defined (Preisig (2021, 2014)). The topology captures all the fundamental structural assumptions. It represents the physical process as a network of primitive capacities (base entities, control volumes), a minimal set of extensive quantities and a set of mechanisms on how entities interact. Since it is a hierarchical representation, it also shows the assumed mereology of the model. The topology reflects all the main assumptions: what parts of the system are considered, which ones exhibit capacity effects, if they show significant gradients in the intensive quantities, if they are constant or dynamic or event-dynamic or not, how they link up to other capacities and exchange extensive quantity, what exchange is considered relevant, the exchange mechanism, and the resource environment. It should be noted, that the topology does not include geometrical information, like the shape of a capacity, also called a control volume.

Our Process Modelling software (ProMo) defines an application-focused ontology that defines the fundamental entities and their mathematical representation. The mereological information is encoded into the model's hierarchical tree. The internal nodes in the tree, we term composite entities. Replacing parts of the topology, usually composite entities, by surrogates generates the skeleton of the workflow. In the case of the coating process, as shown in Fig. 2 it is the two gray boxes that may be replaced by a surrogate solving the input relations in a separate task, a PDE solver like OpenFoam, while the leaching process is solved involving molecular modelling on which we do not expand in this exhibition. When running a workflow, the different tasks do exchange data, in our case this would be reading salinity, but also transferring information about the simulation of the coating and the simulation of the material and its inhibition. The same applies with the interaction of the coating and the leaching process simulations. Since ProMo maps information of the variables and the relations via OWL into a triple store, and the model provides the vectors of information exchange between the blocks, the interoperability problem can be solved using the approach described below. Finally, the task factory generates the simulation code for a solver environment like MatLab or an orchestrator that executes the developed workflow.

The green area in Figure 2 is reflecting only part of the story. In ProMo the access to a data taken from an external database is a controlled process. Figure 3 shows how ProMo maps the variables and equations into an EMMO-extended ontology. Once the model is established, the intertask communications is established over the SOFT/DLITE interoperability framework. ProMo maps the variables and equations into an EMMO-extended ontology. Once the model is established, the intertask communications is established over the SOFT-DLITE framework.

4. Interoperability

SOFT/DLITE (Hagelien et al., 2017) (Mir et al., 2020) represents information exchange between a data source and sink as a pipeline that process output data from the producer into a form that is

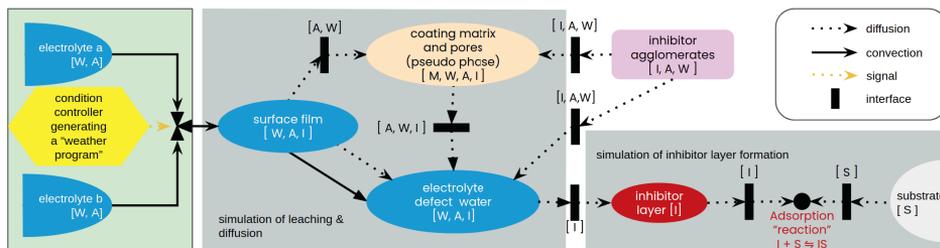


Figure 2: A simplified topology of a coating with inhibitor-loaded nano particles. Replacing the two grey boxes by surrogates providing an input/output solution in a separate task, one has the structure of the workflow.

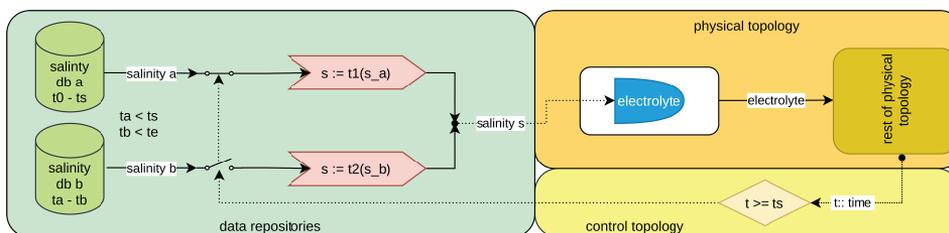


Figure 3: The ProMo view on how the salinity is read from two data sources. The database a covers the first part of the time period, while the database b is used to fill in the remainder.

suitable input to the consumer. Metadata schemas and ontologies capture the description of the data's intent and their meaning (semantics), which can be used in conjunction with specialised data readers for data extraction (Schembera, 2021) (Fig. 4). The next step is to augment high-level semantic artefacts with domain-specific knowledge. EMMO aligned mid-level and domain ontologies can be employed to document concepts such as processes and properties/quantities semantically (Horsch, 2021). Data and metadata need to be represented as knowledge graphs, employing semantic technology standards such as RDF, RDFS, and OWL.

EMMO-aligned mid-level and domain ontologies can also support platform interoperability between VIPCOAT and other infrastructures, particularly those developed within *H2020 NMBP* projects and, in the future, projects supported from the *Horizon Europe CLA resilience and data* lines of funding. Moreover, the EMMC ASBL focus areas and task groups help coordinate a series of ongoing development efforts in this direction. This technology operates at a comparably high level and requires at least the developers of platforms and tools to be familiar both with ontology-based research data management and with the philosophy underlying the EMMO. In contrast, the second line of development of semantic artefacts, equally endorsed by EMMC ASBL, operates at a lower level and is more accessible to domain experts: Based on the terminology of the Review of Materials Modelling (RoMM), the MODA (Model Data) describe simulation workflows, which permits textual descriptions, and is supported by the OSMO's ontology version of MODA (Horsch et al., 2020, 2021). An alignment between semantic artefacts from these two lines of work is not always straightforward because they rely on knowledge graphs that are differently structured. Previous work by Klein et al. (2021) discusses how graph transformation crosswalks can be used to transpose MODA/OSMO representations of simulation workflows, for which annotation and data ingest are comparably easy, into EMMO-aligned mid-level ontologies that are best suitable for platform interoperability in line with EMMC ASBL recommendations.

In the present framework, the mapping property *mapsTo* ensures that the semantics of metadata

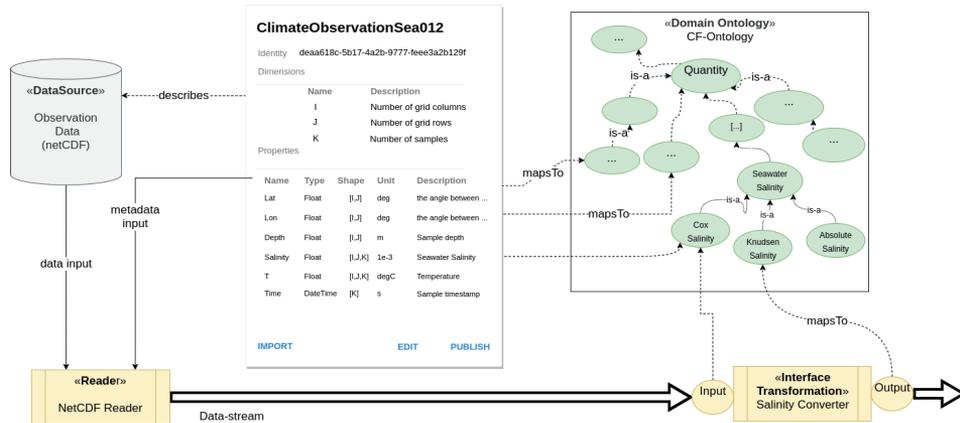


Figure 4: Schematic overview. The key components of the semantic interoperability platform are business data, a business data reader or writer, a metadata representation, domain ontologies, and interface transformations (equations) that connect different ontologies. The interoperability platform will analyse the use case and give the user options to select between multiple semantic pipelines or manually manage the pipeline pathways.

properties are aligned with domain ontology concepts, cf. Fig. 5. Similarly, a data model will represent the schema of application-specific data input on the consumer side. A data ingest system or file format generator using the metadata produces information that the data consumer can interpret. Mapping the metadata schema to the domain ontology concepts captures the semantics of the information. The SOFT/DLite generates the pipeline for the data transfer between the systems, recognising the semantic interpretation and syntactic representations.

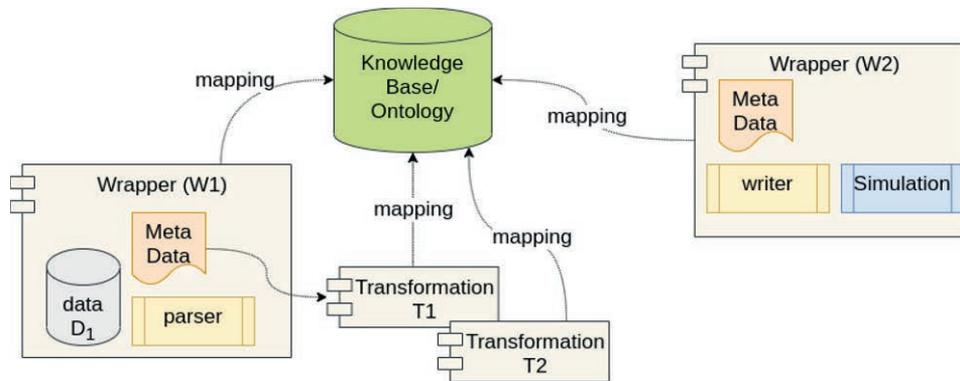


Figure 5: Data resources, data sinks, transformations and simulation packages connect to the Knowledge Base via mappings between the metadata representation of the business data and domain ontology concepts. Dependent on the data model on the receiving side of the information, a pipeline is constructed based on the semantic knowledge of the data and transformations. A pipeline can be constructed from Data resource D_1 in Wrapper W_1 through the interface transformations (T_1 and T_2) to the construction of the syntactic input representation for the simulation tool in W_2 : $W_1 \rightarrow T_1 \rightarrow T_2 \rightarrow W_2$

In the trivial case, the pipeline will only transfer data only. In a more complex scenario, semantic differences between data sources and data sinks require transformations to be taken from an ontology. The ProMo software (see Section 3) is already aligning the equations and variables to

EMMO domain ontologies and can therefore be plugged directly into the SOFT/DLite pipelines. Graph-search algorithms find all possible paths that define the pipeline. The end-user is responsible for selecting the most appropriate pathway if multiple exist, closing the connection through mappings and defining new transformations. In addition, the interoperability platform also manages the representation of the information in terms of differences in units (unit systems) and data shape/dimensionality. Unit conversions transformations can usually be inferred automatically, but strategies for managing differences in dimensionality (extrapolation, interpolation, averaging, machine learning-based methods, etc.) needs to be manually determined.

5. Conclusion

We demonstrate a new approach to a complex business-integrated product design on the example of active-protective coatings for the aeroplane industry. We introduce a new approach to interrelate three daisy-chained industries, two producers and one end-consumer. The new business-to-business environment implements a complex two-level workflow that links the business workflow with the physical simulation workflow, involving all three business partners in forming the final product performance. On the business level, we employ business-process-modelling (BPM) software Camunda and the NTNU's Process Modelling suite (ProMo) to model and generate simulation workflows executed in the MoDeNa platform. The team approaches the data integration of the various software using ontology-based technologies, generating the application interfaces automatically for the different components, providing access to a shared data space and Petri-net based technology for synchronisation.

References

- J. Francisco Morgado, E. Ghedini, G. Goldbeck, A. Hashibon, G. J. Schmitz, J. Friis, A. de Baas, 2020. Mechanical testing ontology for digital-twins: A roadmap based on EMMO. In: R. García Castro, J. Davies, G. Antoniou, C. Fortuna (Eds.), Proceedings of SeDiT 2020. CEUR-WS, Aachen, p. 3.
- T. F. Hagelien, A. Chesnokov, S. T. Johansen, E. A. Meese, B. T. Løvfall, 2017. Soft: a framework for semantic interoperability of scientific software. In: Progress in Applied CFD–CFD2017 Selected papers from 12th International Conference on Computational Fluid Dynamics in the Oil & Gas, Metallurgical and Process Industries. SINTEF akademisk forlag.
- M. T. Horsch, 2021. Mereosemiotics: Parts and signs. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOUST 2021). CEUR-WS, Aachen, p. 3.
- M. T. Horsch, C. Niethammer, G. Boccardo, P. Carbone, S. Chiacchiera, M. Chiricotto, J. D. Elliott, V. Lobaskin, P. Neumann, P. Schiffels, M. A. Seaton, I. T. Todorov, J. Vrabec, W. L. Cavalcanti, 2020. Semantic interoperability and characterization of data provenance in computational molecular engineering. *Journal of Chemical & Engineering Data* 65 (3), 1313–1329.
- M. T. Horsch, D. Toti, S. Chiacchiera, M. A. Seaton, G. Goldbeck, I. T. Todorov, 2021. OSMO: Ontology for simulation, modelling, and optimization. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOIS 2021 Ontology Showcase). CEUR-WS, Aachen, p. 47.
- P. Klein, H. A. Preisig, M. T. Horsch, N. Konchakova, 2021. Application of an ontology based process model construction tool for active protective coatings: Corrosion inhibitor release. In: E. M. Sanfilippo, et al. (Eds.), Proceedings of JOWO 2021 (FOMI 2021). CEUR-WS, Aachen, p. 26.
- Z. M. Mir, J. Friis, T. F. Hagelien, I.-H. Svenum, I. G. Ringdalen, N. Konchakova, M. L. Zheludkevich, D. Höche, jan 2020. Interoperability architecture for bridging computational tools: application to steel corrosion in concrete. *Modelling and Simulation in Materials Science and Engineering* 28 (2), 025003.
URL <https://doi.org/10.1088/1361-651x/ab6209>
- H. A. Preisig, 2014. Visual modelling. *Computer Aided Chemical Engineering* 34 (ISBN 978-0-4444-63433-7), 730–734, (FOCAPD 8 - Foundation of Computer-Aided Process Design).
- H. A. Preisig, 2021. Ontology-based process modelling – with examples of physical topologies. *Processes* 9 (4), 592.
- B. Schembera, 2021. Like a rainbow in the dark: Metadata annotation for HPC applications in the age of dark data. *Journal of Supercomputing* 77, 8946–8966.

Investigation of optimal blending of livestock manures to produce biocrude via hydrothermal liquefaction

Mohammad Alherbawi ^a, Prakash Parthasarathy ^a, Gordon McKay ^a, Hamish R. Mackey ^a, Tareq Al-Ansari ^{a,*}

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

Abstract

Livestock manure is one of the most abundant agro-wastes worldwide. Most of the manures are landfilled or composted to produce fertilizers, which raises concerns associated with the possible pollution of air and contamination of groundwater. Meanwhile, Hydrothermal liquefaction (HTL) is attracting growing attention to valorise wet wastes as an alternative to anaerobic digestion to produce value-added energy products. Livestock manures differ in terms of their composition, availability, and cost; therefore, an optimal blending of multiple livestock manures for liquefaction is expected to yield biocrude with distinct properties and contribute to process efficiency improvement. As such, this study evaluates the effect of different blending ratios of five typical livestock manures on the yield of biocrude and biochar, as well as the process environmental burden in terms of water consumption and CO₂ emissions. Aspen Plus® software is used to simulate the liquefaction process and to develop a mathematical optimization model to determine the optimal blending scenarios that satisfy adopted technical and environmental preferences. A weighted normalised decision metric is then used to minimise the gap between the solutions for the multiple objectives. The model suggested the optimal manures blend to be (50%, 26%, 14%, 5%, 5%) for cattle, camel, horse, poultry, and sheep manures, respectively. This study provides an insight into enhancing HTL process efficiencies through a careful selection of feedstock blends.

Keywords: Livestock manure, Liquefaction, Blending ratio, Biocrude, Simulation.

1. Introduction

Millions of tonnes of livestock manure are generated each year worldwide. While the uncontrolled decomposition of manure results in severe pollution of the environment and contamination of water resources. Meanwhile, Hydrothermal liquefaction (HTL) is becoming an increasingly attractive alternative to anaerobic digestion in terms of converting wet wastes into multiple fuels. In HTL, carbonaceous materials are converted to various energy products at high temperatures and pressure, just like crude oil is formed underground. The HTL process can directly liquefy manure into biocrude oil without the need for an energy-intensive drying procedure. HTL is typically performed at a temperature ranging from 250 °C to 400 °C, for a residence time of 15-60 minutes, and with an operating pressure ranging from 4 to 25 MPa (Chen et al., 2019). Biocrude is the major product of hydrothermal liquefaction. It is an oil with a lower oxygen content (10-20%) and a higher net heating value than pyrolysis oil (Peterson et al., 2008). In addition

to bio-crude, the HTL process also produces an aqueous phase-rich liquid product, biochar, and syngas.

There are significant differences in the composition of livestock manures, which have a direct significant impact on the biocrude characteristics and yield (Li et al., 2018). Livestock manures differ not only in composition, but also in availability and cost; therefore, an optimal blend of multiple livestock manures for liquefaction will likely yield biocrude with distinct properties and enhance processing efficiency.

It has been reported previously that manure can be used to produce biocrude (Chen et al., 2018; Xiu et al., 2010; Yin et al., 2010). However, the HTL of livestock manure blend is relatively new and requires thorough investigations to improve the process outcomes. As such, the purpose of this study is to explore the effects of different blending ratios of five typical livestock manure such as cattle, camel, horse, poultry, and sheep manures on the yield of biocrude and biochar, as well as the process' environmental burden in terms of water consumption and CO₂ emissions. Using Advanced System for Process Engineering (ASPEN PLUS) software, the liquefaction process is simulated, and a mathematical optimisation model is developed to determine the best blending scenarios that satisfy the adopted technical and environmental preferences.

This research sheds light on possible means for the improvement of HTL process efficiencies by carefully selecting feedstock mixtures. The synergistic effect of animal manure blends in HTL demonstration projects would serve as a roadmap for future use of livestock manure blends.

2. Methodology

2.1. Process Modelling

ASPEN PLUS V.10 is used to model a 250,000 tonnes/year hydrothermal liquefaction plant, considering isothermal and steady-state conditions. All manures are defined based on their proximate and elemental analyses as presented in **Table 1** (Akyürek et al., 2021; Al-Ansari et al., 2020; Chong et al., 2019; Whitely et al., 2006; Zhou et al., 2020). The manure streams are initially fed into a blending reactor, in which they are converted into conventional components based on their moisture and elemental composition using a Fortran code to ensure a mass balance of the conversion. The blended stream is then fed into a hydrolysis reactor along with additive water to create a slurry. For this purpose, a Fortran code is developed to estimate the amount of water required to dilute the solids mass concentration in the slurry to 20%. The mixture is then pumped to 100 bar using two consecutive high-performance pumps into the main HTL reactors, where the process is conducted at 100 bar and 350°C.

Table 1. Proximate and elemental analyses of studied livestock manures.

Analysis	Camel manure	Cattle manure	Horse manure	Poultry manure	Sheep manure
Moisture content (%)	58	85	40	40	50
Fixed carbon (%)	22.74	10.11	12.06	13.31	15.56
Volatile matter (%)	60.51	57.18	77.80	63.58	57.30
Ash (%)	16.73	32.69	10.12	23.09	27.13
Carbon (%)	37.11	27.61	39.53	36.67	33.14
Hydrogen (%)	4.07	3.48	4.99	5.33	4.48
Nitrogen (%)	2.27	1.88	0.70	3.54	2.65
Sulphur (%)	0.24	0.41	0.47	0.73	0.37
Oxygen (%)	39.54	33.91	44.16	30.61	32.21

In the first HTL reactor, the maximum biocrude yield and the minimum hydrochar yield are restricted using a Fortran code based on a correlation adapted from Zhong and Wei (2004) considering feedstock's composition, while the composition of biocrude and hydrochar is adapted from Pedersen et al. (2017) and Lentz et al. (2019) respectively. In the second HTL reactor, the composition of syngas is calculated based on thermodynamic approach via minimising Gibb's free energy. Besides, a hydrocyclone is used to recover hydrochar, while fluids are separated using a three-stage flash unit obtaining syngas, biocrude, and an aqueous phase. A simplified HTL process flow diagram is illustrated in **Figure 1** (Alherbawi et al., 2021).

The syngas composition is then calculated in an "RGibbs" reactor via the minimisation of the Gibbs' free energy. The process is conducted at 350 °C and 100 bar (Al-Ansari et al., 2020). In addition, solids are collected using a hydro-cyclone, while a three-stage flash drum is utilised to split the remaining stream into three phases: gas, biocrude and an aqueous phase.

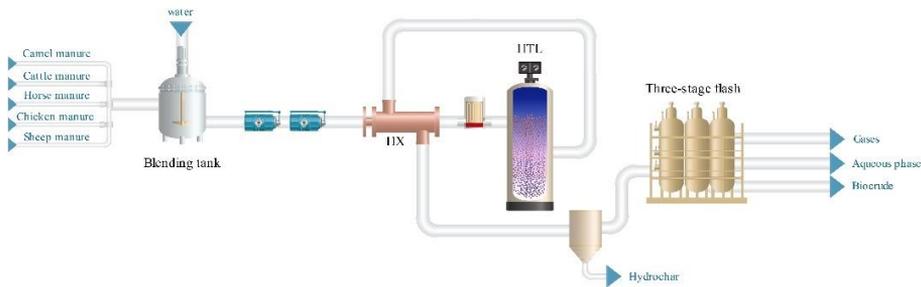


Figure 1. A simplified HTL process flow diagram using manures blend.

2.2. Optimisation Model

A mathematical optimisation model is developed in ASPEN Plus using the objective functions illustrated throughout Equations (1-4) and the constraints in Equations (5-6). Four objectives are adopted to facilitate selecting the optimal blending ratios of the five different manures for the HTL process. The first two objectives concerning the technical aspect of the process aim at maximising the yields of biocrude and hydrochar, while the other two objectives concerning the environmental aspect aim at minimising the water consumption and CO₂ emissions. In addition, the blending is constrained to satisfy 100% of the plant capacity, with a minimum of 5% of each manure utilisation and a maximum of 50% considering the possible limited supply of the manure.

Objective functions:

$$\text{Total biocrude yield } (Z_1) = \text{Max } \sum_{i=1}^5 BC_i X_i \quad (1)$$

$$\text{Total hydrochar yield } (Z_2) = \text{Max } \sum_{i=1}^5 HC_i X_i \quad (2)$$

$$\text{Total water consumption } (Z_3) = \text{Min } \sum_{i=1}^5 W_i X_i \quad (3)$$

$$\text{Total CO}_2 \text{ emissions } (Z_4) = \text{Min } \sum_{i=1}^5 E_i X_i \quad (4)$$

Constraints:

$$\sum_{i=1}^5 X_i = 1 \quad (5)$$

$$0.5 \geq X_i \geq 0.05 \quad (6)$$

Where: X_i : blending ratio % of manure (i).
 BC_i : biocrude relative yield, in tonne per tonne of manure (i).
 HC_i : hydrochar relative yield, in tonne per tonne of manure (i).
 W_i : water relative consumption, in tonne per tonne of manure (i).
 E_i : process emissions, in tonne CO₂ per tonne of manure (i).
 $i=1-5$ represent manure types, whereby: 1: Camel manure, 2: Cattle manure, 3: Horse manure, 4: Poultry manure and 5: Sheep manure.

Each optimisation function is solved independently, whereas the normalised weighted decision metric presented in Equation (7) is used to achieve a solution that relatively satisfies the four adopted objectives. Solving the goal function facilitates minimising the gap between the initial solution and the optimal solution for the different objective functions. Whereby it deals with each objective function independently and generates a single solution, in contrast to multi-objective problems that generates multiple solutions through trade-off between objectives.

$$Goalfunction = Min \sum_{j=1}^4 \omega_j \left| \frac{I_j - Z_j}{Z_j} \right| \quad (7)$$

Where, I_j is the initial solution for objective (j) assuming $X_i=100\%$, while Z_j is the optimal solution for objective (j) represented earlier in Equations (1-4). Besides, ω_i is the relative weight of objective (j), which is assumed to be equal for the four objectives in this study.

3. Results and Discussion

3.1. Model outputs

The products yield and environmental burden of the process considering a single manure feed in each run are illustrated in **Figure 2**. The highest biocrude yield is achieved using horse manure feedstock with a relative yield of 0.23 tonne biocrude per tonne of wet feedstock, while the lowest biocrude yield is achieved using cattle manure, which is possibly due to its low volatile matter and the extremely high ash and moisture contents. However, the highest hydrochar yield is obtained when camel manure is used as feedstock as its fixed carbon is the highest overall. Nevertheless, the HTL of cattle manure does not require any additive water, whereby its moisture content is sufficient to create a slurry. Whereas both, horse and poultry manures required the highest amount of water with 2:1 water to feedstock ratio. Besides, the HTL of horse manure is responsible of the highest CO₂ release, which is possibly due to its high carbon and oxygen contents.

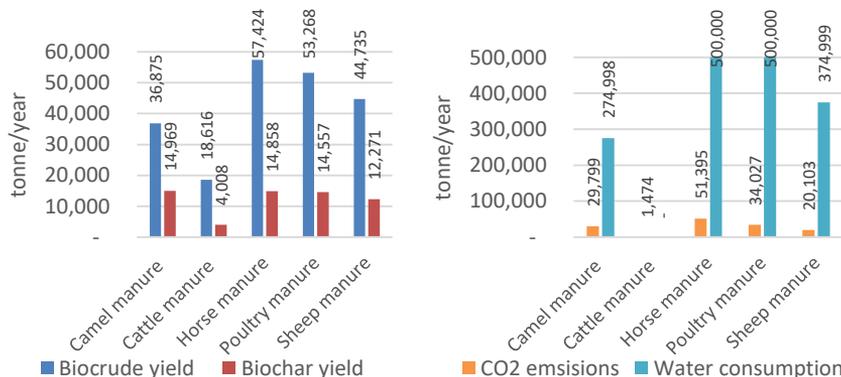


Figure 2. HTL process yield and environmental burden for different manure feedstocks.

3.2. Optimisation model

As indicated in the previous section, different livestock manure feedstocks are associated with distinct product distributions and exert different levels of burden on the environment. As such, considering the technical and environmental aspects of the HTL process, the optimisation model to investigate the optimal blending ratios of manures generated the results presented in **Figure 3**.

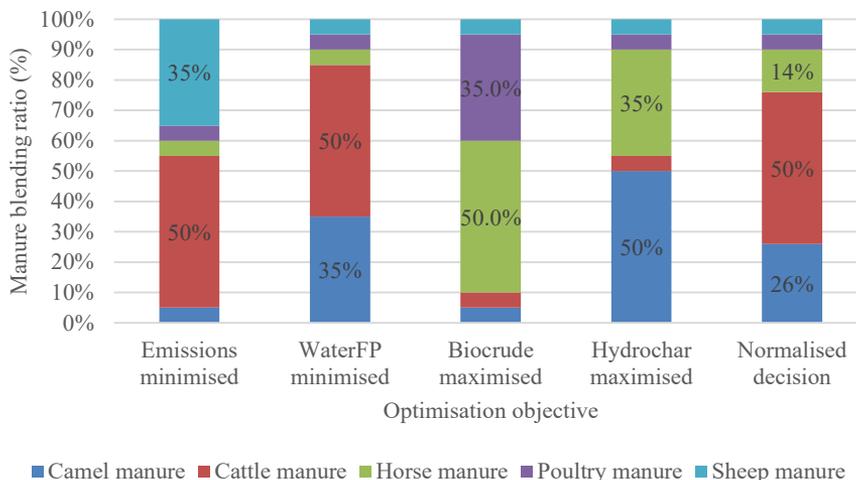


Figure 3. Optimal manures blending ratio considering different objectives.

As far as emissions reduction is concerned, cattle manure and sheep manure are granted 50% and 35% shares, respectively. Whereas the remaining three manures are given the minimum 5% threshold blending ratio. While to minimise the water consumption, both cattle manure and camel manure were selected with 50% and 35% blending ratios, respectively. However, to fulfill the yield objective, the model suggested horse and poultry manures be mainly used to maximise biocrude yield, while the model suggested mainly involving camel manures and horse manures for optimal hydrochar generation. Nevertheless, the weighted normalised decision metric indicated an optimal blending ratio of 50%, 26%, 14%, 5%, and 5% for cattle, camel, horse, poultry, and sheep manures respectively, which is believed to relatively fulfill all four objectives. The said blending ratio results in 31,765 tonne/year of biocrude and 9,289 tonne/year hydrochar, while the system in this scenario consumes 184 ktonne of water and produces 18 ktonne of CO₂.

4. Conclusion

In this study, the influence of different livestock manures feedstocks on HTL process yield and the environmental burden is modeled and studied in ASPEN PLUS. Whereas the optimal blend of different manures is investigated with respect to four objectives including, maximising biocrude and biochar yield and minimising water consumption and CO₂ emissions. The final weighted normalised decision metric suggested the optimal manures blend to be (50%, 26%, 14%, 5%, 5%) for cattle, camel, horse, poultry, and sheep manures respectively. While this study provides an insight on the impact of different feedstock on process efficiencies, it is to be expanded by considering the economic aspect of different blending scenarios, as well as the availability of different manures.

References

- Z. Akyürek, R. Singh, V. Strezov, 2021. Synergetic Effects during Co-Pyrolysis of Sheep Manure and Recycled Polyethylene Terephthalate. *Polym.* 2021, Vol. 13, Page 2363 13, 2363. <https://doi.org/10.3390/POLYM13142363>
- T. Al-Ansari, A. AlNouss, N. Al-Thani, P. Parthasarathy, S. ElKhalifa, G. McKay, M. Alherbawi, 2020. Optimising Multi Biomass Feedstock Utilisation Considering a Multi Technology Approach. *Comput. Aided Chem. Eng.* 48, 1633–1638. <https://doi.org/10.1016/B978-0-12-823377-1.50273-1>
- M. Alherbawi, P. Parthasarathy, T. Al-Ansari, H.R. Mackey, G. McKay, 2021. Potential of drop-in biofuel production from camel manure by hydrothermal liquefaction and biocrude upgrading: A Qatar case study. *Energy* 232, 121027. <https://doi.org/10.1016/J.ENERGY.2021.121027>
- C.H. Chen, Lin, Liu, C.H. Chen, Hung, C.H. Chen, Ong, W.H. Chen, Y.Y. Lin, H.C. Liu, T.C. Chen, C.H. Hung, C.H. Chen, H.C. Ong, 2019. A comprehensive analysis of food waste derived liquefaction bio-oil properties for industrial application. *Appl. Energy* 237, 283–291.
- J. Chen, L. Wang, B. Zhang, R. Li, A. Shahbazi, 2018. Hydrothermal Liquefaction Enhanced by Various Chemicals as a Means of Sustainable Dairy Manure Treatment. *Sustain.* 2018, Vol. 10, Page 230 10, 230. <https://doi.org/10.3390/SU10010230>
- C.T. Chong, G.R. Mong, J.H. Ng, W.W.F. Chong, F.N. Ani, S.S. Lam, H.C. Ong, 2019. Pyrolysis characteristics and kinetic studies of horse manure using thermogravimetric analysis. *Energy Convers. Manag.* 180, 1260–1267. <https://doi.org/10.1016/J.ENCONMAN.2018.11.071>
- Z. Lentz, P. Kolar, J.J. Classen, 2019. Valorization of Swine Manure into Hydrochars. *Process.* 2019, Vol. 7, Page 560 7, 560. <https://doi.org/10.3390/PR7090560>
- H. Li, J. Lu, Y. Zhang, Z. Liu, 2018. Hydrothermal liquefaction of typical livestock manures in China: Biocrude oil production and migration of heavy metals. *J. Anal. Appl. Pyrolysis* 135, 133–140. <https://doi.org/10.1016/J.JAAP.2018.09.010>
- T.H. Pedersen, C.U. Jensen, L. Sandström, L.A. Rosendahl, 2017. Full characterization of compounds obtained from fractional distillation and upgrading of a HTL biocrude. *Appl. Energy* 202, 408–419. <https://doi.org/10.1016/J.APENERGY.2017.05.167>
- A.A. Peterson, F. Vogel, R.P. Lachance, M. Fröling, M.J. Antal, J.W. Tester, 2008. Thermochemical biofuel production in hydrothermal media: A review of sub- and supercritical water technologies. *Energy Environ. Sci.* 1, 32–65. <https://doi.org/10.1039/B810100K>
- N. Whitely, R. Ozao, R. Artiaga, Y. Cao, W.P. Pan, 2006. Multi-utilization of Chicken Litter as Biomass Source. Part I. Combustion. *Energy and Fuels* 20, 2660–2665. <https://doi.org/10.1021/EF0503109>
- S. Xiu, A. Shahbazi, V. Shirley, D. Cheng, 2010. Hydrothermal pyrolysis of swine manure to bio-oil: Effects of operating parameters on products yield and characterization of bio-oil. *J. Anal. Appl. Pyrolysis* 88, 73–79. <https://doi.org/10.1016/J.JAAP.2010.02.011>
- S. Yin, R. Dolan, M. Harris, Z. Tan, 2010. Subcritical hydrothermal liquefaction of cattle manure to bio-oil: Effects of conversion parameters on bio-oil yield and characterization of bio-oil. *Bioresour. Technol.* 101, 3657–3664. <https://doi.org/10.1016/J.BIORTECH.2009.12.058>
- C. Zhong, X. Wei, 2004. A comparative experimental study on the liquefaction of wood. *Energy* 29, 1731–1741. <https://doi.org/10.1016/J.ENERGY.2004.03.096>
- Y. Zhou, Z. Chen, H. Gong, X. Wang, H. Yu, 2020. A strategy of using recycled char as a co-catalyst in cyclic in-situ catalytic cattle manure pyrolysis for increasing gas production. *Waste Manag.* 107, 74–81. <https://doi.org/10.1016/J.WASMAN.2020.04.002>

Systematic Pairing Selection for Economic-oriented Constraint Control

Risvan Dirza, and Sigurd Skogestad*

*Dept. of Chemical Engineering, Norwegian Univ. of Science & Technology (NTNU),
NO-7491 Trondheim, Norway
sigurd.skogestad@ntnu.no*

Abstract

This work considers the problem of minimizing economic losses due to system-wide production systems, where different subsystems share hard coupling constraints. The hard coupling constraints need to be tightly controlled, and it is important that it is done in a way that the overall system remains close to optimal in the time it takes for the much slower optimization layer to implement the required input changes. The particular application that we study is a large-scale subsea gas-lifted oil production network, where different subsystems have a local objective and the shared constraint can be a common compressor, but the method has general applicability to any system with time-scale separation between control and optimization layers.

Keywords: Production Optimization, Self-optimizing Control, Active Constraint Control.

1. Introduction

Determining optimal operation of a large and complex process and production system, such as an oil and gas production system, is a challenging task. Decomposing the process into several subprocesses/subsystems is usually recommended since optimizing a small system is practically less complex. Thus, decomposition strategy requires each local process system/cluster/subsystem to have a local optimizer to ensure local optimal process operation. This decomposition strategy is also responsible for coordinating these subsystems to achieve system-wide optimal process operation. The optimal process operation involves making decisions in real-time to meet production goals. This is typically done in the context of real-time optimization (RTO) using process models and real-time measurements. RTO is developed based on mathematical concepts, and with it, production performance improved.

In the 80s, there was an increasing interest in replacing model-based numerical solvers with a simple feedback loop, named feedback-optimizing control. The idea is to translate the economic objective into process control objective by finding a function of the controlled variables (CVs), and when it is held constant, it leads to the optimal adjustment of the manipulated variables (MVs) (Morari et al., 1980). Twenty years later, Skogestad (2000) introduced the concept of self-optimizing control (SOC). In SOC, when the optimum lies at some constraints, we use active constraint control where the available MVs tightly control the constrained variables. The idea of tight active constraint control is one of the primary motivations of this work to ensure the feasibility and obtain a (near-) optimal process operation.

2. Problem Statement

Consider the following steady-state optimization problem of N different subsystems.

$$\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d}) = \sum_{i=1}^N J_i(\mathbf{u}_i, \mathbf{d}_i) \quad (1a)$$

$$s. t. \quad g(\mathbf{u}, \mathbf{d}) \leq \mathbf{0} \quad (1b)$$

where $\mathbf{u}_i \in \mathbb{R}^{n_{\mathbf{u}_i}}$ denotes the MVs for subsystem i , $n_{\mathbf{u}_i}$ is the number of MVs in subsystem i , and $\mathbf{u} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_N]^T$, $\mathbf{d}_i \in \mathbb{R}^{n_{\mathbf{d}_i}}$ denotes the disturbances in subsystem i , $n_{\mathbf{d}_i}$ is the number of disturbances in subsystem i , and $\mathbf{d} = [\mathbf{d}_1 \ \dots \ \mathbf{d}_N]^T$, $J_i: \mathbb{R}^{n_{\mathbf{u}_i}} \times \mathbb{R}^{n_{\mathbf{d}_i}} \rightarrow \mathbb{R}$ is a function denoting the local objective of subsystem i , $g: \mathbb{R}^{n_{\mathbf{u}}} \times \mathbb{R}^{n_{\mathbf{d}}} \rightarrow \mathbb{R}^{n_g}$ is a function denoting the inequality constraints. n_g is the number of constraints.

The Lagrangian function of problem (1) is

$$\mathcal{L}(\mathbf{u}, \mathbf{d}, \lambda) = \sum_{i=1}^N J_i(\mathbf{u}_i, \mathbf{d}_i) + \lambda^T g(\mathbf{u}, \mathbf{d}) \quad (2)$$

where $\lambda \in \mathbb{R}^{n_g}$ is the shadow price/ Lagrange multiplier of active constraints $g(\mathbf{u}, \mathbf{d})$.

The goal of problem (1) is to determine optimal MVs to achieve system-wide steady-state optimal operation. Our motivation is to solve problem (1) using a feedback control structure that handles changing active constraints.

One possible approach is primal-dual feedback-optimizing control that can eliminate the need for a numerical solver (Krishnamoorthy, 2020; Dirza et al., 2021). Moreover, this approach is flexible in handling active constraint changes. This method has a central coordinator acting as a central constraint controller in a slow timescale in the upper layer. However, this approach has no near-optimal performance strategy due to the non-performing upper layer. There are many practical reasons why the central constraint controller may fail to update the Lagrange multipliers. For example, when the disturbance occurs much faster than the sampling time of the central constraint controller. Another example is when constrained variables from a local system are not updated on time since the optimizer of the other local system may need a numerical solver. This solver requires time to solve the optimization problem. Having different types of local optimizers is normal since every subprocess is unique. In addition, having a central constraint controller in a very slow timescale leads to a longer transient. Constraint violation may occur during this transient. These imply that it is essential to have a good pairing of a (primal) MV to a constrained variable. Thus, “*a systematic pairing procedure is necessary to determine which MV should be paired with a constrained variable such that the pairing minimizes the loss most (near-optimal performance)*”. Finally, this procedure is necessary for selecting the primal MV in primal-dual with direct constraint control proposed by Dirza et al., 2022.

3. Systematic Pairing Formulation

To pair the constrained variables with the right MV, we propose a pairing procedure based on MV’s sensitivities to its local disturbances, assuming no saturation issues in the possible MVs, no back-off problem, and equal value of constraints - MVs gain.

Meanwhile, the remaining MVs control their self-optimizing control variables. To describe this proposal, we consider an Indirect control problem formulation.

Without losing the generality, we consider a case where we have two MVs (i.e., $u_1 \in \mathbb{R}^1$ and $u_2 \in \mathbb{R}^1$), and we want to control the gradient of the Lagrange function w.r.t its input, denoted by $\mathcal{L}_{\mathbf{u}}(\lambda, \mathbf{u}) \in \mathbb{R}^{2 \times 1}$, and the active constrained variable, denoted by $g(\mathbf{u}) \in \mathbb{R}^1$, where λ is Lagrange multiplier for constraint function $g(\mathbf{u})$, and $\mathbf{u} = [u_1 \ u_2]^T$. $\mathcal{L}_{\mathbf{u}}(\lambda, \mathbf{u})$ consists of $\mathcal{L}_{u_1}(\lambda, u_1) \in \mathbb{R}^1$ and $\mathcal{L}_{u_2}(\lambda, u_2) \in \mathbb{R}^1$.

Assume that we want to control the constrained variables tightly with u_2 , and we consider the disturbance, $\mathbf{d} \in \mathbb{R}^{2 \times 1}$, influences input u_2 . It could be the local disturbance of subsystem 2 or a change of u_2 caused by setpoint changing. This setpoint change may occur due to the changes in subsystem 1. Since λ is constant (due to the non-performing upper layer), and $g_{u_1}(u_1)$ is also in many cases (i.e., resource allocation), we only need to control $J_{u_1}(u_1)$. This formulation can be written as an indirect control problem as follows,

$$J_{u_1}(u_1) = G_{11}u_1 + G_{12}u_2 \quad (3a)$$

$$g(\mathbf{u}) = G_{21}u_1 + G_{22}u_2 \quad (3b)$$

$$u_2 = G_d d + \hat{u}_2 \quad (3c)$$

where G_{11} is the gain from u_1 to $J_{u_1}(u_1)$, G_{12} is the gain from u_2 to $J_{u_1}(u_1)$, G_{21} is the gain from u_1 to $g(\mathbf{u})$, G_{22} is the gain from u_2 to $g(\mathbf{u})$ and G_d is the disturbance gain that influences u_2 .

Fig. 1 illustrates this formulation, where we want to ‘tightly’ control $g(\mathbf{u})$ to reference r_2 directly using a direct constraint controller (DCC). In addition, we also want to find the right u_2 such that u_2 can also contribute to controlling $J_{u_1}(u_1)$ to reference r_1 indirectly or by pairing u_2 with $g(\mathbf{u})$. This control structure has a better ability to control $J_{u_1}(u_1)$ than the other possible structure.

We assume that G_{22} is square and invertible. Otherwise, we can replace the solution with the pseudoinverse. By rearranging Eq. 1 and assuming a ‘perfect’ control $g(\mathbf{u}) \approx r_2$, we obtain $J_{u_1}(u_1) \approx G_{12}G_{22}^{-1}r_2$. Thus, we must choose r_2 such that $r_2 \approx G_{22}G_{12}^{-1}r_1$. According to Skogestad and Postlethwaite (2005), $G_{12}G_{22}^{-1}$ should be small. Usually, it implies that we need to select the pairing that gives the largest G_{22} , where $G_{22} = \nabla_{u_2}g(\mathbf{u})$. However, based on this formulation, selecting the pairing based on G_{22} is insufficient. This formulation shows that we should also consider small G_{12} in addition to large G_{22} in the framework of primal-dual approach with direct constraint control, especially when we have faster disturbances (or non-performing upper layer). Selecting based on G_{12} is then essential and complementary to the common rule.

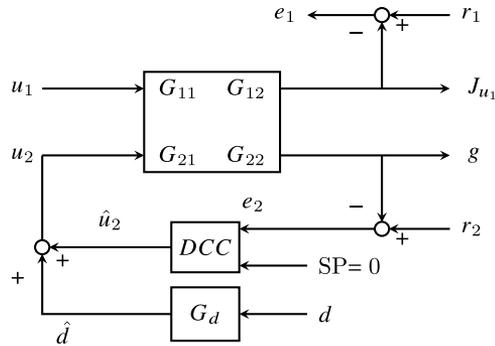


Figure 1: Indirect control problem formulation for systematic pairing

We consider these rules as a near-optimal performance strategy for the primal-dual with direct constraint controls framework (Dirza et. al., 2022).

Defining $\hat{d} = G_d d$, then $G_{12} \approx \frac{\Delta J_{u_1}}{\Delta u_2} \approx \frac{\Delta J_{u_1}}{\Delta(\hat{d} + \hat{u}_2)}$. If one keeps \hat{u}_2 at the same value to control $g(\mathbf{u})$, then a change in \hat{d} can represent any change. Considering Eq. 1c, then $G_{12} \approx \frac{\Delta J_{u_1}}{\Delta \hat{d}}$. Furthermore, assuming that the stationary point is at the local optimum and knowing that J_{u_1} is controlled by u_1 , and J_{u_2} is uncontrolled, then any disturbance on J_{u_1} leads to ΔJ_{u_2} (\mathcal{L}_{u_2} being drifted away from 0). It implies that any disturbance on J_{u_1} leads to the total profit loss ΔJ . Therefore, we can estimate G_{12} as $\frac{\Delta J}{\Delta \hat{d}}$ ($G_{12} \approx \frac{\Delta J}{\Delta \hat{d}}$).

4. Numerical Results

We demonstrate the presented rules in a subsea gas-lifted oil production optimization problem with a fixed gas lift compressor described in Dirza et al. (2022). Moreover, we consider a subsea gas-lifted oil production well network that consists of two wells to provide a better demonstration. Fig. 2 illustrates the case study.

The objective function is to maximize the oil production income while minimizing the cost of the gas lift. The optimization problem is as follows,

$$\min_{\mathbf{w}_{gl}} J = \sum_{i=1}^N (-p_{o,i} w_{po,i} + p_{gl,i} w_{gl,i}) \quad (4a)$$

$$s. t. \quad \mathbf{f}(\mathbf{x}, \mathbf{w}_{gl}, \mathbf{d}) = \mathbf{0} \quad (4b)$$

$$\mathbf{g}(\mathbf{x}, \mathbf{w}_{gl}, \mathbf{d}) \leq \mathbf{0} \quad (4c)$$

$$\mathbf{g}_s(\mathbf{x}, \mathbf{w}_{gl}, \mathbf{d}) = Pow_{gl} - Pow_{gl}^{max} \leq 0 \quad (4d)$$

where $p_{o,i}$, $p_{gl,i}$, and $w_{po,i}$ are the price of produced oil, the cost of gas-lift, and the oil production rate of well i , respectively. Pow_{gl} is the total power consumed by a fixed compressor to inject the sum of gas-lift rate i , and Pow_{gl}^{max} is the maximum available power. The vector $\mathbf{x} \in \mathbb{R}^{n_x}$, and $\mathbf{d} \in \mathbb{R}^{n_d}$ are the vectors of states, and disturbance (i.e., gas-oil ratio) for the entire system. n_x is the number of states. $\mathbf{w}_{gl} \in \mathbb{R}^{n_{w_{gl}}}$ is the vector of inputs for the entire system, where $\mathbf{w}_{gl} = [w_{gl,1} \ \dots \ w_{gl,N}]^T$. Constraint (4b) and (4c) represent model and physical constraints, respectively. We assume that Constraint (4c) is locally managed to maintain the focus of the discussion. Eq. (4a) is additively separable, and eq. (4d) is a linear and hard constraint. The total gas lift rate is supplied by a fixed-efficiency gas lift compressor.

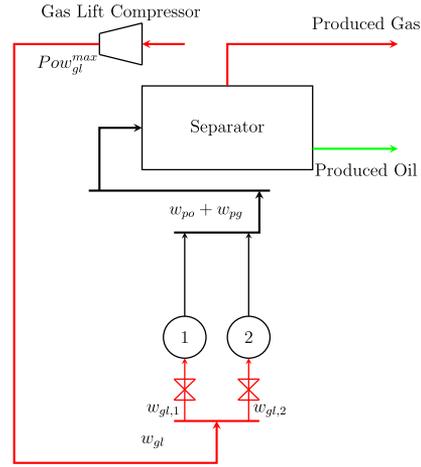


Figure 2: Field illustration

The simulation considers a case where we have a non-performing upper layer (λ is not updated). We show the numerical results of the presented near-optimal performance strategy (Structure 1). As a benchmark, we also show the results of the asynchronous protocol (Structure 0), that the local controllers keep controlling the gradient of the Lagrange function w.r.t input to 0, given any value of the

Lagrange multipliers from the central constraint controllers. In addition, we also show the results of another possible structure (Structure 2).

We solve the steady-state optimization problem (4) to obtain the ‘true’ optimal cost for any considered disturbance cases. We assume that based on historical data, the largest possible error of the disturbance is $\pm 5\%$. The (profit-) loss is the difference between the steady-state cost of structure j to the optimal cost, which can be expressed mathematically as $\Delta J_j = J_j - J^*$, where j is the index of the structure (*i. e.*, $j \in \{0,1,2\}$).

First, we simulate for any possible largest error for Structure 0. The simulation shows that the above case study experiences the largest possible disturbance that happens sequentially starting from $GOR_1 + 5\%$, $GOR_1 - 5\%$, $GOR_2 + 5\%$, $GOR_2 - 5\%$, $Pow_{gl}^{max} + 5\%$, and finally $Pow_{gl}^{max} - 5\%$.

As it can be seen in Fig. 3, Structure 0 fails to satisfy steady-state constraint when either GOR_1 , GOR_2 , or Pow_{gl}^{max} decreases 5% (see time window 18-32 hr, 48-62 hr, and 78-90 hr), which validates the necessity to have a near-optimal strategy in the primal-dual approach.

As mentioned in Section 3, the first general rule is pairing input and active constraint with the largest $G_{22,j} = \nabla_{w_{gl,j}} \mathbf{g}_s(\mathbf{x}, \mathbf{w}_{gl}, \mathbf{d})$. Based on this definition, $G_{22,1} = 3.6740$, and $G_{22,2} = 3.6740$, which corresponds to the assumption of equal value of constraints - MVs gain. This also validates the necessity to have an additional rule to select the pairing that gives the most economic-oriented result.

The additional rule is pairing input and active constraint with the smallest $G_{12,j}$, which one can estimate by calculating $\frac{\Delta J_j}{\Delta GOR_j}$ using the finite difference method. The obtained result is that the smallest $G_{12,1}$ is 1.4441, and the smallest $G_{12,2}$ is 1.4642. According to

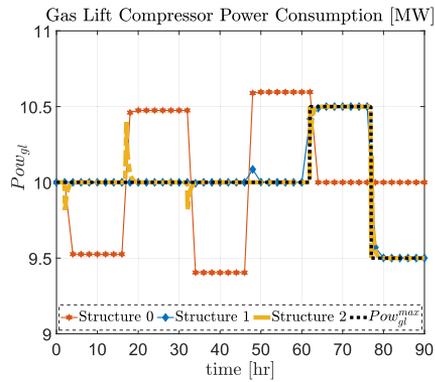


Figure 3: Steady-state constraint satisfaction

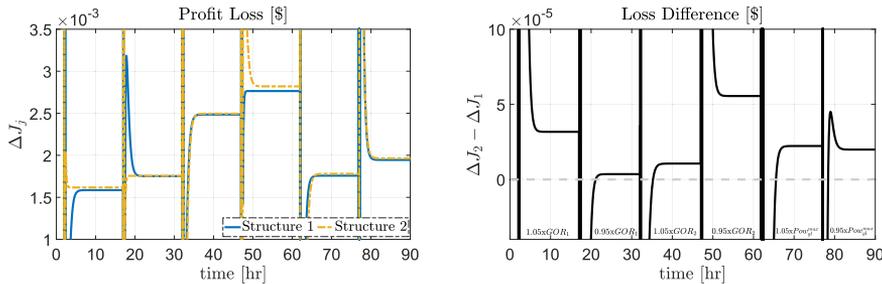


Figure 4: Left figure: Profit loss comparison. Right figure: Loss difference between Structure 1 and Structure 2.

the presented method in Section 3, this result indicates that the most economic-oriented pairing is Structure 1, where we pair the active constraint with $w_{gl,2}$.

Fig. 4 shows the profit loss comparison between Structure 1 and Structure 2, and the right figure shows that, at any possible extreme disturbance, Structure 1 can minimize more the steady-state loss than Structure 2. Additionally, Tab. 1 shows the steady-state profit loss for 24 hours with different extreme disturbance cases.

Table 1: Steady-state profit loss for 24 hours

Structure	GOR_1		GOR_2		Pow_{gl}^{max}	
	+5%	-5%	+5%	-5%	+5%	-5%
0	630.7200	sscv*	751.6800	sscv*	656.6400	sscv*
1	136.3423	151.6720	214.7048	238.8445	151.4331	168.2676
2	139.7161	151.6730	215.0000	243.8771	153.3618	169.9576

sscv*: steady-state constraint violation.

5. Conclusion

In this paper, we have shown that the proposed rule (smallest G_{12}) is complementary to the existing pairing rule (largest G_{22}), especially in the framework of the primal-dual approach. This systematic pairing selection procedure can assist the designer in pairing for economic-oriented constraint control in the primal-dual with direct constraint controls. In addition, this procedure can minimize steady-state loss in the primal-dual framework when we have a non-performing upper layer.

6. Acknowledgment

The authors gratefully acknowledge the financial support from SUBPRO, which is financed by the Research Council of Norway, major industry partners, and NTNU.

References

- R. Dirza, S. Skogestad, D. Krishnamoorthy, 2021. Optimal resource allocation using distributed feedback-based real-time optimization. IFAC-PapersOnLine 54 (3), 706–711, 16th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2021.
- R. Dirza, S. Skogestad, and D. Krishnamoorthy, 2022. Primal-dual Feedback-optimizing Control with Direct Constraint Control for Oil Production. Proceedings of the 14th International Symposium on Process Systems Engineering – PSE 2021+ (Accepted).
- D. Krishnamoorthy, 2020. A distributed feedback-based online process optimization framework for optimal resource sharing. Journal of Process Control 97, 72-83.
- M. Morari, Y. Arkun, G. Stephanopoulos, 1980. Studies in the synthesis of control structures for chemical processes: Part i: Formulation of the problem. process decomposition and the classification of the control tasks analysis of the optimizing control structures. AIChE Journal 26, 220 – 232.
- S. Skogestad, 2000. Plantwide control: the search for the self-optimizing control structure. Journal of process control 10, 487 – 507.
- S. Skogestad, and I. Postlethwaite, 2005. Multivariable Feedback Control. John Wiley & Sons.

Quality assessment for dynamic, hybrid semi-parametric state observers

Isabell Viedt,^a Jonathan Mädler,^a Leon Urbas^a

^a*Technische Universität Dresden, 01062 Dresden, Germany*

isabell.viedt@tu-dresden.de

Abstract

In modular plants, quality of methods and models must be specifiable, automatically testable and certifiable, if they are increasingly integrated into equipment as a product and an added value. Since the underlying processes are often not completely understood, hybrid and data-driven methods are a promising approach to combine process knowledge and process data for more reliable simulation models. In this paper, the conducted quality assessment utilizes the framework proposed by Mädler et al. (2021) for quality assurance applied to hybrid models. The quality model is revised to include quality factors, criteria and metrics for dynamic, hybrid semi-parametric simulation models. A state observer for the estimation of key process parameters during fermentation is presented as a use case. For this three hybrid models of the fermentation with differing levels of detail are identified and coupled with an extended Kalman filter (EKF). It was found that the quality model can successfully be used to assess quality differences in different types of state observers. The quality model allows a structured and quick assessment and is therefore able to show e.g. the performance improvement of the different hybrid models coupled with an EKF. With the transfer of the quality model to hybrid state observers a broader range of simulations models can be assessed within the framework.

Keywords: Quality assessment, hybrid semi-parametric models, state observer.

1. Introduction

Simulation models as part of software and further as part of smart equipment are envisioned to become increasingly important for smart, modern plant as part of the 4th industrial revolution. Soft sensors or model predictive controllers will be essential parts of the process control and optimization (Kadlec et al., 2009). The quality of those simulation models must be specifiable, automatically testable and certifiable, if they are increasingly integrated into smart equipment. In this context, hybrid modeling approaches are an increasingly popular trend to combine knowledge- and data-driven modeling which makes the models more reliable and precise (Glasse & von Stosch, 2018). With this, it is not only important to assure the quality and correctness of the first principles approach but also process data has to be quality assured.

The remainder of the paper is structured as follows. Section 2 introduces current methods for quality assessment of simulation models and applies it to hybrid, semi-parametric models. In section 3 the case study for the quality assessment in accordance to Mädler et al. (2021) is presented and evaluated in section 4. Section 6 provides a conclusion and further research potential.

2. State of the Art

2.1. Quality assessment of simulation models

To assess the quality of simulation models, different strategies exist in literature (Murray-Smith, 2015). The most commonly used ones are the methods for verification and validation (V&V) presented by Balci and Sargent (Sargent & Balci, 2017). Although the V&V methods provide a wide spectrum of assessment methods, they mainly focus on model accuracy. Sargent and Balci (2017) find that many simulation studies do not consequently apply V&V methods or even disregard them. Therefore, Mädler et al. (2021) discuss the applicability of quality assurance methods from software development to simulation models. To make software quality itself measurable and testable, quality assessment strategies from software development like test-driven-development and quality models in form of FCM models are applied. FCM stands for factor (F), criterion (C) and metric (M). Factors describe the desired abstract attributes of the software.

2.2. Dynamic, hybrid semi-parametric models

With the expansion of the framework presented in Mädler et al. (2021) to the field of hybrid state observers, a short introduction to gray-box modeling, and here specifically dynamic hybrid, semi-parametric models, is given. An overview on hybrid, semi-parametric simulation is shown in Figure 1. Mechanistic models represent a broad class of more transparent models that are usually based on e.g. conservation laws (Glassey & von Stosch, 2018). In contrast, data-driven modeling represents a less transparent modeling framework based exclusively on process data (Glassey & von Stosch, 2018).

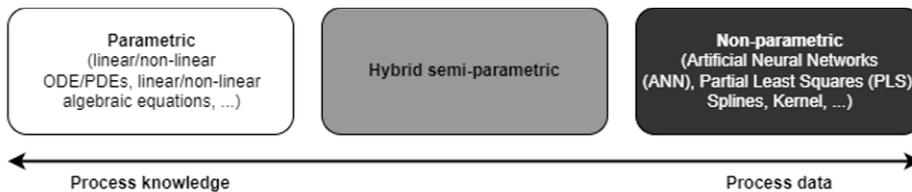


Figure 1: Hybrid, semi-parametric models.

Hybrid semi-parametric models are an example of those gray-box models and combine model structures that are parametric and non-parametric sub-models based on different knowledge sources. Hybrid semi-parametric models have several advantages over traditional mechanistic or data-driven modeling (Bae et al., 2020). Some widely recognized advantages in academia and industry are a broader knowledge base, the transparency of the modeling approach and cost-effective model development (Glassey & von Stosch, 2018). Especially for complex process simulation, a hybrid approach is able to achieve a more accurate and robust estimation (Bae et al., 2020). Existing process data can be used in gray-box models as ML or AI algorithms bridge gaps in the mechanistic modeling approach (Glassey & von Stosch, 2018).

2.3. Functional quality assessment of hybrid, semi-parametric models

When regarding the quality of hybrid semi-parametric models, the structure that combines aspects of parametric and non-parametric modeling has to be considered. So therefore not only quality and correctness of the first principles approach but also the data for non-parametric components of the model have to be equally quality assured. In case of simulation model for smart equipment, the simulation models will rely on continuous integration and continuous deployment strategies (CI/CD) (c.f. Bruckner et al., 2020), which require a high level of transparency and an overview over the quality advancement with time. As presented in Mädler et al. (2021) the first step within the model

development should always be the specification of the purpose and the corresponding requirements. In the next step, a generic FCM model must be adapted or modified regarding the stated purpose. For this framework, the factors and criteria specified within ISO/IEC 25010 are the basis for the generic FCM model. For hybrid models, the adapted FCM model must contain factors, criteria and metrics to equally assess the parametric and non-parametric sub-models (Bae et al., 2020). After determining the target ranges and values for the metrics, the model building can begin. Using the FCM model, the development and quality advancement of the model is tested continuously. If the simulation model fulfills the targets, the user can exit model development successfully.

3. Case study

With the following case study, the framework for quality assessment of hybrid, semi-parametric simulation models is evaluated. The underlying process of the simulation model for the case study is a baker's yeast fermentation and is used for online estimation of biomass concentration c_x and substrate concentration c_s during the process (c.f. de Azevedo et al., 1997) in form of a hybrid state observer.

3.1. Baker's Yeast Fermentation

The underlying process for the use case is a baker's yeast fermentation. The estimated process parameters are in this case the biomass concentration. This process parameter is good quality indicator for the biochemical process of fermentation (De Azevedo et al., 1997). The mechanistic model for the baker's yeast fermentation is obtained from mass balances for all components. The relevant set of model equation for the process of fermentation is shown in the following:

$$\frac{dc_x}{dt} = -\frac{\dot{V}}{V}c_x + (\mu_s^o + \mu_s^r + \mu_E^o)c_x \quad (1)$$

$$\frac{dc_s}{dt} = \frac{\dot{V}}{V}(c_{s,in} - c_s) + \left(-\frac{\mu_s^o}{Y_{X/S}^o} - \frac{\mu_s^r}{Y_{X/S}^r}\right)c_x \quad (2)$$

$$\frac{dc_E}{dt} = -\frac{\dot{V}}{V}c_E + \left(\frac{\mu_s^r}{Y_{X/E}^r} - \frac{\mu_E^o}{Y_{X/E}^{oE}}\right)c_x \quad (3)$$

$$\frac{dc_{O_2}}{dt} = -\frac{\dot{V}}{V}c_{O_2} - OUR + OTR \quad (4)$$

Further, the model includes the oxygen transfer rate (OTR) and the oxygen uptake rate (OUR). For the hybrid, semi-parametric approach of the use case of the case study two different hybrid semi-parametric models are identified. These models combine the parametric approach of the model equations (1) – (4) for the fermentation with a non-parametric approach for the estimation of the specific growth rates $(\mu_s^o, \mu_s^r, \mu_E^o)$. The growth rates are estimated via existing process data.

3.2. Use case

The estimation of the biomass concentration is implemented as a soft sensor which is structured as a non-linear state observer with the state tracking by means of an EKF with the state transition achieved through the hybrid models presented (Auger et al., 2013) in section 3.1. This soft sensor is envisioned to be used in a smart fermenter module during the learning and application phase of the equipment and is shown in figure 2. The use case is the process monitoring (PM) or possibly also online prediction (OP) (Kadlec et al., 2009) of critical process parameters. To achieve the online estimation, the state estimation algorithm of the EKF uses the online measurements of the oxygen

concentration (c_{O_2}) and the carbon dioxide evolution rate (CER). The hybrid state observer structure is shown in figure 2.

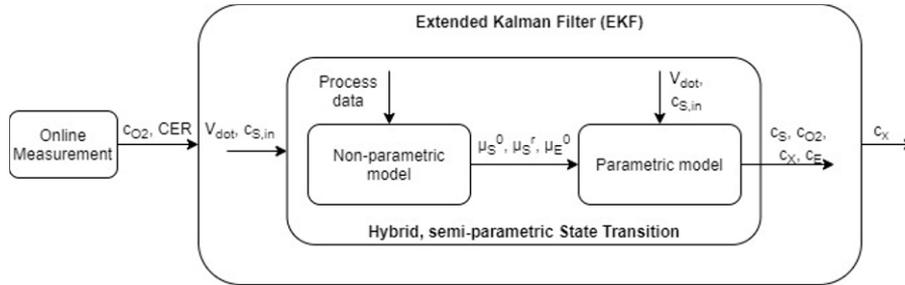


Figure 2: Structure of the hybrid state observer for estimation of biomass concentration.

To show how CI/CD in modular plants (MPs) is applied, the case study shows a fitting of the simulation model in an iterative identification process that represent different quality levels of the state observer. For this, the hybrid model is identified and further quality assessed. After that the actual performance of the soft sensor is compared to the prediction of the quality metric for the simulation model and adapted accordingly. The metrics are able to predict how good the soft sensor will be during operation. The individual cases of the case study are designed to emulate quality differences in the evaluated simulation models and are used to assess the usefulness and accuracy of the semi-automated assessment framework for hybrid semi-parametric simulation models. For the case study, Case 1 will be the version of the state observer that uses the first iteration of the hybrid model for the state transition in the EKF algorithm. To emulate quality differences in the models, different aspects of the model can be modified in the hybrid model. This can be achieved by increasing the noise in the data but also in reducing the size of the data sets or injecting faults into the sets (Kadlec et al., 2009), which affects the accuracy of the non-parametric estimation of the specific growth rates. Inaccurate model parameters, which in this use case is intentional, lead to a less accurate prediction of the model (Case 2). In comparison, Case 3 shows the best quality hybrid model, after two iterations.

3.3. Specified quality model and relevant metrics

To evaluate the quality of the hybrid simulation model, the generic FMC model has to be adapted to the model use case (Table 1). The quality factor functional suitability is described by the criteria functional correctness, appropriateness and completeness. Relevant metrics for the criterion functional correctness are the coefficient of determination (R^2) and the number of domain violations. Another important metric for state observers and is the difference between estimated and measured values (RMSE - Root Mean Square Error), which covers the criterion functional appropriateness. The criterion functional completeness is evaluated by the coverage metric and the model structure metric. The coverage metric describes how the modeling data covers the validity domain of the black-box sub-model and therefore the data quality (Mädler et al., 2021). For the factor reliability the criterion fault tolerance is considered. The corresponding metrics evaluate the model prediction. The metric outlier frequency describes the consistency of the estimated values, which is classified through MAD (Median Absolute Deviation) filter with moving window for more accurate detection. The metric tolerance to fault injection describes how the model reacts to incorrect inputs, which emulates fluctuations during the plant operation. For hybrid models, the metrics have to be able to evaluate both parts of the model. For example, the coverage metric and the data consistency metric can only be applied to the part of the model that is data-driven. Metrics

like the RMSE or R^2 assess the overall quality of the simulation model by evaluating the entire model prediction.

Table 1: FCM model for the quality assessment of hybrid simulation models.

Factor	Criterion	Metric	Equation	Classification	Source
Functional Suitability	Functional Correctness	Prediction consistency	R^2 (determination coefficient)	Prediction quality	Roungas et al, 2017
		Domain violations	Characteristic process boundaries	Prediction quality	Sargent & Balci, 2017
	Functional Appropriateness	Prediction accuracy	RMSE (Root Mean Square Error)	Prediction quality	Montgomery & Runger, 2010
	Functional Completeness	Coverage	Convex hull criterion	Data quality	Mädler et al, 2021
		Model components	Model requirements	Model structure quality	Sargent & Balci, 2017
Reliability	Fault Tolerance	Outlier frequency	MAD filter with moving window	Prediction quality	Roungas et al., 2008
		Data consistency	Error ratio	Data quality	Heinrich et al., 2018
		Fault injection tolerance	Tolerance to fault injection	Prediction quality	Roungas et al., 2008

4. Results and discussion

The quality assessment framework was able to distinguish between the quality differences in the simulation model of the case study. The results, ratings from 0 (insufficient) to 10 (excellent), for the considered metrics during assessment are shown in Table 2.

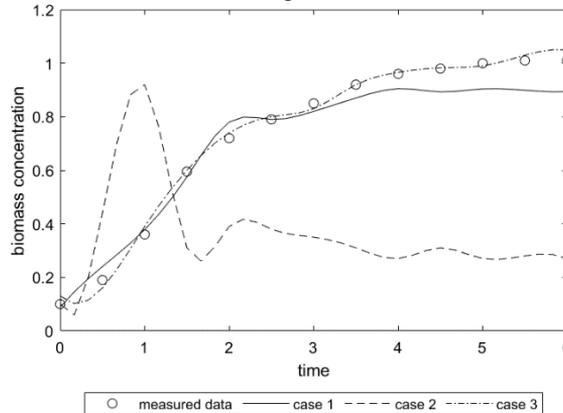


Figure 3: State estimation of the four soft sensors resulting from Case 1 -3.

The assessment shows that the quality model allows easy and automated quality assessment of different model types. With the assessed quality differences in the hybrid, semi-parametric models, the quality of the resulting soft sensor can be predicted. Figure 3 shows the estimation of the biomass concentration with the resulting soft sensors in comparison to measured data. For now, the quality model for the quality assessment only considers the quality factors functional suitability and reliability. In further work it must be expanded to contain factors and criteria to assess attributes like portability, maintainability and performance efficiency of the simulation models to fit the use case of simulation models in smart equipment better. It is envisioned that those simulation models or software components are to be reused in different process scenarios with only minimal necessary adaption.

Table 2: Results of the case study for the model quality.

Factor	Criterion	Metric	Case 1	Case 2	Case 3
Functional Suitability	Functional Correctness	Prediction consistency	8	7	9
		Domain violations	10	3	10
	Functional Appropriateness	Prediction accuracy	8	6	10
	Functional Completeness	Coverage	9	8	9
Reliability	Fault Tolerance	Model components	10	10	10
		Outlier frequency	8	6	10
		Data consistency	8	4	9
		Fault injection tolerance	5	3	8
Aggregated model rating			7,92	5,50	9,38

5. Conclusion

In this paper, the authors show how the in Mädler et al. (2021) presented framework for functional quality assessment can be transferred to dynamic, hybrid semi-parametric state observers. Additional requirements that are necessary for the assessment of dynamic, hybrid semi-parametric models were introduced into the framework. It therefore shows that the generic quality model can easily be adapted and expanded to a new use case and is able to show e.g. the performance improvement of a hybrid state observer over time when the EKF algorithm is coupled with hybrid, semi-parametric model. The flexibility of the generic structure of the quality model allows an easy combinations of metrics that target different parts of the simulation model and make an inclusive view on quality possible. In future work, the quality model must be expanded by further quality factors. The quality assessment must also include factors like portability, performance efficiency and compatibility to meet the requirements for the flexible MP structure.

References

- J. M. Ali, N. H. Hoang, M. A. Hussain, and D. Dochain, 2015, Review and classification of recent observers applied in chemical process systems, *Comput. Chem. Eng.*, 76, 27-41
- F. Auger, M. Hilaret, J. M. Guerrero, E. Monmasson, ... and S. Katsura, 2013, Industrial applications of the Kalman filter: A review, *IEEE Trans. Ind. Electron.*, 60, 12, 5458-5471
- J. Bae, H. J. Lee, D. H. Jeong, and J. M. Lee, 2020, Construction of a Valid Domain for a Hybrid Model and Its Application to Dynamic Optimization with Controlled Exploration, *Industrial & Engineering Chemistry Research*, 59, 37, 16380-16395
- D. Bonvin, C. Georgakis, C. C. Pantelides, M. Barolo, M. A. Grover, D. Rodrigues, ... and D. Dochain, 2016, Linking models and experiments, *Ind. Eng. Chem. Res.*, 55, 25, 6891-6903
- L. Bruckner, M. Oppelt, L. Urbas, and M. Barth, 2020, The current and future use of simulation in discrete and process industries
- S. F. De Azevedo, B. Dahm, and F. R. Oliveira, 1997, Hybrid modelling of biochemical processes: A comparison with the conventional approach, *Comput. Chem. Eng.*, 21, 751-756
- J. Glassey, and M. Von Stosch, (Eds.), 2018, *Hybrid modeling in process industries*, CRC Press
- B. Heinrich, D. Hristova, M. Klier, A. Schiller, and M. Szubartowicz, 2018, Requirements for data quality metrics, *Journal of Data and Information Quality (JDIQ)*, 9, 2, 1-32
- P. Kadlec, B. Gabrys, and S. Strandt, 2009, Data-driven soft sensors in the process industry, *Comput. Chem. Eng.*, 33, 4, 795-814
- D. C. Montgomery, and G. C. Runger, 2010, *Applied statistics and probability for engineers*. John Wiley & Sons
- D. J. Murray-Smith, 2015, *Testing and validation of computer simulation models*, Cham: Springer, 10, 978-3
- J. Mädler, I. Viedt, and L. Urbas, 2021, Applying quality assurance concepts from software development to simulation model assessment in smart equipment, *Comput. Aided Chem. Eng.*, European Symposium on Computer Aided Process Engineering (ESCAPE), Istanbul
- B. Rongas, S. A. Meijer, and A. Verbraeck, 2018, A framework for optimizing simulation model validation & verification, *Int. Journal on Advances in Systems and Measurements*, 11, 1-2
- R. G. Sargent, and O. Balci, 2017, December, History of verification and validation of simulation models, In 2017 Winter Simulation Conference (WSC), IEEE, 292-307

ANN-assisted optimization-based design of energy-integrated distillation columns

Kai Fabian Kruber^{a,b}, Anna Miroschnitschenko^b and Mirko Skiborowski^{a,b,*}

^a*Hamburg University of Technology, Institute of Process Systems Engineering, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany*

^b*TU Dortmund University, Department of Biochemical and Chemical Engineering, Laboratory of Fluid Separations, Emil-Figge-Strasse 70, 44227 Dortmund, Germany*
mirko.skiborowski@tuhh.de

Abstract

The optimal design of chemical processes is of essential importance for an increased sustainability. However, the resulting non-convex mixed-integer nonlinear programming (MINLP) problems cannot directly be solved to global optimality. Therefore, different alternatives have been proposed, which either build on the application of a simulation-based optimization by means of a metaheuristic or the global optimization of a surrogate model, both requiring extensive simulations. The current work proposes a novel alternative approach for a surrogate-assisted hybrid optimization, which exploits a local deterministic optimization of a full MINLP problem to generate a compact artificial neural network (ANN) model that allows for the direct optimization on a reduced search space. In order to provide a sufficient accuracy of the ANN while targeting the global optimum of the design problem, a tailored mixed adaptive sampling is introduced. Application of the algorithm is illustrated for the optimal design of a distillation-based separation of benzene, toluene, and xylene with different means for energy integration.

Keywords: optimization, artificial neural networks, distillation, energy integration, sampling

1. Introduction

Since downstream processes account for 40 - 90% of the overall expenditures of most chemical plants (de Haan et al., 2020), notably optimization-based design methods for distillation-based processes have received considerable attention. However, the resulting non-convex mixed-integer nonlinear programming (MINLP) problems are usually solved by local optimization techniques, especially in case of rigorous thermodynamic models. To overcome this limitation, different strategies were recently reported, applying machine learning methods for the generation of accurate surrogate models, such as Gaussian processes and artificial neural networks (ANNs), to replace complex parts of the original process model (Nentwich and Engell, 2019) or the overall column models (McBride et al., 2020). Global optimization of ANNs in a reduced-space formulation allows for direct evaluation of full process flowsheets (Schweidtmann et al., 2019). Yet, the application of such methods mandates numerically robust simulation models, some penalizing strategy for constraint violations, and an efficient sampling of the search space, which usually requires a significant number of samples.

The current contribution proposes a different application of ANNs to effectively exploit the reduced search space that is offered by the integration of an efficient local deterministic optimization approach. The ANN creates a mapping of the initial values for a subset of the overall design vari-

ables to verified local optima, which already satisfy all equality and inequality constraints. This results in a significant contraction of the design space and allows for an effective search of the global optimum. The applicability of the ANN-based approach with an adaptive sampling strategy is first verified on several test functions, while the hybrid algorithm is successfully applied to the design of a distillation-based separation of benzene, toluene, and xylene considering different options for energy integration.

2. Methodology

The novel surrogate-assisted optimization procedure comprises a two-level hybrid approach. The lower level is a local deterministic optimization, which solves a full MINLP problem. The top level is an ANN including an adaptive sampling procedure to ensure valid results as well as few sampling points. The structure of the hybrid approach builds on the concept of the memetic algorithm proposed by Skiborowski et al. (2015a), which connects an evolutionary algorithm (EA) and a local deterministic optimization. An overview of the current procedure is depicted in Figure 1, while the main parts are described in the following subsections.

2.1. Artificial Neural Network

As ANN, a multilayer perceptron (MLP) is used to guide the optimal selection of initial values for the local deterministic optimization. Therefore, the MLP maps the initial values to the objective function of the local optimization. In general, the MLP can treat continuous as well as ordinal and nominal integer variables as a simple input vector and processes these through the hidden layers towards a desired output. Since there is no ordered relationship for nominal decision variables, a binary input variable b_i for each decision is considered, according to the concept of one-hot encoding.

For the currently considered case study, which addresses the optimization of a potentially energy-integrated direct column sequence, the input variables comprise the initial number of stages and the feed stage of each column in the superstructure model for the local optimization as well as the decision for the energy integration option, which is treated as four different binary decision inputs (cf. Figure 2). The output represents the total annualized costs (TAC) of the locally optimized process.

The MLP consists of two hidden layers with each neuron having a hyperbolic tangent sigmoid transfer function. As training algorithm, the Levenberg-Marquardt algorithm with Bayesian regularization is used. These specifications were derived based on preliminary tests with a test data set for the column sequence case study. The training and sampling mechanisms, i.e., the top-level of the hybrid optimization approach, are implemented in MATLAB® 2020b with the use of the Deep Learning Toolbox™14.1.

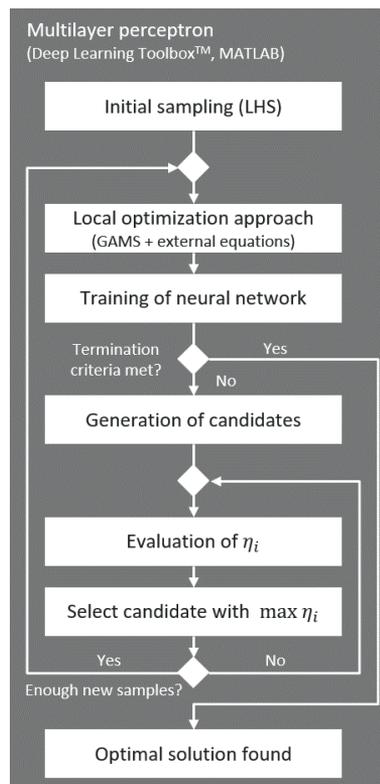


Figure 1: Overview of the surrogate-assisted optimization approach.

2.2. Mixed adaptive sampling

In this work, a mixed adaptive sampling method is applied to generate a reliable model based on the evaluation of the local optimization problems. The overall goal is to gain confidence in the generated results by the local optimization and, at the same time, reduce the number of necessary function evaluations while keeping an adequate model accuracy.

The initial sample points are determined by Latin hypercube sampling (LHS). To ensure that feed locations are within the selected number of column stages, the latter are prioritized and the feed location is considered as a function of the total number of stages. Furthermore, the vertices of the design space are always included in the training set while all other data points are randomly distributed to the training, validation, and test set, in ratios of 80%, 10%, and 10%, respectively.

The subsequently applied mixed adaptive sampling approach successively adds a relative amount of sampling points compared to the already existing samples to the set. In order to select new sampling points, 10,000 candidates are generated by LHS and evaluated based on a combination of different evaluation criteria. The current approach combines the mixed adaptive sampling approach of Eason and Cremaschi (2014), which considers the normalized Euclidean distance ($d_i/d_{i,max}$) as a space filling metric and the normalized jackknife variance ($s_i/s_{i,max}$) as a measure for the model uncertainty, with a normalized expression of the expected improvement criteria EI^{norm} according to Jones et al. (1998), with

$$EI_i^{norm} = \frac{EI_i - \min_i EI_i}{\max_i EI_i - \min_i EI_i}, \quad (1)$$

$$EI_i = (f_{min} - \hat{y}_i) \cdot \Phi\left(\frac{f_{min} - \hat{y}_i}{s_i}\right) + s_i \cdot \phi\left(\frac{f_{min} - \hat{y}_i}{s_i}\right), \quad (2)$$

where f_{min} is the current minimal value of the sampling set, \hat{y}_i is the MLP output at point i , and s_i is the jackknife standard deviation. $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density function, respectively. The first contribution increases with decreasing predicted values \hat{y}_i while the second contribution increases with increasing standard deviations s_i for the predicted values (Jiang et al., 2020). The resulting combined evaluation criterion

$$\eta_i = \frac{d_i}{d_{i,max}} + \frac{s_i^2}{s_{i,max}} + \cdot EI_i^{norm} \quad (3)$$

provides an effective sampling with a balance between exploration and exploitation of the design space.

The EI criterion favors minima of the predicted target function while maintaining a measure for the prediction uncertainty. This ensures a sufficient accuracy of the surrogate model even if the first two contributions of η_i get outweighed by the EI criterion. After each iteration of the adaptive

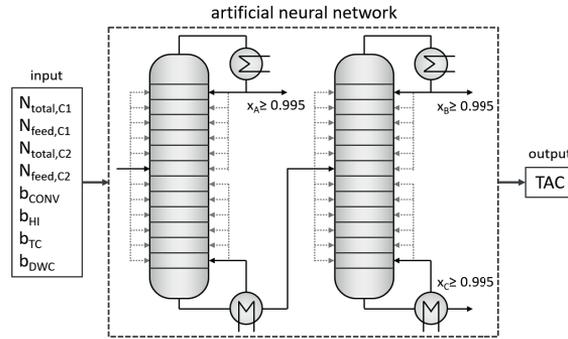


Figure 2: Superstructure of a direct column sequence for three-product separation.

sampling, the termination criteria of the algorithm are checked, i.e., the change in the objective function of the sampling set, the maximum number of iterations, and the maximum number of sample points.

2.3. Local deterministic optimization

By solving a local deterministic optimization problem instead of a simulation problem, the design space of the upper ANN is effectively reduced. For this purpose, a superstructure optimization problem needs to be solved to local optimality. For the current case study, an equilibrium-stage column model with additional correlations for column sizing and costing is solved. The according superstructure for the column sequence is illustrated in Figure 2. The corresponding MINLP problem is solved as a series of successively relaxed nonlinear programming (NLP) problems Skiborowski et al. (2015b) with additional modifications for heat integration, thermal coupling, and an integration of the latter as dividing wall column Waltermann et al. (2020). The local optimization model is implemented in GAMS 34.3.0 and the resulting NLP problems are solved with SNOPT.

3. Results

Prior to the application for the process design example, the functionality of the mixed adaptive sampling is evaluated for simple test functions. Subsequently, the results for the purification of a mixture containing benzene, toluene, and xylene are presented.

3.1. Validation of sampling procedure

To evaluate the functionality of the mixed adaptive sampling approach, three test functions for global optimization, i.e., the Ackley function (Ackley, 1987), MATLAB's Peaks function (Eason and Cremaschi, 2014), and the Six Hump Camel Back function (Floudas et al., 1999), are considered. All test functions have two input variables $x, y \in [-2, 2]$ and one output variable z . For this purpose, a simple MLP is used with a single hidden layer with 15 neurons.

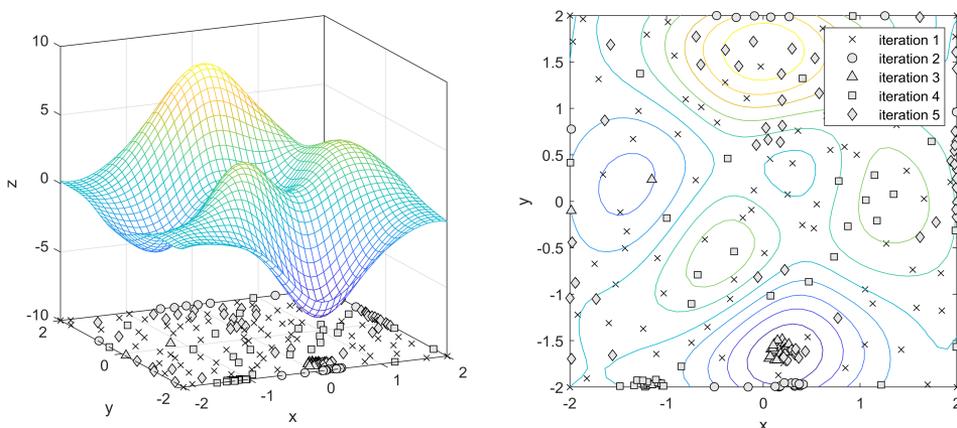


Figure 3: Response surface of the artificial neural network (ANN) representing the Peaks function after five iterations including the sampling locations (left). To increase the clarity, the sample points were added in the $z = -10$ plane and not on the response surface. Corresponding design space representation with contour lines including sampling locations (right).

The representative results for the Peaks function are illustrated in Figure 3. Similar to the work of Eason and Cremaschi (2014), the initial sampling comprised 80 samples followed by a sampling rate of 30% for all consecutive iterations. After five iterations, the response surface of the MLP has an overall mean squared error of about 0.01 but shows a sufficient model accuracy. Especially, the area of the global minimum is tightly covered, which is further highlighted in the contour plot in Figure 3 focusing on iteration three and five. The global minimum is found with a deviation of only 0.015%. Interestingly, a lot of samples are also located on the edges of the design space, which is most likely caused by the tendency to reduce the model variance, while there is also a persistent distribution of samples in a space-filling manner. In conclusion, the mixed adaptive sampling procedures gives a good overall model regression while investigating areas around the potential global optimum to increase the model accuracy in a close vicinity.

3.2. Purification of benzene / toluene / xylene

For demonstration of the approach to optimal process design, the purification of a mixture of $10 \text{ mol}\cdot\text{s}^{-1}$ with 30 mol% benzene, 30 mol% toluene, and 40 mol% xylene into streams with purities of at least 99.5 mol% is investigated. The top-level ANN comprises eight input variables (cf. Figure 2) with $N_{total,C1}, N_{total,C2} \in [20, 80]$ and $N_{feed,C1}, N_{feed,C2} \in [5, N_{total} - 5]$, two hidden layers with five and three neurons, and one output, resulting in 67 parameters (weights and biases). The sampling procedure is carried out with 100 initial samples and a subsequent sampling rate of 25%. For the sake of clarity, Figure 4 (left) shows only the response surface in a reduced design space, illustrating the best solutions in terms of the initial total number of stages in column 1 ($N_{total,C1}$) and column 2 ($N_{total,C2}$) for the heat-integrated process, which is determined as the most economic process configuration. Figure 4 (right) illustrates the overall best solution found in the locally optimized sampling set as well as the feature values and the output of the MLP.

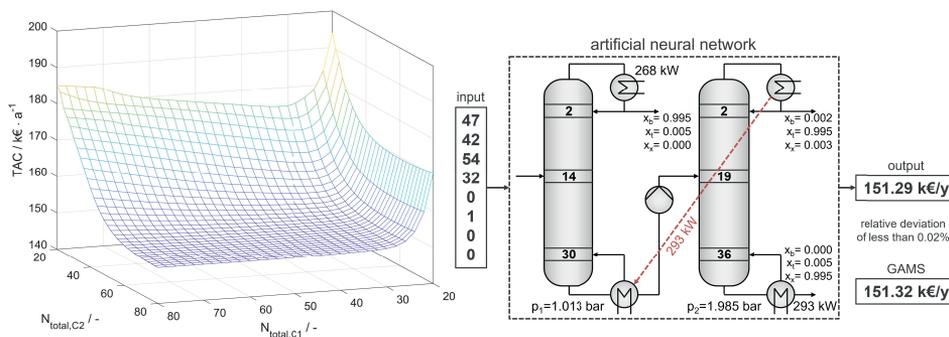


Figure 4: Minimal response surface of the artificial neural network (ANN) for the heat-integrated column sequence in a reduced design space of $N_{total,C1} \times N_{total,C2}$ (left) and the corresponding inputs and output of the ANN for the best found sampling candidate and its locally optimal objective value from GAMS (right).

As indicated in Figure 4 (right), the TAC of 151.32 $\text{k€}/\text{y}$ determined from the ANN is only 0.02% lower than the respective result of the local optimization. The global optimum of the response of the MLP can be determined by a full enumeration of the design space, indicating a potentially reduced objective function value of 149.94 $\text{k€}/\text{y}$, which is, however, less than 1% lower than the previously described local optima in GAMS. In addition to this confirmation of the local optimum, the response surface of the ANN also illustrates a flat optimum in the region of high initial tray numbers, indicating a comparably good convergence of the local optimization approach. While the presented approach does not provide a validation of the global optimum, the combination of local

optimization and the ANN improves the coverage of the design space and strengthens confidence in the quality of the determined solution. Furthermore, the evaluation of a total of 195 converged samples represents only a tiny fraction of the overall number of combinatorial options, which sum up to over 26 million considering the discrete design degrees of freedom that are represented by the features of the MLP.

4. Conclusion

The current work presents a new approach for surrogate-assisted hybrid optimization, which applies an artificial neural network for the regression of a local deterministic optimization approach. Combined with an effective mixed adaptive sampling based on the method of Eason and Cremaschi (2014) and the expected improvement criterion (Jones et al., 1998), an effective search for the global optimum can be conducted for process design problems with MINLP characteristics. The approach is successfully demonstrated for the optimization of a two-column process including different means for energy integration.

Further work will focus on the tuning of the employed hyperparameters, e.g., the architecture of the ANN, the termination criteria, the initial sampling as well as the sampling rate. Additionally, the three parts of the evaluation criterion of new samples can be weighed to actively adjust the ratio of exploration and exploitation in the progress of the iterations. An application to more complex case studies can reveal further advantages and disadvantages of the method and will be part of future research.

References

- D. H. Ackley, 1987. *A Connectionist Machine for Genetic Hillclimbing*. Springer, Boston, MA.
- A. B. de Haan, H. B. Eral, B. Schuur, 2020. *Industrial separation processes: Fundamentals*, 2nd Edition. De Gruyter graduate. Walter de Gruyter GmbH, Berlin and Boston.
- J. Eason, S. Cremaschi, 2014. Adaptive sequential sampling for surrogate model generation with artificial neural networks. *Computers & Chemical Engineering* 68, 220–232.
- C. A. Floudas, P. M. Pardalos, C. S. Adjiman, W. R. Esposito, Z. H. Gumus, S. T. Harding, J. L. Klepeis, C. A. Meyer, C. A. Schweiger, 1999. *Handbook of test problems in local and global optimization*. Vol. 33 of *Nonconvex optimization and its applications*. Kluwer Academic Publishers, Dordrecht.
- P. Jiang, Q. Zhou, X. Shao, 2020. *Surrogate Model-Based Engineering Design and Optimization*. Springer, Singapore.
- D. R. Jones, M. Schonlau, W. J. Welch, 1998. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* 13 (4), 455–492.
- K. McBride, E. I. Sanchez Medina, K. Sundmacher, 2020. Hybrid semi-parametric modeling in separation processes: A review. *Chemie Ingenieur Technik* 92 (7), 842–855.
- C. Nentwich, S. Engell, 2019. Surrogate modeling of phase equilibrium calculations using adaptive sampling. *Computers & Chemical Engineering* 126, 204–217.
- A. M. Schweidtmann, D. Bongartz, W. R. Huster, A. Mitsos, 2019. Deterministic global process optimization: Flash calculations via artificial neural networks. *Computer Aided Chemical Engineering* 46, 937–942.
- M. Skiborowski, A. Harwardt, W. Marquardt, 2015a. Efficient optimization-based design for the separation of heterogeneous azeotropic mixtures. *Computers & Chemical Engineering* 72, 34–51.
- M. Skiborowski, M. Rautenberg, W. Marquardt, 2015b. A hybrid evolutionary-deterministic optimization approach for conceptual design. *Industrial & Engineering Chemistry Research* 54 (41), 10054–10072.
- T. Waltermann, T. Grueters, D. Muenchrath, M. Skiborowski, 2020. Efficient optimization-based design of energy-integrated azeotropic distillation processes. *Computers & Chemical Engineering* 133, 106676.

Application of multiplicative homomorphic encryption in process industries

Zarina Chokparova^a, Leon Urbas^b

^a*Technische Universität Dresden, Boysen-TU Dresden-Research Training Group, Dresden 01187, Germany*

^b*Technische Universität Dresden, Chair of Process Control Systems and Process System Engineering Group, Dresden 01069, Germany*
zarina.chokparova@tu-dresden.de

Abstract

Flexibility and mobility of modern value chains have created a need for continuous information exchange between involved chain parties. In such process industries, many companies share their manufacturing assets with each other and thereby allow partners to have control over them. In this way, some sensitive essential information about operations and control methods could be leaked from the asset supplier to the user and vice versa. Therefore, such information sharing raises confidentiality concerns between the service provider and its operator. The goal of this work is to apply and evaluate a confidentiality-preserving information sharing model for a time series use case in process industries. There are various ways to maintain privacy of the sensitive information, such as anonymization and encryption. To preserve control for the user and to allow data gathering by the vendor about asset operation, homomorphic encryption methods could be implemented. Homomorphic encryption allows for the preservation of confidentiality of the data while enabling computations on the encrypted data. The main focus of this study is an investigation of homomorphic encryption schemes with multiplicative properties, such as RSA and ElGamal, which can be applied to process data within information exchanges. This research investigates the probabilistic and the deterministic homomorphic algorithms with respective differences in encryption and decryption speeds. This approach is based on the simulation of the use case between asset vendor and asset operator. The confidentiality model of the information exchange sustains the zero-knowledge proof between involved value chain partners. The result implies the adaptability of both methods within the privacy-preserving sharing model. This study is limited to a use case with the application of partial homomorphic cryptosystems in process industries. The outcome highlights the statistical justification of the application of multiplicative homomorphic encryption (MHE) for confidentiality preservation.

Keywords: confidentiality, homomorphic encryption, value chains, process systems, secure information exchange

1. Introduction

Privacy and information security are crucial requisites in industrial systems. Many process systems, including networked and distributed control, could be subject to cyber security attacks. Through such attacks, the perpetrators seek to alter the measurements, steal transmitted information, cause harm to the system, or destabilize it. The perpetrators may also attempt to eavesdrop on time series data, what could help to identify the specific process and to develop an attack mechanism. Moreover, significant flows of sensitive

operation data transferred within supply chains via cloud networks between stakeholders can lead to breaches of information confidentiality.

The purpose of the privacy protection methods is to make the data unreadable and incomprehensible to certain readers or the intruders, whereas the intended recipient is able to decrypt the message. Information confidentiality can be secured using cryptosystems. Key management, computational complexity and resource limitations, random number generation, and the encryption of dynamic data could be recognized as the main challenges of current cryptographic schemes. One group of encryption methods is homomorphic encryption, which can perform computations on the encrypted data without their prior decryption. Here the focus is placed on public key cryptography algorithms, which can solve the problem of key arrangement between the system actors present in symmetric cryptosystems. Two examples of homomorphic encryption with multiplicative properties are ElGamal and RSA cryptosystems. ElGamal public key cryptosystem is based on the complexity of computing logarithms (ElGamal, 1985), whereas RSA (Rivest et al., 1983) is based on the complexity of factoring large integers.

2. Theoretical background

2.1. Multiplicative homomorphic encryption

Homomorphism refers to an encryption's ability to combine the ciphertexts so that they decrypt as if the plaintexts were combined. The procedures for the two homomorphic encryption algorithms with multiplicative properties are described below.

The **Rivest-Shamir-Adleman (RSA)** is an asymmetric cryptographic algorithm that is widely implemented in privacy preserving applications and for authentication of the digital information (Boneh, 1999), such as web traffic security, email encryption and signatures, electronic payment systems, etc. The algorithm is as follows.

Key generation: $n = p \cdot q$, where p and q are random prime numbers.

Totient: $\phi(n) = (p - 1)(q - 1)$.

$\text{gcd}(e, \phi(n)) = 1$ such that $1 < e < \phi(n)$ and e is co-prime to $\phi(n)$.

Congruence relation: $de \equiv 1 \pmod{\phi(n)}$, or $d = (1 + x \cdot \phi(n))/e$ to be an integer.

The *public key* is comprised of (n, e) .

The *private key* consists of p, q , and d .

Encryption: $c = m^e \pmod{n}$.

Decryption: $m = c^d \pmod{n}$.

The **ElGamal** encryption is a probabilistic algorithm of public key cryptography and is based on Diffie-Hellman key exchange. The protocol steps are listed below.

Key generation: publish a large prime p and the generator g of the group Z_p^* .

Compute $A = g^a \pmod{p}$, where $1 \leq a \leq p - 1$. The *public key* is (p, g, A) .

Encryption: choose k from $\{1, \dots, p - 1\}$ to compute the ciphertext:

$(c_1 = g^k \pmod{p}, c_2 = m \cdot A^k \pmod{p})$, where m is a message and $1 \leq m \leq p - 1$.

Decryption: compute inverse x^{-1} of $x = c_1^a \pmod{p}$, from which follows that $m = x^{-1} \cdot c_2 \pmod{p}$.

2.2. Related work

One of the applications of ElGamal encryption was implemented in a networked control system for protecting confidential information passed through devices (Kogiso, 2019). The parameters and signals of a discrete-time linear controller were encrypted to prevent reverse engineering, damage, or information theft. As a result, due to the homomorphic properties of the cryptosystem, the controller could compute the output using encrypted input parameters. Asymptotic stability improvements were performed by adding a dynamic encoder and decoder to the cryptosystem (Teranishi et al., 2020). Another

proposal for implementation of an ElGamal cryptosystem for voting in elections is presented in (Mikhail et al., 2014). The main idea is the secrecy of individual votes while allowing each individual only a single vote. Moreover, the system can exclude non-eligible voters and verify the real outcome. Thus, the tallying of votes can be performed using homomorphic encryption. Another application of MHE was found in protecting DNA data (Thangavel and Varalakshmi, 2018). The ElGamal cryptosystem is used to encrypt the data transmitted via cloud and manage the secure passing of keys between data owner and user. This enhanced implementation of the cryptosystem has an improved authentication and higher resistance to attacks. The RSA cryptosystems are used in cloud networks (Tebaa et al, 2012) and, despite its large computational time, the algorithm can maintain high levels of security.

2.3. Current contribution

The application of multiplicative homomorphic encryption could prove useful for information exchange within value creating chains. Such transfers of information can be observed within collaborations between production process partners. In such interactions, the physical assets and production data could be shared. However, not all parties are willing to fully disclose their process data, operation conditions, and process dynamics. Therefore, to establish a trusted environment and maintain privacy protection of process data or confidential computational formulas, the application of homomorphic encryption is proposed. The implementation of RSA and ElGamal cryptosystems for a use case in process industries is described below.

3. Methodology

3.1. Use case for MHE

Homomorphic cryptographic algorithms can be employed for confidentiality preservation of the information being computed, passed process inputs and outputs, and time-series data generated within production chain.

Much of the data and information transmitted in the process industries can be described by models containing the multiplicative relationship between parameters and variables. The working use case was developed based on a simple multiplicative relationship between the volume of liquid in a tank and its density changing with temperature. To represent the application of homomorphic encryption with multiplicative properties, the use case of computing the liquid volume in the tank from the encrypted values of density is studied. The process equipment is provided by the supplier in a process chain, and it is operated by the user who desires to keep information about the physical properties of the

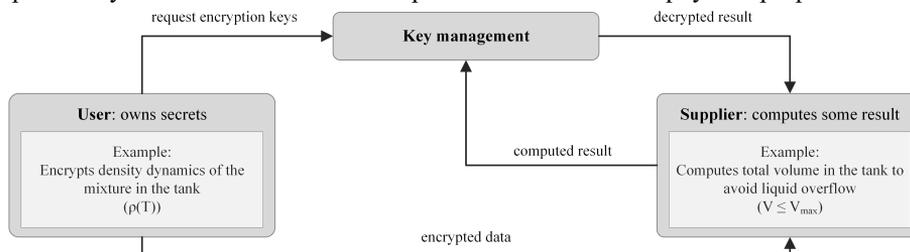


Figure 1. Use case example: information exchange

process materials confidential. The supplier needs to collect essential operation information regarding the process equipment; however, the equipment user can share the information in an encrypted form. Thus, encrypted dynamic density values of the tank mixture are transmitted by the user, and the supplier utilizes them to compute the

corresponding volume in the tank, which is necessary to calculate in order to avoid an overflow in the tank. The protocol of information exchange (Figure 1) between the user and the supplier must sustain a zero-knowledge proof. Using a key management party for generation of a one-time key for each dataset facilitates the confidentiality preserving transfer of data. Therefore, none of the parties are able to reveal the secrets of the other partners, for instance, density dynamics encrypted by the user or computation of the tank volume performed by the supplier.

3.2. Simulation

The data used for the use case are shown in Figure 2 and represent original values of the inverse density dependencies of four substances over temperature. The data used for the generation of the plots are provided by the Dortmund Data Bank (DDBST, 2022). The inverse of density was used to simplify the implementation of the encryption algorithm,

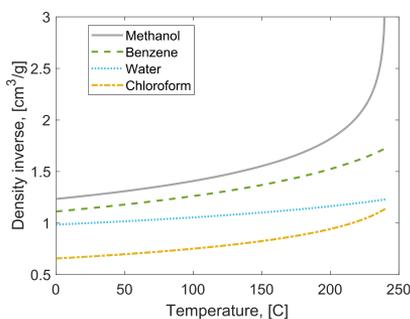


Figure 2. Density inverse profiles

and by directly multiplying the inverse of density by the mass of the liquid in the tank, the operation of division was avoided. To ensure fair evaluation of the encrypted distributions, the algorithms were run for 100 cycles for each liquid and both encryption algorithms using different key generation and encryption constants. To conduct a comparison of the two encryption methods, the encryption parameters were chosen in a manner to have similar ranges of output magnitude.

4. Results and discussion

The goal of this analysis is to perform a statistical evaluation of the encrypted data by identifying patterns from the attacker's point of view. Thus, this does not consider a proof of the strength of encryption algorithms from the perspective of randomness assessment. Instead, the applicability of MHE algorithms to value chains and process related information is analyzed, and the confidentiality level achieved through the implementation of these methods is evaluated.

4.1. Analysis of encryption

The liquids chosen for analysis have different ranges and distributions of density at different temperatures. Methanol has a strong increasing trend, whereas the profile of water has a normal pattern. The profiles of benzene and chloroform also demonstrate similar increasing trends, with a vertical shift between curves. The strategy for analysis of encoded inverse density profiles includes the use of statistical features that are used in similarity assessments of time series data (Nanopoulos et al, 2001). The chances of drawing correlation between encrypted and non-encrypted datasets and identifying the substances is examined based on statistical features.

The comparison between the encrypted datasets for both RSA and ElGamal algorithms was performed using first and second order statistical features. Analyzed first order features include mean value, skewness, and kurtosis. Skewness depicts an asymmetry relative to mean, and kurtosis characterizes the tails of a distribution. Evaluated second order features contain co-occurrence features, such as energy, correlation, and homogeneity. The energy feature reflects the uniformity of the distribution, correlation measures the relationship between neighboring elements of the matrix, and homogeneity describes the probability of closeness to the diagonal in a co-occurrence matrix.

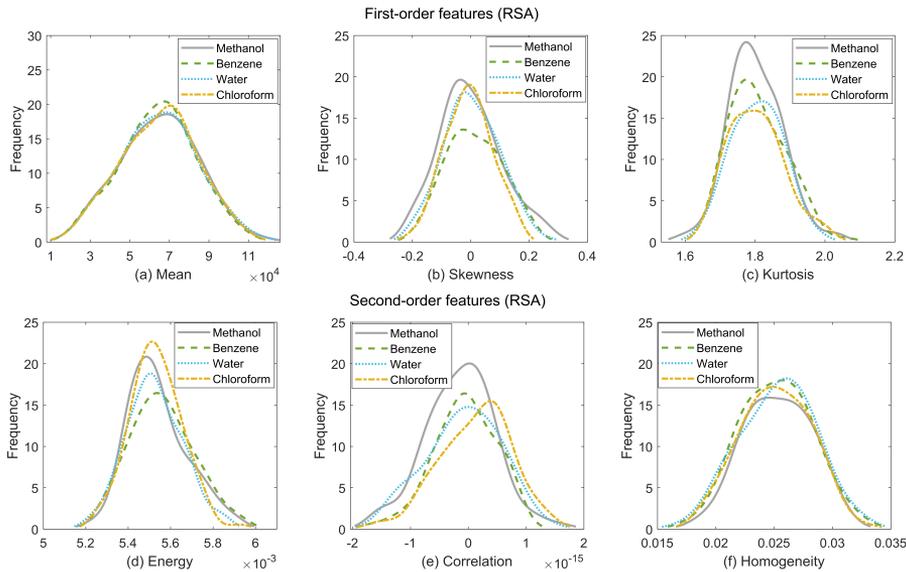


Figure 3. Statistical features of encrypted datasets with RSA algorithm

Figures 3 shows the mentioned first and second order features for datasets encrypted with RSA and Figure 4 depicts the homogeneity of ElGamal implementation. Mainly, the first step in identification of the substances would be to correlate the sequences of distribution peak sizes with the sequence of original datasets from Figure 2. For the RSA encryption,

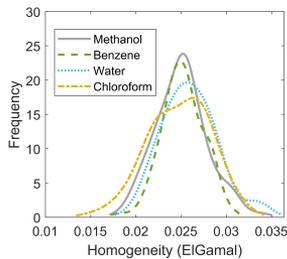


Figure 4. Homogeneity feature

Figure 3 (c) implies that a kurtosis feature applied to the encrypted datasets could reveal the order of liquids, meaning that lower density liquids have higher peaks at the kurtosis distributions. A similar observation could be made from Figure 3 (e) based on the area under the curves of the correlation feature. In comparison, the patterns for the features of datasets encrypted with ElGamal are observed to be different from those with RSA. However, the homogeneity feature in Figure 4 also shows the same arrangement of substances as in the original distribution of inverse densities.

4.2. Inferential statistical comparison

To conduct further evaluations of obtained distributions of statistical features for different substances, a t-test (with $\alpha = 0.05$) was applied to pairs of samples on a per-feature basis. Since the datasets describing statistical features are mostly characterized by normal distribution, a t-test would allow comparison of distributions in each feature to determine whether they could be related to the same population. Table 1 shows the results of null hypothesis (h) and probabilistic significance value (p) for all datasets. When the value of h is 0, the null hypothesis is valid; otherwise, the value is 1. Value of p has a range $[0, 1]$ and higher number designates validity of the null hypothesis.

The t-test concluded that correlation feature distributions of datasets of chloroform density inverse values encrypted with RSA and homogeneity distributions of datasets of water profile encrypted with ElGamal reject the null hypothesis. However, unity values of h have very low probability p . Based on the overall results of t-test, the null hypothesis

was confirmed and it could be inferred that the statistical features are not distinguishable from each other.

Table 1. Results of t-test for statistical features

		<i>M-B*</i>		<i>M-W*</i>		<i>M-C*</i>		<i>B-W*</i>		<i>B-C*</i>		<i>W-C*</i>	
		<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>
Mean	<i>RSA</i>	0	0.76	0	0.96	0	0.94	0	0.79	0	0.81	0	0.98
	<i>ElGamal</i>	0	0.33	0	0.11	0	0.92	0	0.44	0	0.38	0	0.12
Skewness	<i>RSA</i>	0	0.33	0	0.76	0	0.66	0	0.48	0	0.11	0	0.40
	<i>ElGamal</i>	0	0.18	0	0.70	0	0.82	0	0.36	0	0.12	0	0.54
Kurtosis	<i>RSA</i>	0	0.56	0	0.64	0	0.82	0	0.89	0	0.72	0	0.81
	<i>ElGamal</i>	0	0.10	0	0.58	0	0.08	0	0.26	0	0.91	0	0.22
Energy	<i>RSA</i>	0	0.27	0	0.46	0	0.71	0	0.06	0	0.12	0	0.68
	<i>ElGamal</i>	0	0.17	0	0.05	0	0.38	0	0.52	0	0.72	0	0.37
Correlation	<i>RSA</i>	0	0.34	0	0.53	1	0.00	0	0.79	1	0.04	1	0.03
	<i>ElGamal</i>	0	0.60	0	0.55	0	0.99	0	0.25	0	0.58	0	0.52
Homogeneity	<i>RSA</i>	0	0.76	0	0.98	0	0.68	0	0.80	0	0.91	0	0.71
	<i>ElGamal</i>	0	0.70	1	0.03	0	0.59	1	0.04	0	0.35	1	0.01

* *M* – methanol, *B* – benzene, *W* – water, *C* – chloroform

5. Conclusion

The aim of this work was to assess the applicability of the multiplicative homomorphic cryptosystems for confidentiality-preserving secret sharing of process-related information. Since RSA and ElGamal cryptosystems are considered to be secure tools for information exchange, they were implemented for the presented use case that includes multiplicative relations within mathematical models of the processes. The statistical features of the encrypted inverse density profiles for four liquids were compared with the original distribution patterns and several correlations between them were found. Some features were identified as revealing for both encryption methods. However, the result of inferential statistical analysis shows that the feature datasets of encrypted values could not be differentiated from each other, thereby proving the usability of multiplicative homomorphic encryption in process industries.

References

- D. Boneh, 1999. Twenty years of attacks on the RSA cryptosystem. *Notices of the AMS*, 46(2).
 DDBST, 2022. Saturated Liquid Density. Calculation by DIPPR105 Equation. Accessed online on January 1, 2022 from <http://ddbonline.ddbst.de/>
 T. ElGamal, 1985. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. *IEEE transactions on information theory*, 31(4), 469-472.
 K. Kogiso, 2019. Encrypted control using multiplicative homomorphic encryption. In *Privacy in Dynamical Systems*, 267-286. Springer, Singapore.
 M. Mikhail, Y. Abouelseoud, G. Elkobrosy, 2014. Extension and application of El-Gamal encryption scheme. *2014 World Congr. Comput. Appl. Inf. Syst.*, 1-6. IEEE
 A. Nanopoulos, R. Alcock, Y. Manolopoulos, 2001. Feature-based classification of time-series data. *International Journal of Computer Research*, 10(3), 49-61.
 R. L. Rivest, A. Shamir, L. Adleman, 1983. A Method for Obtaining Digital Signatures and Public- Key Cryptosystems. *Communications of the ACM*, 26(1), 96-99.
 M. Tebaa, S. E. Hajji and A. E. Ghazi, 2012. Homomorphic encryption method applied to Cloud Computing. *National Days of Network Security and Systems*, 86-89.
 K. Teranishi, N. Shimada, and K. Kogiso, 2020. Stability-guaranteed dynamic ElGamal cryptosystem for encrypted control systems. *IET Control Theory & Applications*, 14(16).
 M. Thangavel, P. Varalakshmi, 2018. Enhanced DNA and elgamal cryptosystem for secure data storage and retrieval in cloud. *Cluster Computing*, 21(2), 1411-1437.

PPOPT - Multiparametric Solver for Explicit MPC

Dustin Kenefake^{a,b} and Efstratios N. Pistikopoulos^{a,b}

^a*Texas A&M Energy Institute, Texas A&M University, College Station, Texas, 77843, United States*

^b*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, Texas, 77843, United States*
stratos@tamu.edu

Abstract

In this paper, we describe the PPOPT (**P**ython **P**arametric **O**ptimization **T**oolbox) solver, written in Python. A general-purpose multiparametric package written in Python that features: (i) Efficient and parallel implementations of common multiparametric programming algorithms (strong scaling ≥ 25 cores), (ii) Problem reformulation to reduce the computational overhead and numerical stability, (iii) Multiparametric solution code generation to export solutions in C++, JavaScript, Python, and MATLAB. The speed and scaling behavior of PPOPT are explored with computational studies on the POP problem libraries and with explicit MPC controller design. PPOPT is compared to current state-of-the-art multiparametric solvers. Additionally, an exported solution is benchmarked on various platforms, including embedded processors and desktop computers

Keywords: Multiparametric Programming, Parallel Programming, Explicit Model Predictive Control, Optimization, Operations Research

1. Introduction

We consider the following class of multiparametric programming problem 1, multiparametric Quadratic Programs (mpQP) and multiparametric Linear Programs (mpLP), where $x \in \mathbb{R}^n$, $\theta \in \mathbb{R}^k$ and $A_E, A, A_\theta, F_E, F, H$ are matrices of appropriate dimension and b_E, b, b_θ are vectors of appropriate dimension. Constraint set eq. 1b defines the general equality constraints, similarly eq. 1c defines the inequality constraints, and eq. 1d defines the inequality constraints on the uncertain parameter θ , with eq. 1a defining the multiparametric objective.

$$\min_x f(x, \theta) = \frac{1}{2} x^T Q x + \theta^T H^T x + c^T x \quad (1a)$$

$$\text{s.t. } A_E x = b_E + F_E \theta \quad (1b)$$

$$A x \leq b + F \theta \quad (1c)$$

$$A_t \theta \leq b_t \quad (1d)$$

Multiparametric programming has found renewed interest due to the fact that the model predictive controller (MPC) optimization problem can be reformulated into a multiparametric program (Pistikopoulos et al., 2021). This allows for the generation of explicit MPCs that have found much interest in process control applications (Kiparissides et al., 2011; Sakizlis et al., 2004; Dua et al., 2006; Pappas et al., 2021). Moreover, applications have also been found in scheduling (Burnak et al., 2019), simultaneous design and control (Diangelakis et al., 2017), data science (Tso et al.,

2020), multilevel optimization problems (Avraamidou and Pistikopoulos, 2019), and financial applications (Romanko et al., 2012).

Nearly all of the multiparametric algorithms in the literature are based on serial execution, i.e., the tasks are processed sequentially only one execution thread. Current state of the art multiparametric solvers do not utilize parallel programming techniques, and thus not taking advantage of modern computational hardware. The development of parallel multiparametric algorithms and software implementations has been investigated in the literature, with an initial study being conducted by (Oberdieck and Pistikopoulos, 2016) showing that it is possible to create parallel multiparametric algorithms. However, the scaling with core count was not promising for large core counts.

Motivated by these limitations, PPOPT was developed to scale on large core count systems (≥ 25 cores). Much care has been taken to develop the new parallel algorithms and implementations to achieve good computational scaling. PPOPT has implemented parallel versions of the Combinatorial algorithm, Connected Graph, and Geometric algorithms developed by (Gupta et al., 2011; Oberdieck et al., 2017; Spjøtvold et al., 2006), respectively.

2. PPOPT Features

2.1. Key Points

2.1.1. Parallel Multiparametric Programming

The central feature of PPOPT is the efficient and scalable implementations of multiparametric programming algorithms. Due to limitations of the serial nature of literature algorithms, these needed to be reformulated in a way as to allow for the sub-tasks to be solved independently of each other as well as modifying the data structures being used. The multiparametric programming problem test sets MPQP1, and MPLP1 are used to assess the performance of the considered solvers, a more detailed description for these problem sets can be found in (Oberdieck et al., 2016). A performance comparison is performed in subsection 2.3 on the MPQP1 problem datasets. The scaling of the combinatorial algorithm is studied in Subsection 2.2. In addition, a control application, the generation of explicit MPC, is studied in Section 3.1.

2.1.2. Problem reformulation and Constraint Processing

Performance and robustness considerations require the removal of all redundant constraints at solve time, so PPOPT has a robust set of tools for constraint processing. The primary focus is the removal of strongly and weakly redundant constraints, as the solve time for multiparametric programs with any algorithm scale with the number of constraints, and removing redundant and repeated constraints improves computational stability. Condition number of optimization problems influence the problem (Ordóñez, 2002). As the problem is stated ahead of time, various conditioning methods can be applied that are not applicable for the online case, such as reformulating mpQPs into equivalent multiparametric min-norm problems.

2.1.3. Solution Export

An important application of multiparametric programming is generating Explicit Model Predictive Controllers for optimal control applications. However, there are very few solutions to export explicit multiparametric solutions. This is addressed with PPOPT, as the package can generate code based on the solution of the multiparametric problem. Currently, the solver can generate C++, Python, JavaScript, and MATLAB code. The performance of the exported solutions is explored in Subsection 3.2.

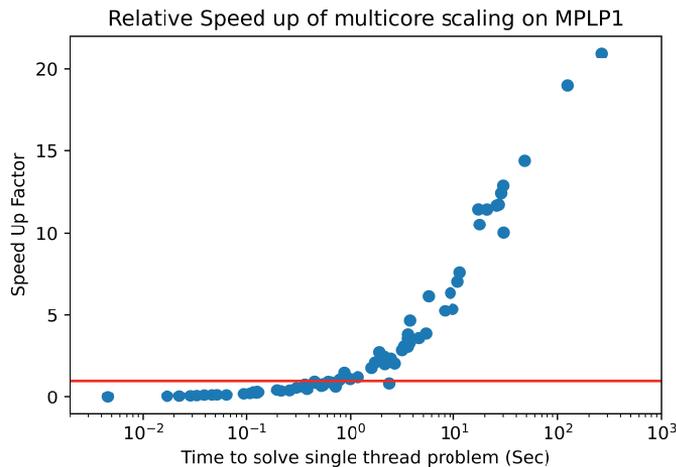


Figure 1: Comparison of single core and multicore solve times

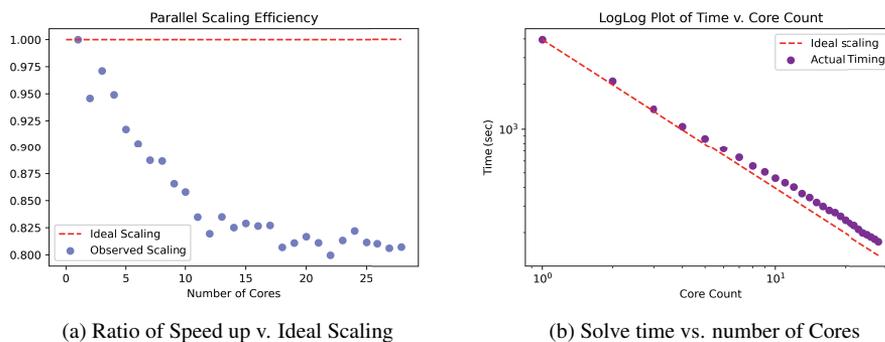


Figure 2: Scaling Study of PPOPT on large scale mpQP

2.2. Case Study: Scaling Analysis of the Combinatorial Algorithm

The scaling behavior of the PPOPT parallel implementation of the combinatorial algorithm (Gupta et al., 2011) is studied for the scaling behavior. A quadratic multiparametric problem was randomly generated such that a solving with a single core becomes computationally burdensome, in the order of an hour. This experiment was run on a single compute node in the Terra Supercomputer (dual Intel Xeon E5-2680 v4 2.40GHz 14-core, 64GB 2400MHz DDR4 RAM, CentOS 7). Here we can see the strong scaling of PPOPT both in terms of scaling efficiency and in a speed-up factor as a function of core count. The multiparametric programming problem considered here has eight optimization variables, four uncertain parameters, and 52 non-redundant affine constraints. This problem took 66 minutes to solve with the PPOPT solver utilizing a single core. In Figure 2b, we see the time scaling behavior of the parallel implementation, scaling nearly ideally with good strong scaling performance, and in Figure 2a, we see the ratio of ideal scaling and observed scaling for this problem, this again show nearly constant strong scaling behavior after 12 cores. With an overall speed-up of 22.7x, the total solve time was reduced to 2.9 minutes from 65.8 minutes, with all 28 cores being used.

A scaling study was carried out on the MPLP1 test set to show scaling performance on solving multiparametric problems with respect to problem-solving time. In Figure 1, It can be seen that problems take less than 1 second to solve that the speed-up factor is less than one. This is due to a fixed overhead of creating multiple python processes to solve the problem in parallel. In the case of solving many simple multiparametric programs, it might make sense to increase parallelization granularity over entire problems instead.

2.3. Case Study: Benchmark Performance on mpQP problems

The performance of PPOPT is tested on the MPQP1 problem sets. Each problem set contains 100 multiparametric problems, each of wide degrees of computational difficulty, with different numbers of constraints, parameters, and optimization variables. A full description of the problem sets can be found in Oberdieck et al. (2017). Here PPOPT will be compared against the POP and MPT3 solvers for the relevant algorithms (Oberdieck et al., 2016; Herceg et al., 2013). This experiment was run on a single compute node in the Terra Supercomputer (dual Intel Xeon E5-2680 v4 2.40GHz 14-core, 64GB 2400MHz DDR4 RAM, CentOS 7). With PPOPT parallel algorithms using all 28 cores, Strong Performance can be observed for PPOPT in Figure 3.

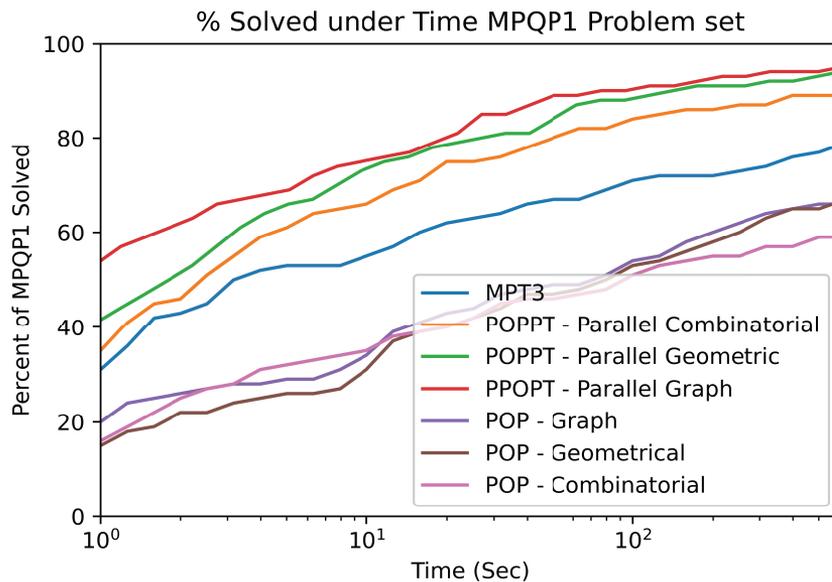


Figure 3: MPQP1 Test Set Performance of PPOPT, MPT3, and POP

3. Applications

3.1. Case Study: Explicit Model Predictive Control

Explicit model predictive control (eMPC) is one of the central applications of multiparametric programming, and thus focus is applied to this application. A computational study is where an explicit model predictive controller is generated for a system with four states, one input, and two outputs with parametric set points with constraints on the state, output and input. The computational study

Table 1: Timing Results of Exported Solution

	C++	Python	Matlab	JavaScript
ESP32	47 uSec	N/a	N/a	N/a
Teensy 4.0	5 uSec	N/a	N/a	N/a
i5-8600K	.3 uSec	8.2 uSec	410 uSec	5 mSec

is the time to solve increasingly larger eMPCs by increasing the control horizon of the eMPC. As can be seen in Figure 4, for larger problems, PPOPT is significantly faster, with a speed-up factor of approximately 15x. Indicating that for the generation of eMPCs, the PPOPT solver would be the faster choice. This computational study was performed on a four core desktop with an i7-4770 with 16 GB of RAM on a Python 3.8 environment for PPOPT and MATLAB 2019a environment for POP and MPT3.

3.2. Case Study: Benchmarking the Real Time Performance of the Exported Solution

PPOPT can export the solution to a multiparametric program as source code with a small included library to run the code. The languages currently supported are C++, Python, JavaScript, and MATLAB. This method was chosen to maximize the portability of the exported solution and allow for modifications in how the solution integrates into the specific environment. A problem with 4 optimization variables, 2 uncertain parameters, 45 constraints, and 87 critical regions was used to benchmark the performance of the exported solution. The exported solution is evaluated on the following hardware: an ESP32, 240Mhz Xtensa LX6 with 520 KiB SRAM, a Teensy 4.0, 600MHz ARM M7, and 1 MiB Ram, and on a desktop 6 core Intel i5-8600k with 32GB DDR4 RAM. As a reference, with Gurobi in Python 3.8 on the desktop system, it takes 800 uSec to solve. Results are summarized in Table 1.

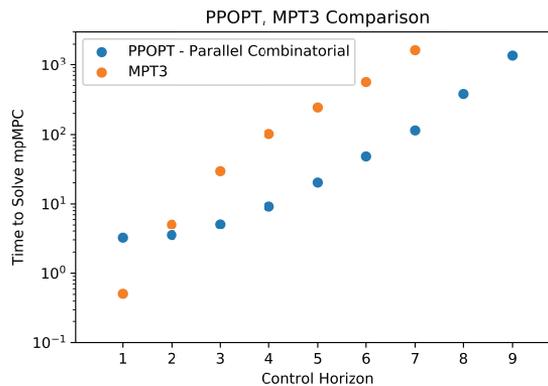


Figure 4: eMPC comparison with PPOPT and MPT 3.1.2

The C++ exported solution is quite fast across all of the hardware evaluated. This is applicable for hard real-time computing optimization problems, even with embedded processors. The indices in the table labeled as N/a cannot be run on the platform, such as MatLab on the embedded processor. While a performance regression was observed for the exported JavaScript solution, the indented application is integrated into a web framework, allowing the end-user to utilize optimization methods without needing an optimization package or license.

4. Conclusions

In this paper, we described PPOPT, the **Python Parametric OPTimization Toolbox**. PPOPT has state-of-the-art implementations of the multiple algorithms implemented to utilize many-core computers to speed up solving multiparametric programming problems. With strong results com-

pared to the state-of-the-art multiparametric solvers. Computational studies on the mpQP1 problem set show superior performance than POP and MPT3. An explicit MPC case study shows an approximate $15\times$ speed up with solving explicit MPCs compared to MPT3. The parallelism scaling of the combinatorial algorithm was explored on large-scale problems showing excellent strong scaling with core count with large scale problems. Additionally, the performance of the exported solution for multiparametric programs was studied on an example solution with different hardware systems showing strong results.

We hope to contribute to many aspects, such as developing and then implementing new scalable algorithms for mpLPs, mpQPs, and the mixed-integer variations of these problems. In addition to completely solving multiparametric programs, a problem type of interest is partially solving the multiparametric program and using this information to speed up the online computational results for large-scale problems.

References

- S. Avraamidou, E. N. Pistikopoulos, 2019. B-pop: Bi-level parametric optimization toolbox. *Computers & Chemical Engineering* 122, 193–202, 2017 Edition of the European Symposium on Computer Aided Process Engineering (ESCAPE-27).
- B. Burnak, N. A. Diangelakis, J. Katz, E. N. Pistikopoulos, 2019. Integrated process design, scheduling, and control using multiparametric programming. *Computers & Chemical Engineering* 125, 164–184.
- N. A. Diangelakis, B. Burnak, J. Katz, E. N. Pistikopoulos, 2017. Process design and control optimization: A simultaneous approach by multi-parametric programming. *AIChE Journal* 63 (11), 4827–4846.
- P. Dua, F. J. Doyle, E. N. Pistikopoulos, 2006. Model-based blood glucose control for type 1 diabetes via parametric programming. *IEEE Transactions on Biomedical Engineering* 53 (8), 1478–1491.
- A. Gupta, S. Bhartiya, P. Nataraj, 2011. A novel approach to multiparametric quadratic programming. *Automatica* 47 (9), 2112–2117.
- M. Herceg, M. Kvasnica, C. Jones, M. Morari, July 17–19 2013. Multi-Parametric Toolbox 3.0. In: *Proc. of the European Control Conference*. Zürich, Switzerland, pp. 502–510, <http://control.ee.ethz.ch/~mpt>.
- A. Kiparissides, M. Koutinas, C. Kontoravdi, A. Mantalaris, E. N. Pistikopoulos, 2011. ‘closing the loop’ in biological systems modeling — from the in silico to the in vitro. *Automatica* 47 (6), 1147–1155, special Issue on Systems Biology.
- R. Oberdieck, N. A. Diangelakis, M. M. Papathanasiou, I. Nascu, E. N. Pistikopoulos, 2016. Pop – parametric optimization toolbox. *Ind. Eng. Chem. Res.* 55, 103–112.
- R. Oberdieck, N. A. Diangelakis, E. N. Pistikopoulos, 2017. Explicit model predictive control: A connected-graph approach. *Automatica* 76, 103–112.
- R. Oberdieck, E. N. Pistikopoulos, 2016. Parallel computing in multi-parametric programming. In: Z. Kravanja, M. Bogataj (Eds.), 26th European Symposium on Computer Aided Process Engineering. Vol. 38 of *Computer Aided Chemical Engineering*. Elsevier, pp. 169–174.
- F. Odonez, 2002. On the explanatory value of condition numbers for convex optimization: Theoretical issues and computational experience. Ph.D. thesis, Massachusetts Institute of Technology.
- I. Pappas, D. Kenefake, B. Burnak, S. Avraamidou, H. S. Ganesh, J. Katz, N. A. Diangelakis, E. N. Pistikopoulos, 2021. Multiparametric programming in process systems engineering: Recent developments and path forward. *Frontiers in Chemical Engineering* 2, 32.
- E. N. Pistikopoulos, N. A. Diangelakis, R. Oberdieck, 2021. *Multi-parametric optimization and Control*. Wiley.
- O. Romanko, A. Ghaffari-Hadigheh, T. Terlaky, 2012. Multiobjective optimization via parametric optimization: Models, algorithms, and applications. In: T. Terlaky, F. E. Curtis (Eds.), *Modeling and Optimization: Theory and Applications*. Springer New York, New York, NY, pp. 77–119.
- V. Sakizlis, N. M.P. Kakalis, V. Dua, J. D. Perkins, E. N. Pistikopoulos, 2004. Design of robust model-based controllers via parametric programming. *Automatica* 40 (2), 189–201.
- J. Spjøtvold, E. C. Kerrigan, C. N. Jones, P. Tøndel, T. A. Johansen, 2006. On the facet-to-facet property of solutions to convex parametric quadratic programs. *Automatica* 42 (12), 2209–2214.
- W. W. Tso, B. Burnak, E. N. Pistikopoulos, 2020. Hy-pop: Hyperparameter optimization of machine learning models through parametric programming. *Computers & Chemical Engineering* 139, 106902.

Improved Sequential Least Squares Programming– Driven Feasible Path Algorithm for Process Optimisation

Yingjie Ma^a, Nan Zhang^a, Jie Li^{a*}

*^aCentre for Process Integration, School of Chemical Engineering and Analytical
Science, The University of Manchester, Manchester M13 9PL, UK*

Abstract

Sequential least squares programming (SLSQP) algorithm has been shown to be useful for driving the feasible path algorithms for process optimisation. However, the existing SLSQP algorithms still need many function evaluations for computationally challenging process optimisation problems. In the current work, we propose several novel strategies to improve both the efficiency and convergence of the SLSQP algorithm. Solving a large-scale process optimisation problem indicates that the algorithm can save computational times by 10-90% with better solutions generated.

Keywords: process optimisation, SLSQP, line search, feasible path, numerical errors.

1. Introduction

Optimisation is a powerful tool to design the best chemical process with the lowest cost or highest profit while satisfying the production requirements and restrictions at the same time (Biegler, 1993). To get optimisation results that can match the real-world production well, it is highly desirable to use rigorous unit operation models. However, this often leads to strongly nonlinear, non-convex or even ill-conditioned large-scale nonlinear programming (NLP) problems, which is hard to solve (Pattison and Baldea, 2014).

The feasible path algorithm is widely used to solve process optimisation problems due to its good convergence, which divides the entire process optimisation problem into a small-scale optimisation problem in the outer level and a large-scale process simulation problem in the inner level (Parker and Hughes, 1981). However, for feasible path algorithms, the simulation needs to be conducted in each iteration, which may suffer from divergence if the models are strongly nonlinear. To tackle with this problem, the pseudo-transient continuation (PTC) modeling approach has been introduced for process simulations, which was reported to be able to improve the convergence of simulations and optimisations significantly (Pattison and Baldea, 2014). However, PTC simulations are much slower than steady-state simulations, making feasible path algorithms less efficient. To improve the efficiency, we proposed a hybrid steady-state and time-relaxation-based optimisation algorithm in our previous work, which combines PTC simulations and steady-state simulations to reduce the times of implementing PTC simulations without degrading the convergence of simulations and optimisations (Ma et al., 2020).

The hybrid algorithm in our previous work (Ma et al., 2020) is driven by a sequential least squares programming (SLSQP) algorithm (Kraft, 1988) that is potential to be more stable than sequential quadratic programming (SQP) algorithm under numerical noise (Schittkowski, 1982). However, due to the lack of wide research, the existing SLSQP

algorithm is less efficient and may generate ascent directions due to the numerical errors in the least squares (LSQ) solver (Lawson and Hanson, 1995) adopted.

With above in mind, in this work, we will develop an improved SLSQP algorithm to drive the hybrid steady-state and time-relaxation-based algorithm. The proposed algorithm uses a formula to initialize the step length during the line search instead of always using full step length as initial values to achieve higher efficiency. Also, some strategies are proposed to improve the convergence of the SLSQP algorithm, such as revising the search directions with a new LSQ solver and relaxing the line search criteria in some iterations. The computational results from solving a challenging large-scale process optimisation problem in the literature indicates that the improved SLSQP algorithm usually generates a better solution with 10-90% less computational time required, compared with an existing SLSQP algorithm.

2. Problem Statement

The NLP problem solved in the current work can be stated as follows:

$$\begin{aligned} \min f(x) & \quad \text{(NLP)} \\ \text{s. t. } h(x) &= 0 \\ g(x) &\geq 0 \\ x &\in R^n, \end{aligned}$$

where $f: R^n \rightarrow R$, $h: R^n \rightarrow R^{m_E}$ and $g: R^n \rightarrow R^{m_I}$ are all twice continuously differentiable functions. We assume both the function values and first-order derivatives are available. When applying the developed SLSQP algorithm to the feasible path algorithm, function values and gradients can be got from equation-oriented simulators.

3. Improved Sequential Least Squares Programming Algorithm

3.1. Sequential least squares programming algorithm

The SLSQP algorithm was proposed by (Schittkowski, 1982), which is to solve a constrained LSQ subproblem in each major iteration to generate a descent direction. The LSQ problem solved at iteration k is shown as follows,

$$\begin{aligned} \min \frac{1}{2} \|R^k d^k - q^k\|^2 & \quad \text{(LSQ)} \\ \text{s. t. } \nabla h^{kT} d^k + h^k &= 0 \\ \nabla g^{kT} d^k + g^k &\geq 0, \end{aligned}$$

where d^k is the search direction to be solved. g^k and h^k are values of constraint functions at iterate x^k , while ∇g^k and ∇h^k are current gradients of inequalities and equalities. R^k and q^k are least squares matrix and observation vector respectively, which are updated during optimisation according to the BFGS formula and an LDL updating algorithm (Fletcher and Powell, 1974).

Line search can be used to guarantee the global convergence of SLSQP algorithm, which is to get a step length α providing a sufficient decrease for the merit function $M(x)$ along the search direction d^k . After a α is obtained, the iterate can be updated using Eq. (1),

$$x^{k+1} \leftarrow x^k + \alpha d^k. \quad (1)$$

A frequently used condition to guarantee the sufficient decrease for constrained NLP is the Armijio condition (Armijio, 1966) as shown in the following Eq. (2),

$$M(x^k + \alpha d^k) \leq M(x^k) + \alpha \rho DM^k(0), \quad (2)$$

where $DM^k(0)$ is the directional derivative of $M(x^k + \alpha d^k)$ along d^k at $\alpha = 0$, and $\rho \in (0, 0.5)$ is a constant. The region of α satisfying the Eq. (2) is shown in Fig. 1.

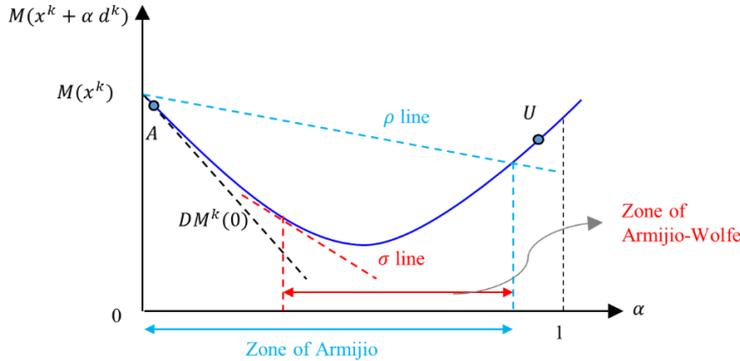


Figure 1 Zones of α in Armijio conditions and Armijio-Wolfe Conditions

The framework of a line search and merit function-based SLSQP algorithm can be described as follows:

Algorithm 1: SLSQP

- Step 1: set $k \leftarrow 0$; give x^0 and evaluate $f^0, g^0, h^0, \nabla f^0, \nabla g^0, \nabla h^0$; assign initial R^k and get q^k ;
- Step 2: solve LSQ subproblem for d^k and check NLP convergence condition. If the solution is found, go to Step 5;
- Step 3: conduct line search to get α satisfying Eq. (2) and to get x^{k+1} ;
- Step 4: evaluate $f^{k+1}, g^{k+1}, h^{k+1}, \nabla f^{k+1}, \nabla g^{k+1}, \nabla h^{k+1}, R^{k+1}, q^{k+1}$ at x^{k+1} ; set $k \leftarrow k + 1$, then go back to Step 2;
- Step 5: return x^k, f^k .

3.2. A new line search algorithm

The backtrack method is usually used for the line search in Step 3 of Algorithm 1, where a full step length is adopted at first and then reduced if it doesn't satisfy Eq. (2). This can achieve superlinear convergence around the optimal point. However, when x^k is far from the solution, the step length satisfying Eq. (2) may be much smaller than one especially for strongly nonlinear problems, so it possibly leads to many unnecessary function evaluations to use full step lengths as initial values for the line search in all iterations. In this work, a formula Eq. (3) from (Fletcher, 2008) that is shown to be efficient for unconstrained optimisation is used here for the constrained NLP problem,

$$\alpha^0 = \min\left(1, \frac{2\Delta M^k}{-DM^k(0)}\right), \quad (3)$$

where ΔM^k is the predicted decrease of the merit function in the current iteration k , which is estimated as $M^{k-1} - M^k$, the decrease of the merit function in the preceding major iteration. The second term in the right hand side of formula (3) is obtained by assuming that the merit function is quadratic with respect to the step length α , while its first term

is to avoid a step length larger than 1. For unconstrained optimisation, α^0 from Eq. (3) will be always 1 when x^k is sufficiently close to solution x^* for a superlinear convergence algorithm, so the superlinear convergence rate of the original algorithm will not be degraded (Fletcher, 2008). For constrained optimisation, although there haven't been a similar proof yet, α^0 are indeed frequently one when x^k is close to optima and not dominated by numerical noise according to our observation.

However, for strongly nonlinear problems, the predicted α^0 from Eq. (3) may be over small and far from the local minimum of the merit function along d^k when x^k is distant from the solution (such as point A in Fig. 1). In such case, if the frequently-used backtrack line search is used and the step length is accepted immediately when it satisfies the Eq. (2), the optimisation may be slow down and/or the approximate Hessian matrix may be polluted by large numerical noise, deteriorating the optimisation. Thus, the Wolfe condition Eq. (4) (Wolfe, 1971) is introduced to avoid an over small α to be accepted,

$$\nabla f(x^k + \alpha d^k)^T d^k \geq \sigma \nabla f(x^k)^T d^k, \quad (4)$$

where $\sigma \in (\rho, 1)$ is a constant. The region of α satisfying both Eq. (2) and Eq. (4) is shown in Fig. 1, which indicates a α satisfying Armijio condition but not Wolfe condition needs to be increased (Step 3.6 in the following Algorithm 2). However, it needs more function evaluations to find a point satisfying both conditions compared with the algorithm enforcing the Armijio condition only. Theoretically, the Wolfe condition doesn't have to be satisfied in the SLSQP algorithm as it is not used in the proof of the global convergence. Thus, we will not check the Wolfe condition if the α satisfying the Armijio condition is not over small judged by Eq. (5),

$$\alpha \geq \tau \alpha^{up}, \quad (5)$$

where $\alpha^{up} \in (0, 1]$ is the smallest α not satisfying Eq. (2) (such as the point U in Fig. 1) and will be updated during the line search. $\tau \in (0, 1)$ is a constant to allow some gap between α and α^{up} , which is set as 0.1 in this work. Overall, the proposed line search algorithm applied to a major iteration k of Algorithm 1 is shown as follows. To avoid the abuse use of indices, we get rid of the index k and use the index i only to represent the iteration in the following Algorithm 2.

Algorithm 2: line search with α^0 predicted and Wolfe condition checked

- Step 3.1: given x^0 ; set $i \leftarrow 0$, $\alpha^{up} \leftarrow 1$, $\rho \in (0, 0.5)$, $\sigma \in (\rho, 1)$, $\tau \in (0, 1)$, $\gamma > 1$;
get an initial step length α^0 from Eq. (3);
- Step 3.2: set $x^{i+1} \leftarrow x^i + \alpha^i d$, and evaluate the merit function $M(x^{i+1})$;
- Step 3.3: if Eq. (2) is satisfied, go to Step 3.5; otherwise, set $\alpha^{up} \leftarrow \alpha^i$;
- Step 3.4: generate an α^{i+1} smaller than α^i ; set $i \leftarrow i + 1$, then go back to Step 3.2;
- Step 3.5: if $\alpha^i \geq \tau \alpha^{up}$ or α^i satisfies Eq. (4), go to Step 3.7;
- Step 3.6: generate $\alpha^{i+1} = \gamma \alpha^i$, set $i \leftarrow i + 1$, then go back to Step 3.2;
- Step 3.7: return α^i , x^{i+1} .

The constants used are $\rho = 0.1$, $\sigma = 0.9$, $\tau = 0.1$, $\gamma = 10$.

3.3. Other improvements

In the existing SLSQP algorithms (Kraft, 1988; Schittkowski, 1982), a powerful algorithm called LSEI (Lawson and Hanson, 1995) are used to solve LSQ subproblems. Although the LSEI algorithm usually works very well, it may generate ascent directions due to numerical errors, which would cause failure of the optimisation. To avoid

divergence, we solve the LSQ subproblem based on a robust commercial solver Gurobi (Gurobi Optimization, 2020) once the LSEI algorithm generates an ascent direction, while in all the other cases, the LSEI algorithm are still used as we find it leads to better performance compared with many other QP and LSQ solvers.

Second, even if a descent direction is generated from an LSQ subproblem, the line search may still fail to find a α satisfying the Armijio condition as function values generated from process simulations suffer from numerical noises. In such case, to continue the algorithm and step out of the noisy region for a better solution, we allow the largest step length α satisfying the following Eq. (6) is accepted,

$$\frac{|M(x^k + \alpha d^k) - M(x^k)|}{M(x^k)} \leq ntol, \tag{6}$$

where $ntol$ is a small value to tolerate noise, which is 10^{-4} here. Eq. (6) is to guarantee the accepted step length will only increase the merit function slightly while the algorithm can still have some progress to have the chance to leave the numerically noisy region. Such condition will only be applied for limited times (here, five times) to avoid the algorithm to run forever when the optimisation cannot escape the region anyway.

4. Case Study

The optimization of a dimethyl ether (DME) production process intensified by dividing wall column (DWC) is used to validate the proposed algorithm, which was first optimized by (Pattison et al., 2016) and later by (Ma et al., 2018; Ma et al., 2020) with better solutions. The problem has 12151 equalities, 17 inequalities and 13661 decision variables in the full model, while it has 122 decision variables, 17 inequalities and one equality in the reduced space when solved by the feasible path algorithm. The objective function to be minimized is a penalty of the cost without a specific unit. The problem is rather challenging to solve not only because of its nonlinearity but more because of its intrinsically ill-conditioning evidenced by that the condition number of the approximate Hessian matrix ranges from $10^{10} - 10^{14}$ in most iterations. In (Ma et al., 2018), an SQP solver was used and the problem has to be scaled by trial and error to be solvable. However, in (Ma et al., 2020), the problem could be solved directly without scaling using an existing SLSQP solver (Kraft, 1988). For a detailed comparison between the existing SLSQP algorithm (Kraft, 1988) (notated as `slsqp`) and our improved algorithm (notated as `i-slsqp`), we conduct the optimisation using `slsqp` and `i-slsqp` from six different initial points on a desktop with 3.20 GHz Intel i7 CPU and 16 GB RAM running 64-bit Windows operating system. The computational results from both algorithms are shown in Table 1.

Table 1 Comparative results from the hybrid algorithm of Ma et al. (2020) and the proposed algorithm

Initial point		1	2	3	4	5	6
Time/s	<code>slsqp</code>	4415	2003	6264	1496	2563	2408
	<code>i-slsqp</code>	469	1173	976	1263	989	2032
Number of simulations	<code>slsqp</code>	2259	923	2373	754	1337	969
	<code>i-slsqp</code>	287	701	527	871	696	1190
Objective	<code>slsqp</code>	125689	125713	125715	125711	125695	125701
	<code>i-slsqp</code>	125708	125687	125712	125705	125683	125690

As can be seen from Table 1, our improved SLSQP algorithm always needs 10-90% less computation time from different initial points than SLSQP. Especially, for the optimisation starting from the first initial point, the computational time is reduced by around one order of magnitude. This is partially because less function evaluations are used in our algorithm as shown in Table 1. The time reduction is also because small step lengths are used during the line search, leading to better convergence of steady-state simulations and hence less number of expensive PTC simulations required. Furthermore, it can be seen that our algorithms usually get better solutions except for the optimisation starting from the first initial point, and the best solution with an optimum of 125683 is also generated from the new algorithm.

5. Conclusion

In this work, we proposed an improved SLSQP algorithm for process optimisation using rigorous models. In the improved algorithm, the initial step length during the line search was calculated from a formula to reduce the number of function evaluations. The Wolfe condition was introduced into the line search to avoid over small step length, but it was not enforced when the step length satisfying Armijo condition was large enough to avoid additional function evaluations. For a better convergence, the LSQ subproblem was solved by our customized solver when the existing LSQ algorithm generated ascent directions. Also, the line search is allowed to accept some step lengths which increase the merit function slightly to tolerate numerical noise from process simulations. One large-scale process optimisation problem from literature shows that the improved algorithm can save computational time by 10-90% with better solutions generated.

References

- Armijo, L., 1966. Minimization of functions having Lipschitz continuous first partial derivatives. *Pac. J. Math.* 16(1), 1-3.
- Biegler, L.T., 1993. From nonlinear programming theory to practical optimization algorithms: A process engineering viewpoint. *Comput. Chem. Eng.* 17, S63-S80.
- Fletcher, R., 2008. *Practical methods of optimization*. Wiley, Chichester, West Sussex, U.K.
- Fletcher, R., Powell, M.J.D., 1974. On the Modification of LDL^TS Factorizations. *Math. of Comput.* 28, 1067-1087.
- Gurobi Optimization, L., 2020. *Gurobi Optimizer Reference Manual*.
- Kraft, D., 1988. A software package for sequential quadratic programming. Report DFVLR-FR 88-28 (Deutsche Forschungs- und Versuchsanstalt für Luftund Raumfahrt).
- Lawson, C.L., Hanson, R.J., 1995. *Solving least squares problems*. SIAM, Philadelphia.
- Ma, Y., Luo, Y., Ma, X., Yang, T., Chen, D., Yuan, X., 2018. Fast Algorithms for Flowsheet Simulation and Optimization Using Pseudo-Transient Models. *Ind. Eng. Chem. Res.* 57, 14124-14142.
- Ma, Y., McLaughlan, M., Zhang, N., Li, J., 2020. Novel feasible path optimisation algorithms using steady-state and/or pseudo-transient simulations. *Comput. Chem. Eng.* 143, 107058.
- Parker, A.L., Hughes, R.R., 1981. Approximation programming of chemical processes—1: Optimization of flowtran® models. *Comput. Chem. Eng.* 5, 123-133.
- Pattison, R.C., Baldea, M., 2014. Equation-Oriented Flowsheet Simulation and Optimization Using Pseudo-Transient Models. *AIChE J.* 60, 4104-4123.
- Pattison, R.C., Gupta, A.M., Baldea, M., 2016. Equation-oriented optimization of process flowsheets with dividing-wall columns. *AIChE J.* 62, 704-716.
- Schittkowski, K., 1982. The nonlinear programming method of Wilson, Han, and Powell with an augmented Lagrangian type line search function. *Numer. Math.* 38, 115-127.
- Wolfe, P., 1971. Convergence Conditions for Ascent Methods. II: Some Corrections. *SIAM Rev.* 13, 185-188.

Parallel Simulated Annealing approach for optimal process plants instrumentation

José Hernández^a, Gabriela Minetti^b, Carolina Salto^b, Mercedes Carnero^{a,*} and Mabel Sánchez^c

^a*Facultad de Ingeniería, Universidad Nacional de Río Cuarto, Río Cuarto 5800, Argentina*

^b*Facultad de Ingeniería, Universidad Nacional de La Pampa, General Pico 6360, Argentina*

^c*Plapiqui, Universidad Nacional del Sur, Bahía Blanca 8000, Argentina*

**mcarnero@ing.unrc.edu.ar*

Abstract

In a smart chemical plant, knowing the process state at any moment is crucial with impact on aspects such as economic, safety, or control. The information is collected by sensors distributed throughout the plant, responsible for measuring and transmitting the values of temperature, humidity, pressure, among others. The set of devices used in the measurement is called the sensor network (SN). Consequently, their design optimization implies importance and huge capital cost. We propose an intelligent optimization solver based on a Parallel Hybrid Simulated Annealing (PHSA) to solve it. The parallelism is applied at the algorithmic level, following a cooperative model. Among the migration parameters, the replacement criteria have an important role in the PHSA performance. Our main objective is to analyze the PHSA behavior by considering different replacement criteria. The results obtained by PHSA achieve the best-known solution for large and complex SN cases. Furthermore, the parallel HSA exhibits efficient scalability to solve the SN distribution problem.

Keywords: Simulated Annealing, parallel computing, sensor network design; combinatorial optimization; metaheuristics

1. Introduction

In the last decade, the integration of the process control, the operation of the entire plant, and the corporation's business system have been a tendency, which is possible thanks to the rapid advances in cybernetic infrastructure and communications technology that allow the flow of information in real-time. The so-called smart plant aims to maximize the economic profitability of the activity with social responsibility, complying with strict standards on environmental care, occupational health, and safety. Christofides et al. (2007) mention the design of the sensor network (SN) as a key issue to make the smart plant a reality since it provides comprehensive knowledge of the process's current state. The systematic design of the SN in a plant, formulated as an optimization problem, is known as sensor network design problem (SNDP) and consists in determining whether or not each variable should be measured while optimizing the performance. Performance criteria can be classified as criteria associated with the performance of the sensor system itself (Carnero et al. (2018); Kotecha et al. (2008)) or related to indicators of the process for which the instrumentation is designed (Sen et al. (2016); Paul et al. (2015); Sambito (2021)).

The number of possibilities regarding how many variables to measure and the type of the desired performance can achieve a very high order even for small-size plants. Therefore, an important

aspect is related to the choice of the resolution methodology, which can be classified into two main groups: exact methods and heuristics. The analysis of the performance exhibited by exact algorithms (Zhang and Chmielewski (2017); Nguyen and Bagajewicz (2011)) shows that their behavior is very dependent on the particular design problem and its size. In this sense, they are not robust enough to solve a wide range of specifications, nor do they have the possibility of scaling. The most notorious disadvantage is that the optimality guarantee cannot be achieved in many cases, given the excessive computation time they require. This limitation comes mainly from the combinatorial nature of the problem, which results in the impossibility of having polynomial-time algorithms for the cases of practical interest. Therefore, heuristic optimization methods emerge as the most feasible solution alternative to address higher dimensional designs (Hernandez et al. (2019); He and Ma (2014); Panizo et al. (2018)). When many variables have to be measured, the SN design process optimization runtime has been reduced. Consequently, parallelization techniques on the heuristic methods (Talbi (2009)) are used, achieving a substantial improvement in the optimization performance.

In this work, a parallel Simulated Annealing (SA) (Kirkpatrick et al. (1983); Cicirello (2017)) algorithm is proposed to optimize the SNDP. In this model, many hybrid SA algorithms (Hernández et al. (2020)) are launched in parallel and exchanges information to improve the quality solutions and enhance efficiency. Consequently, the main contribution of our research is to develop an intelligent solver based on a parallel HSA algorithm, named PHSA, which intends to support the decision-making during the design of complex sensor networks in a chemical plant. The focus is to increase the efficiency of this solver when large and complex SN cases are addressed. Hence, we formulate the following research questions (RQs): *RQ1*) Does the parallelization of HSA increase its efficiency when we deal with large SN cases?; *RQ2*) What parametric configurations should be used in the proposed parallel algorithm to enhance the solution quality?; and *RQ3*) Is our PHSA design efficient to solve the SNDP?. To answer these RQs, we conduct experiments by applying PHSA to different case studies of incremental size, including a well-known problem in chemical engineering literature.

The remainder of this article is structured as follows. Section 2 explains the SNDP. Section 3 explains our algorithmic proposal. Then, we describe the experimental design and the methodology used in Section 4. We analyze and compare the HSA behavior when solving the SNDP problems in Section 5. Finally, the most important conclusions and future research lines are presented in Section 6.

2. Sensor Network Design Problem

The SNDP is summarized as a problem of finding the minimum cost network that satisfies precision and estimability constraints. Formally, a SNDP solution has to satisfy these constraints for a set of key variable estimates, as stated by Eq 1, where \mathbf{q} is an n -dimensional vector of binary variables such that $q_i = 1$ if variable i is measured, and $q_i = 0$ otherwise, \mathbf{c}^T is the cost vector; $\hat{\sigma}_k$ is the estimated standard deviation of the k -th variable in S_σ obtained by a data reconciliation procedure (Bagajewicz and Sanchez (2000)), and E_l stands for the degree of estimability of the l -th variable in S_E . Furthermore, S_σ and S_E are the set of key process variables with requirements in precision and ability to be estimated, respectively.

$$\min(\mathbf{c}^T \mathbf{q}) \quad \text{s.t.} \begin{cases} \hat{\sigma}_k(\mathbf{q}) \leq \hat{\sigma}_k^*(\mathbf{q}), & \forall k \in S_\sigma \\ E_l \geq 1, & \forall l \in S_E \\ \mathbf{q} \in \{0, 1\} \end{cases} \quad (1)$$

In this formulation, it is assumed that a linearized algebraic model represents plant operation, measurements are subject to noncorrelated random errors, there is only one potential measuring

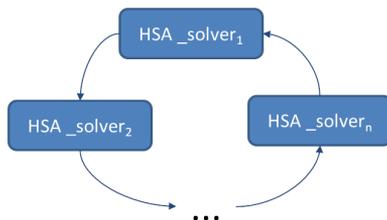


Figure 1: PHSA scheme under ring topology.

device for each variable, and there are no restrictions for the localization of instruments. Regarding degree of estimability constraint, if $E_I=1$, the feasibility of the constraint can be checked by executing a variable classification procedure, which can be accomplished by matrix projection, QR decomposition, or matrix co-optation (Romagnoli and Sanchez (2000); Narasimhan and Jordache (2000)).

3. PHSA solver

In this work, we propose an intelligent parallel solver to optimize the Sensor Network Design problem in chemical plants. This solver is based on the hybrid SA (HSA) introduced in Hernandez et al. (2019). HSA works as the main metaheuristic with a subordinated ad hoc local search, inspired in tabu search with strategic oscillation technique (SOTS), giving rise to the Hybrid Simulated Annealing algorithm. We refer readers to Hernandez et al. (2019) for more HSA details. The proposed parallel HSA (PHSA) follows an Algorithmic-Level Parallel Model (see Talbi (2009)). The PHSA consists of n HSA solvers that are launched in parallel and exchange information to improve the quality solutions and enhance efficiency. Hence, the PHSA design follows a cooperative search strategy. Each HSA solver generates its initial solution S_0 and run an independent HSA. The HSA solvers exchange information related to the search to compute better and more robust solutions regarding migration frequency. Different topologies can be used to do this interchange, but we focus on the ring topology, arising the PHSA_ring (see Figure 1), therefore the HSA solvers are arranged in a logic unidirectional ring. The i^{th} HSA solver sends its current solution S_1 to their $(i+1)^{th}$ neighbor solver in the ring with a certain frequency and following an asynchronous exchange. When the target HSA solver receives a solution (S_2), a replacement criteria is used. In this work, we consider three different approaches: *RC1*) S_2 is accepted with the Boltzmann probability, *RC2*) S_2 is accepted if it is better than the current solution (S_1), *RC3*) S_2 is accepted if it is better than the best-found local solution (S_b).

4. Experimental Design

The PHSA_ring performance is evaluated considering chemical processes of high complexity and size. As a first case study, a simplified ethylene plant was addressed, consisting of 47 units and 82 streams, whose operation is only represented through global mass balances. Secondly, the Tennessee Eastman Process (TEP), Downs and Vogel (1993), a widely considered problem in the chemical process monitoring and control literature was used. Both global and component mass balances are considered in the TEP. The complete system comprises 42 equations that were linearized around the operating point, and a total of 78 variables. Interested readers can gain access to the file containing information about the case studies and the complexity of the set of constraints imposed on all case studies at https://www.ing.unrc.edu.ar/archivos/sndp_cases.doc. In case study 1, a set of 14 required variables with precision constraints on 6 of them is studied. The second case study considers a set of 24 required variables, all of them with precision restrictions. The standard deviation of flow meters is 2% of the corresponding true flow rates for both cases.

Table 1: Solution cost values obtained by PHSA_ring, considering the replacement criteria and #HSA solvers.

Replac. Criteria	#HSA Solvers	Case Study 1			Case Study 2		
		Min	Mean $_{\pm SD}$	%Hits	Min	Mean $_{\pm SD}$	%Hits
RC1	4	50845.16	53322.67 $_{\pm 2056.84}$	13.33%	25280.00	26624.92 $_{\pm 575.98}$	1.69%
	8	50845.16	51809.13 $_{\pm 1775.61}$	23.33%	25590.00	26495.79 $_{\pm 506.05}$	0.00%
	12	50845.16	51121.21 $_{\pm 1047.35}$	30.00%	25280.00	26334.80 $_{\pm 527.87}$	4.00%
	16	50845.37	51121.41 $_{\pm 1047.30}$	13.33%	25290.00	26139.17 $_{\pm 451.25}$	0.00%
RC2	4	50845.16	52497.10 $_{\pm 2056.96}$	26.67%	25470.00	26502.98 $_{\pm 543.85}$	0.00%
	8	50845.16	51946.90 $_{\pm 1856.73}$	20.00%	25430.00	26458.53 $_{\pm 584.39}$	0.00%
	12	50845.16	51259.02 $_{\pm 1259.56}$	10.00%	25690.00	26557.24 $_{\pm 420.38}$	0.00%
	16	50845.16	51258.93 $_{\pm 1259.59}$	20.00%	25600.00	26355.22 $_{\pm 425.04}$	0.00%
RC3	4	50845.16	51120.79 $_{\pm 1047.33}$	56.67%	25280.00	25707.11 $_{\pm 367.96}$	2.22%
	8	50845.16	51396.31 $_{\pm 1427.24}$	23.33%	25280.00	25756.00 $_{\pm 340.43}$	2.86%
	12	50845.16	51258.48 $_{\pm 1259.51}$	43.33%	25290.00	26008.00 $_{\pm 354.51}$	0.00%
	16	50845.16	51258.61 $_{\pm 1259.58}$	36.67%	25280.00	25941.67 $_{\pm 397.60}$	2.78%

To study the performance of this PHSA_ring, we configure each HSA_solver $_i$ as is suggested in Hernandez et al. (2019). Consequently, we use an initial seed temperature $T_s = 900$, the RAND cooling scheme, and the adaptive Markov Chain Length MCLa1. The stop condition of the PHSA_ring is to reach 1,250 iterations. Four PHSA_ring scenarios with 4, 8, 12 and 16 HSA solvers are assessed. Moreover, the three replacement criteria explained in Section 3 are used, and the migration frequency is calculated as a 10% of the HSA_solver $_i$ iterations (1250/#HSA solvers). At least 30 independent executions for each instance and PHSA's configuration and parallel scenarios are necessary due to the stochastic nature of the HSA solver, giving a total of 720 executions (30 runs \times 2 instances \times 3 replacement criteria \times 4 parallel scenarios). This big experimentation allows us to gather meaningful experimental data and apply statistical confidence metrics to validate our results and conclusions. Before performing the statistical tests, we first check whether the data follow a normal distribution by applying the Shapiro-Wilks test. Where the data are distributed normally, we later apply an ANOVA test. Otherwise, we use the Kruskal–Wallis (KW) test. This statistical study allows us to assess whether or not there are meaningful differences between the compared algorithms with $\alpha = 0.05$. To determine these algorithm pairwise differences is by carrying out a post hoc test, as is the case of the Wilcoxon test if the KW test is used.

5. Result Analysis

The solution quality and efficiency of the proposed parallel HSA are assessed taking into account two different stopping criteria: one is based on a maximum number of evaluations (predefined effort), while the other consists in running the PHSA_ring until a given solution is found or the maximum number of evaluations is reached (predefined solution quality). Considering a predefined effort first, we study the PHSA_ring's performance under each replacement criterion in the four parallel scenarios. For a fair comparison, all algorithms used the predefined effort as a cut point. Second, we analyze the parallel performance. We focus on the number of optimum values the PHSA configurations found. Then, we measure the speedup between the serial time (PHSAs using one processor) against the parallel time (PHSAs using 4, 8, 12, and 16 processors) because they run the same underlying algorithm, following the speedup definition given in Alba (2002).

At the beginning, we present the results for both case studies. Table 1 shows the minimum and average solution cost values, the respective standard deviation (SD), and the percentage of times that the best-known solution is found (%Hits) for the PHSA_ring, considering the three replacement criteria and the four parallel scenarios (number of HSA solvers). We observe significant different PHSA_ring's behaviors since only for case study 1 every PHSA_ring variant finds the best-known solution. Furthermore, for this case, the %Hits varies in [10, 56.67], while for the second case, the range is [0, 4]. This difference suggests that the second case needs a bigger predefined effort

Table 2: Total execution times consumed by PHSA_ring, considering the replacement criteria and #HSA solvers.

Replac. Criteria	#HSA Solvers	Case Study 1		Case Study 2	
		Min	Mean \pm SD	Min	Mean \pm SD
RC1	4	180.55	197.23 \pm 23.64	157.14	163.92 \pm 3.06
	8	56.16	68.06 \pm 12.06	47.22	49.47 \pm 0.92
	12	33.37	39.12 \pm 6.69	27.75	28.99 \pm 1.31
	16	25.03	31.82 \pm 4.28	21.43	25.67 \pm 4.05
RC2	4	185.88	197.57 \pm 25.57	154.97	162.86 \pm 3.38
	8	57.33	66.70 \pm 11.01	47.70	49.29 \pm 0.99
	12	33.12	42.2 \pm 6.81	27.60	28.92 \pm 0.55
	16	24.76	31.52 \pm 4.40	21.79	25.22 \pm 3.99
RC3	4	185.28	216.40 \pm 36.96	154.42	160.42 \pm 3.31
	8	57.42	65.16 \pm 9.79	47.43	48.96 \pm 0.78
	12	33.37	41.98 \pm 7.08	27.89	28.87 \pm 0.55
	16	24.95	29.81 \pm 4.72	21.58	24.44 \pm 3.27

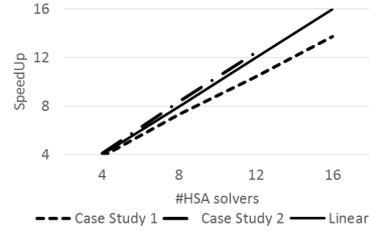


Figure 2: SpeedUp values for PHSA_ring, considering both case studies.

to reach the best-known solution. We follow with the study about the influence of replacement criteria and parallel scenarios from a solution quality point of view. On one side, a statistic and significant difference is observed when the replacement criterion *RC3* is used to solve the two case studies, since the incoming solution is only accepted if it is better than the best local solution. Under this criterion, the lowest cost averages of the solution are found, observing the lowest standard deviations and the highest %Hits. On the other hand, statistically similar behavior is detected when the four parallel scenarios are evaluated. Consequently, the *RC3* enhances the solution quality in every scenario, becoming the best replacement choice for PHSA_ring. In this way, we answer the second research question (*RQ2*).

Now, we proceed with the computational effort through the run-time analysis. Table 2 presents the minimum and average of total execution times and their standard deviations (SD) consumed for the PHSA_ring considering the three replacement criteria and the four parallel scenarios. For both case studies, no statistically significant differences are observed when the total execution time is analyzed, considering the replacement criteria. In other words, the replacement selection only depends on the required solution quality. However, when the parallel scenarios are regarded for this analysis, the reduction in total execution time is noticeable when the number of HSA solvers increases. This reduction loses inertia when going from 12 to 16 solvers since the communication times for 16 HSA solvers are more time-consuming than the search process. A good trade-off between quality and run-time is achieved by PHSA_ring with the *RC3* executed using 8 or 16 HSA solvers to solve complex SNDP cases, answering affirmatively the *RQ3*.

Until now, a predefined effort is given to each algorithm to determine the best one. But from now on, a different scenario is faced in which the computational effort of the algorithms is measured to locate a preset solution. Therefore, the same target solution quality is used as the stop criterion in all algorithms for any given case study. The target cost value depends on the SN case, and their best-known costs are selected for each one as the stopping value for the PHSA_ring variants, to offer a nonbiased scenario produced by an ad-hoc value. This allows to measure the PHSA efficiency through the speedup values, shown in Fig. 2, which are close to linear, showing the efficiency of PHSA to solve the SNDP and answering affirmatively *RQ1*.

6. Conclusions

In this work, we have presented a Parallel Hybrid Simulated Annealing (PHSA) to optimize the design of sensor networks in chemical plants, facilitating the decision-making by providing comprehensive knowledge of the process's current state. The basic underlying search model of the devised PHSA_ring is a set of HSA solvers connected in a ring topology. The availability of

different HSA solvers by cooperating in the process search allows obtaining a powerful parallel algorithm that leads to accurate solutions.

The objective behind the experimentation carried out using sensor networks of different characteristics was to answer our research questions. The objective of *RQ1* was to study the SNDP solution quality. We observed that our PHSA solver found the best-known solution costs. As a consequence, a model based on cooperative self-contained HSA solvers at the algorithmic level resulted in an excellent option to solve this problem. The *RQ2* suggested an analysis of the effect of several parametric configurations on the solution quality. In this sense, different statistical tests proved that the *RC3* allowed significantly enhancing the solution quality. Finally, the *RQ3* led us to assess the PHSA efficiency, considering hit rates and speedup measures. Hence, we confirmed that PHSA held the scalability property because its performance was improved by increasing the number of HSA solvers and the speedup values were close to linear.

Future work will address the resolution of SNDP, where the plant operation model is represented by a set of nonlinear equations. This introduces an additional difficulty to the problem since the feasibility of the solutions is determined by a simulation procedure that involves solving nonlinear optimization problems at a second level. In this scenario, a parallelization strategy turns out to be one of the most promising solving tools.

References

- E. Alba, 2002. Parallel evolutionary algorithms can achieve super-linear performance. *Inf. Process. Lett.* 82, 7–13.
- M. Bagajewicz, M. Sanchez, 2000. Reallocation and upgrade of instrumentation in process plants. *Computers & Chemical Engineering* 24 (8), 1945–1959.
- M. Carnero, J. L. Hernández, M. Sánchez, 2018. Optimal sensor location in chemical plants using the estimation of distribution algorithms. *Industrial & Engineering Chemistry Research* 57 (36), 12149–12164.
- P. Christofides, J. Davis, N. El-Farra, D. Clark, H. K., J. Gipson, 2007. Smart plant operations: Vision, progress and challenges. *AIChE J* 53, 2734–2741.
- V. A. Cicirello, 2017. Variable annealing length and parallelism in simulated annealing. *CoRR* abs/1709.02877.
- J. Downs, E. Vogel, 1993. A plant-wide industrial process control problem. *Comp. Chem. Eng.* 17, 245–255.
- Y.-J. He, Z.-F. Ma, 2014. Optimal design of linear sensor networks for process plants: A multi-objective ant colony optimization approach. *Chemometrics and Intelligent Laboratory Systems* 135, 37–47.
- J. Hernandez, C. Salto, G. Minetti, M. Carnero, M. C. Sanchez, May 2019. Hybrid simulated annealing for optimal cost instrumentation in chemical plants. *Chemical Engineering Transactions* 74, 709–714.
- J. L. Hernández, C. Salto, G. F. Minetti, M. Carnero, C. Bermúdez, M. Sánchez, 2020. Tuning a hybrid sa based algorithm applied to optimal sensor network design. *Journal of Computer Science & Technology* 20 (1).
- S. Kirkpatrick, C. G. Jr, M. Vecchi, 1983. Optimization by simulated annealing. *Science* (220), 671–680.
- P. Kotecha, M. Bhusan, R. Gudi, M. Keshari, 2008. A duality based framework for integrating reliability and precision for sensor network design. *Journal of Process Control* 18 (2), 189–201.
- S. Narasimhan, C. Jordache, 2000. *Data Reconciliation and Gross Error Detection*. Gulf Publishing Company.
- D. Nguyen, M. Bagajewicz, 2011. New efficient breadth-first/level traversal tree search method for the design and upgrade of sensor networks. *AIChE Journal* 57 (5), 1302–1309.
- Á. Panizo, G. Bello-Orgaz, M. Carnero, J. Hernández, M. Sánchez, D. Camacho, 2018. An artificial bee colony algorithm for optimizing the design of sensor networks. In: *Intelligent Data Engineering and Automated Learning – IDEAL 2018*. Springer International Publishing, Cham, pp. 316–324.
- P. Paul, D. Bhattacharyya, R. Turton, S. E. Zitney, 2015. Sensor network design for maximizing process efficiency: an algorithm and its application. *AIChE Journal* 61 (2), 464–476.
- J. Romagnoli, M. Sanchez, 2000. *Data Processing and Reconciliation for Chemical Process Operations*. Academic Press: San Diego, CA.
- G. Sambito, M. and Freni, 2021. Strategies for improving optimal positioning of quality sensors in urban drainage systems for non-conservative contaminants. *Water*. 13, 1–14.
- P. Sen, K. Sen, U. M. Diwekar, 2016. A multi-objective optimization approach to optimal sensor location problem in IGCC power plants. *Applied Energy* 181, 527–539.
- E. Talbi, 2009. *Metaheuristics: From Design to Implementation*. Wiley Publishing.
- J. Zhang, D. J. Chmielewski, 2017. Profit-based sensor network design using the generalized benders decomposition. In: *2017 American Control Conference (ACC)*. pp. 3894–3899.

A sparse polynomial surrogate model for the shrinking core model in phosphate ore digestion

Sanae Elmisaoui^{a,b,c}, Saad Benjelloun^a, Abdellah Chkifa^a and Abderrazak M. Latifi^{a,b,*}

^a*Mohammed VI Polytechnic University, 43150, Benguerir, Morocco*

^b*LRGP, CNRS-ENSIC, Université de Lorraine, 54000, Nancy, France*

^c*L3GIE, Mohammed V University in Rabat, Mohammadia School of Engineers, Rabat, Morocco*
abderrazak.latifi@univ-lorraine.fr

Abstract

In this paper, a surrogate model is developed from a first-principles model constructed for phosphate ore digestion with a phosphoric acid solution. This model is based on sparse multivariate polynomial interpolation. The main motivation is to reduce the computational time of the first principles model while preserving its properties, namely the monotonicity and positivity of the outputs. The temporal profiles of the concentrations of the different components involved in the liquid phase and in the ore particles, the particle radius and the thickness of the liquid film surrounding the particles are estimated by the developed surrogate model. The inputs to the model are the particle size distribution, the initial acid concentration and the hydrodynamic conditions. A design of experiments method is used to generate the sample points required for the surrogate model and several simulations are performed in the MATLAB environment. Comparison of the predictions of the surrogate model with those obtained by the first-principles model demonstrates the high performance (accuracy and computation time) of the developed surrogate model.

Keywords: Digestion, Shrinking core model, Surrogate model, Sparse polynomial interpolation.

1. Introduction

In process optimization, the mathematical formulation of the problems is very often based on a model described by momentum, heat and mass balance equations as well as thermodynamic and kinetic equations. This first-principles model must be identified and validated with experimental measurements in order to accurately predict the performance of the process to be optimized. However, the process model usually involves many ODEs and PDEs that take time to integrate, making the optimization algorithm very slow when it converges. Therefore, to overcome these problems, high-fidelity surrogate models that compute the required outputs from the inputs faster than the phenomenological model are needed.

In this paper, the objective is to develop an accurate surrogate model from a first-principles model of phosphate ore dissolution in the digestion tank of an industrial phosphoric acid production process (Elmisaoui et al. (2021b)). More precisely, the model allows us to calculate the time-varying profiles of (i) the concentrations of the different components involved in the liquid phase and in the ore particles, (ii) the radius of the particles, (iii) and the thickness of the liquid film surrounding the ore particles. The spatio-temporal profiles of the concentrations in the liquid film are also calculated from the main inputs of the model, i.e., particle size, initial acid concentration and hydrodynamic conditions. The approach used to develop the surrogate model is based on

sparse multivariate polynomial interpolation (Rabhi et al. (2018)) generalized to approximate the variables of interest with vectors of time-dependent values.

2. First-principles modelling of phosphate ore digestion

In order to evaluate the performance of the surrogate modelling approach, the phosphate ore dissolution model previously developed in (Elmisaoui et al. (2021a)) is used. From an industrial perspective, two main phenomena are involved in the digestion tank of the phosphoric acid process: the dissolution of the phosphate ore particles, and the crystallization of gypsum. The last phenomenon is not considered since only pure tri-calcium phosphate (TCP) particles are used. They are attacked with a phosphoric acid solution to produce mono-calcium phosphate (MCP) according to the following reaction:



The conversion rate of the reaction depends on the operating conditions, namely the temperature of the reaction medium, the initial concentration of phosphoric acid, the porosity and size of the phosphate ore particles, the solid/liquid ratio, the hydrodynamic conditions, and the residence time in the digestion tank. In this paper, we set three target parameters: particle size, initial phosphoric acid concentration, and tank stirring speed. It should be noted that a shrinking core model is adopted to describe the dissolution phenomenon where three phases are involved, i.e., the liquid bulk, the liquid film surrounding the particles and the solid phase (Elmisaoui et al. (2021a)). The model is based on the following assumptions: (i) diffusional limitation of the MCP product, (ii) spherical shape of the TCP particles, (iii) reaction occurs on the surface of the particles, (iv) only diffusion takes place in the liquid film. The model equations are presented in Table 1.

Table 1: System of equations describing the dissolution model

Solid phase	$\frac{dX}{dt} = \frac{3M_{TCP}D_{MCP}}{x_{TCP}\rho_s R_0 \delta} (1-X)^{\frac{2}{3}} (C_{MCP} _{r=R} - C'_{MCP})$	
Liquid film	$\frac{\partial C_i}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(D_i r^2 \frac{\partial C_i}{\partial r} \right); i = H_3PO_4, MCP$	
	$r = R$	$-D_i \frac{\partial C_i}{\partial r} \Big _{r=R} = v_i k_r C_{H_3PO_4} \Big _{r=R}$
	$r = R + \delta$	$C_i \Big _{r=R+\delta} = C'_i$
	Film thickness δ	$\delta = R \left[1 + \alpha \left(\frac{R}{R_0} \right)^{8/9} D_{MCP}^{-1/3} \right]^{-1}$
Liquid bulk	$C'_{MCP} = \frac{3n_{TCP}^0}{V_L}$ and $C'_{H_3PO_4} = C_{H_3PO_4}^0 - \frac{4n_{TCP}^0}{V_L}$	

m_s and R are the mass and the radius of phosphate ore particles, respectively. x_{TCP} is the mass fraction of TCP in the phosphate ore used, M_{TCP} is the molecular weight of TCP, $C_{MCP}|_{r=R}$ and C'_{MCP} are the concentrations of MCP at the solid surface and in the liquid bulk, respectively. k_r is the rate constant, D_i and v_i are the diffusion coefficient and the stoichiometric coefficient of component i , respectively. n_{TCP}^0 is the initial number of moles of TCP in the solid, V_L is the liquid volume in the tank, α is a hydrodynamic parameter calculated from the stirring speed of the tank. Several simulations are performed under the same operating conditions as those used in (Sinirkaya et al. (2010)). Assuming a uniform particle size distribution, the surrogate model will express the conversion rate $X(t)$ as a function of the target variables, namely the initial radius of the phosphate ore particles R_0 , the hydrodynamic parameter α , and the initial concentration of the phosphoric acid solution C_0 as detailed in the next section.

3. Surrogate modelling framework

Black box (non-intrusive) methodology of surrogate modeling consists in performing as few simulations of the system as possible, at well-chosen instances of the three input parameters, then construct a high-fidelity response surface for the outputs. This is done here by means of hierarchical multivariate polynomial interpolation. For example, we denote $\tilde{X}(t)$ (a function of R_0 , α , and C_0) the response surface for the conversion rate output $X(t)$. Basically, $\tilde{X}(t)$ is a linear combination of $R_0^{p_1} \alpha^{p_2} C_0^{p_3}$ for prescribed exponents p_1, p_2, p_3 of small values $0, 1, 2, \dots$. The process we used is adaptive because the exploration of the parameters space is not predefined in advance. We describe it in its generality for d input parameters in $[-1, 1]^d$. Application to general domains is a simple matter of scaling.

We let $S = (s_j)_{j \geq 0}$ be abscissas in $[-1, 1]$ defined by the following recursion:

$$s_{2j-1} = \sqrt{\frac{s_j + 1}{2}}, \quad s_{2j} = -s_{2j-1}, \quad j \geq 2. \tag{2}$$

and $(s_0, s_1, s_2) = (1, -1, 0)$. The sequence S is called an \mathfrak{R} -Leja sequence (Chkifa et al. (2014)). In particular, it satisfies $2s_{2j-1}^2 = s_j$ for all $j \geq 2$. Thanks to this "binary" feature, it is highly relevant for hierarchical polynomial interpolation. Given f continuous on $[-1, 1]$ that we can query on s_0, s_1, \dots one query at a time, we can approximate f using interpolation polynomials $I_0[f], I_1[f], \dots$, with

$$I_n[f] = \sum_{k=0}^n d_k[f] W_k, \quad d_k[f] := \sum_{j=0}^k \tau_{k,j} f(s_j), \tag{3}$$

where W_k are Newton's polynomials associated to S properly normalized, $d_k[f] \in \mathbb{R}$ are the associated Newton coefficients, and $\tau_{k,j} \in \mathbb{R}$ are the so-called barycentric weights, all defined below. Polynomial $I_n[f]$ is the unique polynomial of degree less than n interpolating f at s_0, \dots, s_n . We note that this interpolation scheme is hierarchical, i.e., going from $I_n[f]$ to $I_{n+1}[f]$ requires only the additive update $d_{n+1}[f] W_{n+1}$, which involves a new query of f at s_{n+1} . For instance, $I_0[f] = f(1)$, $I_1[f](x) = f(1) + (f(1) - f(-1))(x - 1)/2$, $I_2[f](x) = I_1[f](x) + (f(0) - I_1[f](0))(1 - x^2)$. It is easily established that this scheme is numerically stable and insensitive machine precision.

Newton polynomials W_k are defined by $W_0 = 1$ and

$$W_k(x) := \prod_{j=0}^{k-1} [2(x - s_j)], \quad k \geq 1, \tag{4}$$

The double sequence of barycentric weights $(\tau_{k,j})_{0 \leq k, 0 \leq j \leq k}$ is defined by

$$\tau_{k,j} = \frac{2}{W'_{k+1}(s_j)}, \quad k \geq 0, \quad j = 0, \dots, k, \tag{5}$$

where W'_{k+1} is the derivative of W_{k+1} . Such polynomials and weights can be easily computed by recurrence (see Appendix).

The generalization of the interpolation scheme to a multivariate setting is straightforward. We simply use tensor product constructions. More specifically, we are given $f : [-1, 1]^d$ continuous which can be queried on any x in $[-1, 1]^d$, and then we approximate using $I_\Lambda[f]$, with

$$I_\Lambda[f] = \sum_{k \in \Lambda} d_k[f] W_k, \quad d_k[f] := \sum_{j \leq k} \tau_{k,j} f(s_j), \tag{6}$$

where $\Lambda \subset \mathbb{N}^d$ are sets of indices. The interpolation nodes s_j , the Newton polynomials W_k , and the barycentric weights $\tau_{k,j}$ are now indexed in \mathbb{N}^d and defined by:

$$\begin{aligned} s_j &:= (s_{j_1}, \dots, s_{j_d}) & j &= (j_1, \dots, j_d) \\ W_k(x) &:= W_{k_1}(x_1) \times \dots \times W_{k_d}(x_d) , & k &= (k_1, \dots, k_d) , \\ \tau_{k,j} &:= \tau_{k_1,j_1} \times \dots \times \tau_{k_d,j_d} & x &= (x_1, \dots, x_d) \end{aligned} \quad (7)$$

where $j, k \in \mathbb{N}^d$ and $x \in \mathbb{R}^d$. The nodes s_j are in $[-1, 1]^d$, the polynomials W_k are multivariate, and for $\tau_{k,j}$, it is assumed that $j \leq k$ coordinate-wise.

The multivariate polynomial $I_\Lambda[f]$ is the unique polynomial of the form $\sum_{k \in \Lambda} a_k x^k$ (with $a_k \in \mathbb{R}$ and $x^k := x_1^{k_1} \dots x_d^{k_d}$) interpolating f on the grid $\{s_j\}_{j \in \Lambda}$. Only one condition has to be assumed on Λ , it is *lower*, i.e. if $k \in \Lambda$ and $j \leq k$ then $j \in \Lambda$ (Figure 1). The interpolation scheme is hierarchical, the change from $I_\Lambda[f]$ to $I_{\Lambda \cup \{l\}}[f]$ only requires the additive update $d_l[f] W_l$ which involves querying f at the new interpolation node s_l . Of course, we assume that the lower structure of Λ is preserved by considering only eligible $l \notin \Lambda$. The design of experiments and the response surface are entirely determined by Λ that, in practice, we construct one multi-index k at a time. We start with $\Lambda = \{\mathbf{0}\}$, $d_0[f] = f(1, \dots, 1)$ (we recall that $s_0 = 1$) and $I_\Lambda[f] = f(1, \dots, 1)$, then enrich Λ by exploring the eligible indices l and adding the most appropriate one.

Eventually, the $I_\Lambda[f]$ proxy of f will consist on a list of values $\{d_k[f]\}_{k \in \Lambda}$ associated to keys k in Λ . Any query $f(x)$ of f is simply replaced by the proxy $I_\Lambda[f](x) = \sum_{k \in \Lambda} d_k[f] W_k(x)$.

This scheme is fast and easy to implement for the following reasons: (i) the $\tau_{k,j}$ weights (see Eq. (5)) can be computed in an offline step, (ii) the weights $\tau_{k,j}$ are fast to compute once $\tau_{k,j}$ are known, (iii) the queries $f(s_l)$ of f are computed as needed, then stored, (iv) the coefficients $d_l[f]$ are computed as needed, then stored, (v) the Newton polynomials are fast to compute for any x . The generalization to vector-valued functions f is straightforward. Given $f : x \in [-1, 1]^d \rightarrow \mathbb{R}^n$, the analysis and formulas above are unchanged, the only difference being that $d_k[f]$ is a vector in \mathbb{R}^n and no longer a scalar.

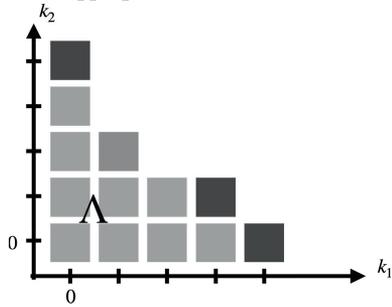


Figure 1: Schematic representation of the interpolation algorithm

4. Application of sparse polynomial interpolation in dissolution modelling

To build the surrogate model, sparse polynomial interpolation is implemented within the algorithm developed by (Chkifa et al. (2014)). Having previously fixed the bounds of the parameters, which constitute the domain of the surrogate model, the first step corresponds to the generation of the sampling points. These are chosen to be sufficiently distributed by means of a design of experiment (*DoE*) approach. Thus, the sparse polynomial interpolation model is fitted to the first-principles model.

To ensure the accuracy of the surrogate model, a comparison between its predictions and those produced by the first-principles model is performed using the root mean square error (RMSE). The rest of the construction of the surrogate model depends heavily on the accuracy of the validation. Indeed, if the accuracy of the surrogate model is good enough, then we can stop the procedure. Otherwise, an increase in the number of sample points is necessary, and the previous steps must be repeated. It is noteworthy that sampling points are generated and distributed by a max-min approach, while the min and max values correspond to the limits set previously (Quirante et al. (2015)).

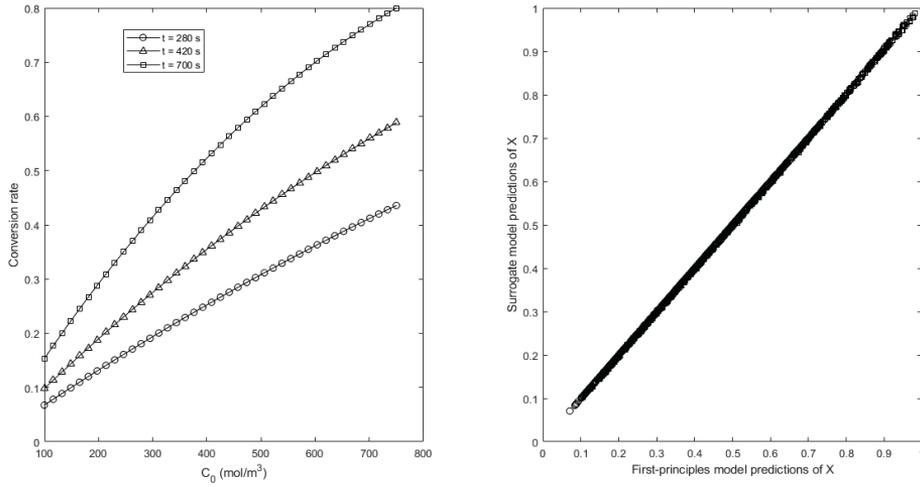


Figure 2: (a) Conversion rate versus initial concentrations at different dissolution times, (b) Comparison of the models' predictions

5. Results and discussions

To use the *DoE* to generate the sampling points, the input data intervals R_0, C_0 and α were defined, with respect to the operating conditions as follows: $[R_{min} = 500, R_{max} = 1000]m$, $[C_{0min} = 100, C_{0max} = 750]mol/m^3$ and $[\alpha_{min} = 2.7, \alpha_{max} = 4.3]$. The values of the target variables are then converted to be between -1 and 1, with respect to the ranges set previously. Thus, the surrogate model developed to describe the conversion rate can be expressed in a $[-1, 1]^3$ domain as $X(R_0, C_0, \alpha) = f(\xi_{R_{min}, R_{max}}(R_0), \xi_{C_{0min}, C_{0max}}(C_0), \xi_{\alpha_{min}, \alpha_{max}}(\alpha))$, where ξ is the affine function used to reduce the variable intervals to the interval $[-1, 1]$, i.e., $\xi_{a,b}(t) = -1 + 2(t - a)/(b - a)$. The MATLAB environment is used to construct the polynomials of the substitution model. Figure 2 (a) shows the conversion rate profiles at various initial concentrations and different dissolution times predicted by the surrogate model using ten sampling points. The values of the remaining target variables are taken as the mean values of the considered intervals, i.e., $\alpha = 3.9225$, and $R_0 = 750\mu m$. Figure 2(b) shows the comparison between the first-principles predictions and those of the surrogate models.

It is found that the surrogate model predicts the conversion rate very well at various initial concentrations and time points. Moreover, the predicted conversion rate increases with the initial concentration of phosphoric acid, and as the dissolution mechanism progresses, the conversion rate also increases. Furthermore, the average error between predicted values is negligible, demonstrating that the surrogate model accurately captures the behavior of the first-principles model.

To assess the accuracy of the surrogate model, the RMSE is calculated for different numbers of sampling points generated and used in the interpolation. It is worth noticing that for the three target variables considered, the RMSE values decrease with increasing number of sampling points. A high precision of the model can therefore be easily obtained (i.e., large number of sampling points) while guaranteeing a very low computation time. The latter is about 13 times lower for the surrogate model with ten sampling points than for the first-principles model. This ratio can be even higher when all side reactions and components involved in the digestion of ore particles are taken into account.

Table 2: Values of RMSE for the conversion rate

Sampling points	R_0	C_0	α
5	2.0110^{-2}	1.0410^{-2}	6.8810^{-5}
10	2.4510^{-4}	5.4210^{-5}	6.7310^{-5}
100	1.9510^{-4}	2.4410^{-5}	4.8710^{-5}

6. Conclusions

In this work, the development of a surrogate model for a phosphate ore digestion model was presented and simulated in the MATLAB environment. The high-dimensional adaptive hierarchical polynomial interpolation used demonstrated its reliability and efficiency in predicting the target variables, namely, the phosphate ore conversion rate, the concentration of reactants in the liquid phase and in the liquid film surrounding the particles. This modelling approach was able to correctly capture the behavior of the dissolution mechanism with high accuracy and a significant reduction in computational time, demonstrating the power of the surrogate modelling approach. Future work will focus on the development of surrogate models for the overall digestion tank model (coupling of dissolution and crystallization models) and their use in real-time optimization of phosphoric acid process performance.

Appendix : Newton polynomials and barycentric weights

Intrinsically, computing sequence S (see Eq.(2)) is straightforward, $S = (1, -1, 0, \cos(\pi/4), -\cos(\pi/4), \cos(\pi/8), -\cos(\pi/8), \cos(3\pi/8), -\cos(\pi/8), \dots)$. Evaluating the Newton polynomials $W_k(x)$ (see Eq.(4)) and computing barycentric weights $\tau_{k,j}$ (see Eq.(5)) is fast (logarithmic complexity in k). We have $W_1(x) = 2(x-1)$, $W_2(x) = 4(x^2-1)$ and $W_3(x) = 8(x^2-1)x$. In general, we use $W_{2N-1}(x) = W_N(2x^2-1)/(2x)$ and $W_{2N}(x) = 2(x-s_{2N-1})W_{2N-1}(x)$ for any $N \geq 2$. Then, by simple derivation rules, it can be verified that $\tau_{0,0} = 1$, $(\tau_{1,0}, \tau_{1,1}) = (1/4, -1/4)$ and $(\tau_{2,0}, \tau_{2,1}, \tau_{2,2}) = (1/8, 1/8, -1/4)$. Then given $N \geq 2$,

- for $k = 2N - 1$,

$$\left. \begin{array}{l} \tau_{k,0} = \tau_{N,0}(s_k + 1) \\ \tau_{k,1} = \tau_{N,0}(s_k - 1) \\ \tau_{k,2} = \tau_{N,1} 2s_k \end{array} \right| \begin{array}{l} \tau_{k,2j-1} = \tau_{N,j}(s_k - s_{2j}) \\ \tau_{l,2j} = \tau_{N,j}(s_k + s_{2j}) \\ j = 2, \dots, N \end{array} \quad (8)$$

- for $k = 2N$,

$$\left. \begin{array}{l} \tau_{k,0} = \tau_{N,0}/2 \\ \tau_{k,1} = \tau_{N,0}/2 \\ \tau_{k,2} = \tau_{N,1} \end{array} \right| \begin{array}{l} \tau_{k,2j-1} = \tau_{N,j}/2 \\ \tau_{k,2j} = \tau_{N,j}/2 \\ j = 2, \dots, N \end{array} \quad (9)$$

References

- A. Chkifa, A. Cohen, C. Schwab, 2014. High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs. *Foundation of Computational Mathematics* 14, 601 – 633.
- S. Elmisaoui, A. M. Latifi, L. Khamar, M. Salouhi, 2021a. Analysis of the dissolution mechanism in the phosphoric acid manufacturing process: modelling and simulation. *Computer Aided Chemical Engineering* 50, 891–897.
- S. Elmisaoui, A. M. Latifi, L. Khamar, M. Salouhi, 2021b. Shrinking core approach in the modelling and simulation of phosphate ore acidulation. *Chemical Engineering Transactions* 86, 871–876.
- N. Quirante, J. Javaloyes, R. Ruiz-Femenia, J. A. Caballero, 2015. Optimization of chemical processes using surrogate models based on a kriging interpolation. *Computer Aided Chemical Engineering* 37, 179–184.
- A. Rabhi, A. Chkifa, S. Benjelloun, A. Latifi, 2018. Surrogate-based modeling in flotation processes. *Computer Aided Chemical Engineering* 43, 229–234.
- M. Sinirkaya, A. Özer, M. Gulaboglu, 2010. Kinetics of dissolution of mardin-mazidagi (Turkey) phosphate ore in dilute phosphoric Acid Solutions. *Society for Mining, Metallurgy & Exploration* 27(2), 110–115.

On the use of data engineering and machine learning in global optimization applications with cutting plane approximations

Vasiliki Deligianni^a, Asimina Marousi^b, Apostolos Chalkis^c and Antonis Kokossis^a

^a*Department of Chemical Engineering, National Technical University of Athens, A Heron Politechniou 9, Zografou, Athens GR 15 780, Greece*

^b*Department of Chemical Engineering, Centre for Process Systems Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom*

^c*Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens 157 72, Greece*
akokossis@mail.ntua.gr

Abstract

This work explores data analytics in the development of optimization methodology for global optimization, as applied through decomposition methods and cutting plane algorithms. Cutting planes are treated as data populations, generated at each iteration, population elements are renewed based on the incumbent solution. The current contribution explores qualitative aspects studied in the previous, essentially attempting to expand the affinity norm to temporal sets of data in full and low dimensional spaces. The separation problem is examined using clustering techniques and is tested against a library of quadratic and box constrained optimization problems, that feature varying sparsity and density patterns. The affinity metric was formed, to efficiently evaluate overlapping cutting planes, noting significant improvement in performance. In continuation of these results, normal vector clustering examines the direction of the hyperplanes, by utilizing the cosine similarity of the normal vectors. Temporal data approach aims to prevent repetitions of chosen sub spaces within rounds. Temporal data outperformed the affinity metric approach on the largest problem tested, for tight elimination criteria. Normal vector clustering accelerated the algorithm beyond previous work, in the first round, but failed to further close the duality gap. Overall, analytics are found to dramatically improve the duality gap and the quality of the solution, consistently in all the problems tested. In conclusion, the geometrical interpretation of the dual space holds the most promising lines for future work.

Keywords: global optimization, cutting planes, cut selection, data analytics

1. Introduction

Convergence to the global optimum in non-convex quadratic programming, at a large scale is an NP-hard problem, and an ongoing challenge in the field of optimization, with various applications. State-of-the-art global solvers employ the Branch&Cut algorithm, with additional classes of computationally light relaxations. Polyhedral relaxations are commonly fused with strong quadratic ones, such as semidefinite and co-positive relaxations. Innovations by Baltean-Lugojan et al. (2019) focus on a generic and effective outer approximation method, suitable for semidefinite relaxations. Their work tackles the computational complexity of the semidefinite relaxation, through the use of low dimensional cutting planes, while maintaining tight bounds. The key element

of the referenced work is the decomposition of high-dimensional cutting planes into their low-dimensional counterparts, taking advantage of the problem's sparsity pattern. This results in a combinatorial explosion of available cutting planes for the separation problem and thus novel selection measures were introduced, based on feasibility violation and improvement of the objective function.

2. Problem description and mathematical formation

The optimization problem consists of a non-convex quadratic boxed function and its linear constraints.

$$z_{qp} = \min_x \{x^T Q x + c^T x \mid Ax \leq b, x \in [0, 1]^N\} \quad (1)$$

where N -variable vector $x, A \in R^{p \times N}$ and $Q \in R^{N \times N}$ represents an indefinite symmetric matrix. Relaxations and reformulation techniques are employed, to tackle Eq(1), as proposed in Sherali and Fraticelli (2002). A new symmetric matrix $X_{ij} \forall i, j$ (lifted variable) is formed to replace the quadratic term $x_i x_j$. The lifted variable is defined as $X = x x^T$ and the quadratic term transforms to $Q \bullet X = \text{Tr}(Q^T X) = \sum_{i,j} Q_{ij} X_{ij}$. Then z_{qp} is lower bounded by,

$$z_{qp}(B) := \min_{x, X} \{Q \bullet X + c^T x \mid Ax \leq b, x \in [0, 1]^N \text{ and } (x, X) \in B\} \quad (2)$$

parametric on any convex set B , which includes all sets of (x, X) , that add valid constraints to the $z_{qp}(B)$ relaxed quadratic problem. The positive semidefinite constraint (PSD), used to create new cuts, derives from the semidefinite relaxation (SDP) $X = x x^T$ to $X \geq x x^T$, or equivalently $\begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \geq 0$, Sherali and Fraticelli (2002); Qualizza et al. (2012). The reformulation-linearization technique (RLT) compliments the SDP relaxation, by adding linear, triangular constraints to the optimization problem, Anstreicher (2009), for $0 \leq X_i, X_j \leq 1$. In the first round, convex set B , consists of just the RLT constraints. For (X^*, x^*) , being the initial solution of the relaxed problem, the augmented matrix $\begin{bmatrix} 1 & x^{*T} \\ x^* & X^* \end{bmatrix}$ is formed. Applying eigen-decomposition, provides violations of the feasible set. The number of feasibility violating cuts t , is equal to the negative eigenvalues of the matrix and each violation corresponds to a PSD constraint that is added to the Master problem $v_k^T \cdot \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \cdot v_k \geq 0, \forall k \in \{1, \dots, t\}$.

Baltea-Lugojan et al. (2019), introduced low-dimensional approach for lighter linear relaxations, to overcome the problem of generating only one dense cut per negative eigenvalue of full-dimensional matrix, building on prior work of Qualizza et al. (2012). With \wp denoting the power set of the vertex set $V : \{1, \dots, N\}$ and $\rho \in \wp(\rho \subseteq V)$ any arbitrary index subset, let $x_\rho \in \mathfrak{R}^{|\rho|}$ and $X_\rho \in \mathfrak{R}^{|\rho| \times |\rho|}$. For any subset of \wp the following semidefinite relaxation is introduced,

$$(\forall F \subseteq \wp) \wp(F) := \{(x, X) \mid \forall \rho \in F : \begin{bmatrix} 1 & x_\rho^T \\ x_\rho & X_\rho \end{bmatrix} \geq 0, X_{ij} \leq x_i \forall i \in \rho\} \quad (3)$$

For n , ($n = |\rho|$) being the dimension of the subsets, the cardinality of \wp_n is determined in a naive way by the combination $\binom{N}{n}$. In practice, chordal extensions are used, to exploit the density pattern of the problem, for smaller $|\wp_n|$. Still, the number of subproblems is such, that a separation issue arises, for the selection of the most informative underestimators. In feasibility strategy, cutting planes are selected based on most negative eigenvalues of the augmented matrix.

3. Metrics and analytics in selecting cutting planes

The motivation behind integration of data analytics, in the separation problem, through the affinity clustering was the notion that qualitative and geometrical aspects matter. Data technology is subsequently applied to measure fathom and screen planes, also to represent data spaces of the dual problem. Research was expanded, by testing this theory on full and low dimensional spaces with the use of normal vector clustering and temporal sets method.

3.1. Affinity metric

The initial approach to the separation problem employed data analytics, in order to select the most informative cutting planes for the reconstruction of the convex space, with encouraging results. The key element of this approach was the formation of a custom metric; the affinity metric. The rationale behind the affinity metric was to prevent the proliferation of sub-spaces with very similar or identical pattern of variables. Sub-spaces featuring common dimensions (affine sets), are considered to hold similar information and could consequently be curtailed. The initial linear constrains, used in Eq. (2), include just the RLT constrains. When the solution is reached, decomposition follows, providing low-dimensional SDP cuts to choose from. Population of cuts is then ordered by means of feasibility and cuts that satisfy the feasibility condition $\lambda_p \leq 0$ constitute the power set \wp used for the analysis. The metrics that best described the correlations among the population were found to be the *Euclidean* distance and the custom *affinity* norm $d_a(x, y)$, that is defined below. The affinity norm serves the purpose of exploring the projection space by mapping the created cuts and revealing overlaps of the sub-spaces. That is achieved with the use of common clustering technics, such as *Kmeans* and *agglomerative* clustering, for varying number of clusters. Hybrid methods were tested to explore sequential implementations of Euclidean and affinity metrics. The affinity metric is defined by,

$$d_a(x, y) := \sum_{i=1}^N 1 - g(x_i, y_i), \text{ where } g(x_i, y_i) = \begin{cases} 1 & \text{if } x_i = y_i \\ 0 & \text{if } x_i \neq y_i \end{cases} \quad (4)$$

The proposed approach achieves significant gap closures that are especially evident in problems of higher complexity.

3.2. Full-dimensional sparse cutting planes: Normal vector clustering

On a new perspective on the reconstruction of the convex space, the low dimensional projections are transferred to the full-dimensional space in sparse form. Working in the original space instead of projections of it, favors the use of dynamic sets for the recovery of unselected planes, in the following rounds. Cuts not added as SDP constrains, could potentially result in loss of valuable information for the duality gap closure. Dynamic sets enable the retrieval of such cuts, by adding part of unselected cuts to the population, prior clustering. Memory limitations prevent the application of dynamic sets in normal vector clustering. The motivation behind this method is to include diversity in orientation of the hyperplanes, in cut selection. The feature of orientation is portrayed through the formation for the normal vector α . Parallel cutting planes are considered to hold similar information and thus are clustered together, based on the cosine similarity of their normal vectors. The ranking criteria are enhanced by considering the search space each hyperplane excludes.

The SDP relaxation of the referenced work in Eq.(3) is reformulated in full dimension. Sparse solutions $x_{\rho,s} \in \mathfrak{R}^N$, $X_{\rho,s} \in \mathfrak{R}^{N \times N}$ are formed, for every subset of \wp , from vector slices $x_{\rho} \in \mathfrak{R}^{|\wp|}$ and sub-matrix slices $X_{\rho} \in \mathfrak{R}^{|\wp| \times |\wp|}$.

Sparse eigen-decomposition is applied to each sparse augmented matrix $\begin{bmatrix} 1 & x_{\rho,s}^T \\ x_{\rho,s} & X_{\rho,s} \end{bmatrix}$, generating $N + 1$ cutting planes. For feasibility violations, the normal vector of every hyperplane is formed

from the eigenvectors, according to Eq.(8). All violated cuts ($\lambda_{\rho,s} \leq 0$) are added to the population, in contradiction to prior work, where one out of $|\rho|$ cuts was chosen per subset, according to feasibility measure ($|\lambda_{\rho}|_{max}$). The normal vectors are clustered with of-the-shelf clustering algorithm measuring cosine similarity. The PSD relaxation is as follows.

$$v^T \cdot \begin{bmatrix} 1 & x_{\rho,s}^T \\ x_{\rho,s} & X_{\rho,s} \end{bmatrix} \cdot v \geq 0, \forall \rho \in F \mid (F \subseteq \wp) \quad (5)$$

where, $v \in \Re^{N+1}$. The normal vector of each hyperplane is generated from the linear form of Eq.(6).

$$\sum_{i=1}^N 2v_1 v_{i+1} \cdot x_i + \left(\sum_{i=1}^N \sum_{\substack{j=i \\ j \neq i}}^N v_{i+1}^2 + \sum_{i=1}^N \sum_{\substack{j=i \\ j \neq i}}^N 2v_{i+1} v_{j+1} \right) \cdot X_{ij} + v_1^2 \geq 0 \quad (6)$$

$$a = [2v_1 v_2 \quad 2v_1 v_3 \quad \dots \quad v_2^2 \quad v_2 v_3 \quad \dots \quad v_{N+1}^2] \quad (7)$$

Cosine similarity measures the cosine of the angle between two normal vectors in the multi-dimensional space.

$$CS(\alpha_1, \alpha_2) = \cos \theta = \frac{\alpha_1 \cdot \alpha_2}{\|\alpha_1\| \cdot \|\alpha_2\|} \quad (8)$$

The corresponding cutting planes, of the clustered normal vectors are ordered by means of excluded search space. The magnitude of the part of the convex set, that is removed by a hyperplane ($\lambda_{\rho,s} \leq 0$), is connected to the minimization of the constant term v_1^2 in Eq.(7). The first eigenvector for every ordered cluster is selected and added as a PSD constraint to the Master problem.

3.3. Temporal sets in the low-dimensional space

The rationale behind this method is to compliment cutting plane algorithms, by conveying valuable information from previous rounds to the cut selection. In affinity clustering, the selection criteria include feasibility measure, as well as geometrical features of the population. With temporal sets, "memory" of used cuts is introduced, preventing the reselection of triplets, with respect to feasibility measure. This line of work aims to limit the deceleration of the algorithm, when reaching a plateau, hence achieving further closure of the duality gap, in large, dense problems. The proposed algorithm follows these basic steps.

1. **Solution** of Master problem with just *RLT* constrains
2. **Formation, clustering and evaluation** of 3-D cuts, according to feasibility measure
3. **Scanning and removal** of repetitive triplets
4. **Selection** of top-rank element and **renewal** of list of used cuts

Integration of temporal sets in cutting plane selection algorithms was most efficient for conventional *Kmeans* clustering, for euclidean distance, with **weak** elimination criteria, **strong** tolerance and for application **after clustering**. Diverse versions of temporal set algorithms were tested, through various rounds of experiments, without noting similar results.

- *Elimination criteria*
Strong: Eliminate triplets that share 3 dimensions with previously selected 3-D cuts
Weak: Eliminate triplets that share at least 2 dimensions with previously selected 3-D cuts
- *Tolerance*
Tight: Allow the minimal number of repetitions in triplets, so as to comply with the 100 cuts per round requirement
Loose: Allow sufficient number of repetitions, for the activation of the temporal set constrains

- *Elimination step*
 - Prior clustering:** Discard used triplets from a raw population
 - After clustering:** Discard used triplets at each cluster

The intuition behind the experiment design is to test performance of key combinations on basic algorithm alterations, regarding elimination criteria, tolerance and the elimination step. Tolerance was manually adjusted to problem size and to cut round.

4. Results

The test set of box quadratic problems (BoxQP) used in the experiments was provided by Baltean-Lugojan et al. (2019). For all computational experiments the dimensionality of the cutting planes is set to $|\rho| = 3$ and the cut selection measure is feasibility. The number of iterations and cuts are dictated by literature, so as to compare the proposed approach with the referenced algorithm (Rf1). The number of rounds is set to 20, the added cuts for each round are set to 100. In coordination with Baltean-Lugojan et al. (2019), the convergence limit is achieved when rounds are 40 and selection of cuts is 5% of the available sub-problems. The computational experiments are carried in python 3.5 using cplex 12.8 python API solver and scikit-learn v0.2 (Pedregosa et al., 2011) package for k-means and agglomerative clustering. Results are evaluated based on the convergence of the algorithm and the final solution

4.1. Normal vector clustering

Normal vector clustering was tested against conventional *Kmeans* clustering and the original referenced algorithm, for feasibility measure. A hybrid version was formed, to test complementarity with conventional clustering algorithm, where 50% of the cuts were produced by each method, independently. Two examples of small non-convex QPs were studied, due to memory limitations. Evidently, a problem of $N=70$, 50% dense, consists of 6608 sub-problems, while each element of the population has a dimensionality of 70. Eigendecomposition of 71×71 augmented matrix takes place, producing 71 eigenvectors of 71 dimensionality, each corresponding to a 2555-D normal vector.

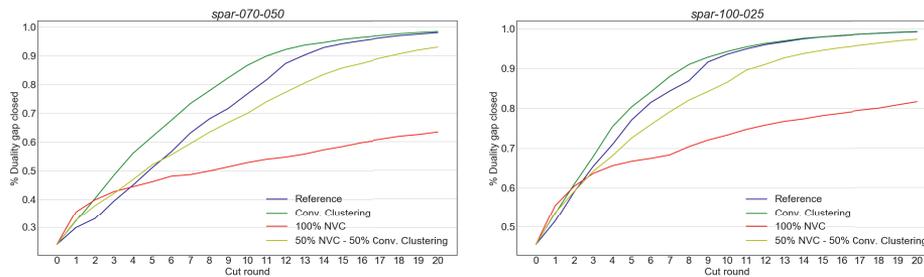


Figure 1: Progression of performance in duality gap closure, across 20 cut rounds, calculated for convergence limit of referenced algorithm. Cuts are selected for $\min(v_1^2)$ and must violate feasibility ($\lambda_{p,s} \leq 0$)

Initially, *NVC* is consistently preeminent for the test sets examined. Acceleration succeeded in the first cut round is dominated in the following rounds, resulting in a gap closure below the standards, set by previous algorithms. *Hybrid* version, is deprived of the for-mentioned edge in the first round and does not appear to complement conventional clustering. The large resource requirements, complicate further research, on *NVC* with temporal sets.

4.2. Temporal sets

Temporal sets noted the best performance, for tight elimination criteria, that loosen as the algorithm progresses, indicating the contribution of geometrical complementarity in the separation problem. Improvement was marginal on the small problems, but significant on the bigger ones, where population of violating cuts is vast and there was a significant duality gap left open from the referenced algorithm. This method appears to complement most, conventional clustering with euclidean metric, among similar algorithms tested.

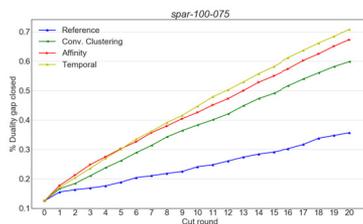


Figure 3: Comparison of % gap closure across 20 cut rounds, for largest test set (N=100, 75% dense)

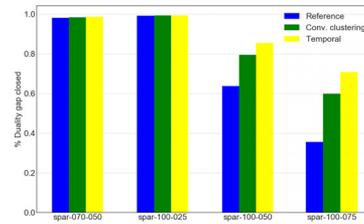


Figure 2: Comparison of % gap closure across different test problems

Performance is enhanced in large sets of sub-problems, where tight elimination criteria can be applied, without exhaustion of available cuts. Temporal sets were also applied to Affinity clustering, without noting comparable results, possibly due to similar approach to the separation problem. The progress made by the utilization of data analytics is most evident for the largest test problem in *Figure 3*. Affinity clustering starts off comparably better, but is dominated by the temporal sets method after cut round 5.

5. Conclusion

Temporal approach on low-dimensional spaces appears to benefit acceleration of bigger and denser problems. In the largest test problem, the temporal approach outperformed Affinity clustering. Algorithms with tighter elimination conditions performed the best, indicating the importance of geometrical diversity over feasibility measure on cut selection. Computational time improves in comparison to conventional clustering algorithm, but still exceeds that of the referenced algorithm. Normal Vector Clustering did not improve the overall performance of the algorithm. From a theoretical point of view, a consistently significant acceleration in the first cut round could potentially benefit a hybrid algorithm. Memory limitations prevent application in bigger and denser problems, as well as further efforts to incorporate dynamic methods. Computational time and resource requirements are restricting.

6. Acknowledgments

Prof. Kokossis is indebted to Prof Kevrekidis at Johns Hopkins University for his continuing encouragement to explore data analytics and machine learning in decomposition algorithms. He is also grateful to Prof Ruth Misener at Imperial College London who shared benchmark problems and her relaxed (sparse and dense) cutting plane approximations used in the experiments.

References

- K. M. Anstreicher, Nov. 2008. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization* 43, 471–484.
- R. Baltean-Lugojan, P. Bonami, R. Misener, A. Tramontani, 2019. Scoring positive semidefinite cutting planes for quadratic optimization via trained neural networks.
- A. Qualizza, P. Belotti, F. Margot, Nov. 2011. Linear programming relaxations of quadratically constrained quadratic programs. In: *Mixed Integer Nonlinear Programming*. Springer New York.
- H. D. Sherali, B. M. P. Fraticelli, 2002. Enhancing rlt relaxations via a new class of semidefinite cuts. *Journal of Global Optimization* 22 (1/4), 233–261.

A systematic approach for the processing of experimental data from anaerobic syngas fermentations

Eduardo Almeida Benalcázar ^{a,b}, Henk Noorman ^{b,c}, Rubens Maciel Filho ^a,
John Posada ^b

^a *Department of Product and Process Development, Faculty of Chemical Engineering, State University of Campinas, Av. Albert Einstein 500, 13083-852, Campinas – SP, Brazil.*

^b *Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, the Netherlands.*

^c *DSM Biotechnology Center, A. Fleminglaan 1, 2613 AX, Delft, the Netherlands.*

Abstract

This study describes a methodological framework designed for the systematic processing of experimental syngas fermentation data for its use by metabolic models at pseudo-steady state and at transient state. The developed approach allows the use of not only own experimental data but also from experiments reported in literature which employ a wide range of gas feed compositions (from pure CO to a mixture between H₂ and CO₂), different pH values, two different bacterial strains and bioreactor configurations (stirred tanks and bubble columns).

The developed data processing framework includes *i*) the smoothing of time-dependent concentrations data (using moving averages and statistical methods that reduce the relevance of outliers), *ii*) the reconciliation of net conversion rates such that mass balances are satisfied from a black-box perspective (using minimizations), and *iii*) the estimation of dissolved concentrations of the syngas components (CO, H₂ and CO₂) in the fermentation broth (using mass transfer models). Special care has been given such that the framework allows the estimation of missing or unreported net conversion data and metabolite concentrations at the intra or extracellular spaces (considering that there is availability of at least two replicate experiments) through the use of approximative kinetic equations.

Keywords: Syngas fermentation, experimental data processing, fermentation data reconstruction, data reconciliation.

1. Introduction

The development of a new fermentation technology starts at the laboratory where the chemostat is one of the main tools for studying microbial behavior. The chemostat is a fermentation setup where the concentrations of substances and cells are maintained nearly constant, and therefore a pseudo-steady state is reached at the intra and extracellular spaces (Noorman et al., 1996; Villadsen et al., 2011). The chemostat serves for gaining insights on the steady-state regulations of, on one hand, the whole metabolism such as the growth rate dependencies on substrates and products concentrations or the substrate requirements for maintenance, and on the other hand, the intracellular reaction rates that

allow cells to exhibit their behavior expressed as net conversion rates, or q-rates, as well as through the rates of growth and decay.

The measurements of substance concentrations is the main source of information about the state of a fermentation process. The net conversion rates are estimated from these measurements. The substances that are generally measured are the carbon and the electron sources, the excreted products, cells, oxygen (in aerobic fermentations), CO_2 , H^+ ions and the nitrogen source, (Stephanopoulos and Tsiveriotis, 1989). Yet, measurements are commonly prone to errors, both random and systematic (van der Heijden et al., 1994b). The reconciliation process is an adjustment of the measured rates aimed at improving their accuracy, such that they fulfill constraints formulated by, for instance a black-box description, which contains the energetic as well as the elemental balances of carbon, hydrogen, oxygen, nitrogen and charge (Noorman et al., 1996, 1991; van der Heijden et al., 1994a). In underdetermined systems, the reconciliation may allow the estimation of unmeasured quantities, whereas in overdetermined systems, the reconciliation allows to assess the consistency of the collected data and the identification of gross measurement errors (Stephanopoulos and Tsiveriotis, 1989; van der Heijden et al., 1994b, 1994a).

The development and up-scaling of the anaerobic syngas fermentor requires the design of mathematical models that are able to reproduce the *i*) characteristics of mass transfer from the gas to the liquid and the *ii*) microbial metabolic responses to the stimulus provided by the extracellular environment. The models on the side of the microorganism are much nurtured by the adequate collection and processing of experimental information. In this document we describe a systematic methodological framework we developed to process the experimental metabolite concentrations and conversion rates from syngas fermentation experiments. We have previously used the proposed framework for assessing the performance of the reported and own fermentation experiments in terms of the distribution of carbon and electrons among the fermentation products, as well as for the parameterization of metabolic models at pseudo-steady state and transient state.

The following pages describe the methodological framework formed by: *i*) data curation and smoothing, *ii*) reconciliation of net conversion rates and *iii*) reconstruction of missing concentrations of the dissolved gases.

2. Data curation and smoothing

In general, the curation and smoothing of the experimental data is useful for the design of kinetic expressions and the early assessment of trends. The use of curated and smoothed data is however not recommended for the parameterization of models because such procedure may introduce errors or may lead to the negligence of exceptional phenomena producing outliers in the data trends.

Curation is here referred to the process of filling missing concentration data in time points where samples have been withdrawn and for any particular reason the concentration of one or more substances have not been able to be measured or the measurements fell below the equipment detection limits. In either of the two cases the empty data points may be filled using the overall trends of the data as reference. For instance, if the missing data correspond to points where low concentrations are expected, as the low points in the oscillatory data collected by (Mahamkali et al., 2020), then the missing data can be replaced by values somewhere around the detection limit of the measuring equipment. If the missing data is instead expected higher than the detection limits of the measuring equipment, their estimation may be supported on the overall trend, where spline interpolation may be far more useful than linear interpolation. The maintenance of the trends in the time derivatives may also be used as a strategy for selecting the interpolation

method. Time derivatives of experimental concentrations (C_j) may be estimated using forward, central and backward differentiation for the first, middle and last points in a time series, respectively. Table 1 shows the equations applicable to the three mentioned differentiation methods.

Table 1 Equations used for the estimation of the concentration gradients with the experimental data

Type of differentiation	Equation	nr.
Central	$\frac{dC_j}{dt} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_{t+1} - C_{t-1}}{2 \cdot \Delta t}$	(1)
Forward	$\frac{dC_j}{dt} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_{t+1} - C_t}{\Delta t}$	(2)
Backward	$\frac{dC_j}{dt} \approx \frac{\Delta C_j}{\Delta t} = \frac{C_t - C_{t-1}}{\Delta t}$	(3)

Smoothing is here referred to the reduction of the noise in the trends of dynamic experiments. The earliest smoothing method to be tried may be moving averages. However, the higher and lower points in oscillatory data, as in (Mahamkali et al., 2020), may be neglected by moving averages. As more refined alternatives, the smoothing may be done using local regression methods that use weighted linear least squares or 1st degree polynomial models that assign lower weight to outliers in the regression. The last two methods may be found coded into MatLab's 'smooth' function as the algorithms 'lowess' and 'rlowess', respectively. In any case, the selection of smoothing algorithm may be, as described for the curation of data, supported on the trends of the time derivatives and on trends from a second experiment, provided it is available.

3. Reconciliation of net conversion rates

In general terms, the reconciliation process consists on the minimization of the mismatch between the reconciled data and the raw experimental data while closing the carbon, hydrogen, oxygen, nitrogen elemental balances as well as the charge balances. Prior to describing how the minimization problem is created, it is necessary to provide an introduction on how the mass balances are structured for the fermentation of syngas.

The fermentation of syngas employs the ability of diverse types of bacteria and archaea, often called as acetogens, to harvest the electrons from CO and H₂ and the carbon from CO and CO₂. CO, H₂ and CO₂ can be the sole sources of energy and carbon for these microorganisms. Acetic acid, ethanol and 2,3-butanediol are the most common native products of acetogenic microorganisms. The first step on the processing of experimental data is the establishment of stoichiometric relations between electron and energy sources, the products and cells. Equations 4 – 9 in Table 2, show the stoichiometries of the reactions leading from CO and H₂ towards acetic acid, ethanol and 2,3-butanediol. In equations 4 – 9, the amount of electron donor required to generate each product is calculated by balancing the degree of reduction (or the amount of electron available for redox exchange (Heijnen, 2002)) between the donor and the catabolic product; CO₂ is added to balance the carbon, H₂O balances the oxygen atoms, and H⁺ ions balance the hydrogen atoms. If the balancing is made correctly, in the mentioned order of steps, the charge should also be balanced with no additional modifications to the stoichiometry.

For simplicity, here we show only apparent species. CO₂, acetic acid and potentially the nitrogen source have different forms depending on the pH. More thorough descriptions can include these species and their equilibrium relations to improve accuracy.

The complexity of the sequence of reactions leading to the production of cells, or anabolism, may also be summarized using two equations, one for each electron donor (see equations 10 and 11 in Table 1). The derivation of the stoichiometry of the anabolic reaction follows the same steps as described previously.

In theory, the catabolic reactions of any syngas fermentation process where CO and H₂ are consumed and where acetate, ethanol and 2,3-butanediol are the main products should be described by a combination between equations 4 – 9. The same is applied to the anabolic reaction, which is derived from a combination between equations 10 and 11. The resulting catabolic and anabolic reactions may be joined to form one metabolic reaction using the biomass yield, which can be estimated using thermodynamics (Heijnen and van Dijken, 1992).

Table 2. Common reactions in syngas fermentations

Part of metabolism	Reaction	nr.
Catabolism from CO as the electron donor	$-4\text{CO} - 2\text{H}_2\text{O} + \text{C}_2\text{H}_3\text{O}_2^- + 2\text{CO}_2 + \text{H}^+$	(4)
	$-6\text{CO} - 3\text{H}_2\text{O} + \text{C}_2\text{H}_6\text{O} + 4\text{CO}_2$	(5)
	$-11\text{CO} - 5\text{H}_2\text{O} + \text{C}_4\text{H}_{10}\text{O}_2 + 7\text{CO}_2$	(6)
Catabolism from H ₂ as the electron donor	$-4\text{H}_2 - 2\text{CO}_2 + \text{C}_2\text{H}_3\text{O}_2^- + 2\text{H}_2\text{O} + \text{H}^+$	(7)
	$-6\text{H}_2 - 2\text{CO}_2 + \text{C}_2\text{H}_6\text{O} + 3\text{H}_2\text{O}$	(8)
	$-11\text{H}_2 - 4\text{CO}_2 + \text{C}_4\text{H}_{10}\text{O}_2 + 6\text{H}_2\text{O}$	(9)
Anabolism	$-2\text{CO} - 0.25\text{NH}_4^+ - 0.5\text{H}_2\text{O} + \text{CH}_{1.75}\text{O}_{0.5}\text{N}_{0.25} + \text{CO}_2 + 0.25\text{H}^+$	(10)
	$-2\text{H}_2 - 0.25\text{NH}_4^+ - \text{CO}_2 + \text{CH}_{1.75}\text{O}_{0.5}\text{N}_{0.25} + 1.5\text{H}_2\text{O} + 0.25\text{H}^+$	(11)

Going back to the structuring of the minimization problem for reconciling the net conversion rates, the mismatch between the reconciled and the experimental rates is the objective function; the mismatch may be quantified by the sum of the squared differences divided by the standard deviation (Villadsen et al., 2011). The reconciled net conversion rates for CO, H₂, CO₂, acetate, ethanol, 2,3-butanediol, biomass growth, water, nitrogen source and H⁺ ions may be used as decision variables. The elemental balances may be used as constraints. Additional constraints may also be applied for improving the precision; for instance, the mass balances of the different species in the two phases of the bioreactor, the liquid and the gas phases. The experimental preference for CO and H₂ uptake and for product generation may also be used as additional constraints.

It is standard procedure to fix the allowed variations of the experimental q-rates according to the error in the experimental data (known using data from multiple experiments at the same conditions); however, if that error is not reported in the source of experimental data, one may assume a percent variation of maximum 10 – 20 % for each reconciled rate. In this case, the use of additional constraints become very helpful to guide the reconciliation to values that are consistent not only with elemental balances but also with time dependent or independent balances of species.

4. Calculation of the dissolved gas concentrations

Techniques for in-line measurement of the dissolved concentrations of CO, H₂ and CO₂ have been developed (Mann et al., 2021; Mislov et al., 2015). However, it is rare to find reports that describe the use of such techniques in syngas fermentations. The calculation of the dissolved concentrations of CO, H₂ and CO₂ is therefore often necessary. This

calculation can be done using two approaches that are different if the experimental data can be assumed at steady state or at transient state.

If the experimental data is likely at steady-state, the time gradients are assumed equal to zero and the mass balances of CO and H₂ in the gas and the liquid phases, one overall CO₂ mass balance and one summation for the composition of the off-gas (see equations 12 - 17) may be enough to estimate the unknown concentrations of CO, H₂, CO₂ in the liquid phase, as well as the molar fractions of CO and H₂ in the off-gas and their flow rates. If CO₂ is a product of the fermentation, it can be assumed that it is saturated in the liquid, thus Henry's law may be applied to find its fraction in the gas phase.

$$\begin{array}{l} \text{CO in the liquid} \\ \text{phase} \end{array} \quad 0 = q_{CO} \cdot C_x + k_L a_{CO} \cdot (C_{CO}^* - C_{CO}) - C_{CO} \cdot D \quad (12)$$

$$\begin{array}{l} \text{H}_2 \text{ in the liquid} \\ \text{phase} \end{array} \quad 0 = q_{H_2} \cdot C_x + k_L a_{H_2} \cdot (C_{H_2}^* - C_{H_2}) - C_{H_2} \cdot D \quad (13)$$

$$\begin{array}{l} \text{CO in the gas} \\ \text{phase} \end{array} \quad 0 = F_{G,in} \cdot y_{CO,in} - F_{G,off} \cdot y_{CO,off} - k_L a_{CO} \cdot (C_{CO}^* - C_{CO}) - C_{CO} \cdot D \quad (14)$$

$$\begin{array}{l} \text{H}_2 \text{ in the gas phase} \end{array} \quad 0 = F_{G,in} \cdot y_{H_2,in} - F_{G,off} \cdot y_{H_2,off} - k_L a_{H_2} \cdot (C_{H_2}^* - C_{H_2}) - C_{H_2} \cdot D \quad (15)$$

$$\text{CO}_2 \text{ overall} \quad 0 = F_{G,in} \cdot y_{CO_2,in} - F_{G,off} \cdot y_{CO_2,off} + q_{CO} \cdot C_x \cdot V_L - C_{CO_2} \cdot D \cdot V_L \quad (16)$$

$$\begin{array}{l} \text{Summation of gas} \\ \text{phase composition} \end{array} \quad 1 = y_{CO,off} + y_{H_2,off} + y_{CO_2,off} \quad (17)$$

The saturation concentrations of the gases (C_{CO}^* , $C_{H_2}^*$ and $C_{CO_2}^*$) may be estimated using Henry's equation (a list of coefficients is given by (Sander, 2015)). Moreover, the estimation of the mass transfer coefficients ($k_L a$) requires making several assumptions about the fermentation broth, the dimensions of the bioreactor vessel and the dimensions of the stirrer; unfortunately, these details are commonly not reported. The calculation of $k_L a$ may be based on the power input by the stirrer (P_S in equation 6) and by the gas sparging (P_{SG} in equation 7). Note that equation 18 and 19 were developed for the transfer of oxygen to pure water at 20 °C; therefore $k_L a$ may be further corrected for the process temperature and the specific gas through their the film diffusivity (\mathfrak{D}) of CO and H₂ in pure water compared to that of O₂ (see equation 20). Equation 20 also contains a correction factor for the mass transfer coefficient ($f_{k_L a}$), which may account for other differences in the experimental set-up and liquid phase composition compared to the ideal case for which equation 18 - 20 were initially constructed and parametrized.

$$P_S = N_p \cdot \rho \cdot N^3 \cdot D_S^5 \quad (18)$$

$$P_{SG} = P_S \cdot \left[0.1 \cdot \left(\frac{N \cdot V_L}{\dot{V}_G} \right)^{\frac{1}{4}} \cdot \left(\frac{g \cdot H_S \cdot V_L^{\frac{2}{3}}}{N^2 \cdot D_S^4} \right)^{\frac{1}{5}} \right] \quad (19)$$

$$k_L a_j = f_{k_L a} \cdot [1.022^{(T-293.15)}] \cdot \left[0.002 \cdot \left(\frac{P_{SG}}{V_L} \right)^{0.7} \cdot v_{Gs}^{0.2} \right] \cdot \frac{\mathfrak{D}_{O_2}}{\mathfrak{D}_j} \quad (20)$$

If the experimental data is instead expected to be at transient state, the concentration gradients represent one more unknown variable per each gas in the balances. The additional equations needed may be formulated using kinetic equations linking the net conversion rates of CO, H₂ and CO₂ to their dissolved concentrations. Mechanistic or

approximative formats may be used (Heijnen, 2005; Rizzi et al., 1997). The combination between both formats has been useful in our work.

To improve the accuracy of the reconciliation of the ten conversion rates, the calculation of dissolved gas concentrations may be done simultaneously by including the unknown C_{CO} , C_{H_2} and C_{CO_2} , $y_{CO,off}$, $y_{H_2,off}$ and $F_{G,off}$ among the decision variables list.

References

- Heijnen, J.J., 2005. Approximative kinetic formats used in metabolic network modeling. *Biotechnology and Bioengineering* 91, 534–545.
- Heijnen, J.J., 2002. Bioenergetics of Microbial Growth, in: Flickinger, M.C., Drew, S.W. (Eds.), *Encyclopedia of Bioprocess Technology*. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 265–291.
- Heijnen, J.J., van Dijken, J.P., 1992. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. *Biotechnology and Bioengineering*, 39, 833–858.
- Mahamkali, V., Valgepea, K., de Souza Pinto Lemgruber, R., Plan, M., Tappel, R., Köpke, M., Simpson, S.D., Nielsen, L.K., Marcellin, E., 2020. Redox controls metabolic robustness in the gas-fermenting acetogen *Clostridium autoethanogenum*. *Proc Natl Acad Sci USA* 117, 13168–13175.
- Mann, M., Miebach, K., Büchs, J., 2021. Online measurement of dissolved carbon monoxide concentrations reveals critical operating conditions in gas fermentation experiments. *Biotechnology and Bioengineering*, 118, 253–264.
- Mislov, D., Cifrek, M., Krois, I., Dzapo, H., 2015. Measurement of dissolved hydrogen concentration with Clark electrode, in: 2015 IEEE Sensors Applications Symposium (SAS). Presented at the 2015 IEEE Sensors Applications Symposium (SAS), IEEE, Zadar, Croatia, pp. 1–5.
- Noorman, H.J., Heijnen, J.J., Ch. A. M. Luyben, K., 1991. Linear relations in microbial reaction systems: A general overview of their origin, form, and use. *Biotechnology and Bioengineering*, 38, 603–618.
- Noorman, H.J., Romein, B., Luyben, K.Ch.A.M., Heijnen, J.J., 1996. Classification, error detection, and reconciliation of process information in complex biochemical systems. *Biotechnology and Bioengineering* 49, 364–376.
- Rizzi, M., Baltés, M., Theobald, U., Reuss, M., 1997. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnology and Bioengineering* 55, 592–608.
- Sander, R., 2015. Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmospheric Chemistry and Physics* 15, 4399–4981.
- Stephanopoulos, G., Tsiveriotis, C., 1989. Toward a Systematic Method for the Generalization of Fermentation Data, in: Fish, N.M., Fox, R.I., Thornhill, N.F. (Eds.), *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*. Springer Netherlands, Dordrecht, pp. 169–178.
- van der Heijden, R.T.J.M., Heijnen, J.J., Hellinga, C., Romein, B., Luyben, K.Ch.A.M., 1994a. Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. *Biotechnology and Bioengineering*, 43, 3–10.
- van der Heijden, R.T.J.M., Romein, B., Heijnen, J.J., Hellinga, C., Luyben, K.Ch.A.M., 1994b. Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnology and Bioengineering*, 43, 11–20.
- Villadsen, J., Nielsen, J.H., Lidén, G., Nielsen, J.H., 2011. *Bioreaction engineering principles*, 3rd ed. Springer, New York.

Process Systems Engineering prospects in Circular Economy implementation in industry

Konstantina G. Stylianopoulou^a, Emilia M. Kondili,^a John K. Kaldellis^b

^a*Optimisation of Production Systems Lab., Mechanical Engineering Department, University of West Attica, Greece*

^b*Soft Energy Applications and Environmental Protection Lab., Mechanical Engineering Department, University of West Attica, Greece*
Email for correspondence: ekondili@uniwa.gr

Abstract

The present study is inspired by our current research work concerning the development of innovative solutions for the optimal use of resources in the industry in the context of circular economy. The outcome of the research is expected to be the development of a framework including prototypes and models for the process and manufacturing industry in order to make the optimal use of resources in the entire life cycle of materials and products. These prototypes and models, including materials reuse, water recovery and reuse, products remanufacturing, waste minimisation, all relevant to minimizing resource consumption and environmental impacts, will be implemented integrating industry 4.0 technologies, optimisation models, feasibility studies and life cycle analysis to support Circular Economy practices.

In the above context, the present paper introduces our wider research issues and parameters and describes the significant role of Process Systems Engineering in the Circular Economy context of the above industrial problems. In parallel, the importance of a new field of PSE tools implementation and expansion emerges as an outcome of the work providing very interesting and challenging perspectives in addition to their existing fields of application.

Keywords: PSE prospects, modelling and optimisation in circular economy.

1. Introduction and Rationale of the work

Many organisations, academics, companies and policymakers have acknowledged the circular economy as an opportunity for more sustainable industrial models and strategies. According to this viewpoint, companies can achieve substantial cost savings and additional revenues by adopting circular business models.

As sustainability gains more attention and is being used by industries, governments and academic cycles, there has been a considerable effort in moving towards sustainable development not only for environmental and societal needs but also as an important economic activity. A comprehensive framework to accomplish this is shifting to the principles of the so-called circular economy (CE), that is attracting more and more attention as a way to face materials scarcity and the growing demand for resources in general (e.g. water, energy, land).

The European Union defines Circular Economic as a model where the value of products, materials and resources is maintained in the production and use cycle for as long as possible, and waste generation is minimised. The Ellen MacArthur Foundation (2013) expanded this definition by stating that the Circular Economy is a systemic approach to economic development designed to benefit businesses, society, and the environment. In contrast to the ‘take-make-waste’ linear model, CE aims to develop design and operation of production systems that gradually decouple growth from the consumption of finite resources.

The concept recognises the importance of the economy needing to work effectively at all scales for big and small businesses, for organisations and individuals, globally and locally, based on three principles (1) minimisation of waste and pollution, (2) extension

of products and materials life and (3) regeneration of natural systems (Ellen MacArthur Foundation, 2013) thus setting the CE concept with the potential to generate economic, environmental and social benefit. In line with the previous principles, the European Union (EU) adopted an action plan in 2015 where the value of products and materials is maintained for as long as possible, bringing thereby significant economic, social and environmental benefits by making its members' states pioneers in making policies and framework to promote the CE in industries, businesses, and services.

In quantitative terms, according to the European Commission more efficient use of raw materials and resources throughout the supply chain materials could reduce the need for new raw material of 17% -24% by 2030, with savings for European industry estimated at 630 billion Euro per year. Several studies on the potential of circular economy indicate that European industry, thanks to substantial savings on the cost of raw materials, could push the growth of European GDP by about 3.9% and create a very large number of new jobs. (<https://www.cesme-book.eu/>)

2. Process Systems Engineering and the Circular Economy

On the other hand, Process Systems Engineering practice started from the need to define and solve increasing complexity problems in technical systems, understand and control these systems' characteristics as a whole. The PSE is focused on the integrated design and operation of complex technical systems. PSE relies on systems thinking, feature very essential to the Circular Economic approach.

Decision-making for supply chains and understanding Circular Economy and the transition from linear to circular models makes PSE a fundamental part of the CE as a solution (Falk et al., 2016, Walmsley et al., 2019, Dantas et al., 2021).

According to Reichel et al. (2016) and Avraamidou et al. (2020), the critical characteristics of CE, which also fall under the Ellen MacArthur Foundation (2021) goals, are (1) minimise the use of natural resources: less import dependence on natural resources by efficient use of natural resources, (2) increase of renewable resources and energy: substituted conventional resources to renewable and minimizing the virgin materials, (3) minimisation of emissions levels: reduction of pollutants through clean material cycles, (4) less material losses/residuals: waste minimisation through the recovery and recycling of materials and products, (5) manage resources to preserve value: promote the development of business models to increase the value of products in redesigning and high-quality recycling. Figure 1 shows indicative research fields with potential use in achieving CE goals (Avraamidou et al 2020).

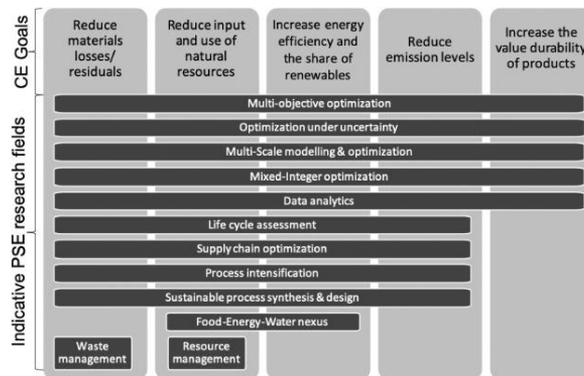


Figure 1: Indicative research fields with potential use in achieving CE goals (Avraamidou et al 2020).

As process systems engineering has become one of the first to acknowledge sustainability, it has become the leading model in process design. Even though the rationale to transform into a sustainable process may have been essentially economic, through the years, this notion changed as more industries wanted their process systems to not only see the economic value but also how the entire system interacts with the

environment (Silk et al., 2019). Their research discovered that in any attempt in modeling or optimizing circular economy, the most common approach to attain sustainability in a process design is through Life Cycle Assessment, which is a way of comparing different process design alternatives. Thus, Life Cycle Assessment (LCA) is being employed within the PSE community for years. The process systems engineering tools may enable a practical way for analyzing performance and improvement in a sustainable way within industries. The reason for this is that LCA quantifies all relevant issues related to the products or services and could evaluate sustainable development, making it a handy tool for CE.

In the present research LCA is one of the main tools to create and evaluate alternatives. Additionally, process optimisation, which relies on mathematical optimisation, gains more attention in the circular economy notion to industries as it can be used as a methodology for minimizing or maximizing suitably defined objectives and goals in the whole life cycle in a sustainable way (Kristoffersen et al., 2020). From the process systems engineering view, the progress of efficient design and state-of-the-art mathematical tools to support quantitative methods in the process optimisation applications with sustainability in mind has been growing in the academic cycles and industries. Besides, when circular economy and sustainability in supply chains are combined, it must include optimisation of supply chain industries to have efficient and circular systems that assist the progress of reducing, reuse and recycling and emphasise environmental, social, and economic demands. This implementation requires adjustment in design management by optimizing all factors to minimise resource consumption and maximise circular economy on materials, products and systems. As a result, circular supply chains management, information and knowledge issues are organised and planned to optimise resource allocation, increase benefits and achieve the circular economy notion. Optimisation model's development is a core issue in our in-progress research.

3. Research gaps and expected added value

In the industrial community a lot of attention has been given to the sustainability issues. According to Bjornbet et al. (2021), there is a promising future in transitioning towards CE in the manufacturing community through circular business models. Thus, acquiring experience in modeling such industries that are yet fully discovered by PSE is essential in modeling and optimizing the CE through supply chains. Making the resources use circular could increase their productivity in recovering an immense amount of those from being employed in manufacturing. Thus, circular economy could contribute to changing even the synthesis of the production system at the first place. In this manner, designing industries for availability, sustainability, and of course including design in reusing and remanufacturing will gain more attention and become vital. However, success in moving towards a circular economy depends on whether industries understand the importance of changing into it for efficient planning and management over time.

Industry is a very wide and complex system and certainly there is no unique or generic approach towards circular economy in this sector. It is important at this point to emphasise that circularity is a wider term and is not identical to sustainability, although most methods and approaches converge to that.

Therefore, more specific models expressing different industrial characteristics, objectives and constraints depending on the type of industry as well as its crucial issues in terms of circular economy principles need to be developed. As Pistikopoulos et al, (2021) mention the Circular Economy will be the framework and the context of Process Systems Engineering in the next years. Therefore, modeling will definitely play a crucial role in achieving CE goals, along with other tools such as Life Cycle Analysis and feasibility studies, that will lead to the economic evaluation of proposed measures and interventions.

4. Research questions, methods and tools

The suggestions – referring to each industrial type under consideration - extend from materials reuse, water recovery, products remanufacturing, waste minimisation, all relevant to minimizing resource consumption and negative environmental impacts.

The main research questions / objectives of the proposed work are (also shown schematically in Figure 2):

- To investigate the potential and the methods for the industry to use less resources and produce less waste when producing materials and products
- To identify the methods and tools and to develop the relevant models that will enable circularity in the manufacturing and process and manufacturing industry
- To focus on the recovering and reprocessing of materials and products that will lead to the optimisation of resources exploitation
- To develop advanced techno-economic modelling, life-cycle assessment and optimisation models, in order to predict/determine which technologies, both established and new, will perform best at scale and enable circularity.
- To investigate the suitability of the above methods and tools depending on the type of industry, i.e. determine the matching between methods/tools and sectors for modelling the industry innovation for circular economy
- To create certain specific models and paradigms and the relevant specific results to the selected sectors,
- To extend the scope of the models and paradigms to other industrial (and non-industrial sectors, such as the energy).

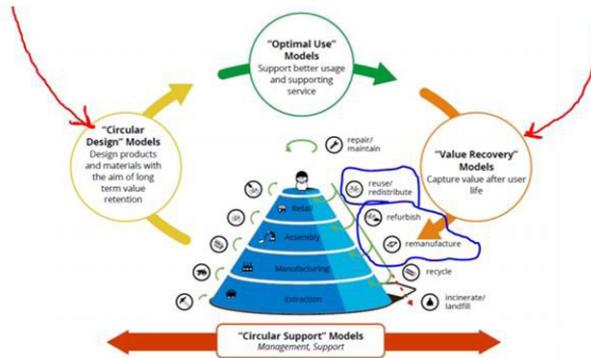


Figure 2: Main scope of the proposed research (figure inspired by The EIB Circular Economy Guide, https://www.eib.org/attachments/thematic/circular_economy_guide_en.pdf)

The expected outcome of the study is the specific recommendation of circular (technoeconomic, optimisation) prototypes and models for the process and manufacturing industry in order to make the optimal use of resources in the entire life cycle of materials and products. This will be implemented exploiting PSE existing tools, optimisation models, feasibility studies and life cycle analysis to support CE practices.

The main methods, tools and models for the assessment of different aspects of the circular economy for the industry, include advanced techno-economic modelling, life-cycle assessment, systems engineering models and tools and optimisation methods and tools (mathematical optimisation, multicriteria analysis). Forecasting work has already taken place concerning the evolution of circular economy on the industry (Angelopoulou, Kondili, July 2021). Furthermore, continuous collaboration with industrial partners for detailed description of cases and paradigms has been established.

The classification of the industries (sectoral classification) according to their suitability in one of the following basic axes of circular economy (as shown in Figure 2 above).

- The waste reprocessing and reuse
- The minimisation of resources utilisation
- The design of materials and products according to the main eco-design principles.

Currently, having completed the existing background work, a classification of the industrial sectors under consideration is taking place according to the resources where

each sector is intensive (metals, energy, water, land, other materials, etc.) as well as its classification according to the potential of materials and products remanufacturing and reuse.

5. Industry 4.0 and PSE

Today the model development based on these new business models is also linked to the Industry 4.0 (I4.0) paradigm, the fourth industrial revolution in production systems using a higher level of digitalisation. The core idea of Industry 4.0 is to use emerging technologies so that business and engineering processes are integrated by making production operate in a flexible, efficient, and sustainable way. (Singh et al., 2019, Rajput & Singh, 2020, Jamwal et al., 2021). Industries cannot prevent themselves from applying I4.0 systems, and the fundamental contribution of I4.0 is digitalizing CE practices through different forefront technologies. Industry 4.0 creates abundant opportunities for industries to improve circular performance and evolve ethical principles of social responsibility by optimizing the usage of resources with the improvement of the lifecycle of products (Rajput & Singh, 2020).

The problem considered here is of implementing Industry 4.0 along with Process Systems Engineering considering circular economy. There is an excess of research studies associated with industry 4.0, smart manufacturing, and the Internet of Things in industries. However, the combination of process systems engineering that focuses on circular economy is limited, with most researchers addressing real-time data availability, computing, and big data. It was proposed that industries adopt more model-driven and object-orientated information-driven models for increased optimisation, better efficiency, and circular economy support in an Industry 4.0 approach.

Thus, in order to be successfully implemented, as far as the supply chain management is concerned, Rajput & Singh (2020) proposed the CE principles that reduce, reuse, recycle, recover, remanufacture and redesign materials and products and life cycle analysis as possible solutions to this integration. Nevertheless, there is a need to quantify the relationships between Industry 4.0 in the industry's management, performance, and sustainability to identify specific stages of product life extension through optimisation. Thus, by adopting and implementing strategies with quantitative and real-time data about Industry 4.0, there is a need for new sustainable manufacturing practices and combining all, Industry 4.0 knowledge and process system engineering could create effective results in not only overall firm performance but also industry performance (Walmsley et al., 2019, Rosa et al., 2020, Bag et al., 2021). This is another important research objective of the present work.

6. Conclusions

Many comprehensive works have been developed concerning the theoretical background and the principles of Circular Economy. Concerning the process and manufacturing industry, there is a wide scope for the development of practical and implementable solutions that will have a real and measurable benefit to the industrial units. There is a need to step forward to establish new frameworks that, together with adequate data and actions in the notion of the CE in industries, could fulfil the CE goals. Enabling circular economy requires changing the business models from selling products to delivering services and from a linear system lifecycle to a circular one. Thus, there is a need for collaboration between the academic community, industries, and society. There has to be a better understanding of the implications of CE and how this can work together with I4.0 by examining case studies and using the appropriate PSE methods and tools as those could help in the whole notion of the CE.

In the present work the basis has been set for the identification of how CE models may be developed and solutions suggested by using PSE methods and tools.

1. The development of relevant integrated models to use less resources or use resources more efficiently when producing materials and products will enable circularity in the manufacturing and process industry focusing on the recovering and reprocessing of materials and products that will lead to the optimisation of resources exploitation. The work will highlight the real potential of the PSE tools for circular economy and at the same

time establish PSE in the context of circular economy, thus revealing an interesting field for expansion, application of PSE tools and developing excellent prospects in the field.

References

- S. Avraamidou, S.G. Baratsas, Y. Tian, E.N. Pistikopoulos, 2020, Circular Economy - A challenge and an opportunity for Process Systems Engineering, *Computers and Chemical Engineering*, 133
- S. Bag, G. Yadav, P. Dhamija, K.K. Kataria, 2021, Key resources for industry 4.0 adoption and its effect on sustainable production and circular economy: An empirical study. *Journal of Cleaner Production*, 281
- M.M. Bjørbet, C. Skaar, A.M. Fet, K.Ø. Schulte, 2021, Circular economy in manufacturing companies: A review of case study literature, *Journal of Cleaner Production*, 294
- CESME White book, <https://www.cesme-book.eu/>, 2021
- T.E.T. Dantas, E.D. de-Souza, I.R. Destro, G. Hammes, C.M.T. Rodriguez, S.R. Soares, 2021, How the combination of Circular Economy and Industry 4.0 can contribute towards achieving the Sustainable Development Goals, *Sustainable Production and Consumption*, 26, 213–227
- Ellen MacArthur Foundation, 2021, Universal Circular Economy Policy Goals, <https://ellenmacarthurfoundation.org/>
- D. Falk, B.G. Steiber, 2016, Systems Engineering for a Circular Economy, 26th Annual INCOSE International Symposium (IS 2016) Edinburgh, Scotland, UK, July 18-21
- A. Jamwal, R. Agrawal, M. Sharma, A. Giallanza, 2021, Industry 4.0 Technologies for Manufacturing Sustainability: A Systematic Review and Future Research Directions. *Applied Sciences*, 11, doi:10.3390/app11125725
- E. Kristoffersen, F. Blomsma, P. Mikalef, J. Li, 2020, The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies, *Journal of Business Research*, 120, 241–261
- Pistikopoulos, E.N., Barbosa-Povoa A., Lee, J.H., Misener, R., Mitsos A., Reklaitis G.V., Venkatasunbramanian V., You F., Gani R., 2021, Process systems engineering – The generation next?, *Computers and Chemical Engineering*, 147, 107252.
- S. Rajput, S.P. Singh, 2020, Industry 4.0 Model for circular economy and cleaner production, *Journal of Cleaner Production*, 277
- P. Rosa, C. Sassanelli, A. Urbinatia, D. Chiaroni, S. Terzi, 2020, Assessing relations between Circular Economy and Industry 4.0: a systematic literature review, *International Journal of Production Research*, 58, 6, 1662–1687.
- S.P. Singh, R.K. Singh, A. Gunasekaran, P. Ghadimi, 2019, Supply Chain Management, Industry 4.0, and the Circular Economy, *Resources, Conservation & Recycling*, 142, 281–282
- D. Silk, B. Mazzali, I.A. Udugama, K.V. Gernaey, M. Pinelo, J. Woodley, S.S. Mansouri, 2019, Systematic decision-support methodology for identifying promising platform technologies towards circular economy, *Proceedings of the 29th European Symposium on Computer Aided Process Engineering June 16th to 19th, 2019, Eindhoven, The Netherlands*.
- S. Sitter, Q. Chen, I.E. Grossmann, 2019, An overview of process intensification methods, *Chemical Engineering*, 25, 87–94
- T.G. Walmsley, B.H.Y. Ongb, J.J. Klemeša, R.R. Tanc, P.S. Varbanov, 2019, Circular Integration of processes, industries, and economics, *Renewable and Sustainable Energy Reviews*, 107, 507–515

Incremental financial analysis of black liquor upgraded gasification in integrated kraft pulp and ammonia production plants under uncertainty of feedstock costs and carbon taxes

Meire Ellen Gorete Ribeiro Domingos^{a,c}, Daniel Flórez-Orrego^b, Moisés Teles dos Santos^a, Silvio de Oliveira Júnior^b, François Maréchal^c

^a*Polytechnic School, University of Sao Paulo, Department of Chemical Engineering, Sao Paulo 05508-000, Brazil.*

^b*Polytechnic School, University of Sao Paulo, Department of Mechanical Engineering, Sao Paulo 05508-030, Brazil.*

^c*École Polytechnique Fédérale de Lausanne, Valais-Wallis, 1951 Sion, Switzerland
meireellengorete@usp.br / meire.ribeirodomingos@epfl.ch*

Abstract

In this work, the conventional scenario of the black liquor (BL) concentration and combustion is compared with the BL upgrading gasification process for ammonia production. The chemical processes synthesis, modeling and simulation are performed by using Aspen Plus® software. The determination of the heat recovery and the solution of the energy integration problem is handled by a mixed integer linear programming model. An incremental financial analysis incorporates the uncertainty related to the acquisition and selling costs of the feedstock and fuels produced and carbon taxation by using the Monte Carlo method. As a result, the incremental financial analysis found that only the integrated pulp and ammonia production route with partial electricity import may economically outperform the conventional kraft pulp mill for moderate carbon taxations (40-90 EUR/t_{CO2}), depending on the interest rate adopted. In this regard, middle-to-high carbon taxations may render ammonia co-production attractive in the Brazilian context of a highly renewable electricity mix.

Keywords: Black liquor, Kraft process, Uncertainty, Ammonia, Decarbonization.

1. Introduction

Ammonia is one of the most demanded bulk chemicals in the world, mainly for the production of fertilizers for the agricultural sector. In 2016, the ammonia production reached 175 million tons, and the trend from 2006 to 2016 shows a growth rate of 1.9% per year (YARA, 2018). The ammonia synthesis is also the largest carbon dioxide emitting chemical industry process, responsible for about 1.8% of global carbon dioxide emissions (Royal Society, 2020). Thus, several efforts are being made towards the decarbonization of the ammonia production supply chain and the mitigation of the environmental impacts of this sector. The thermochemical conversion routes for hydrogen production are among the proposed alternatives, which can also capitalize on the underexploited biomass potential in tropical countries with a well-established biomass conversion expertise, such as Brazil. This work proposes to produce ammonia by using the syngas coming from the gasification of black liquor (BL), a byproduct of the kraft

pulping process. The incremental financial analysis under uncertainty are considered as criteria for assessing the potential to implementing this technology under volatile market conditions.

2. Methods

2.1. Process modeling and optimization problem definition

Figure 1 illustrates the integrated system for ammonia production via black liquor gasification. First, the weak black liquor is dried in a mechanical vapor recompression system. Subsequently, the strong black liquor is gasified in a pressurized system using oxygen. The syngas obtained need to be treated, purified and its composition should be adjusted before enters the ammonia loop. To this end, autothermal reformer, water gas shift, CO₂ capture and methanation systems are required. Finally, ammonia is purified before follow to its end-use. This process is modelled in the Aspen Plus® software and the detailed description of the processes conditions is reported in (Domingos et al., 2021). The integration of the ammonia plant to the existing kraft pulp mill creates new utility demands. So, the determination of the minimum energy requirements (MER) and the solution of the energy integration problem is handled by the OSMOSE Lua platform. Three scenarios are considered: i) conventional: recovery boiler application, with only pulp production; ii) mixed: based on both chips fuel and electricity import, with ammonia and pulp production; and iii) autonomous: only chips import and cogeneration system enabled, with ammonia and pulp production. The market costs and selling prices for feedstock and products that are considered for the optimization problem solution are: wood 0.013 €/kWh; chips 0.016 €/kWh; oil 0.018 €/kWh; electricity 0.06 €/kWh; pulp 0.144 €/kWh; ammonia 0.098 €/kWh; and CO₂ 0.0084 €/kg, as reported in (Domingos et al., 2021).

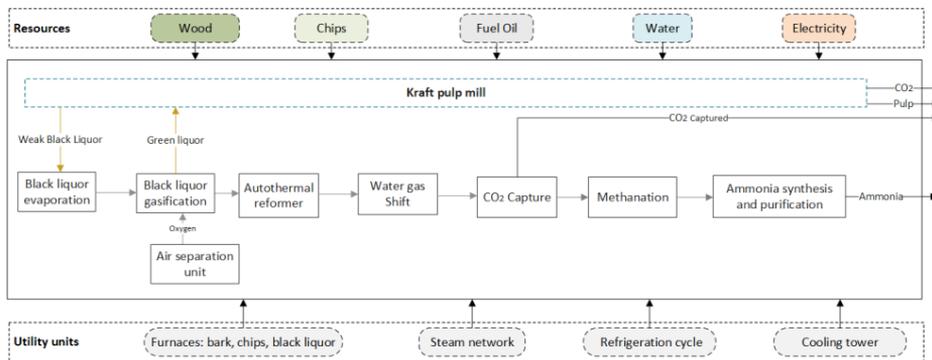


Figure 1. Integrated flowsheet of pulp and ammonia production.

2.2. Economic evaluation

The capital expenditure (CAPEX) is calculated using the methodology described in (Turton et al., 2018), in which actual costs are correlated to the cost of a reference equipment capacity using scaling factors. Equation 1 is used to estimate CAPEX of the black liquor gasifier, air separation unit, autothermal reformer, water gas shift reactors, CO₂ capture unit, and methanation, according to the data reported in (Larson et al., 2000, Williams et al., 1995, Hamelinck and Faaij, 2002, Holmgren, 2015), and corrected for inflation using CEPCI index (Jenkins, 2020).

$$C_1 = C_0 \left(\frac{S_1}{S_0} \right)^r \quad (1)$$

where S_0 represents the reference capacity with known capital cost C_0 ; S_1 represents the actual capacity for which the capital cost C_1 is unknown; and r is the power scaling factor which varies between 0.5 and 0.9 depending on the type of process considered. Also, the kraft pulp mill investment costs are scaled (0.6 factor) as per (Börjesson, 2015). A lifespan of 20 years is assumed for cash flow calculations. Besides, the total capital expenditure is divided between the first (60%) and second (40%) years. A decommissioning cost of 6% of the overall CAPEX is assumed. The operating costs (OPEX) are assessed using the methodology proposed by Turton et al. (2018), except for the kraft pulp mill, in which OPEX is assumed as 4% of the CAPEX as per (Kangas et al., 2014). A contingency cost increment of 20% is considered by reason of the technological risk level (Turton et al., 2018).

2.2.1. Incremental financial analysis

An incremental approach is proposed to compare the economic attractiveness of the integrated ammonia plant (Fig. 1) and the conventional process. This comparison is based on the incremental net present value (INPV), defined by Eq. (2) (Blank, 2011):

$$INPV = \sum_{n=1}^N \frac{[(Rev-Exp)_{n,option B}] - [(Rev-Exp)_{n,option A}]}{(1+i)^n} \quad (2)$$

where $(Rev-Exp)$ is the net cash flow (i.e. revenues minus expenses), calculated at each (n) of the N yearly periods, during which both the new (B) and the reference (A) configurations should operate. Besides, i is the average interest rate.

A sensitivity analysis of the variation of the INPV as a function of the carbon taxation (0-100 €/t_{CO2}) and the interest rate (0-21%) is performed. This analysis aims to consider the scenario of rigorous environmental regulations and the increased perception of the risks associated with these technologies.

Also, an incremental financial analysis that incorporates the uncertainty related to the acquisition and selling costs of the feedstock and fuels produced is performed through Monte Carlo method, by simulating the stochastic variation of the commodities price profiles. A normal distribution with mean prices as those reported in section 2.1 and a standard deviation of 30% is considered for the commodities prices.

As a result, the INPV of the integrated chemical plants can be calculated along the lifespan of those facilities and the so-called ‘likelihood of loss’, which measures the probability of achieving a negative incremental net present value, can be determined. Accordingly, three hypothetical scenarios are considered:

- i) Scenario DCTIR_SC: deterministic carbon taxes (0-100 EUR/t_{CO2}) and interest rates (0-21%), along with stochastic prices of the commodities;
- ii) Scenario LCT_DIR_SC: linearly-increasing carbon taxes over the lifespan (0-100 EUR/t_{CO2}), along with deterministic interest rates (0-21%) and stochastic prices of commodities;
- iii) Scenario SCTC_DIR: stochastic carbon taxes and variable prices of the commodities, along with deterministic interest rates (0-21%).

3. Results and discussion

The breakdown of the capital expenditure for both scenarios of the integrated ammonia plant (Fig. 2) suggests that the operation under the autonomous mode requires larger chips furnaces and Rankine cycle based power plants for supplying the utilities demands.

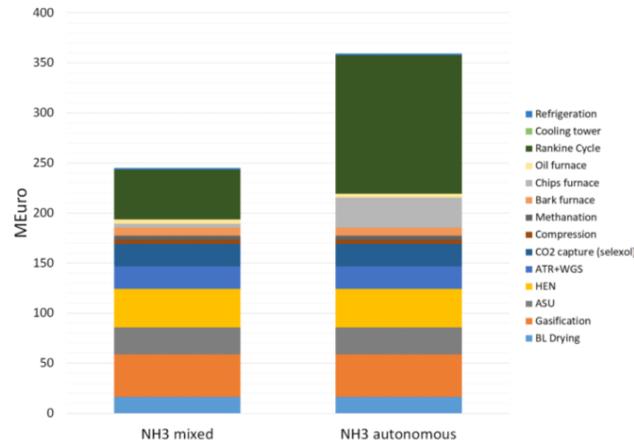


Figure 2. Capital expenditure breakdown.

According to Figure 3, the ammonia production route with partial electricity import (mixed case in Figure 3a) may economically outperform the autonomous setup, depending on the interest rate adopted, for moderate carbon taxations (40-90 EUR/t_{CO2}), such as those reported for Norway or France (World Bank, 2021).

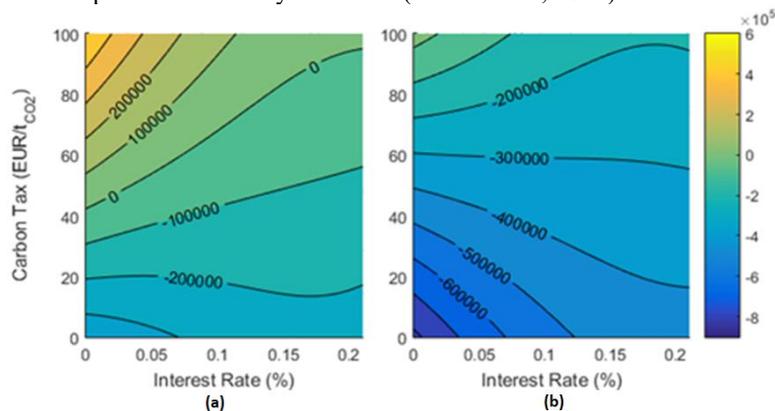


Figure 3. Contour plots of INPV (Euro) variation for integrated ammonia production via BL gasification under: a) mixed and b) autonomous operation modes.

The heat maps given in Tables 2 and 3 summarize the results obtained for the scenario DCTIR_SC. Table 2 suggests that, for the integrated plant co-producing pulp and ammonia while operating under the mixed mode, and for a region of middle to stringent carbon taxations (60-100 EUR/t_{CO2}), there exists a large probability of the INPV being positive, depending on the interest rate adopted. On the other hand, for the autonomous mode (Table 3), presents relatively unfavorable INPVs, regardless of how low interest rates become.

Meanwhile, the heat map shown in Table 4 summarizes the results obtained for scenarios LCT_DIR_SC and SCTC_DIR. For both scenarios, similarly to the previous analysis, only the integrated chemical plant producing NH₃ while operating under mixed mode presented more favorable results in terms of likelihood of loss, which may render NH₃ co-production attractive in the Brazilian context of a highly renewable electricity mix.

Table 2. Probability that INPV is negative (likelihood of loss in %) for scenario DCTIR_SC - ammonia production under the mixed operation mode.

CO ₂ tax (EUR/t _{CO2}) →		0	10	20	30	40	50	60	70	80	90	100
i (%)	0%	100.00	100.00	99.90	95.43	63.84	14.94	0.81	0.01	0.00	0.00	0.00
	3%	100.00	100.00	100.00	99.56	89.46	44.89	7.41	0.21	0.00	0.00	0.00
	6%	100.00	100.00	100.00	100.00	98.59	81.91	34.00	4.64	0.10	0.00	0.00
	9%	100.00	100.00	100.00	100.00	99.86	96.60	71.81	26.74	3.46	0.10	0.00
	12%	100.00	100.00	100.00	100.00	99.97	99.60	93.54	65.03	22.91	2.70	0.13
	15%	100.00	100.00	100.00	100.00	100.00	99.94	99.04	89.60	58.84	19.20	2.31
	18%	100.00	100.00	100.00	100.00	100.00	100.00	99.90	97.50	83.99	50.23	15.59
	21%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.59	96.14	76.60	42.90

Table 3. Probability that INPV is negative (likelihood of loss in %) for scenario DCTIR_SC - ammonia production under the autonomous operation mode.

CO ₂ tax (EUR/t _{CO2}) →		0	10	20	30	40	50	60	70	80	90	100
i (%)	0%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.91	97.26	71.91	22.76
	3%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.96	97.39	74.66
	6%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.94	98.41
	9%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	12%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	15%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	18%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	21%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 4. Probability that INPV is negative (likelihood of loss in %) as a function of the interest rate for LCT_DIR_SC and SCTC_DIR scenarios.

i (%)	LCT_DIR_SC		SCTC_DIR	
	NH ₃ mixed	NH ₃ auto	NH ₃ mixed	NH ₃ auto
	0%	15.84	100.00	20.10
1%	31.53	100.00	28.29	100.00
2%	52.40	100.00	37.50	100.00
3%	73.21	100.00	47.47	100.00
4%	88.21	100.00	58.83	100.00
5%	95.64	100.00	69.59	100.00
6%	98.61	100.00	78.37	100.00
7%	99.57	100.00	86.01	100.00
8%	99.94	100.00	90.56	100.00
9%	99.99	100.00	94.61	100.00
10%	100.00	100.00	96.59	100.00
11%	100.00	100.00	98.34	100.00
12%	100.00	100.00	99.11	100.00
13%	100.00	100.00	99.30	100.00
14%	100.00	100.00	99.57	100.00
15%	100.00	100.00	99.80	100.00

16%	100.00	100.00	99.89	100.00
17%	100.00	100.00	99.97	100.00
18%	100.00	100.00	100.00	100.00
19%	100.00	100.00	100.00	100.00
20%	100.00	100.00	100.00	100.00
21%	100.00	100.00	100.00	100.00

4. Conclusion

The incremental financial analysis under uncertainty of feedstock costs and carbon taxes allowed to understand the behavior of the integrated systems proposed by considering different market fluctuations. As a result, positive INPVs are achieved when moderate carbon taxes (40-90 EUR/tCO₂) are considered, provided that greener electricity import from the grid is enabled. Even when the system is subject to linear-increasing or stochastic carbon taxes, the scenario operating under the mixed mode points towards a great potential of decarbonization. The carbon taxation is reported as an effective measure to reduce CO₂ emissions, however other measures should be adopted, such as more stringent regulatory commitments, attractive fiscal incentives, and investments towards the equipment maturation and deployment.

References

- M. Borjesson, E. Ahlgren. 2015. Pulp and Paper Industry - Energy Technology System Analysis Program - ETSAP, IEA ETSAP - Technology Brief I07 - May 2015. IEA.
- M. E. G. R. Domingos, D. Florez-Orrego, M. Teles dos Santos, H. Velasquez, S. de Oliveira Jr. 2021. Exergy and environmental analysis of black liquor upgrading gasification in an integrated kraft pulp and ammonia production plant. *International Journal of Exergy*, 34.
- C. N. Hamelinck, A. P. Faaij. 2002. Future prospects for production of methanol and hydrogen from biomass. *Journal of Power sources*, 111, 1-22.
- K. Holmgren. 2015. Investment cost estimates for gasification-based biofuel production systems. IVL Svenska Miljöinstitutet.
- S. Jenkins 2020. 2019. Chemical Engineering Plant Cost Index Annual Average [Online]. Available: <https://www.chemengonline.com/2019-chemical-engineering-plant-cost-index-annual-average/> [Accessed 08/02/2021 2021].
- P. Kangas, S. Kaijaluoto, M. Maattanen. 2014. Evaluation of future pulp mill concepts—Reference model of a modern Nordic kraft pulp mill. *Nordic Pulp & Paper Research Journal*, 29, 620-634.
- E. D. Larsson, S. Consonni, T. G. Kreutz. 2000. Preliminary economics of black liquor gasifier/gas turbine cogeneration at pulp and paper mills. *J. Eng. Gas Turbines Power*, 122, 255-261.
- Royal Society. 2020. Ammonia: Zero-Carbon Fertiliser, Fuel and Energy Store. The Royal Society London, UK.
- World Bank. 2021. Carbon Pricing Dashboard [Online]. Available: https://carbonpricingdashboard.worldbank.org/map_data [Accessed November 6th 2021].
- R. Turton, W. B. Whiting, D. Bhattacharyya, J. A. Shaeiwitz. 2018. Analysis, synthesis, and design of chemical processes.
- L. Blank, A. Tarquin. *Engineering Economy*, 7th edition ed., Mc Graw Hill, El Paso, Texas, 2011.
- R. H. Williams, E. D. Larson, R. E. Katofsky, J. Chen. 1995. Methanol and hydrogen from biomass for transportation. *Energy for Sustainable Development*, 1, 18-34.
- YARA. 2018. Yara Fertilizer Industry Handbook [Online]. Available: <https://www.yara.com/siteassets/investors/057-reports-and-presentations/other/2018/fertilizer-industry-handbook-2018-with-notes.pdf/> [Accessed november 2 2020].

A Capex Opex Simultaneous Robust Optimizer: Process Simulation-based Generalized Framework for Reliable Economic Estimations

Kristiano Prifti^a, Andrea Galeazzi^a, Massimo Barbieri^b and Flavio Manenti^{a,*}

^a*Politecnico di Milano, CMIC Department "Giulio Natta", Piazza Leonardo da Vinci 32, Milan 20133, Italy*

^b*Politecnico di Milano, Technology Transfer Office, Piazza Leonardo da Vinci 32, Milan 20133, Italy*

flavio.manenti@polimi.it

Abstract

During the design of a process, the most economically impactful choices are often taken with limited data and details available on the specifics of unit operations. Despite this fact, the effect some of these choices have on the final economic performance of the plant can be dramatic and far more influential than later decisions. The availability of fast and reliable preliminary cost estimations based on few key design parameters for the most common units can go a long way in improving flowsheet design. Due to the early design stage, these estimates don't need to be very accurate since errors in the range of 50% are admissible as long as the estimation requires few input data. Automating the costs computation process and making the interface with the most common simulation packages standardized and easily accessible for users not accustomed to programming languages is very important in speeding up the cost evaluation of the plant and reducing the human resources tied to this task. These concepts stand at the core design of the CAPEX OPEX Robust Optimizer (CORO) developed in this work. Aspen HYSYS serves as the commercial simulation package to estimate the input variable of the economic libraries. Excel is used both as a GUI and as a data extraction tool from Aspen HYSYS due to its widespread diffusion in industry and versatility provided by Visual Basic for Applications. The CORO code is detached both from Excel and HYSYS and interacts only with a standardized xml data sheet to allow for in-house expansions to other simulation packages. The long-term development goal is a generalized CAPE-OPEN interface working with every commercial software that supports the interface. In the current CORO release, the interface can interact only with Aspen HYSYS. This paper will showcase the economic libraries implemented in CORO, the mathematical C++ optimization libraries, the overall structure of the tool, planned future expansions, and customization options.

Keywords: Economics, Aspen HYSYS, Simulation, Optimization

1. Introduction

As pointed out by Martin et al. (2007) during the conceptual design stage decisions have the most impact on the final capital expenditure of the project. At the same time, changing these decisions is rather cheap as there has been no concrete financial commitment yet. For this reason there is a strong incentive to estimate the costs as early as possible in the design phase of a new project to weed out the less competitive solutions. The issue lies with the limited data available on the units and the reliability of their estimation. While this might have been a bottleneck in the past,

modern commercial simulation packages allow for the development of process designs with low human and financial resource commitment and ever growing prediction precision. The automated development of process layout is still in the early phases of development, but the automation of cost estimation has already some commercial solutions implemented in simulation software such as Aspen One. The inherent design flaw with this solutions is that their extreme flexibility leads the correlations to some hard failure condition whenever the characteristic dimension of the unit is outside the boundaries of the correlation used for the cost estimation, or industry specific designs for unit operations are used (LNG, cryogenic equipment). Since the economic estimation package comes in a commercial solution the parameters and correlations are protected and not customizable by the user if not only in a limited way. Moreover, certain units will lead the cost estimation routine to fail and return either an error or a zero cost estimation. In this conditions the deep integration between the cost estimation libraries and simulation software does not allow for new unit operations to be integrated in the cost estimation routines. Here a new structure to deal with the cost estimation is presented as a detached structure from the specific commercial simulator interfacing with it only with a standardized .xml structure. This approach however has its own drawbacks;

- A specific add-on to write and read the .xml file must be developed for each simulation software
- The information required by the .xml file might not be available in the simulation software
- More human resources have to be invested in the development phase of the project

[noitemsep, nolistsep] the detached structure however also offers some opportunities:

- Once the interface between the commercial software and optimization package has been developed no further development is necessary beyond cross-release updates, on the long term, this will lead to less resources used in the cost estimation process
- The standard correlations used in literature can be heavily customized both in terms of fitting parameters and model variables
- Internal company database can be used and automatically digested by the software to refit the correlations for specific applications (Ammonia production, small scale plants, cryogenic operations)
- Unconventional units can be modeled using conventional simulation blocks (such as using a Gibbs reactor to model a Furnace, and later evaluate the furnace cost using the MW leaving the Gibbs reactor)

Critical to the success of such a solution is ease of use and low back-end responsibilities from the end user. In a long term vision this project should be included in a rigorous interface using the CAPE-OPEN protocol logic. This approach would allow the data extraction and writing to be inherently cross platform and use an already standardized protocol. However, the feasibility of this solution is still being investigated and, as a temporary solution for a proof of concept application, this application was specifically developed for the Aspen HYSYS package. The choice of the software comes from the availability of a functional library in VBA that makes interfacing data with HYSYS easy and fast. Moreover, it provides Excel as a versatile interface. From a theoretical point of view the structure of the solution is not impacted as well as its effectiveness. It is however marginally impacted in the performance, since data extraction and writing in Excel is not optimized for effective data transfer between software. At the current state of development the main aim is investigating the feasibility and convenience of such a solution, while optimization is left for later stages.

2. Methods

The automation process is based on three modules; the simulation package, the evaluation package, the optimization library. Each covers a standalone role and operates completely detached from the others except for a single file collecting the data in input and output of the module.

The simulation package has the role of modeling the physical and chemical problem as well as the characteristic dimensions of the unit and process flowsheet with a consolidated structure. Alongside this feature, already included in all commercial packages, the simulation package must also have an add-on to condense the relevant data in a standardized format and read data from the same format. This add-on could be developed using the CAPE-OPEN protocol and work for most of the currently popular commercial simulation packages, but conceptually, even a non-standard, standalone physical model of the system developed outside of the CAPE-OPEN framework could be optimized using the same library.

The evaluation package role is to provide a performance indicator for the optimization routine to use. The evaluation package used as a case study here is a financial key performance indicator: the Payback time of the process. In order to evaluate the payback time of the process an estimation of both CAPEX and OPEX is necessary. Operative expenditure estimation is rather straightforward since it can be estimated looking at input and output energy and material streams of the process and associating a price to each stream. Capital expenditures on the other side involve a much more complex and widespread set of information that will be discussed in a dedicated section.

The optimization library role is the robust search of the minimum or maximum of the performance indicator in the system domain. The function it covers is strictly numerical and does not involve a physical model which is instead kept separated in the simulation package.

2.1. The simulation package: Aspen HYSYS

The simulation package chosen for the conceptual design is Aspen HYSYS V10 release. The ease of use and already included macro-enabled library for Microsoft Excel VBA allowed to develop the data extraction and writing add-on in a limited amount of time exploiting Excel spreadsheets for user data input and plotting of the results. All the data necessary for the evaluation library are organized in and eXtensible Markup Language file (.xml). The xml file is structured data file organized in nodes to facilitate accessibility. For this specific case study the nodes are organized as follows:

- **Options:** this attribute includes the economic library used for the estimation, its parameters, and the units of measure
- **Streams:** this attribute includes child nodes corresponding to the single material stream and the thermodynamic properties of each
- **UnitsList:** this attributed groups all the supported unit operations of the simulation
- **Utilities:** this attribute groups all non material streams to be associated with operative expenditures
- **Specifications:** this attribute groups all the degrees of freedom used for each unit operation in the simulation

The combination of Aspen HYSYS and the Excel workbook managing the data processing is responsible for the description of the physical system.

2.2. The evaluation library: Capex/Opex estimation

The data extracted from the simulation into the xml file are read by the Capex/Opex estimation libraries implemented in C++. The economic libraries included in the software reach a study estimate precision on the final cost estimation with an estimated error ranging from +30% to -25%. For the sake of direct comparison two different libraries are included: the Peter and Timmerhaus (2001) library based on percentage of delivered equipment cost and the Guthrie (1974) and Turton

(2012) approach based on the bare module cost. The detailed discussion on this costing approaches is not the topic of this paper and can be explored in detail in . Both costing methods include in the Capital expenditure estimation:

- **ISBL:** Inside Battery Limits cost include purchasing and shipping costs of equipment, piping, catalysts, and any other material needed for final plant operation, or construction of the plant. ISBL costs also include any associated fees with construction such as permits, insurance, or equipment rental; even if these items are not needed once the plant is operational
- **OSBL:** Off site battery limits are defined as utilities, common facilities, and other equipment and components not included in the ISBL definition
- **Engineering and Construction:** indirect expenses associated with the actual building of the plant such as supervision, engineering and legal expenses.
- **Working Capital:** is defined as the money required to start and run the already constructed plant until income can be obtained from the products
- **Contingency:** allows for variation from the predicted cost estimate

The logic behind the Timmerhaus library is based on the assumption that each term of capital investment can be estimated as a proper percentage of the delivered equipment cost. By applying proper multipliers to the cost of the delivered equipment the overall cost of the plant can be estimated with accuracy in the +/- 25% range, but for similar plant configurations it can go as low as +/- 10%.

The Guthrie method operates by computing the cost of a single equipment in standard conditions and then, using appropriate cost factors, it scales the price of the equipment for actual operating conditions accounting for temperature, pressure, material, piping, contingency etc.

When the conventional approach to cost estimation fails due to out of boundary values then each library falls back to less precise, but more versatile methods such as the sixth tenth's rule. While these estimations have errors that be as high as +100% it is still preferable to completely neglect unit costs. Moreover, different exponents are implemented according to the specific unit as recommended by Turton et al. (2018)

Operative expenditures are managed separately and divided as follows:

- **Direct manufacturing costs:** this are the cost that vary with production rate
- **Fixed manufacturing costs:** this costs are independent from production rate
- **General expenses:** this costs are typically not related to production and include expenses such as marketing, management, financial investments and so on

Operative expenditures are known once cost of labor, cost of raw materials, cost of waste management, cost of utilities and fixed capital investment costs are known.

2.3. Robust Optimization: BzzMath library

The optimization problem is carried out by the BzzMath library developed by Buzzi Ferraris Buzzi-Ferraris and Manenti (2014) in C++. The library, developed in house at Politecnico di Milano offers a wide variety of numerical tools for the solution of linear and non-linear complex problems common in the chemical industry. The optimization variable chosen for the economic optimum design is the Payback time defined as:

$$\text{Payback Time} = \frac{\text{Total Initial Capital}}{\text{Average Annual Cash Flow}} \quad (1)$$

the optimization routine task is to look for the design and operating conditions minimizing this value. The convergence iterative procedure is reported in Fig 1. Excel covers 2 main functions: interacting with Aspen HYSYS, and acting as a GUI in which the user can input and visualize data necessary for the computations.

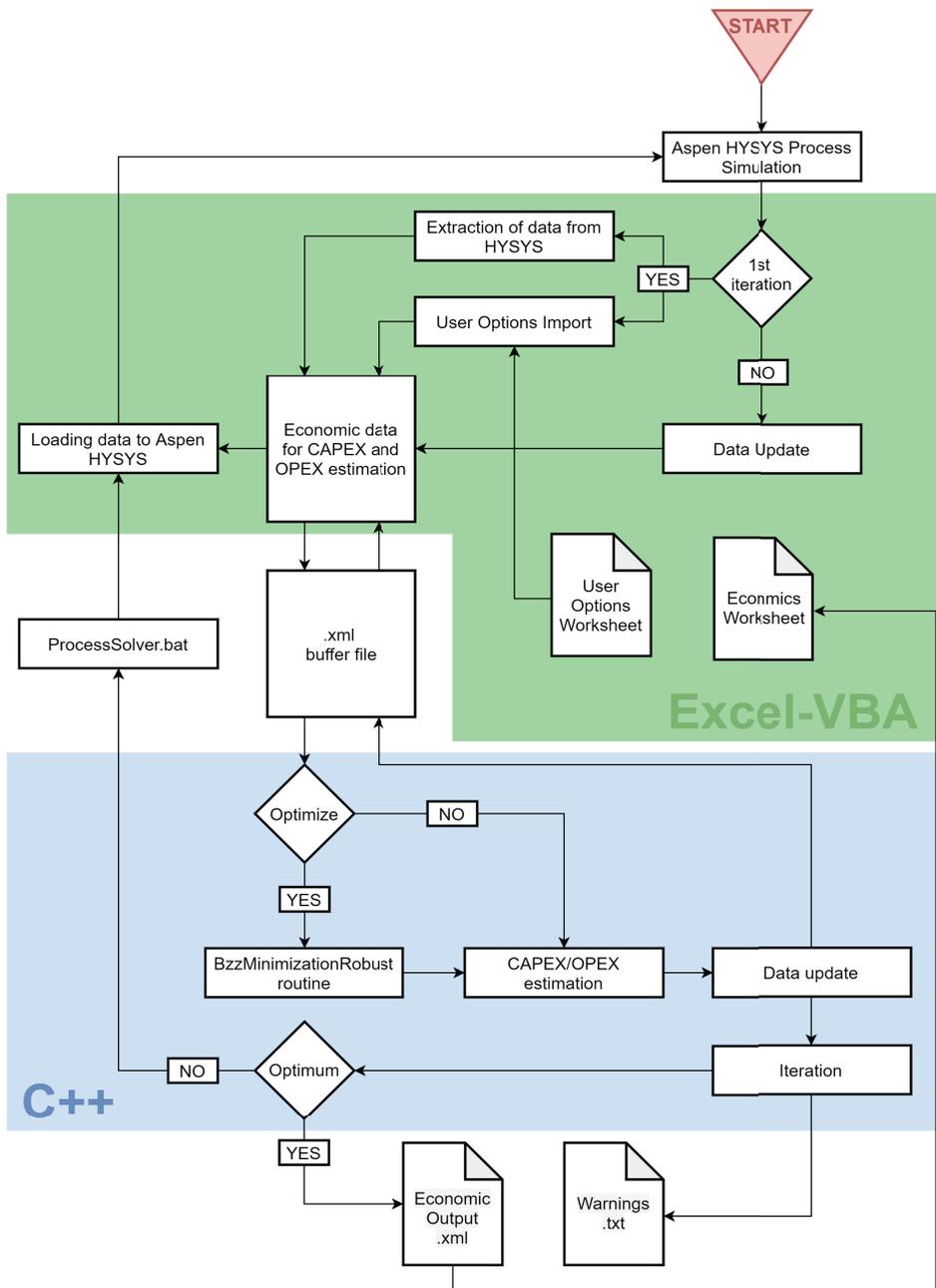


Figure 1: Iterative convergence structure

The xml files act as a buffer between the physical problem and the economic/mathematical one which are both implemented using C++. During the first iteration the data from a converged simulation is extracted into the Excel workbook and divided in Spreadsheets each containing the

information contained in a xml node.

The first spreadsheet contains the directory path to the simulation, the units of measure desired, the economic library of choice. The data extraction routine collects information into separate spreadsheets. The Material streams are organized together and their mass, molar flow as well as temperature and pressure are extracted. In a separate spreadsheet all the material streams are reported again and categorized by the user as raw material stream, process stream, waste stream, product stream or utility stream and associated with a price or value. Moreover, they can be reported as 'degree of freedom' allowing the optimization routine to act on them. If this option is chosen then boundary conditions must be provided by the user.

Following the material spreadsheets are the ones for the unit operations currently supported; compressor, heater, cooler, air cooler, LNG plate exchanger, Pump, Valve, Distillation column, Separator, Absorber, CSTR, PFR, Conversion Reactor, Equilibrium reactor. In each spreadsheet the unit operations of that category in the simulation are grouped and their characteristic dimension is reported and the user can choose the design of unit (i.e. Centrifugal, volumetric, axial for a compressor), the material (Carbon Steel, Stainless Steels of different grades, etc...), whether the unit operation must be optimized, and the optimization variable (output pressure for a compressor) with the possibility to provide upper and lower boundaries. Finally the last spreadsheets contain a set of parameters to customize the economic libraries changing the multiplication factors of both the Turton and the Timmerhaus methods.

3. Conclusions

The overall structure of a robust optimization toolbox was presented in its overarching logical organization discussing the advantages and disadvantages of a compartmentalized optimization routine and presenting each of its modules functions and purpose. This package has been developed using VBA and C++ to interface it with a commercial simulation software as a proof of concept and tested. The program is now entering the testing phase in its alpha version and being tested on custom made simulations. Further development is intended to generate another module to interface with the optimization routine capable of digesting data from the industry to automatically perform a regression on a Guthrie style equation to predict the cost of units for which extended historical data are available.

References

- G. Buzzi-Ferraris, F. Manenti, 2014. Nonlinear systems and optimization for the chemical engineer: Solving numerical problems. Wiley-VCH Verlag, Weinheim, Germany.
- P. Martin, J.-Y. Dantan, A. Siadat, 2007. Cost estimation and conceptual process planning. In: Digital Enterprise Technology. Springer US, Boston, MA, pp. 243–250.
- R. Turton, J. A. Shaeiwitz, D. Bhattacharyya, 2018. Analysis, synthesis and design of chemical processes, 5th Edition. Prentice Hall.

A soybean supply chain model to analyze the greenhouse gas emissions of the transport sector

Milagros Verrengia^a, Aldo Vecchietti^a.

^a*INGAR–CONICET, Avellaneda 3657, Santa Fe (3000), Argentina*
mverrengia@santafe-conicet.gov.ar

Abstract

This article presents a mathematical model of the soybean's supply chain for Argentina where the different stakeholders and the material flows among them are represented. The transport used in this sector are trucks, trains, and river ships. The objective is to analyze the emissions of greenhouse gases (GHG) generated by the transportation in this sector using electric trucks as an alternative to biodiesel ones.

The model generated is a mixed multi-period / multi-objective linear integer model, destined at minimizing operating and GHG emissions cost. The accuracy of the model is compared against two statistical studies made by the Transport Agency of Argentina in 2014 and 2017 regarding the soybean transportation. The results show a good fit with those reports. Two scenarios are compared, in the first one only biodiesel trucks are used for transportation, while in the second one trains, barges and electric trucks are included. Results show the tradeoff between investment costs and reduction of emissions where it is possible to achieve a 60% GHG decrease, which is far to compensate for the investment cost.

Keywords: soybean, supply chain, emissions, transportation

1. Introduction

Argentina has an important role as a privileged supplier of food to the whole world, it is the world's leading exporter of soybean meal, for other grains, such as corn and sunflower, it is among the top five exporters. Therefore, the agri-food sector has an important weight in the Argentine economy, both in production and in employment. It is the most important sector that generates foreign exchange and makes an enormous contribution to geographical equality given its wide territorial presence in the country [1]. Due to its importance, it is of great interest to study agri-food supply chains (SC), it is necessary to generate strategic information on market trends, their operating dynamics, the characterization of their main actors and production strategies [2].

There are works in the literature related to the proposal of this article. For reasons of space, we are only going to mention some of them. Courtonne et al. [3] carry out an analysis of the flow of materials in the cereal supply chain in France. The authors apply a material flow analysis (AFM) which is a systematic evaluation of flows and link them to the stock of materials in space and time. Mogale et al. [4] study the problem of storage and transportation of wheat in bulk in India, in a two-stage supply chain network that depends on the Public Distribution System (PDS) of that country. To do this, they develop an integer mixed non-linear programming model (MINLP), whose objective is to minimize the transportation, storage, and operating cost of food grain from producing states to consuming states. They employ a heuristic algorithm based on Taboo Search to solve the problem. He and Li [5] address the problem of harvesting and

transporting wheat in China, they propose a joint optimization framework and a general algorithm to optimize the wheat harvesting and transportation problem considering fragmented farmland.

Our approach proposes to analyze the soybean supply chain, its logistic from the harvesting to the industrial transformation in conjunction with the generation of greenhouse gases involved in the transport. A real-world representation of the soybean's supply chain is made. The objective is to evaluate economic and environmental issues, for this purpose, different transportation media are included in the model: diesel and electric trucks, trains and barges. The purpose is to provide an analysis tool for the optimization of the chain's operation. A MILP model (a linear mixed-integer program) was generated. Unlike the proposals of other authors mentioned, our model is a deterministic one, and there are no known cases involving an economic and GHG emission analysis of a real-world agro-industrial supply chain.

2. Methodology

2.1. Description of the problem

To reproduce the operation of the soybean supply chain, an important amount of data is collected from different public and private sources regarding the sowing, harvesting, production, transportation and export of soybeans and by-products. For our model, a three-level supply chain is proposed: the first corresponds to the production areas (i), where the soybean grain is harvested, the second level is formed by the stockage of grains in silos (j) from where grains are sent to the third level which corresponds to the destination zones (k) where the soybean grain is demanded for grinding or exportation. In Fig. 1 you can see the operation of the chain and its flows from production to its destination. The largest percentage goes to the industrial sector to produce flours and oils for exportation purposes.

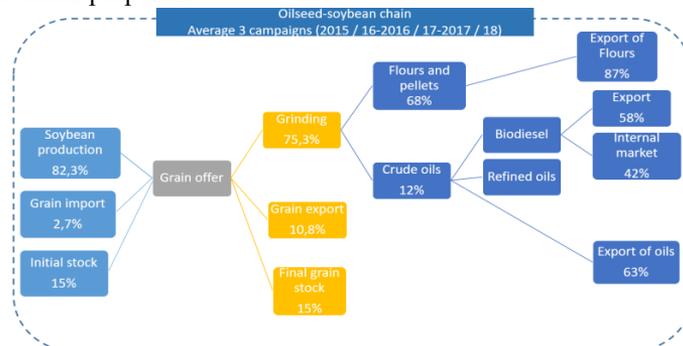


Fig. 1: Soy supply chain and flows between the different levels.

In Fig. 2, the map on the left shows the location of soybean production and the industrial establishments that process it in Argentina, on the right a graph of Argentine soybean production in millions of tons per year between years 2006 and 2018 and the percentage rate of change with respect to the previous year.

The most important assumption made in the model is that both the production, stockage and the final destinations areas are divided into different regions, each one composed of a centroid. To measure the circulation between the different levels of the soybean supply chain, the distances traveled between these centroids and the numbers of trips made between them are measured. The model is multiperiod one with 15 time periods, covering five years (from 2016 to 2020) corresponding each one to a four-month period.

To deal with economic and environmental objectives, the weighted sum method is used, to combine both objectives into one objective function. For this purpose, an economic value of the GHG emissions is set based on the value proposed by the Interagency Working Group on Social Cost of Greenhouse Gases [7].

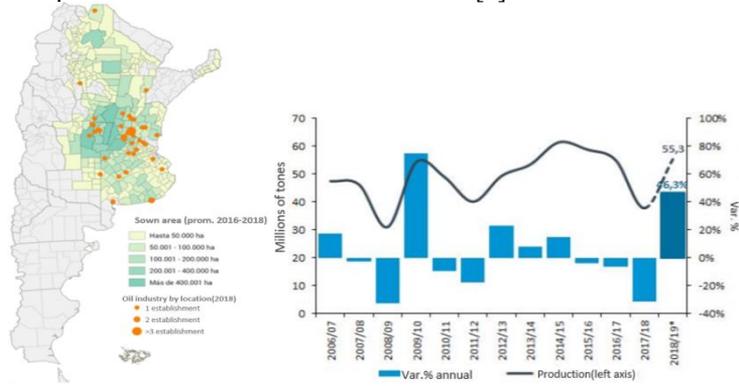


Fig. 2: (a) Map of soybean production, storage, and industries areas. (b) Soy production in annual tons from 2006-2018.

The decision variables of the model are soybean flows, in tons per time t , sent by using trucks from i to j and from j to k the flows delivered by different transports (trucks, trains and river ships), the number of trips made between the different levels and the transports used for each time t , the stock of soybeans (in ton) for each time t at location j and the amount of GHG emissions measured in equivalent tons of CO₂.

2.2. Formulation of the Model

For space reasons, in this section only the most important constraints of the problem are included so that the work done can be visualized, some logical and bounding constraints are omitted.

Eq. (1) and (2) establish that if there is a flow ($EIJcijt$) of soybeans with the type of truck c from production area i to stockpile j at time t , for which the binary variable ($ycijt$) must be equal to 1, the flow must be less than or equal to (Eq. 1) the number of trucks ($NCIJcijt$) times their maximum capacity ($KCMAX$) or greater than or equal to (Eq. 2) to the number of trucks times their maximum capacity ($KCMIN$). Two types of trucks c (1: using biodiesel 80-20 and 2: electric) are considered.

$$EIJcijt \leq NCIJcijt * KCMAX + M * (1 - ycijt) \quad \forall c, \forall i, \forall j, \forall t \quad (1)$$

$$EIJcijt \geq NCIJcijt * KCMIN + M * (1 - ycijt) \quad \forall c, \forall i, \forall j, \forall t \quad (2)$$

Similar restrictions are used for truck, train, and barge shipments from silos storage to consumption areas. These last two transports have limited origin and destination zones compared to those carried out by truck.

Eq. (3) corresponds to the material balance of the silos stockage by time (Ijt), where the inventory at time t is equal to the inventory at time $t-1$ ($Ijt-1$), plus the shipments received of the production area ($\sum \sum EIJcijt - lic$) minus the flows sent by different transports to the consumption areas, both at time $t-1$.

$$Ijt = \sum \sum EIJcijt - lic - \sum \sum EJKcikt - 1kc - \sum EVJKjkt - 1k - \sum EBJKjkt - 1k + Ijt-1 \quad \forall j, \forall t, t > 1 \quad (3)$$

In Eq. (4) the inventory is limited according to the storage capacity of silos.

$$Ijt \leq KACK \quad \forall j, \forall k, \forall t \quad (4)$$

Eq. (5) indicates that the sum of all shipments from all stockpiles to each destination area in time t must equal to or exceed its demand ($DSkt$).

$$\Sigma\Sigma EJKc_jkt_j + \Sigma EVJK_jkt_j + \Sigma EBK_jkt_j \geq cDSkt \quad \forall k, \forall t \quad (5)$$

Eq. (6) indicates that 10% of the truck fleet is changed annually (according to statistics of the trucks market), translated into 3.33% percent for time t . The model allows the purchase of new biodiesel or electric trucks from the first period. At the beginning of the time horizon electric trucks are 0, only biodiesel's are available. The differences are that electric cars are more expensive than biodiesel ones, but CO2 emissions are zero. We also consider that a truck can make up to 20 trips per time.

$$NEWT_{ct} = 0.033 * INITIALT \quad \forall c, \forall t \quad (6)$$

$$NT1_t = NT1_{t-1} + NEWT_{1,t-1} - \Sigma NEWT_{ct-1} \quad \forall t, t > 1 \quad (7)$$

$$NT2_t = NT2_{t-1} + NEWT_{2,t-1} \quad \forall t, t > 1 \quad (8)$$

$$NT_{ct} * 20 \geq \Sigma \Sigma NCJKc_jkt_k \quad \forall c, \forall t \quad (9)$$

2.2.1. Objective function

The objective function (10) is the minimization of the total costs which is function of transportation ($CIJ + CJK$), storage (CS), investment (new trucks) (CI) and emission (CE) costs.

$$CT = CS + CIJ + CJK + CI + CE \quad (10)$$

Eq. (11) represents inventory costs where CI is the financial cost of having the stock.

Eq. (12) is the transportation cost from harvesting areas i to the storage location j , which corresponds to the summation of the transportation cost per km (CAR) times the flow $EIJc_{ijt}$ and the distance between i and j ($DAij$). Similarly, Eq. (13) calculates the costs associated with shipping from the silos j to destination areas k . The first term of Eq. 15 represents a fixed cost of the delivery, the second term is a variable cost depending on the flow and the distance traveled. Then the costs of rail transport are added, corresponding to the train a cost per ton (CR) times the flow of the shipment ($EWJK_jkt$), and the last term is the river transportation, composed by the ton cost per barge (CB) multiplied by the quantity shipped (EBK_jkt).

$$CS = \Sigma \Sigma I_{jtt} * CI \quad (11)$$

$$CIJ = \Sigma \Sigma \Sigma \Sigma CAR * EIJc_{ijt} * DAij_{tj} \quad (12)$$

$$CJK = \Sigma \Sigma \Sigma (CFK * NCJKc_jkt + CVK * EJKc_jkt * DAjk)_{tkj} + \Sigma \Sigma (CR * EWJK_jkt + CB * EBK_jkt)_{tkj} \quad (13)$$

Eq. (14) corresponds to investment cost where CIT is the cost per truck multiplied by the amount purchased per time t . Lastly, Eq (15) calculate the costs of CO2 emission where CTE is the cost per ton of CO2, NTE is the amount of CO2 emitted per truck/(km.ton), similarly for trains (NWE) and barge (NBE).

$$CI = \Sigma \Sigma NT_{ct} * CIT_{cttc} \quad (14)$$

$$CE = CTE * (\Sigma \Sigma NTE * t_{ji} DAij * EIJc_{ijt} + \Sigma \Sigma DAjk * (NTE * EJKc_jkt + t_{kj} NWE * EWJK_jkt + NBE * EBK_jkt)) \quad (15)$$

CAR and CVK costs are \$ 0.072 per ton kilometer of soybeans shipped, data extracted from the Rosario Stock Exchange report [6]. For the cost of CO2 emissions (CTE), a value of 56 dollars per ton is taken from a report of the Social Cost of Carbon, Methane, and Nitrous Oxide, United States Government [7]. The purchase cost of a biodiesel truck for Argentina is about \$ 80,000, while for electric ones the cost reported internationally is about \$ 160,000.

2.3. Results

As a first step, the accuracy of the model was tested by solving two examples. Their results are compared with two statistical works [8-9] carried out in 2014 and 2017 by the Secretary of Cargo Transportation Planning and Logistics of Argentina. In Fig. 3 you can see the comparison between the results of the model and the government's work

for the year 2017, measured in number of trips among centroids. The model presents a very good approximation considering that the government work uses real data from “consignment notes” in the whole year, while in our case the results are provided by the execution of the model

The model is a multiperiod multiobjective mixed integer linear programming problem (MILP) that was posed in GAMS [10] system and solved with CPLEX 12.6.3 on a PC with an Intel i7 processor with 12 GB of RAM. The model consists of 127,249 equations, 204,581 variables, and 105,504 discrete variables solved in 80 seconds of CPU time

To achieve this result, the trucks were loaded with 28 tons. of soybeans on average, the lower bound is 25 tn, while the upper one is 30 tn.

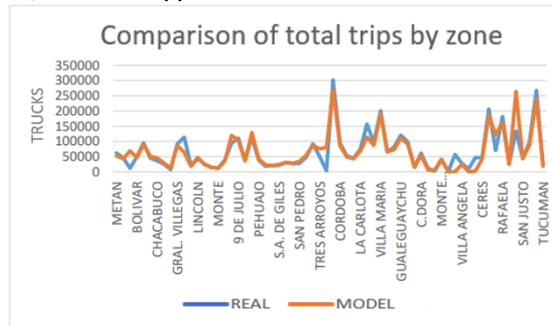


Fig. 3: Comparison of total trips by area between real data and model results for the year 2017.

We have executed two scenarios, in the first one only biodiesel truck are used for transportation and purchasing new ones. Then, scenario 2 includes the use of the railroad and barges for transportation from the silos to the destination areas, and the model considers the possibility of purchasing electric trucks for transportation.

Table 1 shows the results obtained from both scenarios, based on the costs and emissions. In Fig. 4, it can be seen the scenario comparisons.

As can be seen scenario 2 is more expensive in costs than the first one due mainly to the increase in investment costs of electric trucks, although the emissions are lower by about 60%. The cost of the emissions per ton of CO2 equivalent, taken from the USA government report, is too low to compensate for the investment cost in electric trucks. It means that the purchase of electric trucks must have incentives to diminish its price or to force its use.

Table 1: Cost and emissions comparison

COST COMPARISON	SCENARIO 1 [\$]	SCENARIO 2 [\$]
TOTAL COST	24.070.000.000	29.917.813.360
TOTAL TRANSPORTATION COST	4.578.500.000	3.342.866.680
TRUCK TRANSPORTATION COST SECTION 1	996.600.000	745.480.000
TRUCK TRANSPORTATION COST SECTION 2	3.581.900.000	2.473.500.000
RAILWAY TRANSPORTATION COST	-	103.400.000
BARGE TRANSPORTATION COST	-	20.486.680
INVESTMENT COST	15.840.000.000	23.232.080.000
EMISSIONS COMPARISON	SCENARIO 1 [TN]	SCENARIO 2 [TN]
TOTAL EMISSIONS	52.220.000	20.435.000
TRUCK EMISSIONS SECTION 1	11.333.500	5.700.100
TRUCK EMISSIONS SECTION 2	40.891.000	14.659.000
RAILWAY EMISSIONS	-	16.165
BARGE EMISSIONS	-	57.140

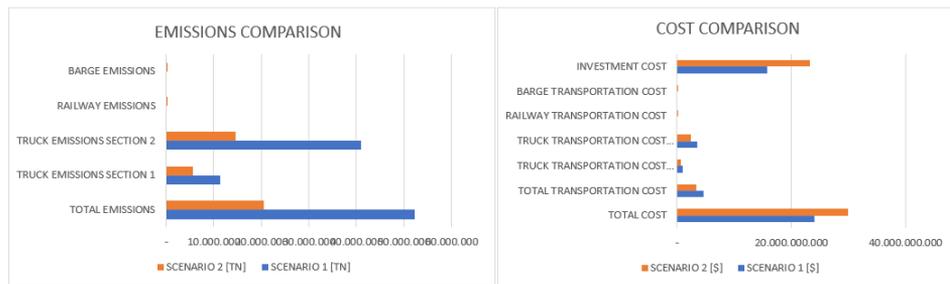


Fig. 4: Cost and emissions comparison

2.4. Conclusion

In this work, the objectives were met in two stages, initially with a model of the supply chain of soybean cultivation for Argentina. The model corresponds to a multiperiod multiobjective linear integer mixed mathematical program (MILP) to represent the supply chain in three levels: production area, silos, and destination areas. Although in our country the soybean industry has been studied from different perspectives, according to the knowledge of these authors, there is no proposal that models the CS and the flows of materials, their transport and GHG emissions with a mathematical program.

The scenarios executed by the models have shown that the investment costs in electric trucks does not compensate for the lower emissions obtained with their use. Therefore, in case of seeking the viability of the electric truck, the government should propose an incentive system that can level the difference in costs and thus promote better care for the environment.

In addition, this work only considers the use of trucks for transporting soybeans only, but the truck fleet is used for many other products that are transported in large volumes throughout the country, its analysis is part of a future work.

References

1. Ódola A., Morra F., Picon N. Año 2020 “Cadenas de valor agroalimentarias”. Recuperado de <https://www.argentina.gov.ar/sites/default/files/cadenasagroalimentarias-febrero2020.pdf>.
2. Giancola S.I, Salvador M.L, Covacevich M, Iturrioz G. “Análisis de la cadena de soja en la Argentina”, Instituto Nacional de Tecnología Agropecuaria, ISSN 1852-4605, (2009).
3. Courtonne J., Alapetite J., Longaretti P., Dupré D. y Prados E. “Downscaling material flow analysis: The case of the cereal supply chain in France”. *Ecological Economics* 118, p.67, (2015).
4. Mogale D.G., Krishna Kumar S., García Márquez F.P. y Tiwari M.K. “Bulk wheat transportation and storage problem of public distribution system” *Computers & Industrial Engineering* 104, p. 80–97, (2017).
5. He P. y Li J. “A joint optimization framework for wheat harvesting and transportation considering fragmental farmlands” *Information Processing in Agriculture*” 8, 1, p- 1-14, (2021).
6. Julio Calzada, Blas Rozadilla. “The heavy weight of bulk trucker freight continues over long distances.”, Year XXXVII, Issue No. 1937, (2020). Retrieved from.
7. Technical Support Document: Social Cost of Carbon, Methane, and Nitrous Oxide Interim Estimates under Executive Order 13990. Intergovernmental Working Group on Social Cost of Greenhouse Gases, United States Government. February 2021.
8. Matrices origen y destinos de cargas 2014. Ministerio de transporte de la Argentina, 2014.
9. Transporte Terrestre de Cereales y Oleaginosas-2017. Ministerio de transporte de la Argentina, 2019.
10. Brooke A., Kendrick D., Meeraus A., Raman R. and Rosenthal R.E. GAMS A User’s Guide; GAMS Development Corporation: Washington,DC, 1998.

Ontology for Enhanced Industrial Process Control

Renata Samara Rodrigues de Sousa,^a Song Won Park ^a

^a *Polytechnic School, University of São Paulo. Av. Luciano Gualberto 380. São Paulo 05508010 Brazil.*

renatasousa@usp.br

Abstract

The cyber-physical integration in the sense of industry 4.0, applied to industrial process control, needs to develop new methodologies. Both top-down hierarchies of commands for Enterprise Resource Planning, Optimization, Advanced Control, Local Control, and bottom-up flux of information from the plant floor to control system, optimization level, and planning strategy are dissolved, aiming to enhance the vertical and horizontal integration and flexible operability. Even though this is a requirement for Cyber-Physical Systems, it would be merely a coexistence of advanced optimization IoT technologies states of art. Therefore, there is a need for a new architecture of a functional system. Since it works essentially as an event-based system, ontology plays a key concept in the practical working of the new process control. First, an overview of the general ontologies framework applied in Industry 4.0, and on the other hand, some ontologies published for chemical process design are discussed extensively here. Then the practical ontology and taxonomy needed for industrial process control are discussed. Here is presented how to recover the functional layers to each specific application of former traditional hierarchy top-down and bottom-up, now named as semantic layers. A description's language development for process control using this ontology is another challenging task, also proposed. A case based on the transition of the traditional batch reactor process to modern industry 4.0 application illustrates the change of the operation mode. Finally, the potential gains and technology limitations are analyzed as a critique as the enabling technologies (or pillars) of Industry 4.0 to show what concepts apply to this enhanced process control. It also compares and analyzes this modern approach in the manufacturing process and the difference with the industrial process control. The construction of a complete system based on ontology and description language ready for application is an impressive task to be developed further. The main objective of the work, for now, is to clarify the concepts and show the methodology with its practical application, discussing limitations.

Keywords: Cyber-Physical Integration, Industry 4.0, Ontology, Process Control.

1. Introduction

The storm of technological advances driven by the German initiative Industrie 4.0 (I4.0), induces process system engineering to absorb a new format, where evolutions will be leveraged by intelligent production systems, vertically and horizontally integrated (Schwab, 2017). With event-driven architecture, microservices, and componentization of software elements there is a need for new methodologies aimed to apply industrial process control for digital transformation, where ontologies are a fundamental part of information

reuse and system interoperability. There are several frameworks and reference architectures proposed for the I4.0 scenario. The architectural proposal IMC AESOP (Karnouskos et al., 2014) is an example that offers an extension to the traditional model found in the ISA95 IEC 62264 standard for well-defined implementation of a network-structured modularized architecture. Kumar et al. (2019) discusses important aspects for I4.0 and survey the ontologies, considering the effort and importance of standardization in the scenario.

To consider the historical evolution of information technology and ontology design for chemical processes is a great path to understand PSE for I4.0 and industrial process control. Batres (2017) provides an ontology review for process systems engineering, discussing the use of different ontologies for various purposes in PSE: Multipurpose Ontologies, such as OntoCAPE (Ontology for Computed Aided Process Engineering), and ISO 15926, and discusses some methodologies for proposals to develop the well-designed ontology. See also Morbach et al. (2007), Schneider et al., (2019), Johnsson and Brandl (2021), Burns et al. (2019), Guarino (1997), and Lin and Harding (2007).

2. What you should know about Ontology Design in PSE

2.1. Before you start

There are no rules when choosing a method to engineer an ontology, and there is no explicit method that is better than others. On the other hand, there are some good practices that one should follow to build an ontology.

Ontology design must go from its idealization, through the definition of the scope, to the definition of classes, subclasses, their hierarchy, the definition of properties and their restrictions and definition of instances, and not just representing something through a drawing. Furthermore, the designer should try it on the domain of application and, using reasoners and validating test the consistency ontology. This session suggests a train of thought aiming to help those who would like to design an ontology for process system engineering.

2.2. What is an Ontology?

An ontology is a formal representation of something that exists in the real world (Gruber, 1995). As it happens when different people try to describe something, particular aspects likely appear in the same domain depending on the expertise of the ontology designer. To design a great ontology, one must be familiar with this concept.

In terms of knowledge domain representation, an ontology defines the concepts for entities present in a given domain and their relationships. As such, a great way to manage knowledge and enable flexibility and interoperability between systems through formal semantics (Yang et al., 2019). Arp et al. (2015) explain interesting concepts for ontology design.

2.3. Know your Domain

In the case of chemical processes, there is a tendency to think of a chemical process as a set of domains grouped by their characteristics (unit operations), which overturns a definition of classes, subclasses, and information, in a certain intuitive way. To build a great ontology, one should remember to define its scope well, raise questions that need to be answered by such ontology, and answer them. Also, one should keep in mind that this is an iterative process and that you can and should use methodologies to find inconsistencies (Abburu, 2012) (Chen, 2020) (Geng, 2018).

2.4. Reuse if it is possible

In addition to knowing your domain, designers should choose to reuse ontologies to avoid redundancies. The idea is to consult known ontologies with potential applications for the desired domain, try to use it, and propose an expansion, if applicable. An ontology aims to promote interoperability between systems with heterogeneous data. Therefore, reuse is something that, by its nature, must be considered (Wang et al., 2020).

2.5. Define Classes and Subclasses and Specify the Taxonomic Hierarchy

As mentioned in item 2.1, there is not a single methodology to structure classes and subclasses. On the other hand, the most common approaches are: top-down, when the reasoning goes from the most general concepts to their specializations, bottom-up, reverse process to the previous one, where you start from the more specific concepts and combination, highlighting the most salient ones and then generalizing and specified according to their nature.

2.6. Formalize Slots and Facets

One must define properties (slots) and slot restrictions (facets) for your ontology. Through the process control perspective, we can denominate properties accordingly to the variables involved in a given domain and its facets according to the boundary conditions (constraints) that each variable must-have for that process to run under its proper operating conditions.

2.7. Create Instances

After formalizing slots and facets, the next step is to set individual instances of classes in the hierarchy. One must choose a class and then create an individual instance in order to fill the slot values and the possible restrictions. Noy and McGuinness (2001) propose a guide for beginnings in the subject matter.

3. Case of Study: Batch Reactor Plant for I4.0

The semantic abstraction layers metamodel foundations of OntoCAPE were applied, see more in section 3.2. As presented by Engel et al. (2018), this article considers three abstraction layers: Upper Layer, Domain Layer, and Application Layer. On the other hand, we aim to focus on the approach of the control processes so that the batch reaction takes place. In the following sessions, the authors describe the design methodology in the mentioned domain.

3.1. Domain Description

The authors chose a classic batch reactor control scheme proposed by Zoss (1979). The intent is to represent this classic scheme in new ways, illustrating the change in the operating model and the transition from the traditional batch process to Industry 4.0. (Figure 1).

3.2. Designing the Ontology

The *Rbatch* domain, which represents the entire batch system (considered in this article as a part of OntoCAPE's Chemical Process System Partial Model), has four subclasses, classified in this article as Chemical Process System functions or CPS_function. Three of them according to the type of control we found in the example: *I/O* (interlocks and security alarms), *Scheduling* (Routine of reactor operation batch), and *Split_Range_Control* (temperature control through the reactor jacket). The reaction process is represented by *Reaction* ($A+B \rightarrow C$). Considering each of these subclasses, we

break the application layer into processes (Chemical Process System Realization) that happen during the operation.

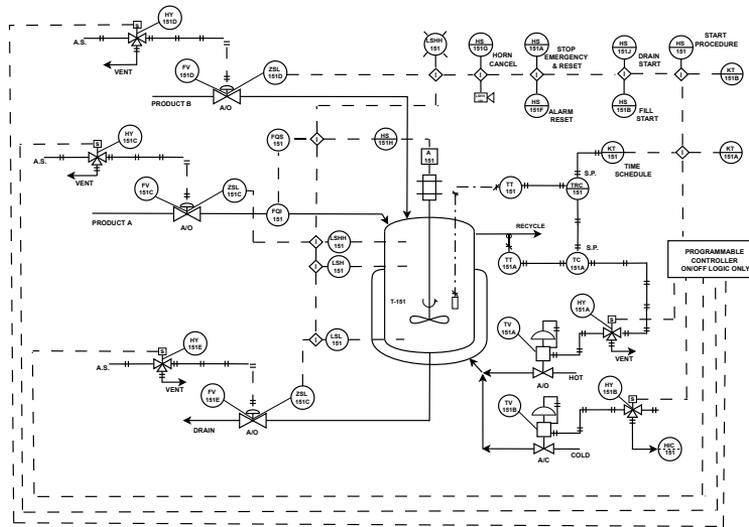


Figure 1: RBatch Domain.

3.3. Ontology Application

In order to apply the ontology in the RBatch domain, the authors chose Protégé, a software for ontology design, as in Figure 2. Furthermore, the ontology format is in OWL (Ontology Web Language) following Antoniou and Van Harmelen (2004).

The recommended practices follow McGuinness and Van Harmelen (2004). OWLViz and OntoGraph were used for ontology visualization as in Figure 3. The ontology modeling contemplates the representation of different control layers associated with the domain. The validation of restrictions of property values and the respective bindings were tested using the HemiT reasoner. The authors are working on an orchestration algorithm so that the ontology could be used in a proper way to facilitate interoperability for level, temperature, and reactor operating scheduling controls, considering alarms for level and temperature.

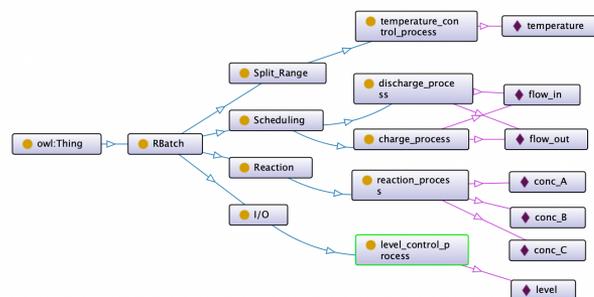


Figure 2: OntoGraph Visualization for class hierarchy view.

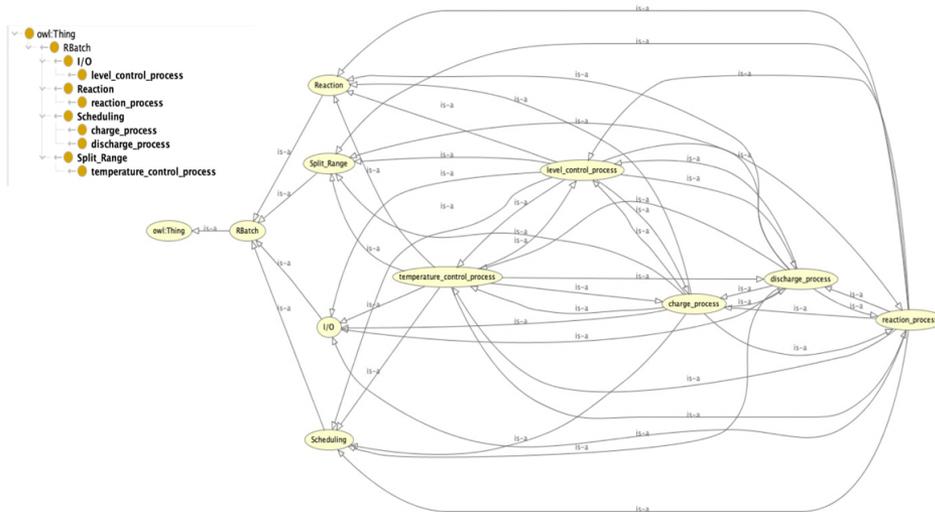


Figure 3: Inferred Hierarchy on OWLViz.

4. Conclusions

The conversion of the classic batch reactor control into an interoperable and flexible environment can be found elsewhere. Industry 4.0 as a concept of the IIoT evolves to Digital Transformation and Cyber-Physical Integration. Beyond the “enabling technologies” (Big Data & Artificial Intelligence, Cloud & Fog Computing, Vertical & Horizontal Integration, Virtual & Augmented Reality, Industrial IoT, Additive Manufacturing & 3D Printing, Autonomous Robots, Digital Twins, Cybersecurity, and so on), nowadays the “pillars” of more organizational concepts such as Technology and Business Transformation, Analytics and Information Management, Intelligent Automation, and Business Planning and Execution emerge with great interest. So, ontology as formal representation plays a key role in Product Design and Development, Process and Plant Design, Safer Processes and Plants, Supervision and Control, Work Processes, Process Systems Modeling.

The construction of a complete system based on ontology and description language ready for application is an impressive task to be developed further. Hence, the main objective of the work is to clarify the concepts and show the methodology with its practical application. There are no fully automatized constructions of Ontology for Process Control. Although OWL provides powerful help for the construction of formal ontologies, the propagation or chaining of properties in the multiple hierarchies interoperating in process control requires more direct experience in the practical applications than theoretical methodologies. Considering the use of the nomenclature Domain, Reusability, Classes and Subclasses, Taxonomic Hierarchy, Properties, Restrictions, and Instances, more familiarity with real control applications will organize the practical differences of ontology application in the PSE domain from Semantic Web Ontology.

Acknowledgments

This study was financed in part by CAPES (Coordination of Superior Level Staff Improvement) - Finance Code 001 process 88887.336319/2019-00 e FAPESP (The São Paulo Research Foundation) process 17/50343-2.

References

- D., Geng, Y., Wen, and Q. Gao, 2018, Research on Semantic Extension Method of Real-time Data for OPC UA.
- G. Antoniou and F. Van Harmelen, 2004, Web ontology language: Owl. In Handbook on ontologies (pp. 67-92). Springer.
- R. Arp, B. Smith, and A. D. Spear, 2015, Building ontologies with basic formal ontology. Mit Press.
- R. Batres, 2017, Ontologies in process systems engineering. *Chemie Ingenieur Technik*, 89(11), 1421-1431.
- T. Burns, J. Cosgrove and F. Doyle, 2019, A Review of Interoperability Standards for Industry 4.0, *Procedia Manufacturing* 38, 646-653.
- G., Chen, T., Jiang, M., Wang, X., Tang, & W., Ji, 2020, Modeling and reasoning of IoT architecture in semantic ontology dimension. *Computer Communications*, 153, 580-594.
- G. Engel and T. Greiner, 2018, Ontology-Assisted Engineering of Cyber-Physical Production Systems in the Field of Process Technology, *IEEE Transactions on Industrial Informatics*, 14(6), 1551-3203.
- N. Guarino, 1997, Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. In International Summer School on Information Extraction, pp. 139-170. Springer.
- T. R. Gruber, 1995, Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies*, 43(5-6), 907-928.
- C. Johnsson and D. Brandl, 2021, Beyond the Pyramid: Using ISA95 for Industry 4.0/Smart Manufacturing. InTech, 14-20.
- S. Abburu, 2012, A survey on ontology reasoners and comparison. *International Journal of Computer Applications*, 57(17).
- S. Karnouskos, A. W. Colombo, T. Bangemann, K. Manninen, R. Camp, M. Tilly and M. Graf, (2014). The IMC-AESOP architecture for cloud-based industrial cyber-physical systems. In *Industrial Cloud-Based Cyber-Physical Systems* (pp. 49-88). Springer, Cham.
- V. R. S. Kumar, A. Khamis, S. Fiorini, J. L. Carbonera, A. O. Alarcos, M. Habib, and J. I. Olszewska, 2019, Ontologies for industry 4.0. *The Knowledge Engineering Review*, 34.
- H. K. Lin and J. A. Harding, 2007, A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration. *Computers in Industry*, 58(5), 428-437.
- D. L. McGuinness and F. Van Harmelen, 2004, OWL web ontology language overview. W3C recommendation, 10(10).
- J. Morbach, A. Yang and W. Marquardt, 2007, OntoCAPE—A large-scale ontology for chemical process engineering. *Engineering applications of artificial intelligence*, 20(2), 147-161.
- N. F. Noy and D. L. McGuinness, 2001, Ontology development 101: A guide to creating your first ontology. Technical Report KSL-01-05 and SMI-2001-0880, Stanford Knowledge Systems Laboratory and Stanford Medical Informatics.
- K. Schwab, 2017, The fourth industrial revolution. Crown Publ.
- G. F. Schneider, H. Wicaksono, and J. Ovtcharova, 2019, Virtual engineering of cyber-physical automation systems: The case of control logic. *Advanced Engineering Informatics*, 39, 127-143.
- R. S. R. Sousa, G. R. Taira, S.W. Park, 2019, Cyber-Physical Integration, Programming, Interlockinh and Control System for a Batch Reactor. *The Journal of Engineering and Exact Sciences*, 5(5), 0424-0432. (in Portuguese)
- X. Wang, H.,Wei, N. Chen, X. He, and Z. Tian, 2020, An Observational Process Ontology-Based Modeling Approach for Water Quality Monitoring. *Water*, 12(3),715.
- L. Yang, K. Cormican, and M. Yu, 2019, Ontology-based systems engineering: A state-of-the-art review. *Computers in Industry*, 111, 148-171.
- L. M. Zoss, 1979), *Applied Instrumentation in the Process Industries*. Gulf Pub Co.

Complementing Natural Gas Driven Syngas with Optimum Blends of Gasified Biomass Waste

Ahmed AlNouss, Gordon Mckay, Tareq Al-Ansari*

*College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation,
Doha, Qatar*

*talansari@hbku.edu.qa

Abstract

Countries around the world strive to diversify their energy portfolio with a suitable and sustainable alternative to fossil fuels, whilst achieving the reduction in environmental impacts from released wastes. The co-conversion of biomass wastes and natural gas (NG) has received much attention due to the potential improvement in downstream power and production of fuels, while minimizing greenhouse gas (GHG) emissions. This study explores the optimal blending of synthesis gas generated from biomass wastes and NG feeds. Aspen Plus is utilized to develop the models of biomass and NG steam gasification considering Qatar's biomass and NG characteristics. Three types of biomass wastes; date pits, sludge and manure are gasified to generate the H₂-rich syngas which is blended later with the NG-driven syngas. The simulated flowsheets have then been used to optimize the blending of downstream generated syngas by means of manipulating the biomass wastes and NG feeds. The optimization problem is constrained by the downstream quality of produced syngas to be utilized for the generation of power and fuels. Typically, the generation of syngas involves high-cost subsequent purification prior to the production of downstream value-added products. However, the optimization attained in this study lowers the requirement of further syngas purification and waste removal, through the blending of NG and biomass-driven syngas and minimization of gasifying agents. This requirement can be further reduced by manipulating reaction agents and process conditions. The result of the optimization problem demonstrates an increase in biomass wastes utilization with the increase in syngas quality constraint. Dates pits biomass dominated the biomass utilization with a lower contribution from sludge and manure wastes.

Keywords: Natural Gas, Biomass, Co-Conversion, Optimization, Syngas, Blending.

1. Introduction

There is continuous growing demand for energy and conversion of natural gas (NG) into downstream value-added products given the large reserves of fossil-fuels worldwide. This is coupled with the growing need of utilizing renewable energy sources in the global energy mix to support environmental protection, and reduce the harmful emissions associated with fossil fuel burning. Therefore, the concept of complementing natural gas utilization networks with renewable energy sources has gained global attention within sustainable development and natural resources preservation targets (M. Wright et al., 2015). Biomass gasification as a renewable process has the potential to contribute in the downstream production of power and fuels and minimization of harmful emissions (Shahbaz et al., 2021). This study investigates the optimal co-conversion blends of biomass wastes and NG to meet the downstream quality of syngas. Aspen Plus is utilized

to develop the models of biomass and NG steam considering Qatar's biomass and NG characteristics. Three types of biomass wastes; date pits, sludge and manure are gasified to generate the H₂-rich syngas, which is blended later with NG-driven syngas. The simulated flowsheets are then used to optimize the blending of downstream generated syngas by means of manipulating the biomass wastes and NG feeds. Moreover, the quantities of gasifying agents required to convert the biomass and NG feedstock into syngas are minimized. The ultimate aim is to maximize the syngas production, while reducing the consumption of NG and gasifying agents and the process environmental emissions. The attention on this topic has been increasing recently with focus on the enhancement of process efficiency and reduction of environmental impacts. Wright et al. (2015) studied enhancing biofuels production from biomass with NG feedstock and using the Fischer–Tropsch (FT) technology. The study demonstrated the contribution of NG in enhancing FT biofuel production through achieving required H₂:CO ratio while reducing the requirement for water gas shift. However, the NG contribution had to be limited to 19% to eliminate the increase in process emissions. Caligiuri et al. (2021) explored experimentally the impact of completing bio-syngas in spark ignition engines with NG on the process efficiency and emissions. The experimental results demonstrated the impact of different feedstock blends on the power derating, where increasing syngas quantity in the mixture decreases power output by 2-6% in addition to NO_x and THC emissions. Hence, only minor power derating is anticipated considering the significant reduction in volumetric energy density. Zhang et al. (2018) investigated the techno, economic and environmental benefits from the co-processing of NG and biomass feedstock to produce liquid fuels. The results demonstrated that coupling biomass with NG increases the economic profitability of producing liquid fuels production, albeit with higher environmental impacts. A maximum NG limit of 28% was determined to meet sustainability targets with a minimum selling price of \$2.75 per gallon gasoline equivalent.

This study applies a unique optimization nature problem to identify the optimum blends of various biomass feedstock and NG to meet multiple end-use downstream options of H₂-rich syngas. The biomass gasification model has been detailed in an earlier study that utilizes oxygen and steam as gasifying agents (AlNouss et al., 2020), utilized to address energy-water-food nexus applications (AlNouss et al., 2019b), generate value-added products while involving captured CO₂ (AlNouss et al., 2021), and solve the superstructure optimization of downstream syngas utilization (AlNouss et al., 2019a). The optimization attained in this study is deployed to reduce the requirement of further syngas purification and waste removal, through the blending of NG and biomass-driven syngas and minimization of gasifying agents.

2. Methodology

Aspen Plus software has been utilized to develop the flowsheets of biomass and NG steam gasification as illustrated in Figure 2. The flowsheets are simulated under steady state and isothermal operation, neglected tar formation, char as a pure carbon (C), atmospheric pressure and negligible pressure drop. The fuel-bounds of chlorine, sulphur and nitrogen are converted to hydrogen chloride, hydrogen sulphide, and ammonia, respectively. The simulation model employs Peng-Robinson equation of state with Boston-Mathias modifications to account for the nonpolar and real components. The various biomass feedstock are blended in a mixer and fed to a drying and decomposition step to yield the conventional components from the non-conventional attributes of the three biomass wastes presented in Table 1. The effluent stream is purified from ash and a small portion

of char utilized to supplement the gasifier with heat through combustion. The purified stream enters the gasifier along with steam to yield syngas which is then mixed with the syngas generated from the steam reforming of NG fed according to the compositions in Table 2.

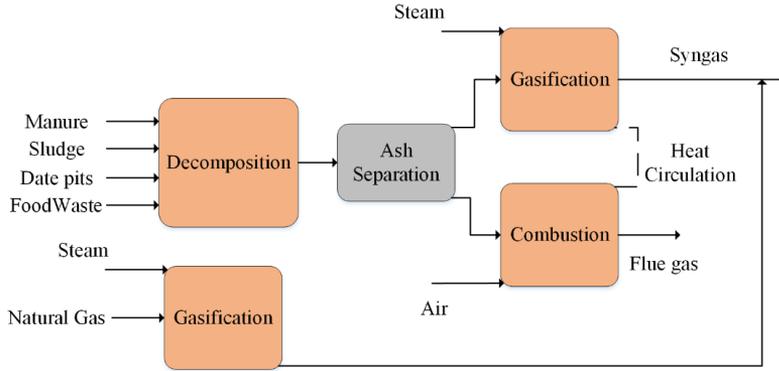


Figure 1: Overall block diagram of NG and biomass co-processing.

Table 1: Attributes of biomass wastes (AlNouss et al., 2018).

Biomass	Date Pit Waste	Dried Sewage Sludge	Manure
Moisture (wt %)	5.0	8.3	27.4
Proximate analyses (dry basis, wt %)			
Ash	1.0	71.8	21.6
Volatile matter	81.8	8.8	65.0
Fixed carbon	17.2	19.4	13.5
Ultimate analyses (dry basis, wt %)			
Ash	1.0	71.8	21.4
S	0	0.1	0.5
Cl	0	0	1.0
N	4.5	1.1	3.7
H	6.8	2.3	5.1
O	37.9	5.7	31.4
C	49.8	19.1	37.1

Table 2: NG feedstock compositions

Component	N ₂	CO ₂	CH ₄
Composition (mol.%)	9	1	90

The simulated flowsheets are used to optimize the blending of downstream generated syngas by means of manipulating the biomass wastes and NG feeds. The optimization is constrained by the potential H₂:CO ratios for further utilization of the generated syngas in different downstream applications. In this study, the downstream applications considered are ammonia/urea production (H₂:CO = 3), Fischer-Tropsch liquid fuels production (H₂:CO = 2), aldehydes production (H₂:CO = 1) and methanol production ($\frac{y_{H_2} - y_{CO_2}}{y_{CO} + y_{CO_2}} = 2$). This is reflected in the formulation of the optimisation problem presented in Eqs. 1-7. Moreover, the quantities of gasifying agents required to convert the biomass and NG feedstock into syngas have been minimized. The ultimate aim is to maximize the syngas production while reducing the consumption of NG and gasifying agents and the process environmental emissions.

$$\begin{aligned}
 & \text{Maximize} && \sum_{i=1}^n x_i \cdot \text{Syngas} - \text{Steam}_{NG} - \text{Steam}_{Biomass} \\
 & && \forall i \in \text{Biomass and NG Sources}
 \end{aligned} \tag{Eq.1}$$

$$\begin{aligned} \text{Subject to } \sum_{i=1}^n x_i &= 1 && \text{Eq.(2)} \\ \frac{y_{H_2} \cdot \text{Syn gas}}{y_{CO} \cdot \text{Syn gas}} &= \text{Application unique} && \text{Eq.(3)} \\ m_i &\leq 2,000 \frac{\text{kg}}{\text{h}} && \text{Eq.(4)} \\ m_{\text{Steam}_{NG}} &\leq 5,000 \frac{\text{kg}}{\text{h}} && \text{Eq.(5)} \\ m_{\text{Steam}_{Biomass}} &\leq 4,000 \frac{\text{kg}}{\text{h}} && \text{Eq.(6)} \\ \frac{y_{H_2} - y_{CO_2}}{y_{CO} + y_{CO_2}} &= 2 \quad (\text{methanol production case}) && \text{Eq.(7)} \end{aligned}$$

Where, steam is the mass flowrate of gasifying agent, Syngas is the mass flowrate of the produced H₂-rich gas, y_{CO}, y_{CO₂} & y_{H₂} are molar fractions of carbon monoxide, carbon dioxide and hydrogen, m is the mass flowrate of NG and biomass wastes, and x is the blending fraction of NG and biomass wastes.

3. Results and discussion

The results considered for the optimization problem and compared with the base case are:

1. Optimization of feedstock blending;
2. Optimization of gasifying agent;
3. Optimization of feedstock blending and gasifying agent.

The results of the base case are summarized in Table 3 with the equal contribution of biomass and NG feedstock presented in Figure 2.

Table 3: Base case data.

S/B Ratio	S/NG Ratio	Syngas yield	H ₂ /CO	CO ₂ % in syngas
0.74	3	2.08	4.59	9%

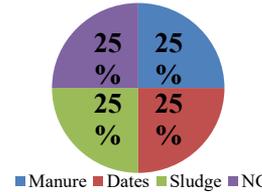


Figure 2: Base case feedstock blending.

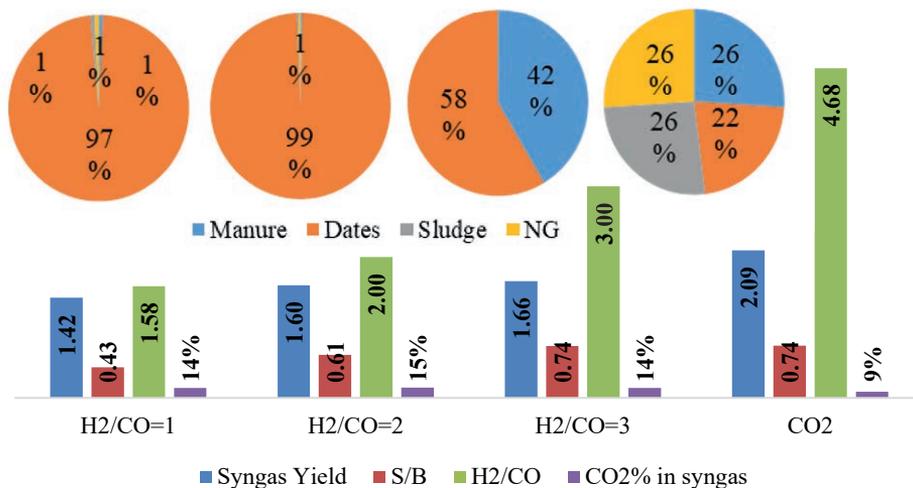


Figure 3: Case 1 results.

The results of case 1 presented in Figure 3 demonstrate the domination of dates biomass feedstock for the generation of H₂:CO ratios of 1 and 2. Whereas, the H₂:CO ratio of 3 presented the contribution of manure and dates mainly in the blended feedstock, while NG is only contributing in the CO₂ ratio case. The syngas yield is reduced slightly from the base case while the CO₂ emissions increased. The requirement of steam for biomass gasification is reduced compared to base case. For case 2, the results presented in Figure 4 demonstrate the contribution from NG and different biomass wastes in the feedstock blends. The syngas yield is increased slightly compared to case 1 while the CO₂ emissions are reduced. The requirement of steam for biomass and NG gasification is increased compared to case 1.

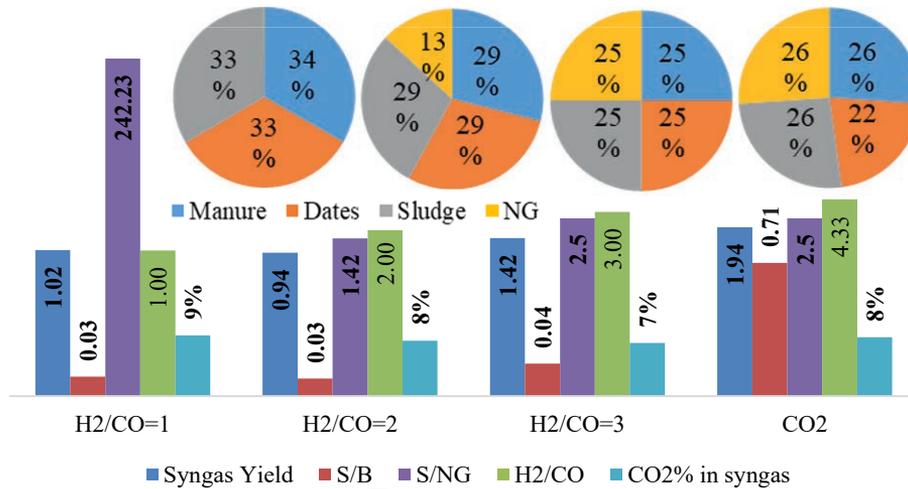


Figure 4: Case 2 results.

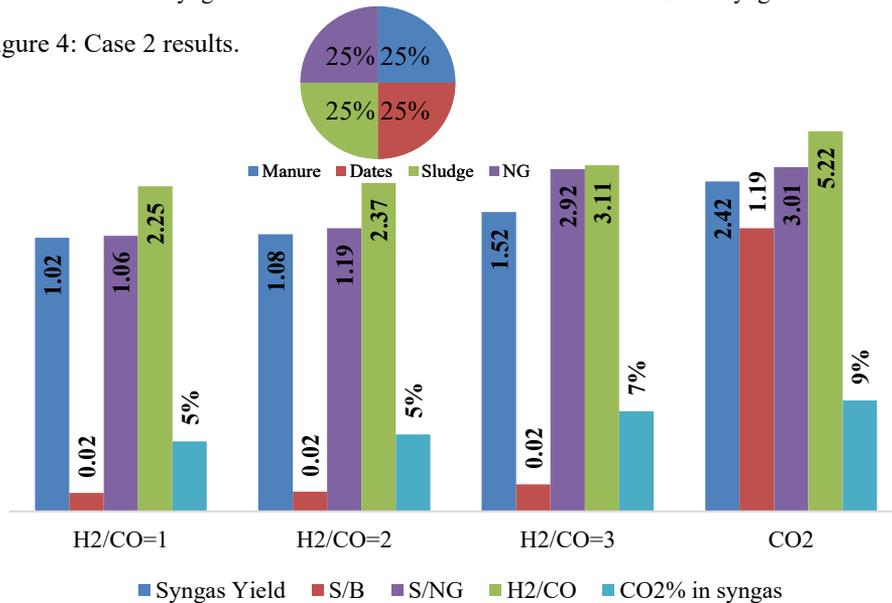


Figure 5: Case 3 results.

The results of case 3 presented in Figure 5 demonstrate equal contribution from NG and different biomass wastes in the feedstock blends. The syngas yield is similar to case 1, while the CO₂ emissions are reduced. The requirement of steam for biomass and NG gasification is decreased as compared to case 2.

4. Conclusion

The concept of complementing the natural gas utilization networks with renewable energy sources contributes to various sustainable development goals in parallel. This study applies a unique optimization problem to identify the optimum blends of various biomass feedstock and NG to meet multiple end-use downstream options of H₂-rich syngas. The outcomes of the optimization problem demonstrates an increase in biomass wastes utilization with the increase in syngas quality constraint. Dates pits dominated the biomass utilization with a lower contribution from sludge and manure wastes.

References

- A. AlNouss, G. McKay, and T. Al-Ansari, 2018, Optimum utilization of biomass for the production of power and fuels using gasification, *Computer Aided Chemical Engineering*, 43, 1481-86.
- A. AlNouss, G. McKay, and T. Al-Ansari, 2019a, Superstructure Optimization for the Production of Fuels, Fertilizers and Power using Biomass Gasification, *Computer Aided Chemical Engineering*, 46, 301-06.
- A. AlNouss, G. McKay, and T. Al-Ansari, 2020, A comparison of steam and oxygen fed biomass gasification through a techno-economic-environmental study, *Energy Conversion and Management*, 208, 112612.
- A. AlNouss, G. McKay, and T. Al-Ansari, 2021, Utilisation of Carbon Dioxide and Gasified Biomass for the Generation of Value Added Products, *Computer Aided Chemical Engineering*, 50, 1567-72.
- A. AlNouss, S. Namany, G. McKay, and T. Al-Ansari, 2019b, Applying a Sustainability Metric in Energy, Water and Food Nexus Applications; A Biomass Utilization Case Study to Improve Investment Decisions, *Computer Aided Chemical Engineering*, 46, 205-10.
- C. Caligiuri, U. Žvar Baškovič, M. Renzi, T. Seljak, S.R. Oprešnik, M. Baratieri, and T. Katrašnik, 2021, Complementing Syngas with Natural Gas in Spark Ignition Engines for Power Production: Effects on Emissions and Combustion, *Energies*, 14, 12, 3688.
- M. M. Wright, N. Seifkar, W.H. Green, and Y. Román-Leshkov, 2015, Natural Gas and Cellulosic Biomass: A Clean Fuel Combination? Determining the Natural Gas Blending Wall in Biofuel Production, *Environ Sci Technol*, 49, 13, 8183-92.
- M. Shahbaz, A. AlNouss, I. Ghiat, G. McKay, H. Mackey, S. Elkhalfā, and T. Al-Ansari, 2021, A comprehensive review of biomass based thermochemical conversion technologies integrated with CO₂ capture and utilisation within BECCS networks, *Resources, Conservation and Recycling*, 173, 105734.
- Y. Zhang, A.H. Sahir, E.C.D. Tan, M.S. Talmadge, R. Davis, M.J. Bidy, and L. Tao, 2018, Economic and environmental potentials for natural gas to enhance biomass-to-liquid fuels technologies, *Green Chemistry*, 20, 23, 5358-73.

How tools and technologies can help drive sustainability

Julien de Beer^a, Mihaela Hahne^b

^aAVEVA, 5 square Félix Nadar, 94300 Vincennes, France

^bAVEVA, 26561 Rancho Parkway South, Lake Forest 92630 CA, USA

Abstract

The chemicals industry is facing a highly dynamic environment. Demand is continually fluctuating, and the pressure for new sustainable processes and products is rising. New environmental regulations are always expanding and becoming less harmonized globally, and customers are demanding sustainable products that are friendly to the environment. To follow all the market changes and remain competitive, companies need to invest to develop new processes and products. However, since it's a capital-intensive industry, new investments need to be carefully managed.

Today, companies need to find ways to be more efficient in the execution of engineering projects. One way is to compress engineering cycles and adapt processes and products to comply with sustainable Key Performance Indicators and new demands. The Unified Engineering methodology is one option enabled by the latest technologies and tools available. The data-centric approach is the first step to apply the Unified Engineering methodology, which will evolve later to the plant Digital Twin. The Unified Engineering methodology uses a single source of information that is available for all the teams involved in the project. Engineers become more efficient and work with reliable information as documents and models are updated in a controlled way as soon as any change is made.

Unified Engineering reduces capital project costs, risks, and delays enabling shorter engineering cycles required to deliver new sustainable projects. By minimizing engineering errors and accelerating project execution, companies can get 5% reduction in Total Installed Cost.

Keywords: sustainability, compressed engineering cycles, unified engineering, process engineering

1. Introduction

As an industrial software provider, we support chemical companies achieve superior performance in their quest to make sustainable products, align with the circular economy and demonstrate product stewardship throughout the product life cycle. By digitally connecting assets, process, and people, our solutions empower companies to run safe and responsible operations, mitigating EHS risks (Environmental, Health, Safety), and moving toward more circular systems, while remaining profitable.

With more than 50 years of industrial software innovation, AVEVA enables 13 of top 15 chemical companies and most of the world's petrochemical crackers, with the most comprehensive portfolio that ties profitability to sustainability goals.

Research typically identifies the following Critical Sustainability Drivers:

- Demand higher transparency on a company's Environmental, Social & Governance (ESG) performance by stakeholders
- Manage a complex environment and provide safety to employees, processes, products and local communities.
- Minimize energy and utilities consumption, so as emissions.
- Manage liquid effluents, waste and suppliers, seeking the circular economy.
- Commitments to global and local regulations
- Portfolio management towards innovation and sustainable products

Today's technology allows suppliers to develop better solutions to those markets that are in constant change. The Digital Twin technologies, initially adopted mainly by the automotive and aerospace industries, are now promoting big changes in how chemical plants are operated and managed. This type of technology can change the decision-making process since more reliable information is available in real time.

Significant step was taken recently in terms of process simulation driven by two major sustainability industry trends. For the circular economy industry trend, chemical companies are developing new chemical processes that yield materials that can be recycled rather than used one time. For the hydrogen economy industry trend, companies will substitute hydrogen fuels to reduce CO₂ emission into the atmosphere. Both the chemical and hydrogen economy trends lead to corresponding advancement in process simulation including modeling and thermodynamics.

Now it is possible for the engineering and operating companies to build the Digital Twin of the process plant. The Digital Twin is built on a simulation platform that will support the entire plant lifecycle, from design to operation, while also addressing new hydrogen and circular economy requirements.

2. The Digital Twin

The new generation of process simulation uses a platform approach that evolves the simulation model from the conceptual engineering to the operation optimization. It allows a new approach to be implemented so companies can transition from the conventional scenario to the use of the process simulation Digital Twin, expanding benefits to the entire plant lifecycle. The same platform is used for process simulation and process utilities (cooling water, flare, steam and others), allowing engineers to further evaluate how each system impacts the other. Heat and material balances can be re-evaluated after equipment and pipeline sizing, since that information is in the simulation from the beginning, as a result, little or no extra engineering effort is required. Once sizing is validated, the simulation is switched to dynamic mode, in which control loops are included to the simulation model to validate the process control strategy. As it is easier to shift the simulation to dynamic mode, rather than build a completely new model using the conventional approach or converting a model that cannot be taken back to the steady state mode, dynamic studies are performed earlier in the project lifecycle. This promotes savings in equipment acquisition and in operating costs, since control logic responses are evaluated in earlier stages. Plus, when something doesn't respond as expected, simulation is taken back to steady state mode, for re-evaluation of heat and material balance and re-sizing. The ability to go back and forth between steady state and dynamic modes is critical

to increase efficiency in the project lifecycle, leading to huge savings in engineering effort.

According to DECHEMA, one major obstacle for digital transformation is the division of process simulators into single-purpose point solutions. Today, separate models are created – often in different tools – for process design, control strategy design, operator training simulation, performance monitoring and online optimization. This drives up the total cost of modelling to a degree that can become prohibitively expensive. One single model should be able to cover the entire plant lifecycle from idea to operations.

3. Unified Engineering

Companies are beginning to make progress on their digitalization journey, finding the right applications for digital transformation and seeing increasingly better returns on their investment. While the age-old market environment challenges (such as supply and demand, cost and price) haven't gone away, competitive pressures are making the digital transformation opportunity more pressing than ever. Many have already started to leverage the latest data-centric technology and work processes for their workforce to collaborate and take control of their data, reducing the risk for errors, delays and increased project cost throughout the asset lifecycle. By doing so they are in a stronger position to become more competitive, increase their margins and win new business.

As outlined by an AVEVA whitepaper (2019), Unified Engineering is a new proposition to break down the silos between FEED and Detailed Design to minimize risk and maximize return on Capital Investment, Unified Engineering enables global multi-discipline teams to work concurrently in a common data-centric environment, controlling and managing change across the entire project. This breaks down the silos between FEED (Front End Engineering and Design) and detailed design. The simulation data created in FEED is readily available for use in detailed design and is checked and validated in real-time, increasing efficiency, minimizing risk, and maximizing return on investment on your Capital Projects.

Never have the stakes been higher for companies when it comes to making improvements to their engineering work processes to maximize Return on Investment (ROI) on Capital Projects. Productivity has not developed in decades – the average Capital Project schedule lags by 20 months and goes over budget by 80%. These are results that have been shared by Jayanth (2017) during the Rice Global E&C Forum Roundtable.

In many of today's Capital Projects, there is a disconnect between FEED and Detailed Design. In response, AVEVA are the first industrial software provider to pioneer a new solution to break down the silos between these project phases and engineering disciplines.

Unified Engineering consists of two main components, the Unified Lifecycle Simulation Platform (one model), and Integrated Engineering and Design (one database). The bi-directional integration of a steady state and dynamic process model with an engineering database makes the process seamless and eliminates the need for Microsoft Excel or other intermediate steps to transfer information between tools, enabling one single version of the truth to be kept ever up to date.

4. Process Simulation Lifecycle

Process simulators are irreplaceable tools for every process engineer. Since the nineteen seventies, process simulators have found widespread adoption within operating companies in oil & gas, refining and chemical industries, as well as the engineering companies and equipment manufacturers that service these industries. The tools available in the market today have incrementally improved over the years to provide more features and functionality. However, they trace their origins to legacy architectures, operating systems and aftermarket user interfaces, which create inherent limitations, see DECHEMA Tutzingen Thesen for more details. Today's simulators typically only support a single phase of the lifecycle and are often based on thermodynamics of different simulation vendors and different calculation methods. This not only leads to lack of trust in the results but causes substantial rework by having to build a new simulation model in each new tool. And the results are hard to compare. Willetts and Depew (2020) describe in detail the current challenges and how a Process Digital Twin will significantly help improve efficiency and drive increased sustainability.

Global competition, pricing pressure and energy alternatives are now driving the need for a new approach. The oil & gas industry has seen high volatility and the lower price level of today is seen as the “new normal”. The chemicals industry has a continuous need to innovate for greater agility and lower costs.

The next generation of workers also expects a modern, scalable and easy to use solution with technology they now take for granted – high speed internet access, mobile devices, touch screens and virtual reality. New concepts like the Industrial Internet of Things (IIoT), Industry 4.0, and Artificial Intelligence have created greater opportunities with a new next generation platform that provides a “Digital Twin” of the plant through the process lifecycle that cannot be provided with today's tools.

A next generation process simulation platform means that one process model is extended throughout the entire lifecycle of the plant, from concept through to operations. This requires a process design mode, a fluid flow/rating mode and a dynamic mode, in combination with the ability to toggle back and forth between modes. Optimization may be provided to any mode.

A single, easy-to-use simulation platform will allow engineers to move seamlessly between questions of design, analysis, and optimization. Engineers will be able to assess the impact of design and specification changes quickly and with a holistic view of multiple disciplines.

Intensive collaboration becomes commonplace. In this environment, organizations will be able to adopt agile engineering workflows based on smaller pieces of work with continuous integrated testing to reduce development cost while eliminating surprises at the end of the project. Further details can be found in C. Depew's Digital Transformation of Process Engineering presented during the 16th GCPS AIChE Spring Meeting in 2020.

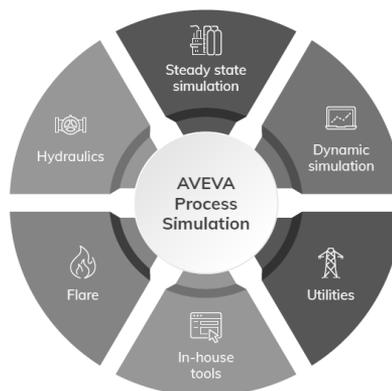


Figure 1: Next Gen Simulation Platform

5. Process Simulations for Sustainable Technology

Physics-based, first-principles models are a critical component of the asset digital twin. As industry develop new process technologies to serve sustainable industry trends such as the circular economy and the hydrogen economy, process simulation programs must adapt to become useful tools for process engineers to design sustainable processes. While academia research how to simulate model reactions or chemical thermodynamics, this modeling and thermodynamics must be integrated within industrial simulation products so that the unified engineering workflow may be applied to develop the digital twin of plants as they are engineered. The AVEVA whitepaper (2021), *Take the first step to the circular economy with AVEVA™ Process Simulation*, Top 5 reasons to adopt a transformational approach, further explains how modern technology allows to truly transform an organization and its outputs by speeding up the FEED stage by 50%, generating more options in less time, allowing for a more optimized process and more informed decision making as well as increasing engineering efficiency by 30%.

There are many Hydrogen-based projects under consideration around the world. Blue, Green, and Gray Hydrogen projects need common process simulation advancement that need to be modeled in processes that support the hydrogen economy.

Green hydrogen is produced by the electrolysis of water. Process simulators will include new electrolyzer equipment models. The Digital Twin for the operation of these plants may use an integrated model to optimize the use of wind and solar power generation for electrolysis.

Blue and gray hydrogen is produced by splitting the methane in natural gas into Hydrogen and CO₂ by steam methane reforming or auto thermal reforming. Development of integrated process plant simulation is required for removal of the CO₂ for capture and storage to produce blue hydrogen. New membrane adsorption models will separate hydrogen for greater purity for blending with natural gas. The carbon dioxide created must be removed using new generation of amines thermodynamics for CO₂ capture by companies that manufacture the amines and companies that use the amines for the separation of CO₂ for carbon capture.

Due to the low volumetric energy density of gaseous hydrogen when used as a fuel, plants must liquify the hydrogen to use it as a transportation fuel. Process simulators will include thermodynamic models appropriate for cryogenic hydrogen. In addition, ammonia is used as a transition fuel with models and thermodynamics required for ammonia pipelines and hydrogen conversion.

Circular economy trends are driving chemical companies to produce materials that may be recycled more readily. This drives the development of new thermodynamic methods and component data to study new process simulation. Chemical and energy companies are replacing fossil based raw materials with renewable feedstocks driving research and development of new reaction technology.

Finally, to make all processes more sustainability by reducing their energy consumption footprint, new simulation products will include the ability to calculate the cost of utilities such as steam, cooling water, and electricity, to minimize operating costs.

6. Conclusion

Lifecycle process simulation has been a vision for process simulation providers and their customers for a long time. Today's simulators cannot leverage the rapid developments occurring in the software industry due to legacy architecture.

Looking at the industry's increasing demand for higher transparency, this can only be achieved using a data-centric Digital Twin approach. This data-centricity enables an ideal platform for new product and process development to create new models and include the management of complex environments bringing together both steady-state and dynamic simulation with constant iteration and constant solving capabilities.

The integration of the process analysis and simulation with other disciplines also allows to breakdown the silos that were typically existing previously. Connecting and remotely controlling previously unconnected processes will increase sustainable operations and improve business efficiency in a sustainable environment. Over the long term the impact will drive resilience and sustainable performance through technologies.

References

DECHEMA, 2016, Digitalisierung in der Chemieindustrie – Whitepaper

Portfolio Marketing group, AVEVA, 2019, Unified Engineering: A new proposition to break down the silos between FEED and Detailed Design to minimize risk and maximize return on Capital Investment – Whitepaper

T.G. Jayanth, Expert – Capital Projects & Infrastructure, McKinsey & Co, 2017, Reinventing Construction - a route to higher productivity, McKinsey, Rice Global E&C Forum Roundtable

DECHEMA, 2018, Tutzingen Thesen

Ian Willetts, Cal Depew, AVEVA, 2020, Building a process Digital Twin for your plant with Unified Lifecycle Simulation – Whitepaper

C. Depew, 2020, Digital Transformation of Process Engineering, 2020 Virtual Spring Meeting and 16th GCPS AIChE Spring Meeting

Portfolio Marketing group, AVEVA, 2021, Take the first step to the circular economy with AVEVA™ Process Simulation, Top 5 reasons to adopt a transformational approach – Whitepaper

The use of game theory in the analysis of marine lubricant markets in ports

Dimitra Tsakona and Antonis Kokossis*

*National Technical University of Athens, School of Chemical Engineering, Iroon
Polytechniou 9, Zografou 157 80, Athens, GR*

**corresponding author: akokossis@mail.ntua.gr*

Abstract

The supply of marine lubricants to vessels can be arranged by multinational companies (majors) or local suppliers (locals) who have interactions and contract agreements between them. The challenge of the paper has been to use game theory to analyze the market share that is attributed to each supplier and who it is connected to the level of interactions, the market constraints and how they can affect the overall profitability for each player. The paper presents the framework of players interactions that it formulates as a bilevel optimization problem with major (leader) and local players (followers) assigned to the outer and inner problems. The model is used to assess threats and promises. The bilevel approach can be used to calculate parameter sensitivity and key aspects to retain market balance and profit share. The approach has been validated with real data in which the analysis achieved to predict the actual market and profit share using the model-based approach. Model extensions subsequently test parameters responsible for system stability and player options that address threatening moves from other players.

Keywords: marine lubricants, multinational companies, local suppliers, game theory, bilevel approach

1. Marine lubricants and game theory

Lubricants are necessary for the proper operation not only of the main engine (cylinder oil, trunk piston oil and system oil) but also of other vessel's systems (hydraulic fluids, gear oil, turbine oils, greases and other). They may be classified into automotive, industrial, process, and marine oils. Even though marine oil lubricants make up a small percentage of lubricants consumed worldwide, the study of this market has a great deal of interest, as the global seaborne trade is projected to rise. According to recent studies, the global marine lubricants market size was valued at USD 8.01 billion in 2018, and it is estimated to reach USD 9.47 billion by 2026, with a CAGR (Compound Annual Growth Rate) of 2.13% over the forecast period.[1] Furthermore, around of 90% of world's trade is carried through maritime transport.

In the 1950-70s the supply of marine lubricants was controlled by the so-called "Seven Sisters" (BP, Shell, Gulf, Chevron, Exxon, Mobil Texaco). [2] Following acquisitions of the above companies and the fact that new local suppliers entered the market gaining market share, the market divided to major suppliers such as BP, Shell, Exxon-Mobil, Total and local suppliers such as Sinopec (China), Nippon (Japan), Petrobras (Brazil), Gazprom (Russia) and others. The level of equilibrium in the market is differentiated depending on the port and is worth studying as it can easily be modified by changing the

players' strategy. These changes can create threats and opportunities to the profitability for both suppliers.

The blending conditions of marine lubricants varies depending on the port. In key ports, such as Singapore, the ports of America or Belgium, both players have their own refineries to blend their lubricant, in which case the locals have a very small share of the market. Lubricants' quality for majors and locals is considered equal as the products of both players hold all the compulsory approvals for the utilization of oils. However, the extensive experience of the multinational companies as well as their strong R&D department helps them to design lubricants with special features such as their slowest consumption, which gives them another lead over the local companies.

The marine lubricants market constitutes a duopoly market, in which the multinational companies are "the leader". Due to their worldwide brand awareness and their long-term agreements with several shipping companies around the world, they can reassure the majority of market share without the followers to be able to compete them. On the other hand, the local firms, "the followers", have a better knowledge of their domestic market, thus they can be more flexible and provide more expeditious services. Furthermore, it is considered that selling price of the lubricant is the same for both suppliers.

The use of game theory holds an apparent and promising potential to better understand the context behind such developments. Still, it has never been tried, mainly due to the overall scepticism on whether such an approach could deliver realistic results and recommendations. The purpose of the paper has accordingly been to test the potential of game theory in real ports with a further view to postulate critical parameters and review stakeholder interactions. The paper first explains the application context, defines the problem, and next applies a mathematical approach for analysis using bilevel optimization.

2. Market and problem description

This study examines cases of peripheral ports such as Greek or Egyptian ports, where multinational companies do not have a physical presence and blend lubricants by utilizing locals' third-party blending plant. Under this conditions, major suppliers are charged from local suppliers with an additional fee know as premium, so to be able to blend their lubricants under their name and their specification standards. Both majors and local firms gain benefit from this cooperation since they manage to reduce productions costs. Major firms are willing to pay a premium fee to locals to avoid installation costs, while local firms reduce their own production costs by achieving economies of scale. Local firms pursued to keep premium mark up at low levels in order maintain the cooperation. As a result, both players can be considered as interconnected. Such players can also be characterized as connected, as they face several common threats and opportunities based on their price levels. By setting a high price level, the players maximize their profit. However, as the price for the same lubricant differs between ports, players risk losing business for vessels that may be supplied at a next port with more competitive price levels. On the other hand, low price levels may be beneficial for both players as the port will increase its demand but may also cause a loss of maximum profits. Although the selling price range is a common challenge for both players, they still compete each other. Strategic decisions leading to a change in each optimal production quantity and the market share gained create opportunities for one player that pose a threat to the other. For instance, a decision of locals to offer more competitive prices than multinational companies or the decision to increase premium price can be an opportunity for them, as they can gain a larger share of the market. This opportunity constitutes a threat to majors,

who should respond to avoid losing market share. The use of game theory is certainly justified even though never applied as the markets are not ideal and the terms of elasticity are discontinuous functions.

The cylinder oil accounts to approximately 50,31% of the requested lubricants. Market data collected for the study concern volumes of cylinder oil consumed in Greece and in Singapore in 2018 and relative price from both major brands (BP, SHELL, MOBIL) and local brands (AVIN, ELIN, ENOC, SINOPEC).

Fig 1 and 2 present market data for Piraeus and Singapore, the two cases studied in this paper. Major firms blend a quantity q_1 and local firms blend a quantity q_2 , while the selling price P is the same for both suppliers. Both suppliers produce lubricants under the same refinery of the local supplier, with the leader being charged a premium by the local. According to the Greek market data, the total domestic marine lubricants market was 31,30 million liters in 2018 [3]. They suggest that leaders and followers play on different terms in the markets: major players benefit from special relationships winning customers at lower volumes; locals compete at larger volumes with lower costs. Modelling market data constitute a challenge as they are scattered, and linear models may certainly not perform reliably.

In the Singapore market figure, the points collected and plotted were 26 for major suppliers and 11 for local suppliers, while in the Greek market, 12 data points were gathered and plotted for major suppliers and 6 data points for local suppliers.

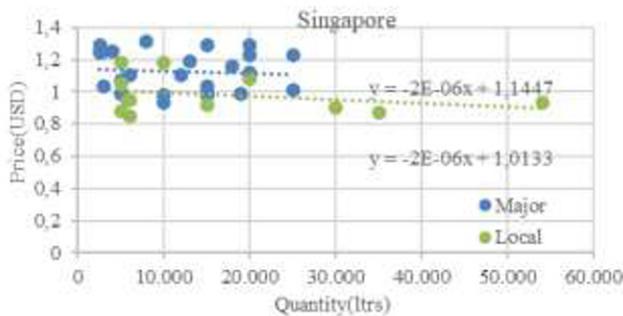


Figure 1: Indicative prices (USD) and Quantities (liters) for cylinder oil in Singapore ports in 2018

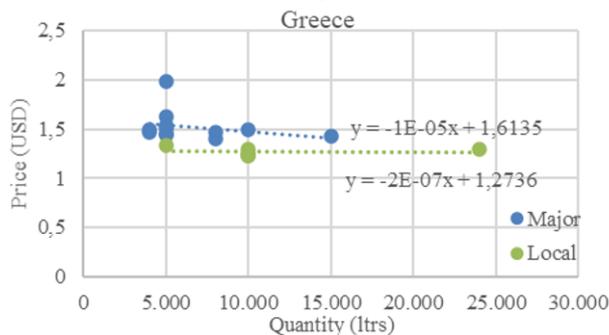


Figure 2: Indicative prices (USD) and Quantities (liters) for cylinder oil in Greek ports in 2018

Let majors and locals be respectively leaders and followers in a context of Stackelberg duopolies [4]. Stakeholders interactions are sketched in Fig 3:

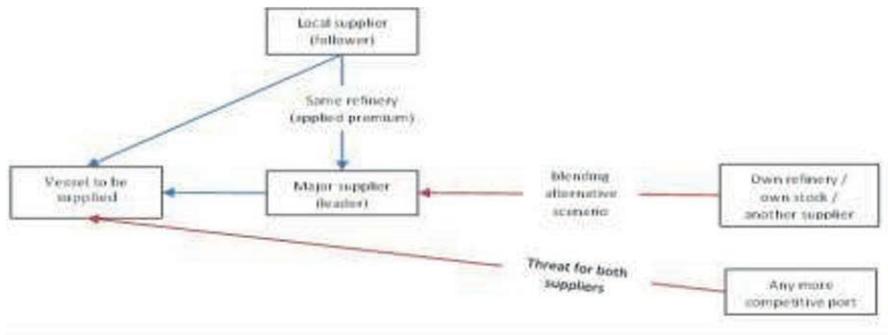


Figure 3: Marine Lubricants supply in peripheral ports

The purpose of the paper would then be to: (a) develop a game model based on Fig 3 accounting for the presumed interactions of stakeholders, (b) use market data of Fig 2 and 3 to assess equilibrium states comparing them with real-life market shares, and (c) evaluate market parameters and interactions holding significance to retain equilibrium.

3. Methodology and mathematical model

A leader-follower relationship is modelled as a bilevel optimization problem where the follower's lower-level program is embedded in the leader's upper-level program. Such problems are formulated as

$$\begin{aligned} & \text{Max}_x F(x, y) \quad \text{s.t. } G(x, y) \geq 0 \\ & \text{where } y \text{ solves } \text{Max}_y f(x, y) \quad \text{s.t. } g(x, y) \geq 0 \end{aligned}$$

x denotes upper-level variables; y : lower-level variables; F : the upper-level objective function; f : the lower-level objective function; G : the upper-level contains, g : the lower-level contains [5].

Given a market of two players, $Q = q_1 + q_2$ (1). Each firm aims to maximize profit. The profit of each player is: $\Pi_i = P(Q) \cdot q_i - c_i \cdot q_i$, where c_i is the cost of each player. The total cost of the leader is the cost of production plus the premium, while that for the follower is only the cost of production.

$$\Pi_1 = P(Q) \cdot q_1 - (C + m) \cdot q_1 \quad (2)$$

An additional amount $q_1 \cdot P$ is added to the follower's profit function (Π_2).

$$\Pi_2 = P(Q) \cdot q_2 - C \cdot q_2 + m \cdot q_1 \quad (3)$$

Concerning $P(Q)$ (4), one could argue that this is not a function but rather a bounded domain that contains market data. In the study we have attempted (a) linear and nonlinear approximations [6], and (b) convex relaxations. Using,

(i) linear approximations, $P(Q) = a \cdot Q + b$; For Greek port $a = -2 \cdot 10^{-15}$, $b = 1.5687$

(ii) Nonlinear elasticity functions, $P = a \cdot Q^{-b}$: Greek ports $a = -4.501$, $b = -0.128$; Singapore port $a = 1.57$, $b = -0.039$

(iii) Convex relaxations for the best fit of the market function $w = P \cdot Q$:

$$\tilde{w} \geq Q^L \cdot P + Q \cdot P^L - Q^L \cdot P^L; \quad \tilde{w} \geq Q^U \cdot P + Q \cdot P^U - Q^U \cdot P^U;$$

$$\tilde{w} \leq Q^U \cdot P + Q \cdot P^L - Q^U \cdot P^L; \quad \tilde{w} \geq P^U \cdot Q + P \cdot Q^L - Q^L \cdot P^U \quad [7]$$

with $\tilde{w} = 8268$; $P^U = 1,23$; $P^L = 1,99$; $Q^U = 24000$; $Q^L = 4000$

Premium (m) is also a parameter affecting results. The production cost (C) is common assuming the use of the same refinery. Sale prices are also identical for both players. Production costs account to 90% of the average sales price of supplies in local firms. Fair bounds on profitability are set at 30%; the maximum premium is set by market practices at 1%. The quantity produced is identical to the quantity sold. The bilevel problem is formulated using EMP of GAMS and solver Conopt. The variables of the upper level problem include the leader's quantity (q_1) and profit (Π_1), the leader's cost of production (C); premium (m) is a sensitivity parameter. The variables of the lower level problem include the follower's quantity (q_2) and profit (Π_2), and the follower's cost of production (C).

4. Results

4.1 Sensitivity analysis

Production cost and the premium are parameters whose values are derived from the market study. Sensitivity analysis for the two parameters led to the conclusion that the most premium is the most sensitive parameter; even a change by 1% may affect equilibrium. Changes in the premium do not affect both players equally. An increase in the premium increases the share of the follower in the market. Although sensitive, the cost of production has an equal impact on the players. An increase in the production cost leads to the reduction of supplied quantities. A decrease in the cost raises the profitability of both players

4.2 Market shares and profitability

Results for different types of approximation have apparently revealed different market shares for the players. Linear approximations concluded to market shares 65% (major), and 35% (local suppliers); profitability shares are respectively estimated at 59% and 41%. Nonlinear approximations concluded to market shares 53% (major), and 47% (local suppliers); profitability shares are respectively estimated at 48% and 52%. Convex relations finally estimated market shares 44% (major), and 56% (local suppliers); profitability shares are respectively estimated at 39% and 61%.

4.3 Equilibrium as key or peripheral port

In key ports both local and major suppliers have their own refineries; peripheral ports use only local suppliers. Game theory has set 30% as a profitability threshold for peripheral ports. In the case of Singapore, the market share corresponding to major suppliers is computed at 83% and 16% for local suppliers. The major's profitability rate is 78% while the local's is 22%. These numbers are not close to the threshold hinting that Singapore will not retain equilibrium unless the premium is 200% higher (leading to splits 67% and 32% for major and local players), a rather uncommon case for existing markets. Instead, the case for Piraeus suggests a peripheral port.

5. Discussion

5.1 Sensitivity analysis

Sensitivity analysis around the premium revealed that its increase would be beneficial to the local. However, this is also a threat to the major, that, to avoid losses, could consider alternative scenarios and options to discontinue the cooperation with the local refinery and opt to (a) import their own stock from a major's refinery located in other regions, (b) alliances with local suppliers of nearby ports utilizing those refineries as third-party

blending plant, or (c) install a new refinery on the site. These options create a threat to the locals and the premium can be set on the feasibility that such other options can prevail. The sensitivity analysis related to production costs leads to common threats and opportunities for both parties. Consequently, they both should aim at reducing the cost price to the minimum possible level.

5.2 Market share and profitability

The equilibrium numbers calculated by the game theory model can be compared with the real-life data for Piraeus. The latter suggest that a split of 54% and 46% for major and local, also a profitability split between 52% and 48%. The numbers are extremely close to the nonlinear approximation (53% and 47%); linear models and convex approximations appear to overestimate the share of the local players: they respectively predict 11% and 7% off the actual data.

5.3 Key ports

The model further confirms that Piraeus suits a profile of a peripheral port as profit/market shares and premium values fall within bounds. It also validates that ports such as Singapore assume profiles of key ports as the share for equilibrium falls outside bounds unless the premium is very high for common practices. However, such high premia are not usually acceptable by multinational companies, which, having the largest market share opt to expand and own refineries.

6. Conclusion

Game theory has proved instrumental to analyse marine lubricants markets. Bilevel optimization, even in considering the basic set of constraints outlined in the paper, constitute a useful resource for suppliers to review their status in the market, understand actions to secure and increase profits, also to prepare for competitors' actions. The model presented in the paper has been validated by the state of a large Greek port (Piraeus), both with respect to its status (peripheral port) as well as its equilibrium share. By a similar token, another large port (Singapore) has been validated as a key port where equilibrium is achieved once major players undertake third party blending plants.

7. Acknowledgement

The authors acknowledge financial support by the H2020-EU.1.3.3. Grant RENESING II (H2020-778332).

References

- [1] Market research report, 2020, Fortune business insights
- [2] Thomas Herold, 2017, Herold Financial Dictionary - Trading edition
- [3] LPC web site, <https://www.lpc.gr/en/the-market/greek-lubricant-market/>, last date of access: 29/11/20221
- [4] Romualdas Ginevičius1, Algirdas Krivka,2008, Application of game theory for duopoly market analysis
- [5] Roger J. Jiao, 2013, Green modular design for material efficiency: a leader–follower joint optimization model
- [6] Saim Özkar and Richard G. Finke,2017, Silver Nanoparticles Synthesized by Microwave Heating: A Kinetic and Mechanistic Re-Analysis and Re-Interpretation
- [7] Shahin Goodarzi1 Mohsen Gitizadeh1 Ali Reza Abbasi2, 2019, Efficient linear network model for TEP based on piecewise McCormick relaxation

Application of Outlier Treatment Towards Improved Property Prediction Models

Adem R.N. Aouichaoui, Seyed Soheil Mansouri, Jens Abildskov, Gürkan Sin*

Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kgs. Lyngby, DK-2800, Denmark

gsi@kt.dtu.dk

Abstract

Property prediction models based on the principle of a quantitative structure-property relation (QSPR) such as the group contribution models are an important tool that provides a quick, simple, and costless evaluation of various thermophysical properties of chemicals for various applications such as P-V-T calculations and product design. These models rely heavily on the interplay between the chosen descriptor (molecular information), the chosen mathematical formulation (to relate the descriptor to the target property), and the data used to produce such models. Therefore, such models suffer heavily if the quality of experimental data is low (inaccurate) or if there are discrepancies in the descriptors used or the mathematical representation chosen. In this work, we apply a systematic methodology to detect and treat outliers on 18 thermophysical properties and showcase the model improvements across various statistical metrics. This results in significant improvements across all property models illustrated through an increase in the coefficient of determination (R^2), the standard deviation (σ), and the mean absolute error (MAE).

Keywords: Outlier Treatment, Group-Contribution models, Thermophysical properties, Property Prediction, Quantitative Structure-Property Relations (QSPR)

1. Introduction

Knowing the properties of a chemical plays an important role in various chemical engineering disciplines and applications such as risk assessment, process design, and product design. In fact, some of these disciplines would be impossible without this knowledge. These properties can either be related to the thermal properties (normal melting point, critical temperature), physical properties (molar volume, critical volume), flammability properties (lower flammability point and auto-ignition temperature), or the enthalpic properties (enthalpy of formation, enthalpy of fusion) of a compound. These properties are usually determined either directly from experiments or derived from other closely related experimental measurements. Some of these experiments can be cumbersome in terms of the resources needed to conduct them especially with regards to time and the experimental setup needed to provide accurate results. Large collections of experimental data for such properties are scarce and few in numbers. This creates a need for alternative approaches to explore the chemical design space to direct future experimental efforts towards the most promising chemicals for a given application. One way to do this is to correlate the molecular structure to the target property of interest through what is known as Quantitative structure-property relations (QSPR)(Austin *et al.*, 2016). These models take a numerical translation of the molecular structure as input (also

known as descriptors) to a mathematical model that relates these inputs to the target output. An example of such models is the group-contribution (GC) type of model where the model is described in terms of small fragments of functional and structural groups. The success of these models heavily depends on the interplay between the chosen descriptor, the mathematical model, and the equality of data used for the regression (Gasteiger, 2016). In this study, we apply a systematic methodology for outlier treatment to increase the models' predictive power. This methodology is applied to a group-contribution type of model and 18 different property datasets and is benchmarked against previously published results to highlight the effect of the outlier treatment.

2. Methods

2.1. Group-Contribution models

The molecular descriptor employed in this study is the group fragments defined by (Marrero and Gani, 2001), where the molecule is described through three levels of complexity with increasing resolution on the molecular structure. In total, the method defines 220 first-order groups, 130 second-order groups, and 74 third-order groups through which, a molecule can be described by an occurrence vector that indicates the number of time each fragment occur in the molecule. The generic form of a GC model is shown in Eq (1), where $f(X)$ is functional transformation in order to make the right hand side of Eq (1) linear, $c_y^{(x)}$ and $n_y^{(x)}$ are the contributions and occurrence of group y of order x respectively. Alternatively, the equation can also be represented in matrix format where G is the group occurrence matrix and θ are the corresponding group contributions. The left hand side of Eq (1) is determined by visual inspection of the trend of the property at increasing carbon number for various homologous series. On overview of the structure of these transformations can be seen in (Hukkerikar et al., 2012).

$$f(X) = \sum_{i=1}^{220} c_i^{(1)} n_i^{(1)} + \sum_{j=1}^{130} c_j^{(2)} n_j^{(2)} + \sum_{k=1}^{74} c_k^{(3)} n_k^{(3)} = G \theta \quad (1)$$

2.2. Parameter Fitting and Outlier Treatment

A systematic method has been developed by (Frutiger et al., 2016) to perform parameter fitting, outlier treatment, and uncertainty estimation. The same methodology with few modifications is applied in this work and is illustrated in Figure 1. The method is explained in the following:

Step 1 (S1): The exact form of the function is determined by inspecting the trend of the property with increasing carbon numbers across the homologous series.

Step 2 (S2): By taking advantage of the matrix representation of the GC model, an initial guess for the parameters can be produced by applying linear Algebra as seen in Eq. (2).

$$\theta^0 = (G^T G)^{-1} \cdot G^T \cdot f(X) \quad (2)$$

Step 3 (S3): The parameters are estimated sequentially based on their order: first-order groups are estimated first. Whatever the model fails to describe using only first-order groups is then rectified by adding the second-order groups and estimating the parameters for these while keeping the parameters for the first-order fixed. The same reasoning goes for third-order groups. For each group order, the parameters are first estimated using the simplex method to provide a good initial estimate for the gradient-based optimization

using the Levenberg-Marquardt (LM) algorithm. The optimization problem is shown in Eq (3), where y_i^{exp} and y_i^{pred} are the experimental and predicted property values.

$$\theta^* = \arg \min \sum_i (y_i^{exp} - y_i^{pred})^2 \quad (3)$$

Step 4 (S4): All parameters are allowed to vary during the simultaneous parameter fitting. In this step, only the simplex method is used. This ensures that the parameter fitting is more stable and less prone to diverge due to the large number of parameters.

Step 5 (S5): Outlier detection is done by constructing the empirical cumulative distribution function (ECDF) of the prediction errors. The ECDF aims to estimate the true underlying distribution of the residuals making it a suitable method for detecting the outliers without the need for any assumption related to the distribution of these residuals (Frutiger *et al.*, 2016). Data for which the residuals are below the 2.5% and above the 97.5% probabilities are considered outliers and are thus removed from the dataset.

Step 6 (S6) & Step 7 (S7): Sequential and simultaneous parameter estimation is performed on the remaining data using the previously optimized parameters as an initial guess.

Step 8 (S8): Although not discussed in this work, uncertainty analysis is also performed using the maximum likelihood estimation (MLE) by assuming that the measurement errors follow a Gaussian distribution white noise and are independently distributed (Frutiger *et al.*, 2016). This is done by linear propagation of error through asymptotic approximation of the parameter covariance matrix shown in Eq (4). SSE is the sum of squared errors, DoF is the degrees of freedom while J is the Jacobian. The $(1-\alpha_t)$ confidence interval for the prediction is calculated using Eq (5) assuming a student t-distribution.

$$COV(\theta^*) = \frac{SSE}{DoF} (J(\theta^*)^T J(\theta^*))^{-1} \quad (4)$$

$$y_{1-\alpha_t}^{pred} = y^{pred} \pm \sqrt{diag(J(\theta^*)COV(\theta^*)J(\theta^*))} \cdot t(DoF, \alpha_t/2) \quad (5)$$

Step 9 (S9): The final model performance is evaluated by calculating various statistics such as the coefficient of determination (R^2), the standard deviation (σ), and the mean absolute error (MAE) shown in Eq (6), Eq (7) and Eq (8) respectively.

$$R^2 = \frac{\sum_i (y_i^{exp} - y_i^{pred})^2}{\sum_i (y_i^{exp} - \bar{y})^2} \quad (6)$$

$$\sigma = \frac{1}{N} \sqrt{\sum_i (y_i^{exp} - y_i^{pred})^2} \quad (7)$$

$$MAE = \frac{1}{N} \sum_i |y_i^{exp} - y_i^{pred}| \quad (8)$$

The main modification compared to the methodology proposed in (Frutiger *et al.*, 2016) is that the simplex algorithm and the Levenberg-Marquardt algorithms are used alternatively during the sequential parameter estimation and that the final simultaneous parameter estimation is only conducted using the simplex methods. This was shown to provide more stability to the parameter estimation procedure as well as faster convergence and robustness towards the initial guess (starting point of the regression).

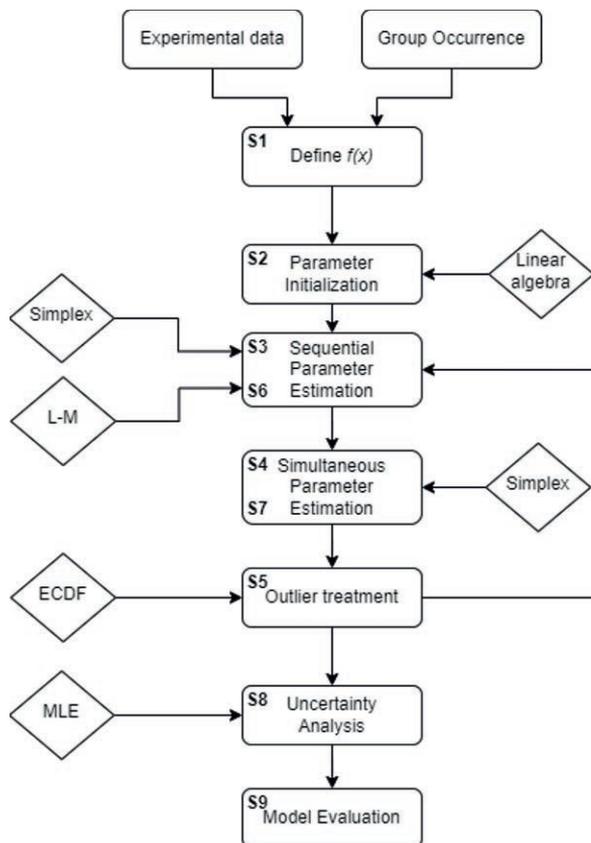


Figure 1: Systematic approach for parameter fitting and outlier treatment (rhombus shape signifies technique or method applied)

2.3. Property Data

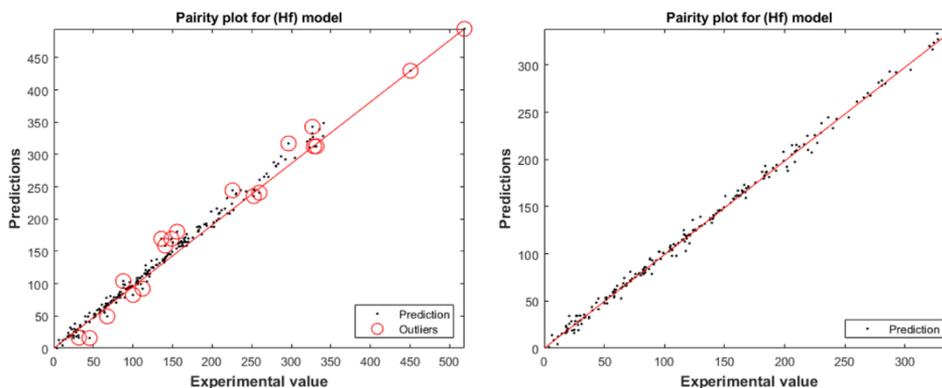
The methodology is applied to 18 property data covering: the normal boiling point (T_b), the critical temperature, pressure and volume (T_c , P_c , and V_c respectively), the normal melting point (T_m), the Gibbs energy of formation (G_f), the enthalpy of formation (H_f), the enthalpy of fusion (H_{fus}), the Octanol-water partition coefficient ($\log Kow$), the Hansen solubility parameters (δ_D , δ_P and δ_H), the enthalpy of vaporization at 298K (H_v), the enthalpy of vaporization at the normal boiling point (H_{vb}), the entropy of vaporization (S_{vb}), the Hildebrand solubility parameter (δ), the acentric factor (ω) and the molar volume (V_m). All data are retrieved from the CAPEC database (Nielsen *et al.*, 2001).

3. Results

An example of model prediction and outlier detection can be seen in Figure 2. An overview of the results and model statistics can be seen in Table 1. The improvement of each metric compared to the result in (Hukkerikar *et al.*, 2012) is also provided in Table 1. Note that for R^2 , the improvement is illustrated by an increase (+), while for σ and MAE is illustrated by a reduction of the metric (-).

Table 1: Performance statistics of the new models and their improvements compared to benchmark model (in parenthesis is the improvement compared to (Hukkerikar *et al.*, 2012))

Property	Unit	N	R ²	σ	MAE
T _b	K	3,510	0.99 (+0.00)	5.77 (-2.13)	4.63 (-0.15)
T _c	K	858	0.99 (+0.00)	6.34 (-4.43)	4.57 (-3.15)
P _c	bar	852	0.99 (+0.02)	1.06 (-1.32)	0.74 (-0.66)
V _c	cc/mol	797	0.99 (+0.00)	9.26 (-2.39)	6.58 (-1.39)
T _m	K	5,183	0.97 (+0.02)	17.25 (-1.91)	14.47 (-1.52)
G _f	kJ/mol	749	0.99 (+0.00)	4.79 (-3.57)	3.02 (-2.22)
H _f	kJ/mol	882	0.99 (+0.00)	4.41 (-3.33)	2.92 (-2.11)
H _{fus}	kJ/mol	764	0.95 (+0.12)	2.38 (-2.78)	1.72 (-1.07)
logK _{ow}	-	12,193	0.92 (+0.05)	0.49 (-0.15)	0.39 (-0.09)
δ_D	MPa ^{1/2}	1,037	0.92 (+0.20)	0.50 (-0.58)	0.37 (-0.23)
δ_P	MPa ^{1/2}	1,017	0.85 (+0.10)	1.51 (-0.69)	1.16 (-0.65)
δ_H	MPa ^{1/2}	1,016	0.93 (+0.06)	1.13 (-1.67)	0.83 (-0.45)
H _v	kJ/mol	705	0.99 (+0.02)	1.04 (-1.30)	0.71 (-0.58)
H _{vb}	kJ/mol	512	0.98 (+0.02)	0.88 (-0.54)	0.63 (-0.32)
S _{vb}	J/mol K	512	0.95 (+0.10)	1.54 (-1.46)	1.05 (-0.67)
δ	MPa ^{1/2}	1,384	0.93 (+0.10)	0.99 (-0.64)	0.72 (-0.36)
ω	-	1,723	0.97 (+0.07)	0.05 (-0.05)	0.03 (-0.02)
V _m	cc/kmol	1,056	0.99 (+0.00)	0.00 (-0.00)	0.00 (-0.00)

Figure 2: Parity plot for H_f before (left) and after (right) outlier removal

4. Discussion

Outlier treatment greatly improved the performance of the GC model across all datasets without exception compared to the results obtained in (Hukkerikar *et al.*, 2012). Most significant improvements are seen for H_{fus}, δ_D , δ_P , S_{vb} and δ , where the R² increased between 0.1 and 0.2. The same effect is also observed on the standard deviations and the mean absolute error. Although the method does indeed improve the model performance, it is also important to note that the method does provide a drawback. The outlier deletion results in the removal of some groups for which a contribution will not be available. This reduces the chemical design space it can explore especially in the absence of any proven

method to predict the contribution of groups not present in the dataset. In fact, a vast majority of groups are not present in every dataset. Despite limiting the domain of exploration of the model, the method reinforces the accuracy for the compounds that remain after outlier deletion and thus increases the reliability of the model within the domain of applicability set by the present groups after outlier deletion. Furthermore, the outliers are defined based on the model performance, thus potentially more informative descriptors or a better mathematical model will result in identifying different outliers and thus the reason for identifying a given compound as an outlier cannot only be attributed to the quality of data but also the model and the descriptor in this case. Important to note is that GC type of models are additive models and thus might struggle for highly non-linear properties as seen e.g. H_{fus} , and various solubility parameters, additionally they do not capture any proximity effects e.g. arrangement of the groups does not affect the prediction. Advanced models capable of capturing highly non-linear trends such as deep neural networks could provide an alternative to the mathematical presentation and thus potentially provide more accurate predictions (Aouichaoui *et al.*, 2021).

5. Conclusion

In this study, we demonstrated that a systematic approach to treating outliers greatly improves the models' predictive performance by narrowing the domain of applicability and increasing the reliability of the model within this domain. One major drawback is, however, that some descriptors are lost in the process of data deletion which reduces the portion of chemical space that the model can explore. Further work will explore the analysis of outliers using different modeling approaches including machine learning and deep learning approaches such as deep neural networks and graph neural networks. This is needed to better understand what causes a data point to be an outlier: a) measurement error or b) a modeling error. In the latter case, the outlier is an observation that provides an opportunity to develop new models.

References

- Aouichaoui, A.R.N., Al, R., Abildskov, J. and Sin, G. (2021) 'Comparison of Group-Contribution and Machine Learning-based Property Prediction Models with Uncertainty Quantification', in *Computer Aided Chemical Engineering*. Elsevier Masson SAS, pp. 755–760.
- Austin, N.D., Sahinidis, N. V. and Trahan, D.W. (2016) 'Computer-aided molecular design: An introduction and review of tools, applications, and solution techniques', *Chemical Engineering Research and Design*, 116, pp. 2–26.
- Frutiger, J., Marcarie, C., Abildskov, J. and Sin, G. (2016) 'A Comprehensive Methodology for Development, Parameter Estimation, and Uncertainty Analysis of Group Contribution Based Property Models-An Application to the Heat of Combustion', *Journal of Chemical and Engineering Data*, 61(1), pp. 602–613.
- Gasteiger, J. (2016) 'Cheminformatics: Achievements and Challenges, a Personal View', *Molecules*, 21(2), p. 151.
- Hukkerikar, A.S., Sarup, B., Ten Kate, A., Abildskov, J., Sin, G. and Gani, R. (2012) 'Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis', *Fluid Phase Equilibria*, 321, pp. 25–43.
- Marrero, J. and Gani, R. (2001) 'Group-contribution based estimation of pure component properties', *Fluid Phase Equilibria*, 183–184, pp. 183–208.
- Nielsen, T.L., Abildskov, J., Harper, P.M., Papaconomou, I. and Gani, R. (2001) 'The CAPEC database', *Journal of Chemical and Engineering Data*, 46(5), pp. 1041–1044.

Data-driven, Image-based Flow Regime Classification for Stirred Aerated Tanks

Corinna Kröger,^a Valentin Khaydarov,^a Leon Urbas^a

^a*Technische Universität Dresden, Dresden 01062, Germany*
valentin.khaydarov@tu-dresden.de

Abstract

Monitoring and optimization of flow regimes in aerated stirred tanks is crucially important for process efficiency and product quality. To date, experimentally generated flow maps and correlations are mostly used to classify flow regimes. However, such an approach is highly limited in terms of scalability and transferability. We propose a model for a soft sensor to classify flow regimes of aerated stirred tanks based only on image data. To select an architecture for the model, we compared various modern architectures including LeNet, VGG16, MobileNetV2, Dense121. Of these, LeNet-5 and custom CNN show the best performance. Furthermore, we tested how disturbances of process and light conditions, and the fill level in the reactor affect the classification performance.

Keywords: stirred aerated tank, flow regime classification, computer vision, CNN.

1. Introduction

Monitoring and optimization of flow regimes in aerated stirred tanks is crucially important for the process efficiency and product quality. The flow regime has a direct impact on the spatial distribution and size of gas bubbles inside the reactor, which affects overall mass transfer rate. Operating the aeration process in an efficient and safe manner demands a comprehensive understanding of the occurring flow regimes (distribution of liquid and gas phases in the mixing process) and their classification. According to (Wang et al., 2017) the recognition and classification of flow regimes in general has a research gap. Stirred tanks in particular require special attention (Khopkar et al., 2005; Liu and Bai, 2019). Currently, manually generated flow maps and correlations are mostly used for this purpose. Boundaries between flow regimes are specified by substance properties, reactor geometry as well as process conditions such as temperature, fill level, flow rate of gas, stirring velocity and geometry of reactor and stirrer. However, such an open-loop approach has decisive drawbacks (Liu and Bai, 2019; Torisaki and Miwa, 2020). Pre-defined flow maps are inflexible regarding uncertainties and variance in substance properties and process conditions. Thus, creation of flow maps itself might be an extremely time intensive task for which effort is proportional to the number of possible parameter combinations. The use of a soft-sensor that could provide the information about the flow regime for a closed-loop control strategy independently of process conditions and the substances used is beneficial in terms of robustness, flexibility, and scalability. Among measurement-based in-line and on-line techniques, image analysis is the only low-cost, non-invasive approach which works with transparent media regardless of their physical properties and process conditions (Bowler et al., 2020). In recent years, artificial intelligence and particularly the use of Convolutional Neural Networks has indicated an enormous potential for solving image classification tasks (LeCun et al., 1998; Liu et al., 2020; Redmon and Farhadi, 2018; Shin et al., 2016; Zhang et al., 2019). The goal of this

paper is to examine the potential of CNNs for solving the flow regime classification problem in aerated stirred tanks with transparent liquids. The input basis for the classification is image data only.

2. Problem description and test set up

Aeration processes in stirred tanks is usually a two-phase problem with a goal to intensify the mass transfer from the gas phase into the liquid phase. Depending on the volume ratio of the two phases, different interface patterns (or flow regimes) can be observed (Pal et al., 2019). According to (Majumder, 2016), a flow regime depends on the following factors: dynamic process parameters, geometry of the reactor and agitators, thermodynamic variables, and physical properties of both phases.

This work focuses on bubbly flows, i.e. a continuous liquid phase with bubble inclusions. We consider only flow regimes in a vertically positioned cylindrical stirred tank equipped with a Rushton turbine, the standard type of the reaction equipment for biochemical and pharmaceutical processes. According to (Paglianti et al., 2000) there are three main flow regimes in such reactors: flooded, loaded and completely dispersed (see Figure 1).

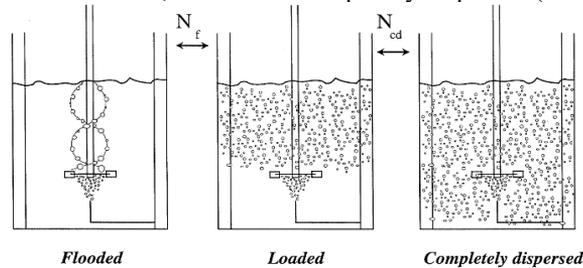


Figure 1: Flow regimes in the stirred tank (Paglianti et al., 2000):

Specifically, the optically visible pattern of gas bubbles in the reactor is proposed as distinguishing criterion for the classification of the flow regime. Stainless steel industrial bioreactors of up to 200 L usually have a single viewing port through which the aeration process can be inspected. Because of this limitation, unlike the cases with glass bioreactors, only image data with a limited view on the process can be used for modelling. We used a 30 L Stedim Biostat® D-DCU bioreactor (D=290 mm, H=640 mm) manufactured by Sartorius. The reactor is equipped with three flat blade turbine impellers (D=95 mm, H=20 mm) and a ring sparging system for aeration using normal air. The vessel is filled with water at the level of 30 L. A single light source is installed in the lid of the bioreactor. For image acquisition we used a Fujifilm X-T20 (aperture F4.5, resolution 1080p, ISO 12800, and shutter speed 1/1250 s).

To cover the operating range of the bioreactor, rotation speed of the stirrer and gas flow rate were varied from 50 to 600 rpm and 5 to 50 nlpm accordingly. Overall, about 10,000 images were collected with 70 combinations of parameters. Figure 2 shows sample images of different flow regimes.

Additional images were collected to test derived models against previously unexposed data where combinations of process parameters, lighting conditions and fill levels were slightly different from those in the training data set. In this way, we wanted to investigate the effect of disturbances that commonly occur in production plants.

3. Data preparation

To ensure that the data set includes only viable images, we manually removed blurred images, or images taken when the stirrer was not running.

The flow regime was determined for each agitator and gas flow setting by visual observation, and images were manually labelled accordingly. To validate the labels, a flow map was created. The position on the flow map, which results from the gas flow number and the Froude number, was used to determine whether the images were marked with the correct flow regime. Clearly incorrect flow regime labels, especially when single images of one class are surrounded by images of another class, were corrected.

Next, image preprocessing techniques were applied. They can improve image detection, for example, by using filters to increase contrast or applying thresholds. Independent A/B tests were carried out for each image preprocessing method and the effect on the model performance was evaluated. The following preprocessing algorithms were tested: Gaussian Blur, Sharpening Filter, Histogram Equalizer, Adaptive Thresholding, Scharr Operator, and Sobel Operator. Additionally, the effect of cropping and converting the color image to grayscale was tested. It was found that the following steps provided the best results: conversion to grey-scale color map, sharpening kernel filter (3x3), and padding with pixels of zero intensity. Data augmentation was carried out using the Keras data generator. During our experiments, we found that random translation, rotation, and horizontal mirroring of the input experimental image provided better results.

4. Modelling

Since there is no previous work about identification of flow regimes with a comparable setup, different CNN models were tested to determine which architecture is the most appropriate. The CNN models were implemented in Python programming language, using the TensorFlow and Keras libraries.

First, we chose the LeNet-5 model (LeCun et al., 1998) due to its simple architecture that has been used for image recognition for years and can be trained very quickly. LeNet5 consists of 5 layers: two convolutional layers, followed by one max pooling layer each. Finally, a fully connected layer and a SoftMax classifier are attached. For the activation we chose the state of the art ReLu function instead of the originally used tanh function. The LeNet model is designed for input data of a dimension of 28x28 pixels. We implemented LeNet additionally for images with 56x56 and 112x112 pixels. All three models were trained for 20 epochs with a learning rate of 0.001 and the Adam Optimizer. Next, we trained more sophisticated networks to compare. VGG16 (Simonyan and Zisserman, 2014), MobileNet (Howard et al., 2017) and DenseNet (Huang et al., 2016) were selected because they have very compact architectures and therefore relatively few parameters to train. For this reason, they have a higher potential for success with the relatively small dataset. For the implementation we used the models from the Keras API with average pooling, Softmax classifier, Adam optimizer and a learning rate of 0.0001. The input shape of the images varied from 56x56 to 224x224 pixels.

Last, we designed custom CNNs. Since smaller models consisting of only a few layers gave better results than larger ones, the LeNet architecture was used as basis: two groups of convolutional layers, ReLu activation, batch normalization and max pooling layer, followed by a fully linked dense layer, a dropout layer and the SoftMax classifier.

In order to find the best set of model parameters, we performed hyperparameter optimization with the hyperband approach using the Keras Tuner tool. The kernel size of the convolutional layers and the max pooling were kept at 3x3 and 2x2. The size of the convolutional layer filters, the size of the dense layer and the learning rate were varied.

We trained the model for 20 epochs with a batch size of 32. This approach resulted in a model with convolutional filters of size 48 and 32, a 768 size Dense Layer and a learning rate of 0.001 (CNN_48_32_768).

In addition, we designed a model consisting of four convolutional layers. The developed model has convolutional filters with the sizes 48, 16, 64 and 128, a 512 large dense layer and a learning rate of 0.001 (CNN_48_16_64_128_512).

The performance of the models and the overall F1-scores are shown in Table 1.

Table 1: Model overview (Prec – Precision, Rec – Recall)

Architecture	Input Shape	Flooded		Loaded		Fully dispersed		Overall F1-score
		Prec	Rec	Prec	Rec	Prec	Rec	
LeNet	(28,28,1)	96%	89%	82%	95%	98%	89%	91%
LeNet	(56,56,1)	99%	100%	96%	92%	94%	97%	96%
LeNet	(112,112,1)	100%	100%	100%	100%	100%	100%	100%
VGG16	(56,56,1)	100%	30%	21%	23%	53%	100%	48%
MobileNetV2	(160,160,3)	0%	0%	36%	100%	0%	0%	37.5%
MobileNetV2	(56,56,1)	0%	0%	30%	100%	0%	0%	29.3%
DenseNet121	(224,224,3)	100%	13%	15%	1%	32%	100%	35.7%
CNN_48_32_768	(56,56,1)	100%	100%	96%	92%	92%	96%	96%
CNN_48_16_64_128_512	(56,56,1)	100%	100%	95%	93%	93%	95%	96%

The LeNet models achieve relatively high F1-scores of over 90% while the other standard CNN models do not exceed values of 90% and therefore do not classify the images with sufficient accuracy. The possible reason is more complex architecture and therefore a higher demand on the training dataset size.

The performance of the implemented LeNet models were compared with an additional test data set. This data set consists of the images taken at different lighting and operating conditions to which the models had not previously been exposed. The results show that the 56x56 resolution images allow for the highest F1-score. Both custom designed CNNs achieved F1-score values of 96%, the same as the LeNet model did.

5. Evaluation and discussion

The best performing models (LeNet with 56x56 resolution and both custom CNNs) were evaluated visually, using saliency maps after (Simonyan et al., 2013). Figure 2 shows a visualization of the training images along with saliency maps for three classes of the 56x56 LeNet model. The relevant pixels are displayed with increasing importance from blue (irrelevant) to red (most relevant). It is apparent that the model focuses on the area where the agitator blades are. This proves that the model detects pixels that are relevant for classification from a chemical engineering point of view. According to (Nienow et al., 1985) exactly the area near agitators provides the most information about current flow regime because cavities build there.

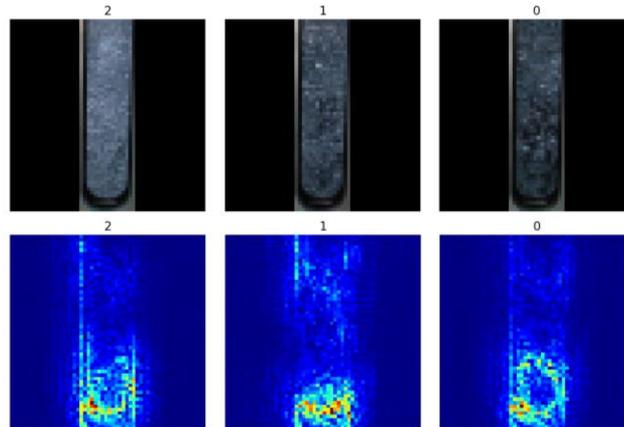


Figure 2: Image samples from collected dataset and corresponding saliency map of trained LeNet 56x56 (2 – completely dispersed, 1 – loaded, 0 – flooded)

The performance of all three models with the test dataset (10% of original dataset) is about 96%. The examination of precision and recall from LeNet for the different flow regimes shows that the flooded regime is identified very reliably (precision 99%, recall 100%). The other two regimes, which resemble each other visually, were more frequently misclassified, where images of loaded regime were classified as completely dispersed. This results in a lower recall for loaded regime (92%) and lower precision (94%) in completely dispersed regime. A similar phenomenon can be observed with the custom CNNs.

Two additional datasets, consisting of images taken under different lighting and operating conditions (different rotational speeds and gas flow rates) show the difference between the models: an F1-score of 73% for LeNet, achieves an 71% for CNN_48_32_768, and 77% for CNN_48_16_64_128_512. In order to compare the performance of the classification without using context data (parameters such as stirrer rotation speed and gas flow rate), 200 images were classified manually in the same test data set. Here, an F1-score of 85% was achieved, only slightly outperforming automatic classification.

Another realistic use case is different filling levels of the reactor, which is why we made an additional data set with levels of 28L and 26L (compared with 30 L for the standard training set). Using the models trained with the standard filling level, the models achieve 96% - 100% on the test dataset. This shows that the filling level has no relevant influence on the accuracy with a relatively small change of 2 to 4L.

6. Conclusion

When comparing various model architectures (LeNet, VGG16, MobileNetV2, Dense121 and custom LeNet alike architecture) for the flow regime classification problem in stirred aerated reactors, our results indicate that shallow networks in general perform better. LeNet and custom CNNs achieved F1-Scores of almost 100% on test dataset and up to 77% on a dedicated test dataset that includes combinations of process parameters to which the models were not exposed during training. Against fluctuation of fill level (up to 4L less) the models showed little or no decrease in performance (96% to 100%).

Acknowledgement: The authors acknowledge the financial support by the Federal Ministry of Economic Affairs and Energy of Germany in the project KEEN (project number 01MK20014T).

References

- A.L. Bowler, S. Bakalis, N.J. Watson, 2020, A review of in-line and on-line measurement techniques to monitor industrial mixing processes, *Chem. Eng. Res. Des.* 153, 463–495, <https://doi.org/10.1016/j.cherd.2019.10.045>
- A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, 2017, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*
- G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, 2016, *Densely Connected Convolutional Networks*
- A.R. Khopkar, S.S. Panaskar, A.B. Pandit, V. V. Ranade, 2005, Characterization of gas-liquid flows in stirred vessels using pressure and torque fluctuations, *Ind. Eng. Chem. Res.* 44, 3298–3311, <https://doi.org/10.1021/ie0493412>
- Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, 1998, Gradient-based learning applied to document recognition, *Proc. IEEE* 86, 2278–2324, <https://doi.org/10.1109/5.726791>
- L. Liu, B. Bai, 2019, Flow regime identification of swirling gas-liquid flow with image processing technique and neural networks, *Chem. Eng. Sci.* 199, 588–601, <https://doi.org/10.1016/j.ces.2019.01.037>
- L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikäinen, 2020, Deep Learning for Generic Object Detection: A Survey, *Int. J. Comput. Vis.* 128, 261–318, <https://doi.org/10.1007/s11263-019-01247-4>
- S.K. Majumder, 2016, Flow Regime and Its Transition, in: *Hydrodynamics and Transport Processes of Inverse Bubbly Flow*. Elsevier, pp. 25–64, <https://doi.org/10.1016/B978-0-12-803287-9.00002-3>
- A.W. Nienow, M. Konno, M.M.C.G. Warmoeskerken, J.M. Smith, 1985, On the flooding/loading transition and the complete dispersal condition in aerated vessels agitated by a Rushton-turbine.
- A. Paglianti, S. Pintus, M. Giona, 2000, Time-series analysis approach for the identification of flooding/loading transition in gas-liquid stirred tank reactors, *Chem. Eng. Sci.* 55, 5793–5802, [https://doi.org/10.1016/S0009-2509\(00\)00125-1](https://doi.org/10.1016/S0009-2509(00)00125-1)
- R. Pal, S. Roy, B. Maji, S. Guin, S. Sarkar, A. Mukhopadhyay, 2019, Flow Regime Identification of Two Phase Flow based on Image Processing Techniques, in: *2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)*. IEEE, pp. 1–6, <https://doi.org/10.1109/INCOS45849.2019.8951401>
- J. Redmon, A. Farhadi, 2018, YOLOv3: An Incremental Improvement
- H.-C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Noguees, J. Yao, D. Mollura, R.M. Summers, 2016, Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, *IEEE Trans. Med. Imaging* 35, 1285–1298, <https://doi.org/10.1109/TMI.2016.2528162>
- K. Simonyan, A. Vedaldi, A. Zisserman, 2013, Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps
- K. Simonyan, A. Zisserman, 2014, Very Deep Convolutional Networks for Large-Scale Image Recognition, 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–14
- S. Torisaki, S. Miwa, 2020, Robust bubble feature extraction in gas-liquid two-phase flow using object detection technique, *J. Nucl. Sci. Technol.* 57, 1231–1244, <https://doi.org/10.1080/00223131.2020.1779145>
- G. Wang, X. Li, C. Yang, G. Li, Z.S. Mao, 2017, New Vision Probe Based on Telecentric Photography and Its Demonstrative Applications in a Multiphase Stirred Reactor, *Ind. Eng. Chem. Res.* 56, 6608–6617, <https://doi.org/10.1021/acs.iecr.7b00871>
- Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, B. Yu, 2019, Recent advances in convolutional neural network acceleration, *Neurocomputing* 323, 37–51, <https://doi.org/10.1016/j.neucom.2018.09.038>

Data-driven modeling for physical property prediction of polypropylene composites using artificial neural network and principal component analysis

Chonghyo J,^{a,b} Hyundo P,^{a,b} Seokyoung H,^a Jongkoo L,^c Hyungtae C,^a Il M,^b Junghwan K^{a,*}

^a*Green Materials and Processes R&D Group, Korea Institute of Industrial Technology, 55, Jonga-ro, Ulsan 44413, Korea*

^b*Department of Chemical and Biomolecular Engineering, Yonsei University, 50, Yonsei-ro, Seoul 03722, Korea*

^c*Research & Development Center, GS Caltex Corporation, 359, Expo-ro, Yuseong-gu, Daejeon, 34122, Korea*

kjh31@kitech.re.kr 82

Abstract

Recent research with artificial neural network (ANN)-based predictive model has emerged as a solution to reduce the number of trial and error effectively. However, it is still challenging to develop a high-performance model using a sparse dataset. Especially, the high-dimension and the small number of polypropylene composite's (PPC's) material data make it difficult to develop a predictive model. In this study, we proposed the ANN-based predictive model using principal component analysis (PCA) to predict the physical property of PPC with the high performance. The optimal dimension reduction of the raw dataset was suggested by the proposed framework to overcome incomplete dataset of PPC materials including the zero values. The dimension reduced dataset was used to develop the ANN-based model for physical property prediction of PPC. As a result, the model accuracy based on the reduced dataset is 0.9061, and 4.6% higher than the model using the raw dataset. This result demonstrates that ANN-based model with dimension reduction improves the prediction performance by reducing the sparsity of PPC material data. Moreover, the proposed model is expected to reduce the number of trial and error in the PPC development process.

Keywords: polypropylene composite, principle component analysis, artificial neural network

1. Introduction

Accompanying the advancement of data science, many data-mining methods have been developed for the purposes of prediction. One such representative data mining method is an artificial neural network (ANN), which has been widely used to develop data-driven models owing to its high performance and effectiveness. An ANN analyzes the relationships between a dependent variable and independent variables, updates the relationships by minimizing the error, and calculates a dependent variable using such relationships when unknown independent variables are given. A data-driven model using an ANN, called an ANN-based model, has been applied to many fields to reduce the trial

and error required, which is one of the major issues in numerous industries, by replacing experimental values with predicted values. However, if an ANN-based model is developed with a sparse dataset where a higher percentage of the variable cells do not have actual data, the predicted values cannot be used owing to the reduction in the model accuracy.

This problem often occurs during the development of polypropylene composites (PPCs). Because many industries require the specific physical properties of PPCs for engineering applications, numerous experiments on their physical properties have been conducted to satisfy such requirements. Although the number of experiments should be reduced to increase the efficiency, it is difficult to reduce it using data-driven modeling because of the sparse dataset of recipes, which indicates the combination and ratio of raw materials used in PPCs. In addition, although the recipe dataset is composed of 90 types of raw materials, not every type of raw material is used in a recipe at the same time. Normally, only 12 types of raw materials at maximum are blended for a PPC. For this reason, the dataset includes many zeros values, becomes sparse, and decreases the performance of the data-driven model. Hence, a dimension reduction has to be applied when an ANN-based model is developed with a recipe dataset for a prediction of the physical properties.

This study proposes a data-driven model for the prediction of the physical properties of PPCs using an ANN and a principal component analysis (PCA). Figure 1 shows an overview of the study. First, data sparsity was found, and the sparsity problem was solved through a dimension reduction process. Among the dimension reduction methods, a PCA, a representative method, was applied. The explained variance ratio was calculated to select the optimal number of dimensions. Second, to validate the dimension reduction method, two ANN-based models were developed using the original and reduced datasets. The performance of the models was then evaluated and compared with R^2 .

The novelty and contributions of this paper are as follows:

- This is the first study to apply an ANN and a PCA to a polypropylene composite (PPC) dataset used in the commercial chemical industry.
- This study improves the model performance of an ANN-based model for a physical property prediction through a dimension reduction.
- This study demonstrates that the ANN-based model with a PCA is a powerful data-driven model for reducing the amount of trial and error required during the PPC development process.

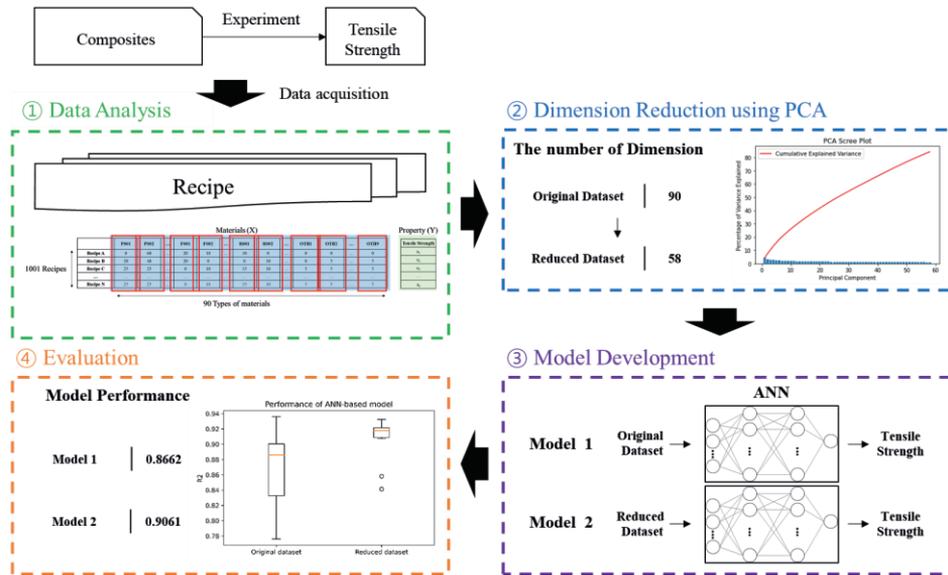


Figure 1 Overview of the present study

2. Method

2.1. Data Analysis

Figure 2 shows the PPC dataset used in this study. The dataset contains information on 1001 recipes, which consist of 90 materials including polypropylene (P), filler (F), rubber (R), and other additives (OTH), as well as the tensile strength. Most importantly, each recipe has many zero weight percentage because only 12 types of materials are used in the recipe at maximum, which means that each recipe has 78 zero values at minimum. The number of zero values determines the data sparsity and has an impact on data-driven modeling. Therefore, the analysis of data sparsity is crucial before modeling.

To analyze the number of zero values, the percentages of zero values in different materials are calculated using. In the PPC dataset, the percentages are distributed from 31.3% to 99.9%. This result shows that both popular and unpopular materials exist in the PPC development. For example, the minimum percentage of zero values (31.3%) is for OTH4, which means that 314 of 1001 recipes do not contain OHT4. By contrast, the maximum percentage of zero values (99.9%) is for P013, which means that 1000 of 1001 recipes do not contain P013. For these reasons, the PPC dataset is sparse owing to the existence of many zero values, and thus the number of zero values must be reduced for data-driven modeling.

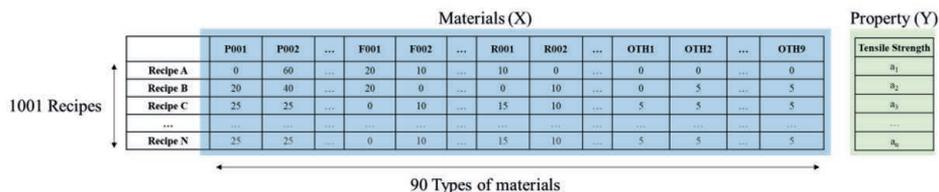


Figure 2 PPC dataset used in this study

2.2. Principal Component Analysis

The sparsity of the dataset can be solved using dimension reduction methods. One popular method is a principal component analysis (PCA), which has been a powerful method since its initial development. PCA transforms a large set of variables into a smaller set that still contains most of the information in a large set. The information is usually defined as a variance of variables, which cannot be maintained when the number of dimensions is too small. Therefore, selecting the optimal number of dimensions is important when reducing the number of dimensions to maintain the variance properly.

2.3. Artificial Neural Network

Among various data mining methods, artificial neural networks (ANNs) have been widely used for data-driven modeling due to their high performance. ANN consists of three layers: an input layer, a hidden layer, and an output layer. Each layer contains nodes and edges that contain mathematical relations with weights and biases. An ANN minimizes the error by updating the weights during model training, and the developed model estimates an unknown dependent variable from the known input and output pairs, including the mathematical relations.

3. Results and discussions

3.1. Selecting the number of dimensions

The variance ratio explained is calculated to select the optimal number of dimensions in the PCA. Figure 3 shows a PCA scree plot of the PPC dataset, with the blue bars representing the explained variance ratio. The ratio represents the variance explained by each of the principal components, starting with the first component, which is the principal component that explains most of the variance. When the ratios are summed, the total value is equal to 1, indicating that the 90 components together explain 100% of the variance of the dataset. More usefully, the variance ratio explained is used as a cumulative sum, such as indicated by the red curve in Figure 3. In this study, we found that 58 components are required to explain at least 85% of the information in the dataset from the cumulative explained variance. Hence, 90 types of materials were replaced with 58 principal components.

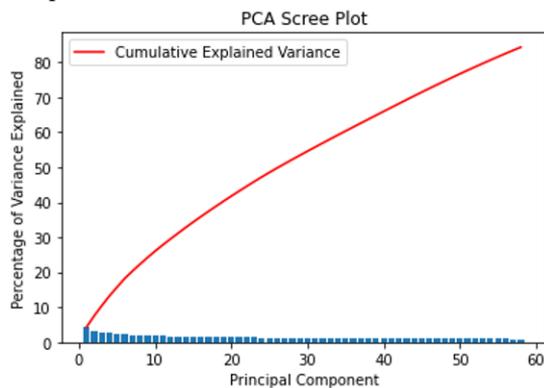


Figure 3 PCA scree plot of PPC dataset

3.2. Development of ANN-based models

After a dimension reduction, we developed ANN-based tensile strength prediction models using two datasets: the original dataset with 90 types of materials and the reduced dataset with 58 components. An ANN-based model uses weight percentage of 90 types of materials as inputs, whereas the other model uses the values of 58 components as inputs. The two datasets were randomly split into 60% training, 20% validation, and 20% test datasets. Although the training and validation datasets were used for data-driven modeling, the test dataset was used to evaluate the models. The performance of the models is evaluated using R^2 , which is a representative evaluation criterion for regression and is calculated through Eq. (1).

$$R^2 = 1 - \frac{\sum_{i=1}^N (S_{ia} - S_{ip})^2}{\sum_{i=1}^N (S_{ia} - \bar{S}_{ip})^2} \quad (1)$$

In the above equation, N denotes the number of data, S_{ia} denotes the i^{th} actual value, and S_{ip} denotes the i^{th} predicted value of the model.

The ANN used in this study has one input layer, one output layer, and two hidden layers consisting of 30 nodes, an Adam optimizer, and ReLU as an activation function for each layer. Namely, the hyperparameters of the two ANN-based models are identical, except for the number of nodes in the input layer.

Although the ANN hyperparameters are fixed, because R^2 depends on the datasets (training dataset, validation dataset, and test dataset), ten case studies were conducted. Each case consists of a comparison of the two ANN-based models after the random data split. Table 1 shows the results of the case studies. The average R^2 of the ANN-based model using the original dataset was calculated as 0.8662, whereas the average R^2 using the reduced dataset was calculated as 0.9061. This difference demonstrates that a dimension reduction contributes to the improvement of the model performance. The R^2 of the model increased by 4.6% after the dimension reduction in this study.

Table 1 Results of ten case studies

	R^2 of ANN-based model using the original dataset	R^2 of ANN-based model using the reduced dataset
Case 1	0.82931	0.93247
Case 2	0.84299	0.85808
Case 3	0.89898	0.91378
Case 4	0.79735	0.91785
Case 5	0.89373	0.90737
Case 6	0.93614	0.93069
Case 7	0.90797	0.91807
Case 8	0.77613	0.91948
Case 9	0.87827	0.92202
Case 10	0.90079	0.84131
Average	0.8662	0.9061

In addition, we analyzed the interquartile range, called IQR, of the results, as shown in Figure 4. The ANN-based model using the original dataset was unstable because the model had a wide range of R^2 values of 0.77613 to 0.93614. However, the ANN-based model using the reduced dataset was more stable than that using the original dataset because the range of R^2 was from 0.84131 to 0.93247. While the model using the original dataset is able to be tested with different inputs because maximum 12 materials are used

in 90 types, the model using the reduced dataset is tested with same inputs because the dataset is not sparse. Therefore, the sparsity in the dataset should be reduced before data-driven modeling when using sparse data. Consequently, a dimension reduction is required not only to improve the performance but also the robustness of the data-driven model.

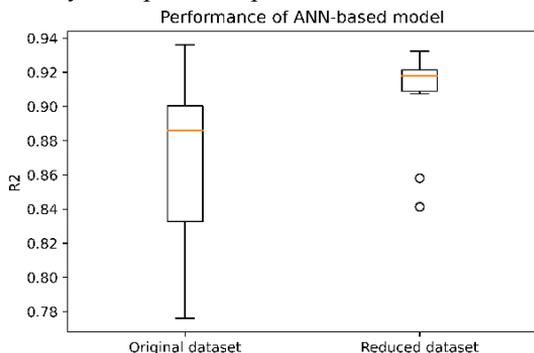


Figure 4 IQR graph of the R^2 of the two ANN-based models using the original and reduced datasets

4. Conclusion

In the PPC development process, an ANN-based prediction model is suggested to reduce the amount of trial and error required. However, the sparsity of the PPC dataset hinders the development of the prediction model. In this study, we proposed an ANN-based model for the property prediction of PPC using a PCA. A total of 58 components were selected as the optimal number of dimensions. To compare the performances of the proposed model using 58 components and the original model using 90 types of materials, two ANN-based models were developed with the same hyperparameters and evaluated based on the R^2 . The results illustrate that the proposed model outperforms the original model and is more robust than the original model. Thus, this study proved that a dimension reduction is essential for data-driven modeling when using sparse datasets.

The proposed model is expected to be applied to the other fields handling sparse datasets as a powerful solution to reduce the amount of trial and error, such as blending processes. In future studies, we will compare the use of a PCA with other dimension reduction methods and extend this study to other properties of PPC.

Acknowledgements

This study has been conducted with the support of the Korea Institute of Industrial Technology as “Development of AI Platform Technology for Smart Chemical Process (kitech JH-21-0005)” and “Development of Global Optimization System for Energy Process (kitech EM-21-0022)”

References

- M. Habeck, “Statistical mechanics analysis of sparse data”, *Journal of Structural Biology*, vol. 173, issue 3, pp. 541-548
- Z. Zhang, A. Castello, "Principal components analysis in clinical studies", *Annals of Translational Medicine*, vol. 5, issue 17, pp. 1-7, 2017, doi: 10.21037/atm.2017.07.12
- C. Joo, H. Park, "Development of physical property prediction models for polypropylene composites with optimizing random forest hyperparameters", *International Journal of Intelligent Systems*, 2021, doi: 10.1002/int.22700

Optimization of the Wastewater Treatment Plant Aeration Using Artificial Neural Networks Models

Norbert-Botond Mihály and Vasile Mircea Cristea

Babeş-Bolyai University of Cluj-Napoca, Faculty of Chemistry and Chemical Engineering, 11 Arany Janos Street, 400028, Cluj-Napoca, Romania

Abstract

The present work aimed the modelling and optimization of the Wastewater Treatment Plant (WWTP) operation based on predicting its energy and quality performance indices using Artificial Neural Networks (ANNs). The best model architecture and structure were searched among three different ANN types, with different topologies. A standard dataset originating from the plant calibrated first-principle model (FPM) data was used to develop the ANN models. Their performance was evaluated by the coefficient of determination and mean squared error (MSE) values, first at testing and subsequently at the prediction performed for a new input dataset. Using the most promising identified ANN types and topologies, two ANN structures were investigated, one with three single output neural networks and another one with a single network with three outputs for predicting WWTP performance indices: aeration energy, effluent quality and pumping energy. The analytical model and the two ANN structures were used in the study of the aeration optimization of the WWTP, for finding the optimal air distribution in the aerated reactors. The obtained results were tested and compared taking into account the performance index values as well as the required computation time. The developed ANN models showed similar results to the FPM in terms of performance indices, while the required computation time was reduced by several orders of magnitude.

Keywords: artificial neural network, wastewater treatment plant model, aeration optimization

1. Introduction

Artificial Intelligence (AI) techniques are gaining popularity in Waste Water Treatment Plant (WWTP) modelling, optimization and control due to their simplicity compared to the analytical models. The well-known Activated Sludge Models (ASMs) describe tens of subprocesses, leading to numerous model components and coefficients [1-4]. Relying on the complex ASM models used in operation optimization of the WWTP proves to be a demanding task for the model development [5], but also for the computing resources and time [6]. AI techniques present an alternative approach, as they have proved their applicability in the field of chemical and environmental engineering [7]. ANNs are widely used to develop data-driven models, and these networks are capable of modeling nonlinear processes, such as those specific to the WWTP, and also show adequacy for forecasting or classification tasks [8].

Several studies focused on developing ANN models for applications in WWTP modelling. Such networks determined the required alum dosage [9], or estimated plant costs [10]. ANN modelling was used in various areas of WWTP modelling, the focus being on effluent modelling and prediction. These studies successfully emerged in models capable of accurate predictions of the effluent variables, such as: Total Suspended Solids

(TSS) [11]; Biological Oxygen Demand (BOD), Chemical Oxygen Demand (COD) [12]; and pH, total nitrogen and COD [13]. Control of the effluent COD level was also implemented using ANN based predictions [14]. The aforementioned steady-state studies revealed that most often a simple ANN structure is sufficient in providing accurate results. The studies regarding the implementation of dynamic neural networks for prediction the dynamics of the WWTP effluent variables are scarcer, a representation being the network developed for the diagnosis of a biological wastewater treatment process that removes organic carbon residuals [15]. Another research showed that recurrent neural network models (RNN) are suitable for working with time series data and temporal changing behavior [16].

This work proposes to design, train and use the most suitable dynamic ANN model for the purpose of WWTP aeration optimization.

2. Methodology

The basic data and methods used for the development of the ANN models is shown next.

2.1. Data collection and ANN structure

Multiple ANN types were studied with the aim of finding the most efficient networks to be used for making predictions. Based on the proposed training methodology and best ANN results, new ANN models can be directly developed using exclusive industrial plant data. The data used for training was obtained from simulations with a calibrated FPM of a municipal WWTP, based on ASM1 (first from its family). Dynamic simulations on 458 days provided the database for the development of ANNs and data were collected with a sampling time of 0.5 hours. Simulation scenarios were selected using a full factorial Design of Experiment method, having 4 levels for each of the three gain factors of the air flowrate distribution in the aerated reactors, and resulting in a total 64 different simulation scenarios. This approach provided a statistically comprehensive training dataset.

The following WWTP influent variables (as mean values during each sampling period) were taken from plant measurements: chemical oxygen demand (COD), NH_4^+ & NH_3 nitrogen (NH) concentration, volumetric flow (Q) and temperature (T). The outputs of the networks were chosen as the three WWTP performance indices: aeration energy (AE), effluent quality (EQ) and pumping energy (PE), computed according to equations (1)-(3).

$$AE_i = \frac{SO_{sat}}{1.8 \cdot 1000 \cdot ts} \cdot \int_{t_{i-1}}^{t_i} \sum V \cdot K_L a(t) dt \quad (1)$$

$$EQ_i = \frac{1}{1000 \cdot ts} \cdot \int_{t_{i-1}}^{t_i} [PU_{TSS}(t) + PU_{COD}(t) + PU_{BOD}(t) + PU_{TKN}(t) + PU_{NO}(t)] \cdot Q_e(t) dt \quad (2)$$

$$PE_i = \frac{1}{ts} \cdot \int_{t_{i-1}}^{t_i} [0.004 \cdot Q_{NR}(t) + 0.008 \cdot Q_{RAS}(t) + 0.05 \cdot Q_w(t)] dt \quad (3)$$

The AE index is calculated based on the oxygen transfer coefficient ($K_L a$), the volumes of the aeration tanks (V) and the saturated oxygen concentration (SO_{sat}). The EQ index considers for computation the effluent concentration of total suspended solids (TSS), COD, biological oxygen demand (BOD), total Kjeldahl nitrogen (TKN), nitrate and nitrite nitrogen (NO) concentration and effluent flow rate (Q_e). The PE is computed based on the nitrates recirculation (Q_{NR}), return activated sludge (Q_{RAS}) and waste (Q_w) flow rates. The performance indices were computed at each point in time t_i by integration over period t_{i-1} to t_i and considering the sampling time ts . Two main structures were designed for each type of investigated ANN, one multiple input multiple output (MIMO) structure with a single network trained to predict all outputs and a multiple input single output (MISO) structure with 3 parallel networks, each trained for predicting one of the ANN outputs. Three ANN types were considered: Time Delay Neural Network, Recurrent Neural Network and Generalized Regression Neural Network.

2.2. ANN type and architecture

All of the investigated networks were trained using past and present data to obtain suitable models for time-series prediction. In each case, the inputs for point t_i in time consisted of the influent variables and output(s) past values, from points t_{i-12} to t_{i-1} in time, and the influent parameters at point t_i in time. They served to compute the output(s) at point t_i .

2.2.1. Time delay neural network

Time Delay Neural Networks (TDNNs) are feed-forward neural networks that have a so-called tapped delay line block between the inputs and the first hidden layer, allowing the networks to work with data from several past time moments. In our study the number of sampling time delays for the input data was chosen as 12. The dataset was divided into three separate subsets: training, validation and testing subsets. The training subset amounted to 70% of the total data, while the other two subsets amounted to 15% each. In order to find the best ANN architecture, the number of hidden layers and the number of neurons in each of the hidden layers were varied.

2.2.2. Recurrent neural network

Recurrent Neural Networks (RNNs) have a similar construction to the above mentioned TDNNs. These networks also contain the tapped delay line. However, the subset of inputs consisting in the past values of the outputs is directly fed by feedback from the network's own outputs. In this case only the number of neurons in the hidden layer was varied. The dataset was also divided in 70% used for training, while 15% and 15% was used for validation and testing.

2.2.3. Generalized regression neural network

The Generalized Regression Neural Networks (GRNNs) have two layers. The first one is a radial basis layer, while the second is called a special linear layer. The weights of the input layer consist in the transposed matrix of the input training dataset, while the weights matrix of the second layer is the matrix composed of the target training vectors dataset. As a result, this type of ANN is easily developed and has very good function approximation performance. The available data was divided in two subsets, 70% for training and 30% for testing. The GRNN spread parameter was varied from 0.1 to 2.

2.3. Selection criteria for the best ANN models

Two criteria were taken into consideration when selecting the best ANNs: the coefficient of determination (R^2) and the mean squared error (MSE). At the testing step they were used to select the best model from the different explored topologies. Then, these networks were used for predicting the performance indices for the next 7 days, using a new and not yet known set of ANN input data. In this process, the developed networks worked similarly to RNNs, i.e. their own computed outputs were used as inputs for the next time step prediction.

For this comprehensive testing procedure, all of the models were evaluated again for the prediction performance by calculating the R^2 and MSE given in equation (4), in order to assess the fit between the ANN predicted outputs and targets:

$$MSE = \sum_{i=1}^N \frac{(x_i - y_i)^2}{N} \quad (4)$$

where N is the number of data points, y_i is the desired output and x_i is the model output.

2.4. Optimization procedure

The optimization aims to find the optimal gain factors values for the airflows entering the three aerated reactors, taking into account the general objective function consisting in the sum of the averages of the three WWTP performance indices (AE, EQ, PE).

ANN design, training and optimization were carried out using Matlab and Simulink software environment.

3. Results and Discussion

3.1. ANN type and architecture

The results of the ANN model performance and selection criteria based on R^2 and MSE are presented in Table 2, both for the testing and prediction steps.

Table 2. Testing and prediction performance of the ANN models

ANN topology	No. of hidden layers	Output	Testing step		Prediction step	
			R^2	MSE* 10^{-3}	R^2	MSE* 10^{-3}
TDNN	1	AE	1.00	10.7	0.91	169
TDNN	1	EQ	1.00	0.16	0.93	1144
TDNN	1	PE	1.00	4.57E-04	0.96	2.83
TDNN	1	All	1.00	3.78	0.99	906
RNN	1	AE	0.99	32.0	0.97	75.3
RNN	1	EQ	1.00	1.77	0.89	2199
RNN	1	PE	1.00	0.21	0.82	21.8
RNN	1	All	1	18.55	0.99	587
TDNN	2	AE	1.00	12.5	0.94	113
TDNN	2	EQ	1.00	0.11	0.93	1199
TDNN	2	PE	1.00	3.21E-04	0.97	2.43
TDNN	2	All	1.00	3.87	0.97	1893
GRNN	2	AE	0.97	128	0.94	159
GRNN	2	EQ	0.97	381	0.92	1319
GRNN	2	PE	0.99	0.54	0.93	13.8
GRNN	2	All	1.00	166	0.99	477

3.1.1. Time delay neural network

In case of the TDNN models with a single hidden layer, the number of hidden neurons was varied from 2 to 20. The 7-16-1 network proved to be the best from all of the MISO models for predicting EQ.

For the TDNN with 2 hidden layers the number of hidden neurons were varied from 1 to 15 for each layer. This 7-15-15-1 MISO architecture showed the best prediction results for PE, among all the designed and trained networks.

3.1.2. Recurrent neural network

The number of neurons of the hidden layer was varied from 2 to 20. In this case the best ANN model had 10 neurons and it presented the best results at the prediction step for AE.

3.1.3. Generalized regression neural network model

At training, the spread design parameter was varied from 0.1 to 2 by increments of 0.1. The best prediction results were achieved for the MIMO model with 0.7 value of the spread.

3.2. Selection criteria for the best ANN

Selection of the best networks was based on the R^2 and MSE results at the prediction step. Accordingly, the best MISO architecture was the one of the RNN type for AE prediction, the TDNN type with one hidden layer for EQ prediction, and the TDNN type with two hidden layers for the PE prediction. When all AE, EQ and PE aggregated indices were considered, the GRNN with multiple outputs showed the best results and this architecture was chosen as the MIMO model to be further used. The ANNs prediction and target series are shown in Figure 1 for the three best MISO (3MISO) and MIMO models respectively. It can be observed that MISO ANNs perform better at predicting the AE and PE values, while both MISO and MIMO ANN structures perform similarly for EQ prediction.

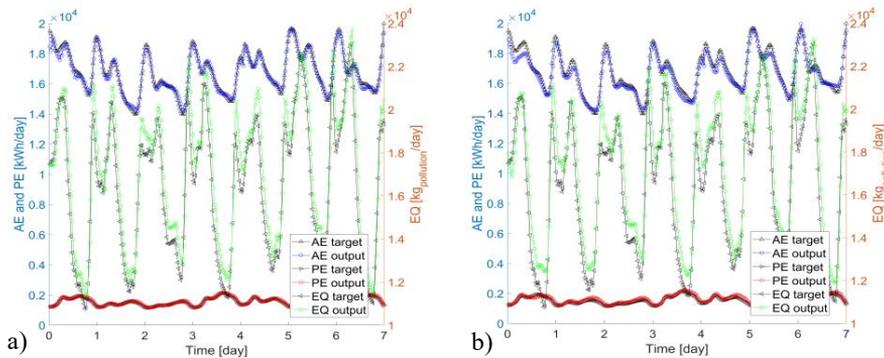


Figure 1. ANN predictions and targets for a) 3MISO and b) MIMO models

3.3. Optimization procedure

Optimization of the WWTP aeration was carried out for three different cases. The objective function was the sum of the averages of the WWTP performance indices for the investigated time period of 7 days, as shown in eq. (5).

$$f_{min} = \frac{\sum_{i=1}^N AE_i}{N} + \frac{\sum_{i=1}^N EQ_i}{N} + \frac{\sum_{i=1}^N PE_i}{N} \quad (5)$$

The optimization was performed for each of the three different cases using NOMAD algorithm [17]; firstly, with the FPM, secondly with the best ANN 3MISO models, and lastly with the best GRNN MIMO model. The WWTP performance simulation results with the optimized air distribution gain factors, for each of the cases, were compared and the computation times (Compt. time) for optimization is also shown in Table 2.

Table 2. Optimization results

Model Case	Gain 1	Gain 2	Gain 3	Compt. time [s]	AE [kWh/day]	EQ [kg/day]	PE [kWh/day]
FPM	0.80	0.41	0.38	56,460	16,616	16,572	1,602
3MISO	0.80	0.40	0.34	3.554	16,791	16,474	1,560
MIMO	0.95	0.40	0.40	1,719	16,794	16,473	1,584

All three types of models led to similar results. The structures based on ANNs succeeded to reduce the EQ and PE indices, while the FPM found the best solution for reducing the AE index. But notably, we may observe that the differences in computation time are much more significant than those of the WWTP performance indices. While the optimization took nearly 16 hours using the FPM, it concluded in a few seconds employing the 3MISO model, and took about half an hour for the MIMO structure. The difference in the computation time between the ANN structures can be attributed to the differences in their construction. TDNNs and RNNs were implicitly designed to work in a recurrent design, while this was added for the GRNNs as a extra task at each iteration. It is also worthy to mention the downside of the TDNNs and RNNs consisting in the typically large time duration needed for their training, when compared to the much faster training of GRNNs.

4. Conclusions

In this paper, the design and training ANN dynamic models for predicting the WWTP AE, EQ and PE performance indices are presented, with the aim to be further used for optimization of the WWTP aeration. Two ANN structures were taken into consideration, a mix of three different types of networks which showed the best results for the single

output predictions and the network that had the lowest error values when predicting all three WWTP performance indices. The R^2 of the designed ANNs was in the range of 0.93-0.99 and demonstrated very good prediction accuracy. Optimization performed using these ANN models revealed similar results to those obtained using the first principle analytical model. But the duration of the optimization time is dramatically reduced when ANN models were used. The MIMO structure was more than 30 times faster, while the optimization based on the 3MISO structure finished more than four orders of magnitude faster. The results prove the potential and applicability of these ANN models for the real time WWTP optimization aiming its operation improvement.

References

1. M. Henze, C.P.L. Jr. Grady, W. Gujer, G.v.R. Marais and T. Matsuo, 1987, Activated Sludge Model No. 1. (IAWPRC Scientific and Technical Report No. 1) London: IAWPRC.
2. M. Henze, W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel and G.v.R. Marais, 1995, Activated Sludge Model No. 2. (IAWQ Scientific and Technical Report No. 3) London: IAWQ.
3. M. Henze, W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel, G.v.R. Marais, and M.C.M. van Loosdrecht, 1999, Activated Sludge Model No. 2d, ASM2d, *Wat. Sci. Technol.*, Vol. 39, No. 1, pp 165–182.
4. W. Gujer, M. Henze, T. Mino and M.C.M. van Loosdrecht, 1999, Activated Sludge Model No. 3, *Wat. Sci. Technol.*, Vol. 39, No. 1, pp. 183–193.
5. G.S. Ostace, V.M. Cristea and P.S. Agachi, 2011, Extension of activated sludge model no 1 with two-step nitrification and denitrification processes for operation improvement, *Environ. Eng. Manag. J.*, Vol. 10, No. 10, pp.1529–1544.
6. M. Simon-Varhelyi, V.M. Cristea and A.V. Luca, 2020, Reducing energy costs of the wastewater treatment plant by improved scheduling of the periodic influent load, *J. Environ. Manage.*, Vol. 262, 110294.
7. V. Venkatasubramanian, 2019, The promise of artificial intelligence in chemical engineering: is it here, finally?, *AIChE J.*, Vol. 65, No. 2, pp. 466–78.
8. K.B. Newhart, R.W. Holloway, A.S. Hering and T.Y. Cath, 2019, Data-driven performance analyses of wastewater treatment plants: a review, *Water Res.*, Vol. 157, pp. 498–513.
9. A. Robenson, S.R. Abd. Shukor and N. Aziz, 2009, Development of Process Inverse Neural Network Model to Determine the Required Alum Dosage at Segama Water Treatment Plant Sabah, Malaysia, *Comput.-Aided Chem. Eng.*, Vol. 27, pp. 525–530.
10. M. Marzouk and M. Elkadi, 2016, Estimating water treatment plants costs using factor analysis and artificial neural networks, *J. Cleaner Prod.*, Vol. 112, Part 5, pp. 4540–4549.
11. A. Tumer and S. Edebali, 2015, An Artificial Neural Network Model for Wastewater Treatment Plant of Konya, *IJISAE*, Vol. 3, No. 4, pp. 131–135.
12. M.S. Nasr, M.A.E. Moustafa, H.A.E. Seif and G.E. Kobrosy, 2012, Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT, *Alexandria Eng. J.*, Vol. 51, No. 1, pp. 37–43.
13. E.B. Hassen and A.M. Asmare, 2019, Predictive performance modeling of Habesha brewery wastewater treatment plant using artificial neural networks, *Chem. Int.*, Vol. 5, No. 1, pp. 87–96.
14. N. Bekkari and A. Zeddouri, 2019, Using artificial neural network for predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant, *Manag. Environ. Qual. Int. J.*, Vol. 30, No. 3, pp. 593–608.
15. M. Miron, L. Frangu, G. Ifrim and S. Caraman, 2016, Modeling of a wastewater treatment process using neural networks, 20th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, October 13-15, pp. 210–215.
16. I. Pisa, I. Santin, J.L. Vicario, A. Morell and R. Vilanova, 2018, A Recurrent Neural Network for Wastewater Treatment Plant effluents' prediction, In: XXXIX. Jornadas de Automatica, pp. 621–628.
17. C. Audet and J.E. Jr. Dennis, 2006, Mesh adaptive direct search algorithms for constrained optimization, *SIAM J. Optim.*, Vol. 17, pp. 188–217.

The truncated Q statistic for Statistical Process Monitoring of High-Dimensional Systems

Marco S. Reis,^{a,*} Ricardo Rendall,^a Tiago J.Rato,^a Cristina Martins,^b Pedro Delgado^b

^a*University of Coimbra, CIEPQPF, Department of Chemical Engineering, Rua Silvio Lima, Pólo II – Pinhal de Marrocos, 3030-790 Coimbra, Portugal*

^b*Bosch Car Multimedia, SA, Rua Max Grundig 35, Lomar, 4705-820, Braga, Portugal*
marco@eq.uc.pt

Abstract

Among the high-dimensional methods for process monitoring, the most widespread and researched are those belonging to the class of latent variable approaches. As more variables are continuously gathered, the existing latent variable methods start to experience difficulties in maintaining their good monitoring properties. Of particular importance is the decreased sensitivity in detecting localized faults through the Q statistic as the accumulation of residuals over a very large number of variables masks the deviations due to faults affecting one or a small number of sensors. To overcome this limitation, we propose the use of the truncated Q statistic. This statistic has the ability to identify the relevant residuals before summing them, by comparing their magnitude against a statistically motivated allowance threshold. Only the residuals beyond the threshold are kept and used to compute the truncated Q statistic (this step is then repeated for each new observation). The performance of the proposed statistic was analyzed through Monte Carlo simulations of several latent variable processes with 5 000 to 20 000 variables. Process monitoring was performed by replacing the Q statistic in the standard combination of T^2 & Q of MSPC-PCA, but other latent variables monitoring methods can benefit from this proposal as well. The results obtained clearly show that MSPC-PCA with the truncated Q statistic becomes much more sensitive than the original formulation; for instance, it can detect faults with a magnitude of 2 standard deviations more than 90 % of the times, whereas MSPC-PCA with the standard Q statistic only detected faults with a magnitude of 4 standards deviations less than 40 % of the times. This methodology was also applied in a real world process to monitor more than 17 000 variables associated to printed circuit boards (PCB) produced by Bosch Car Multimedia Portugal. For this case study, the proposed truncated Q statistic was able to identify 17 abnormal PCBs while the standard Q statistic only detected 5 PCBs.

Keywords: Statistical Process Monitoring; Latent variables; High-dimensional processes.

1. Introduction

As the number of monitored variables increases, the classic Statistical Process Control (SPC) methodologies start to lose sensitivity in detecting localized faults. This is more critical for the residual statistics since they accumulate the model's residuals over a very large number of variables, which effectively raise their basal level during normal operation conditions. Thus, faults are only detectable if they surpass not only the normal operation variation of the faulty variables, but also the sum of all irrelevant residuals,

which may be considerably high in very high-dimensional systems. To address this limitation, we propose the use of a truncated Q statistic that, for each new observation, screens for the relevant residuals before summing them.

This problem is here illustrated for the Q statistic of principal component analysis (MSPC-PCA) (Jackson et al., 1979, Jackson, 1959) since it is the most well-know and widespread methodology. Nevertheless, the proposed methodology can also be applied to other methods such as those based on artificial intelligence (Yin et al., 2019, Yan et al., 2019) or machine learning (Lee et al., 2020, Ge et al., 2017, Rato et al., 2017). Furthermore, the Q statistic of MSPC-PCA is equivalent to other residual statistics, such as the T_r statistic of canonical variate analysis-based monitoring (Jiang et al., 2015), the CVD index of Canonical variate dissimilarity analysis (Pilario et al., 2018), the SPE statistic for (non-linear) manifold-based SPC (Li et al., 2015), and the SPE statistic for nonlinear process monitoring based on kernel global-local preserving projections (Luo et al., 2016). As the proposed methodology effectively screens for the relevant residuals before applying the standard monitoring statistic, any monitoring methodology based on the analysis of residuals may potentially benefit from this proposal.

The rest of this paper is organized as follows. In Section 2, the proposed truncated Q statistic is described. In Section 3, the results for two case studies are presented and discussed. Finally, a summary of the conclusions is provided in Section 4.

2. Methodology

The standard Q statistic of MSPC-PCA directly sums the squared residuals over all variables. Thus, for very high-dimensional systems, deviations in a single sensor or a small set of sensors can be masked by the noisy residuals generated by normal variables. Instead of summing all residual, it is here proposed to screen, in each observation, for the relevant residuals, namely using a proper threshold, τ . Afterwards, only the residuals that were found to be relevant are used to build the so-called truncated Q statistic.

To define the subset of relevant residuals, they are first standardized as:

$$\check{e}_j = \frac{e_j - \bar{e}_j^{NOC}}{s_j^{NOC}}, \quad (1)$$

where, \check{e}_j is the standardized residual for the j -th variable, \bar{e}_j^{NOC} and s_j^{NOC} are the sample mean and standard deviation of the j -th residual under normal operation conditions.

Afterwards, the absolute value of the standardized residuals is compared against a threshold and those found to be below it are set to zero. Therefore, they will not impact the value of the truncated Q statistic. The threshold can be defined by the user. We propose to set $\tau = 3$, representing the typical “3-sigma” limits (Cinar et al., 2007). As the significant residuals (in absolute value) are always larger than τ , they are further corrected by subtracting this constant. This operation is analogous to that performed by the cumulative sum (CUSUM) control chart (Montgomery, 2008), and corresponds to save only the deviations beyond an allowance threshold. The aforementioned operations can be performed through a soft thresholding function $\Phi_\tau(x)$ with parameter τ :

$$\Phi_\tau(x) = \begin{cases} x - \tau & \text{if } x > \tau \\ x + \tau & \text{if } x < -\tau, \\ 0 & \text{if } |x| \leq \tau \end{cases} \quad (2)$$

Following these definitions, the truncated Q statistics is computed as:

$$Q_{\Delta} = \sum_{j=1}^m \left\{ \Phi_{\tau}(\tilde{e}_j) \right\}^2, \quad (3)$$

where m is the number of variables. The control limits for Q_{Δ} can be determined by adjustment of the empirical distribution obtained from normal operation conditions or by a kernel density estimator.

3. Results

3.1. Simulated Case Study

To evaluate the performance of the proposed truncated Q statistic, a series of very high-dimensional processes were simulated. For each process, a matrix of measured data (\mathbf{X}) is generated through a latent variable structure defined by,

$$\mathbf{X} = \mathbf{T}\mathbf{A}^T + s\mathbf{E}, \quad (4)$$

where \mathbf{T} is a $(n \times p)$ matrix of latent variables, \mathbf{A} is a loadings matrix with orthogonal columns randomly generated, \mathbf{E} is a $(n \times m)$ matrix of residuals and s is a scaling factor. Furthermore, n is the number of simulated observations, m is the number of variables and p is the number of latent variables. In this study, \mathbf{T} and \mathbf{E} are generated through random realizations of the standard normal distribution. The scaling factor (s) is tuned so that the error component represents about 5 % of the total variation of \mathbf{X} .

The data generation procedure was run with $p = 10$ latent variables and for four very high-dimension process with 5 000, 10 000, 15 000 and 20 000 variables (m). For each of the four very high-dimension scenarios, the data generation procedure was repeated 100 times. In each replicate, a calibration dataset with 1 000 observations is generated to train the PCA model. Similarly, a validation dataset with another 1 000 observations is generated to establish the control limits of the monitoring statistics.

Faulty datasets with 1 000 observations each were generated by introducing a step deviation in the first variable of the measurement matrix (*i.e.*, on the first column of \mathbf{X}). The magnitude of this fault is defined as k times the in-control standard deviation of the affected variable. Note that since the loading matrix is randomly generated, the index of the affected variables has no consequence. Faults are generated for k between 1 and 10. For each fault, the true positive rate (TPR) was computed as the number of faulty observations correctly declared as an alarm by either of the monitoring statistics, divided by the total number of faulty observations.

Two monitoring schemes are considered in this study. The first monitoring approach corresponds to the standard implementation of MSPC-PCA and is composed by the standard T^2 & Q statistics. In turn, the second monitoring scheme uses the T^2 to monitor the score subspace and the proposed truncated Q statistic to monitor the residuals subspace. Both monitoring schemes were set to an overall false detection rate of 1 %.

The average TPR along with their interquartile range are represented in Figure 1. These results clearly show that the truncated Q statistic leads to significantly higher TPR, regardless of the number of variables involved. In all cases, the proposed monitoring scheme can detect faults with a magnitude higher than 2 standard deviations. In contrast, the standard monitoring scheme only detects faults with 4 standards deviations less than 40 % of the times. In other words, faults with a magnitude twice as large are detected less than half of the times. It is also observed that the sensitivity of the standard monitoring scheme decreases with the increase of the number of monitored variables.

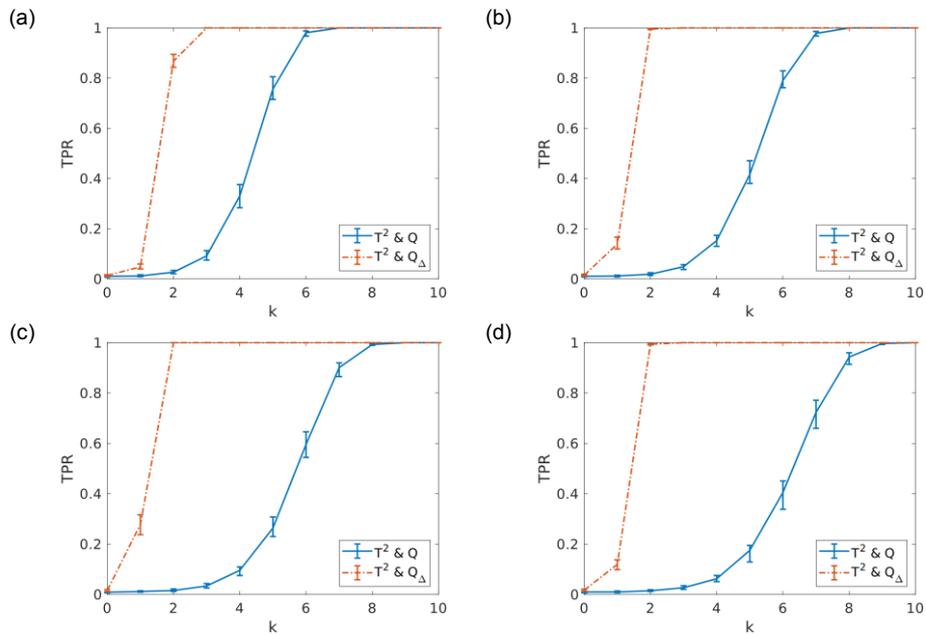


Figure 1 True positive rate (TPR) for processes with: (a) 5 000 variables; (b) 10 000 variables; (c) 15 000 variables; (d) 20 000 variables.

To exemplify the masking effect caused by summing several (irrelevant) residuals, the residuals obtained for a process with 20 000 variables with a fault of $k = 1$ standard deviation in one variable is represented in Figure 2. From visual inspection of Figure 2 (a) it is verified that the residual of the faulty variable is completely undistinguishable since several unaffected variables have either similar or even higher residuals. By summing the squared residuals, it is further verified that the contribution of this variable to the Q statistic is less than 0.1 %.

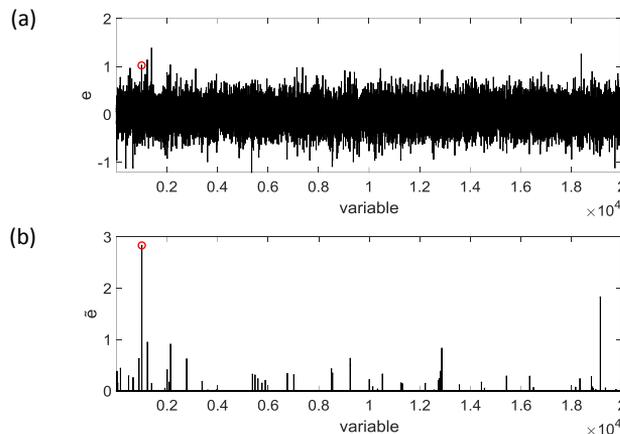


Figure 2 Residuals for a deviation of $k = 1$ standard deviation in the 1 000th variable (marked with a circle) for a process with 20 000 variables: (a) original residuals used by the Q statistic; (b) truncated residuals used by the truncated Q statistic.

On the other hand, with the proposed truncated Q statistic, the truncated residuals presented in Figure 2 (b) are obtained. In this case, the PCA residuals are firstly auto-scaled by their in-control sample mean and sample standard deviation. Afterwards, the allowance threshold is subtracted from the scaled residuals. The truncated residuals lower than zero (*i.e.*, bellow the allowance threshold) are set to zero. Following these operations, the irrelevant residuals are removed from analysis, while the fault's effect is highlighted. In this case, the contribution of the faulty variable to the truncated Q statistic is 45.6 %. Therefore, the fault has a significantly higher contribution to the monitoring statistic, which in turn improves the detection capability.

3.2. Real Case Study: Monitoring of Printed Circuit Boards

To further validate the proposed methodology, it was applied to an industrial Surface Mount Technology (STM) process of Bosch Car Multimedia Portugal. The process concerns the printing of solder paste deposits in copper pads of printed circuit boards (PCB). In this process, each PCBs has 3 507 copper pads and measurements for the (i) area, (ii) height, (iii) volume, (iv) offset in the x-coordinate and (iv) and offset in the y-coordinate are taken for each solder paste deposit. Therefore, each PCB comprises 17 535 variables.

The dataset used in this case study is composed by 3 747 PCBs. Most of these PCBs are within specifications, but a few of them have significant deviations. A training dataset was built by randomly selected 2 000 normal PCBs from the raw dataset. This dataset was split into a calibration dataset and a validation dataset with 1 000 PCBs each. The calibration dataset was used to train the PCA model, while the validation dataset was used to set the control limits of the monitoring statistics. The control limits were determined by taking the $(1 - \alpha/2) \times 100$ % upper percentile of each monitoring statistic computed on the validation dataset. The significance level α was set to 0.01. The remaining 1 747 PCBs were used as a test dataset.

The test dataset was monitored by the standard of T^2 & Q of MSPC-PCA, and the proposed T^2 & Q_{Δ} . The obtained Q and truncated Q statistics are presented in Figure 3. From these results it is verified that the standard Q statistic only detects 5 PCBs as faulty. On the other hand, the proposed truncated Q statistic signals 17 PCBs as abnormal. These detections include 3 PCBs there were already detected by the Q statistic and 14 additional PCBs that were not detected before. In turn, 2 of the PCBs detected by the Q statistic are considered normal by the truncated Q statistic. These PCBs have relatively low deviations in a few pads, and even the standard Q statistic is only slightly above the control limits (see PCBs #518 and #1046 in Figure 3 (a)). Therefore, it is concluded the truncated Q statistic led to a significant improvement in the detection of faulty PCBs.

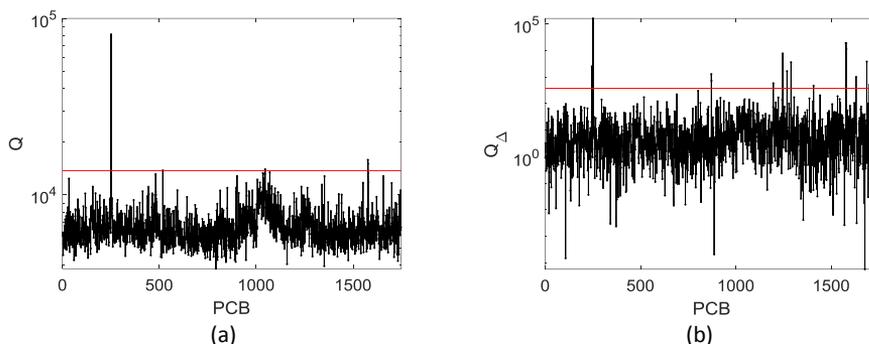


Figure 3 Monitoring statistics for the real case study: (a) Q statistic; (b) truncated Q statistic.

4. Conclusions

The standard Q statistic of MSPC-PCA is prone to sensitivity deterioration as the number of monitored variables increases. This is a consequence of summing the contributions of irrelevant residuals, raising the basal level of the Q statistic and masking the deviations due to faulty variables. To address this limitation, it is proposed to screen for the relevant residuals in each observation and then build a truncated Q statistic using only the relevant residuals. This methodology was tested on a Monte Carlo simulation study and a real case study from an industrial Surface Mount Technology (STM) process. The results show that MSPC-PCA with the truncated Q statistic is much more sensitive than the original formulation, leading to the detection of 14 additional PCBs that were not signaled by the standard Q statistic in the real case study.

The proposed methodology is expected to also improve the detection performance of other monitoring methodologies that monitor the residuals of high-dimensional processes.

Acknowledgements

The authors acknowledge support from the Chemical Process Engineering and Forest Products Research Centre (CIEPQPF), which is financed by national funds from FCT/MCTES (reference UID/EQU/00102/2019).

References

- A. Cinar, A. Palazoglu, F. Kayihan, 2007, *Chemical Process Performance Evaluation*, Boca Raton, CRC Press.
- Z. Ge, Z. Song, S. X. Ding, B. Huang, 2017, *Data Mining and Analytics in the Process Industry: The Role of Machine Learning*, IEEE Access, 5, 20590-20616.
- J. E. Jackson, 1959, *Quality Control Methods for Several Related Variables*, Technometrics, 1, 4, 359-377.
- J. E. Jackson, G. S. Mudholkar, 1979, *Control Procedures for Residuals Associated With Principal Component Analysis*, Technometrics, 21, 3, 341-349.
- B. Jiang, X. Zhu, D. Huang, R. D. Braatz, 2015, *Canonical Variate Analysis-based Monitoring of Process Correlation Structure using Causal Feature Representation*, Journal of Process Control, 32, 109-116.
- H. Lee, C. Kim, S. Lim, J. M. Lee, 2020, *Data-driven fault diagnosis for chemical processes using transfer entropy and graphical lasso*, Computers & Chemical Engineering, 142, 107064.
- N. Li, W. Yan, Y. Yang, 2015, *Spatial-Statistical Local Approach for Improved Manifold-Based Process Monitoring*, Industrial & Engineering Chemistry Research, 54, 8509-8519.
- L. Luo, S. Bao, J. Mao, D. Tang, 2016, *Nonlinear process monitoring based on kernel global-local preserving projections*, Journal of Process Control, 38, 11-21.
- D. C. Montgomery, 2008, *Introduction to Statistical Quality Control*, John Wiley & Sons.
- K. E. S. Pilario, Y. Cao, 2018, *Canonical Variate Dissimilarity Analysis for Process Incipient Fault Detection*, IEEE Transactions on Industrial Informatics, 14, 12, 5308-5315.
- T. J. Rato, M. S. Reis, 2017, *Markovian and Non-Markovian Sensitivity Enhancing Transformations for Process Monitoring*, Chemical Engineering Science, 163, 223-233.
- S. Yan, X. Yan, 2019, *Using labeled autoencoder to Supervise Neural Network Combined with k-nearest Neighbor for Visual Industrial Process Monitoring* Industrial & Engineering Chemistry Research, 58, 9952-9958.
- J. Yin, X. Yan, 2019, *Mutual Information - Dynamic Stacked Sparse Autoencoders for Fault Detection*, Industrial & Engineering Chemistry Research, 58, 21614-21624.

Development of a predictive emissions monitoring system using hybrid models with industrial data

Ataide S. Andrade Neto^a, Argimiro R. Secchi^a, Bruno D. O. Capron^b, Antônio Rocha^b, Luciana N. Loureiro^c and Patrícia R. Ventura^c

^a*Chemical Engineering Program, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

^b*Chemical Engineering Department, School of Chemistry, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

^c*CENPES, PETROBRAS, Rio de Janeiro, Brazil*
ataide@peq.coppe.ufrj.br

Abstract

In this paper we present a study for the development of a predictive emission monitoring system (PEMS) through hybrid modeling. The system was built by coupling a thermodynamic equilibrium model based on the minimization of the Gibbs energy and machine learning (ML) models, such as artificial neural networks (ANN) with different architectures, supporting vector machines (SVM), decision-tree ensembles using histogram-based gradient-boosting (HGB) algorithms, among others. The ML models were trained with real operation data of a gas turbine from one of the three largest thermoelectric power plants in Brazil. The main goal of this study was to generate accurate NO_x prediction models in order to replace the current continuous emission monitoring system (CEMS) in operation. In the hybrid framework, the data-driven models were used to estimate the error of the theoretical model, then, the prediction of pollutant levels was made by the sum of both results. An analysis was made with 6,513 samples split between training and validation sets. Three key performance indicators: (i) maximum absolute error (MAE), (ii) average relative error (ARE) and (iii) the correlation coefficient were evaluated and the results have shown that the hybrid approach can be overall more accurate than the single models alone, at least for the gas turbine in study.

Keywords: PEMS, Machine learning, Hybrid models, Gas turbine

1. Introduction

The climate change debate has promoted a growing collective awareness about the impacts of emissions on the quality of life in general population. Because of that, many governments have created legislation mechanisms in order to monitor, reduce or even prevent the accumulation of environmental hazardous gases in the atmosphere. In this context, stationary emission sources such as thermoelectric power plants are targeted and rigorous monitoring methodologies based on expensive gas analyzers, known as continuous emissions monitoring system (CEMS), have been implanted on site. In order to mitigate those economic impacts, nowadays many countries have regulations for the implementation of virtual monitoring approaches such as the predictive emissions monitoring system (PEMS), a model-based software that can estimate emission levels using common measured process variables, as a less expensive alternative to the CEMS technology (Hadjiski et al., 2005).

Thermoelectric power plants, which, in most cases use natural gas as fuel in gas turbines, still represent around 8.3% of the overall Brazilian power generation. The main concern of local environmental agencies is the release of greenhouse gases to the atmosphere, such as the CO_2 , and other pollutants as the NO_x . Because of that, the national air quality control program (PRONAR) establishes the emission tolerance levels and requires their constant monitoring.

Many successful case studies of PEMS application, mostly using data-driven models, have already been presented in literature, a full-detailed review can be found in Si et al. (2019). However, the major challenge that still persists is the (high) dependence on industrial data for the model development. In many plants, for example, it is impossible to collect an ideal and rich data-set due to safety and operational constraints or even a flawed controlling system.

Despite being a low-complexity equipment, gas turbines have several design specifications that vary according to each manufacturer. Because of that, most of the commercially available PEMS use data-driven models in their conception (Allaire et al., 2007). It is undeniable that data-driven models such as neural networks are, by far, easier to obtain and to use in real-time industrial processes than first-principles ones. However, the quality of the prediction system becomes strongly dependent on the quality of the data used in the modeling. This could be a problem, especially in poorly instrumented factories where data reliability is a real issue.

We present, in this work, a hybrid approach for the NO_x emission modeling in a gas turbine fed with natural gas from a large thermoelectric industrial unit. This approach is based on the coupling of: (i) a thermodynamic equilibrium model obtained through the minimization of Gibbs free energy of the combustion reaction system; and (ii) machine learning models such as neural networks of several architectures, supporting vector machines, and decision-tree ensembles used to predict the phenomenological modeling error. Those ML models were built with a limited set of real operation data, as described in the following section. A more detailed review on hybrid modeling can be found in Sansana et al. (2021).

2. Methodology

2.1. The case study

The present case study focuses on a single gas turbine from the *Termorio* power plant, which has a total production capacity of 1,058 MW. The gas turbine in analysis operates around 100 MW through a combined cycle (see Figure 1), that is, the residual heat from the natural gas combustion is recovered in a boiler to produce steam. Then, the steam is sent to a steam turbine generating additional power. In this particular case, there is no supplementary combustion in the boiler.

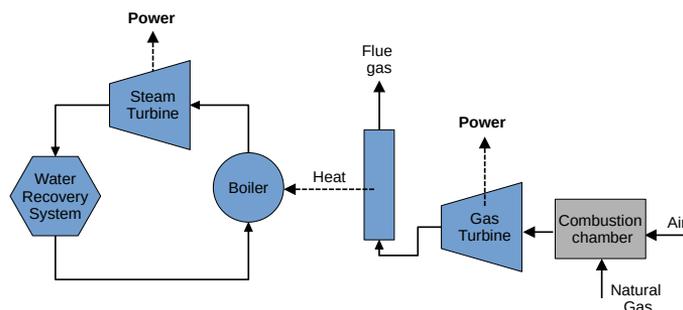


Figure 1: Combined cycle for power generation.

2.2. Thermodynamic model

The Gibbs free energy minimization is a classic approach for equilibrium calculations in chemical reactions. The problem can be formulated as

$$\min_{\mathbf{n} \in \mathcal{N}} G(T, P, \mathbf{n}) \quad (1)$$

subject to the following equality constraints, corresponding to the atomic balances:

$$\sum_j a_{i,j}(n_j - n_{f,j}) = 0 \quad (2)$$

where G is the Gibbs free energy, T and P are the temperature and pressure in the combustion chamber, \mathbf{n} is the equilibrium composition (in moles), $n_{f,j}$ is the feeding composition, and $a_{i,j}$ are the coefficients of the atomic balances matrix; i represents an atom in a species j . The optimization problem in Equations (1) and (2) can be solved with the Lagrange multipliers method, resulting in the non-linear algebraic problem (Smith et al., 2001):

$$\begin{bmatrix} \Delta G_{f,j}^o + RT \ln(\phi_j) + \sum_i \lambda_i a_{i,j} \\ \sum_j a_{i,j}(n_j - n_{f,j}) \end{bmatrix} = 0 \quad (3)$$

where $\Delta G_{f,j}^o$ is the Gibbs free energy of formation of the species j calculated from the database available in McBride et al. (2002), ϕ_j is the fugacity coefficient calculated by the SRK equation of state, and λ_i are the Lagrange multipliers. The system in Equation (3) was implemented in the process simulator EMSO (Soares and Secchi, 2003).

2.3. Machine learning models

Several machine learning models were used within the hybrid modeling framework. They were built considering the same ten input variables (features): (i-ii) temperature and pressure; (iii-vii) the fuel composition (methane, ethane, propane, nitrogen and CO_2); (viii) fuel flow rate; (ix) overall flue gas flow rate; and (x) output turbine power. The predicted output variable (label) is the NO_x concentration in the flue gas. It is worth pointing out that each architecture was trained 10 times, and only the best results of each kind were used for results comparison.

A total of 6,513 sampling points split between training and validation sets were used. The data were obtained directly from the industrial unit data historian within the period from Jan/2015 to Dec/2015. A minor data cleaning was performed to remove inconsistent samples. The training set was created with the first 70% of the samples (January to mid-September), and the validation set with remaining 30% (mid-September to December).

Among the different ML models, we highlight the structured neural network (SNN) in Figure 2. The SNN has two hidden layers, each one with 10 and 9 nodes, respectively. Because the turbine output power is strongly correlated with the NO_x concentration, we chose to bypass the hidden layers and connect it straight to the output layer, thus, giving more weight and relevance to the neural network parameters associated with this input. The SNN was built in Python with the PyTorch library.

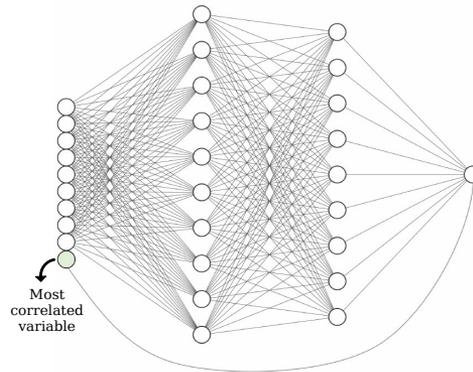


Figure 2: Structured neural network used within the hybrid modeling framework.

2.4. Hybrid model

The hybrid model architecture is presented in Figure 3. In this framework, the ML models are used to make predictions on the phenomenological modeling error, instead of the NO_x concentration itself. Then, the output of both models are combined in order to make the actual prediction:

$$y_{\text{predicted}} = y_{\text{model}} + \sigma_{ML} \quad (4)$$

The model was built in two steps: (i) the thermodynamic simulations to obtain the baseline prediction (using only the information on the input variables); and (ii) the ML training with the same input variables, and the modeling error as the output variable.

2.5. Accuracy evaluation

Three main performance indicators that are required by the US legislation were used to evaluate overall accuracy of the models, which are: the maximum absolute error (MAE)

$$MAE = \max_{i \in \{1, N\}} (|y_{p,i} - y_{m,i}|) \quad (5)$$

The average relative error (ARE)

$$ARE = \frac{1}{N} \sum_{i=1}^N \frac{|y_{p,i} - y_{m,i}|}{y_{m,i}} \quad (6)$$

and the correlation coefficient (CC)

$$CC = \frac{\sum_{i=1}^N (y_{p,i} - \bar{y}_p)(y_{m,i} - \bar{y}_m)}{\sqrt{\sum_{i=1}^N (y_{p,i} - \bar{y}_p)^2} \sqrt{\sum_{i=1}^N (y_{m,i} - \bar{y}_m)^2}} \quad (7)$$

where y_p and y_m are the predicted and the measured values, respectively, and N is the number of samples.

3. Results and Discussion

The time window with the training and test data sets used for the ML models is presented in Figure 4. The data was split congruently because our main concern was the extrapolation potential of such models, and not the interpolation one. The prediction results for all of the models using the test set are presented in Figures 5 and 6 for: (a) the Gibbs model; (b) a neural network like the SNN, but without the bypass; (c) the SNN; and (d) the hybrid model coupling Gibbs and SNN models. The performance indicators are presented in Table 1. It can be seen, in the time series of Figure 5, that the hybrid model (5d) can capture better the data tendency in comparison with the other models (5a-5c). The accuracy can be more easily perceived in the histogram of Figure 6, as one can observe that the distribution of the absolute errors for the hybrid model is more concentrated in the range of 0 to 1 ppm than the other models.

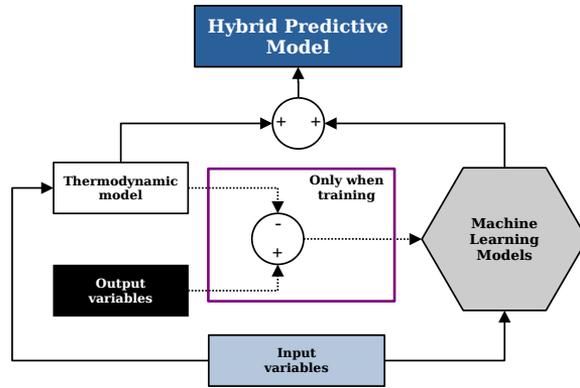


Figure 3: Architecture of the hybrid model predictor.



Figure 4: Training and test data sets.

Table 1: Performance indicators for some of the built models¹.

Model	MAE	ARE (%)	CC (%)
Gibbs	5.26	6.30	21
SVM	5.77	5.94	19
HGB	4.20	6.19	64
NN	4.08	4.20	27
SNN	3.87	4.01	54
Hybrid	3.62	3.91	69

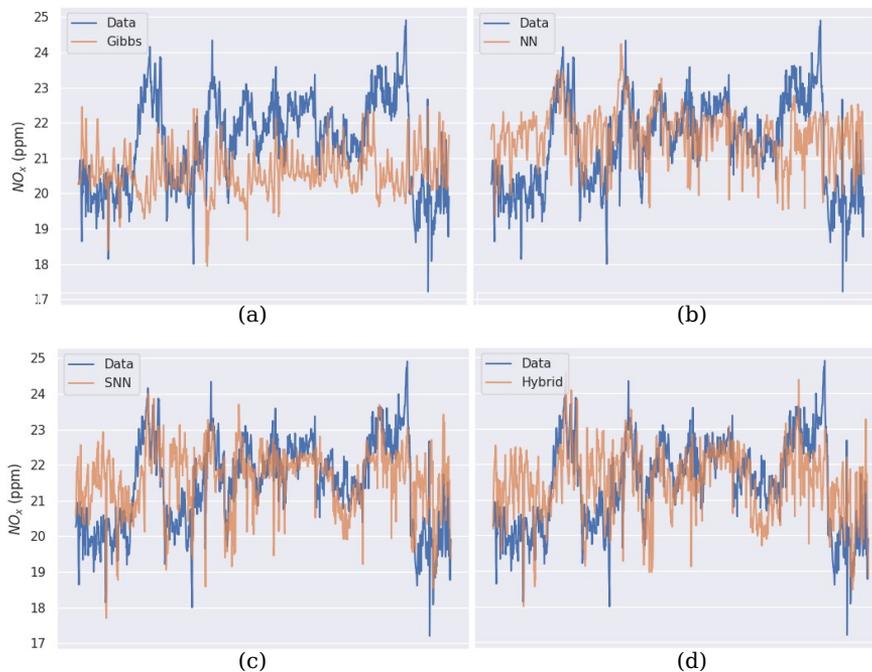


Figure 5: NO_x predictions: Time series for (a) Gibbs model, (b) Neural network model, (c) Structured neural network model, and (d) Hybrid model (Gibbs + SNN).

Those results show that the Gibbs model, despite being oversimplified, can describe reasonably well the levels of NO_x concentration, even though it has no estimated parameters, with an average error of 6.3%. However, it cannot capture the modulation of the output variable (it has a correlation of only 21%). This is probably due to the existence of three-dimensional temperature gradients inside the combustion chamber, forming different equilibrium regions, while only one temperature is specified on the model (the turbine output temperature).

Analyzing the two neural networks, we can see that the bypass used in the SNN improved its prediction capacity (its error is 5 to 6% smaller). The main reason for SNN's increased performance is that the parameters associated with the most correlated variable has more meaning, as the physical relation between the input and output is more preserved, improving the network awareness, as

¹Due to the lack of space, additional results for the SVM and HGB models will not be discussed in this paper.

shown by the correlation coefficient that is twice the NN's value.

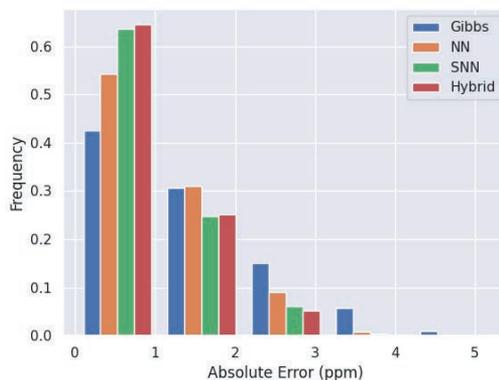


Figure 6: Histogram of the absolute errors for the analyzed models.

Then, we have the hybrid model using the SNN, which was chosen for being the best tested ML model. It can be seen that it is more accurate than the models alone, reducing the SNN's MAE by 6% and Gibbs's by 32%. However, the greatest improvement is in the correlation coefficient, that is three times higher than Gibbs's and 27% better than the SNN's. Although those correlation values can still be considered low, taking into account the high quantity of noise in data and the low variability in the turbine's range of operation, the results obtained so far are very promising, as the errors are also low, and visually the model can represent data tendency.

4. Conclusion

A hybrid approach for the modeling of NO_x emissions on gas turbines was presented. The hybrid predictor built with neural networks and thermodynamic equilibrium equations proved to improve the overall accuracy of each model considered alone, having errors up to 37% smaller and correlation coefficient up to three times higher. The quality of the industrial data is a key factor to improve the prediction potential, as the low variability and high quantity of noise reduce the quality of the data-driven modeling.

References

- D. L. Allaire, I. A. Waitz, K. E. Willcox, 2007. A comparison of two methods for predicting emissions from aircraft gas turbine combustors. In: Proceedings of ASME Turbo Expo 2007: Power for Land, Sea, and Air. Vol. 2. Montreal, Canada, pp. 899–908.
- M. Hadjiski, K. Boshnakov, N. Christova, 2005. Simulation-based predictive emission monitoring system. In: European Council for Modelling and Simulation (Ed.), Proceedings 19th European Conference on Modeling and Simulation. Riga, Latvia.
- B. J. McBride, M. J. Zehe, S. Gordon, 2002. NASA Glenn coefficients for calculating thermodynamic properties of individual species. Tech. rep., NASA Glenn Research Center.
- J. Sansana, M. N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L. H. Chiang, M. S. Reis, 2021. Recent trends on hybrid modeling for industry 4.0. Computers & Chemical Engineering Vol. 151, pp. 107365.
- M. Si, T. J. Tarnoczi, B. M. Wiens, K. Du, 2019. Development of predictive emissions monitoring system using open source machine learning library – keras: A case study on a cogeneration unit. IEEE Access, Vol. 7, pp. 113463–113475.
- J. M. Smith, H. C. V. Ness, M. M. Abbott, 2001. Introduction to Chemical Engineering Thermodynamics. Irwin/McGraw-Hill Series in Marketing. McGraw-Hill.
- R. P. Soares, A. R. Secchi, 2003. EMSO: A new environment for modelling, simulation and optimisation. Computer Aided Chemical Engineering, Vol. 14, pp. 947–952.

On The Interpretability of Graph Neural Networks in QSPR Modeling

Fan Fan, Adem R.N Aouichaoui, Gürkan Sin

Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kgs. Lyngby, DK-2800, Denmark
gsi@kt.dtu.dk

Abstract

Artificial intelligence-based models (AI) and in particular Graph Neural Networks (GNN) are considered a more promising approach for modeling molecular properties compared to the use of traditional descriptor-based models due to their enhanced ability to express the structural information and their ability to better generalize to unseen data. However, their 'black box' nature and the lack of transparency and interpretability could hinder their wider acceptance and usage. In this work, we combine knowledge-based molecular descriptors and two AI-based concepts, the junction tree model and the attention mechanism to produce an interpretable model. The model is trained on the enthalpy of formation of organic compounds and the insights gained are highlighted and compared to the insights gained from models with a higher level of interpretability such as the group-contribution models. The results obtained show consistency with the insight gained in the form of the relative importance of the molecular sub-structures to the overall property.

Keywords: Molecular Property Prediction, Quantitative Structure-Property Relations (QSPR), Machine-Learning (ML), Graph Neural Networks (GNN), Interpretability

1. Introduction

Integrating AI-based approaches into molecular property prediction models such as deep neural networks (DNN) has recently gained popularity. These models relate the structural information to desired property numerically are called quantitative structure-property relations (QSPR). An example of such a model is the graph neural network (GNN), which represents atoms and bonds by a set of nodes connected through edges respectively in the graph. However, despite these models achieving state-of-the-art for many properties, a major drawback is their black-box nature and the absence of interpretability. In this context, interpretability can be defined as insight into the relative importance of the molecular substructures towards the target property.

The aspect of interpretability should be considered significant to the scientific understanding of the parts of the molecule (substructure or functional group) that influence a specific property and evaluate whether the insights provided from the model coincide with those obtained from knowledge. (Coley *et al.*, 2017) compared molecular representation obtained from DNN with another known descriptor extended circular fingerprints (ECFP4), claiming there is a mapping relation between atoms and indices in representation. Accordingly, the most critical fingerprint indices can be found by estimating these mapping coefficients. But in classic GNN, the architecture would aggregate all nodes, eliminating the contributions from individual atoms.

The attention mechanism has since become a significant concept in neural networks, where importance weights are learnable through the backpropagation, enabling the model to focus on a specific part of the knowledge gained. (Veličković *et al.*, 2018) successfully introduced a self-attention mechanism into a classical GNN, updating the target node by summing up neighborhood nodes with attention weights. The attention mechanism has also been incorporated on a graph level (molecular level) to produce a learnable readout function by considering the molecule as a virtual node into which all node information is aggregated, e.g. AttentiveFP developed by (Xiong *et al.*, 2020).

In this work, we combine the original AttentiveFP with group-fragments from the well-known MG group-contribution models (by (Hukkerikar *et al.*, 2012)) to build a new architecture with an increased prediction accuracy and with the ability to show the importance of each fragment simultaneously. The insights gained are then compared to those gained by group-contribution models where the magnitude of the group contribution reflects the relative importance of the groups towards the overall property.

2. Method

2.1. AttentiveFP

The *Attentive Fingerprint* (AFP) by (Xiong *et al.*, 2020) is employed to illustrate the importance of single atoms to the target property in molecule-scale. The application of the attention mechanism makes this model capable of calculating the weights of information and assigning importance to each node (atom) w.r.t. the prediction. The aggregation or pooling employed differs from those usually used in other GNNs (based on summation or min-max pooling), the predicted context C^l at layer l in AttentiveFP is derived from a weighted summation readout operation as seen in E.q.1.

$$C^l = \sum \alpha_i^l \cdot h_i \quad (1)$$

In E.q.1 α_i^l refers to the coefficient weight of atom feature h_i at layer, l contributing to the property prediction. The implementation of the *softmax* function as an activation function to compute attention weights, normalizes the values into a range from 0 to 1, representing the degree of importance of every single atom.

2.2. Group contribution models and molecular Fragments

Group-contribution models use an occurrence vector/matrix to represent the molecule where each element indicates the number of times a specific fragment/substructure is present in the molecule. The fragments or groups considered in this work consists of the first-order groups developed by (Hukkerikar *et al.*, 2012). These fragments present no to little overlap and are mostly defined based on chemistry and thermodynamic knowledge rather than heuristics. Property models based on the group contribution model have a generic form as shown in Eq.2, where $c_y^{(x)}$ and $n_y^{(x)}$ are the contributions and occurrence of group y of order x respectively.

$$\hat{y} = \sum_i^{220} n_i^{(1)} c_i^{(1)} + \sum_j^{130} n_j^{(2)} c_j^{(2)} + \sum_k^{74} n_k^{(3)} c_k^{(3)} \quad (2)$$

As mentioned earlier, the first-order groups defined are the most knowledge-rooted groups out of all three scales (orders). If one specific motif (or group) has higher contributions to the target property, it is fair to assume that this group is overall more important compared to other present groups. The values of contributions are accessible in

previous work presented by (Hukkerikar *et al.*, 2012). Since sequential regression of the group orders is used to obtain the group contribution, the property is described well enough using the first-order groups only and the higher orders can be considered as a correction to the approximations gained by the first-order term.

2.3. Junction Tree

(Vats and Nowak, 2014) proposed the junction tree algorithm (also called 'Clique Tree') consisting of a method for extracting modified graphs (or sub-graphs) from an original graph by branching the supernodes into different smaller nodes, similar to a tree. (Jin *et al.*, 2021) used this concept to build a tree-structured architecture to aggregate features of some particular nodes into bigger super-nodes. In this work, the fragment-based networks use this tree structure to build a new graph for further training, as shown in Figure 1, the tetranitromethane molecule can be split into four supernodes (S_1, S_2, S_3, S_4) consisting of their corresponding structures in the original graph.

2.4. Architecture

The architecture of the new model is depicted in Figure 1. The original molecular graph is broken down into fragments and then fed in as features into the prediction model. Here some prior knowledge for division is introduced to guide the fragmentation process. Therefore, a junction tree (tree-structured scaffold over fragments) can be generated automatically, where the features in supernodes and original nodes are regarded as the root and branches respectively during back-propagation. The prepared super nodes features from the first block are used as the initial states for the node features in the last block. Note that the features of the edge connecting each of the fragments in the original graph are kept as attributes of the new edge between supernodes in situ. In the last step, the new graphs are forwarded to an *AFP* block to compute the importance of each fragment to the predicted property. Due to the employment of prior knowledge on group contribution, the model is named *Attentive Group Contribution (AGC) Model*. Similar to the attentions in *AFP*, the atom feature h_i in Eq.1 can be replaced by the fragment feature S_i obtained from the junction tree, thus the new attention weights α_i^1 can then represent the degree of importance of fragments instead.

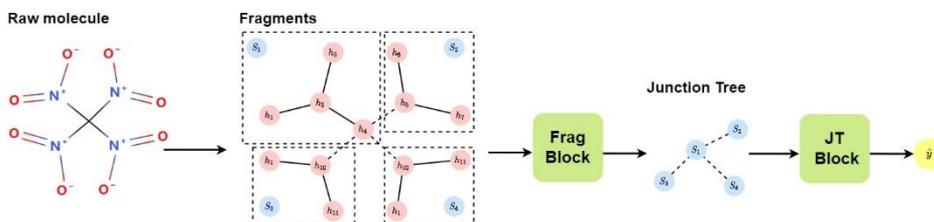


Figure 1: Illustration of Attentive Group Contribution (AGC) workflow for tetranitromethane

2.5. Data

The models were trained on the enthalpy of formation of 741 organic compounds obtained from the AICHE DIPPR database (Wilding *et al.*, 2017). The property was chosen as it exhibits a linear trend for the homologous series and thus the relative importance of the group fragments obtained here can be readily compared with the relative contribution of these fragments in the group-contribution model developed in (Hukkerikar *et al.*, 2012).

2.6. Visualization

In addition to visualizing the relative comparison between different atoms/fragments, the relative attention $\hat{\alpha}$ is introduced instead of weights obtained directly to overcome the

potential issue of bad distinction when all attention weights are lower than 0.5 and to account for the relative size of the molecule.

$$\hat{\alpha}_i = \frac{\alpha_i - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}} \quad (3)$$

α_{\min} and α_{\max} are the maximum and minimum of attention weights computed in a single molecule. A blue-red sequential color map is used to visualize the relative importance of atoms/ fragments to molecules, the motifs marked with red refers to high significance ($1 > \hat{\alpha}_i > 0.5$) while blue tells the specific motifs are not critical ($0 < \hat{\alpha}_i < 0.5$). Darker colors represent the importance closer to 0 (least important) or 1 /most important).

3. Results and Discussion

The models obtained all provide good predictions on the data. The R2 and mean absolute error (MAE) of *AGC* are 0.99-0.99-0.99 and 18.4-27.4-22(KJ · mol⁻¹) respectively across training, validation, and testing (8:1:1 ratio). Meanwhile, the *AFP* achieves 0.99-0.99-0.99 and 21.2-23.2-28.7(KJ · mol⁻¹) respectively for the three data folds.

To provide insights into the interpretability of the models, five case studies were selected to inspect and compare the computed importance of fragments obtained from the AttentiveFP model and the *AGC* model. These case studies include a diverse set of molecular families such as carboxylic acids, halohydrocarbons, benzylamines, acetylenes, and nitrate-esters. The visualization of atoms' importance (relative attention weights $\hat{\alpha}$ calculated by AttentiveFP is demonstrated in Figure 2.I, while the relative attentions of fragments with *AGC* are shown in Figure 2.II.

The importance weights considered are those from the final layer (in terms of the AFP model terminology, the last time-step), the corresponding visualization can be seen in Figure 2.I. The area surrounded by a dashed ellipse is the most significant group based on the scheme defined by (Hukkerikar *et al.*, 2012) and will be considered in this work as the benchmark with respect to the importance of the fragment towards the property value. As seen in Figure 2.I, the *AFP* model fails to reproduce the fragment importance of the GC models for all cases. Furthermore, the atom importance of the same functional groups is inconsistent and changes with increasing carbon chain length for carboxylic acids. For instance, C=O in the carboxyl group shown in butyric acid is more important than the same substructures in acetic acid and propanoic acid, despite only adding CH₂ to the molecules, unlike the true insights of the property. Similar issues also appear when it comes to the rest of the straight-chain aliphatic hydrocarbons, where their corresponding importance/coefficients are consistent neither in the same molecule nor in different compounds among the same family. In both cases of acetylene and nitrate-ester, this model succeed to detect the different substructures like triple bonds and ONO₂ respectively, but the prediction of importance does not adhere to insight obtained from **GC** models. When predicting hexafluoroethane and chloropentafluoroethane, the relative importance of carbon to fluorine are in contradiction.

The *AGC* model developed seems to provide insights that are overall in line with those obtained from the GC models. Considering the first set of molecules are carboxylic acids, as shown in Figure 2.II, the red fragments refer to the acid group *COOH*. In all carboxylic groups tested, the importance of this fragment always shows the highest significance. The raw attention weights reported in Table 1 also confirm this. Note that the parenthesis

marked by † denotes there are multiple of the same groups in the molecule. It is a remarkable fact that the attention weights are not the proportion of groups in a molecule, they instead represent how important these fragments are to the overall structure. Therefore, it is not straightforward to compare them parallel to others in different molecules, especially molecules of different sizes. However, the results obtained here are enough to extract the important part considering every single molecule individually. In the GC method, the absolute contribution of $COOH$ is higher than the other two groups CH_2 and CH_3 , which is consistent with the result we obtained here.

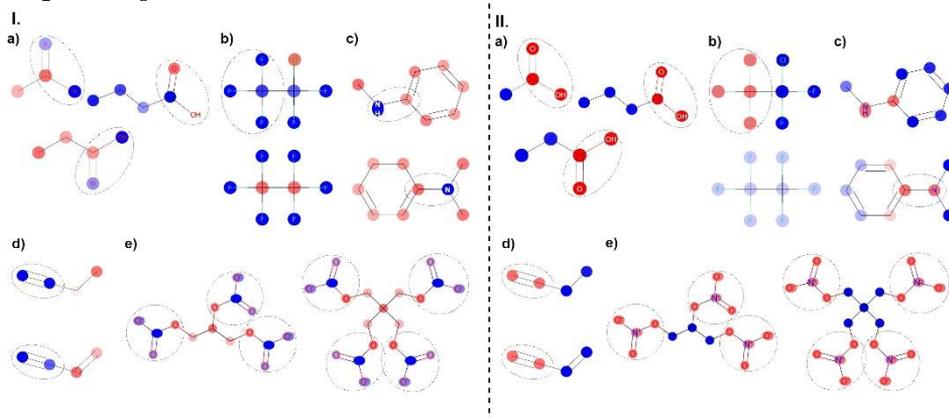


Figure 2: Visualization of atom significance by AFP (I) and AGC (II) for Carboxyl acids (a), Halohydrocarbons (b), Benzylamine (c), Acetylene (d), and Nitrate-esters (e)

Table 1: Raw attention weights of each fragment to final prediction on carboxylic acids

First-Order Groups	Acetic acid	Propionic acid	n-Butyric acid
COOH	0.6275	0.5365	0.4830
CH ₂	-	0.2397	(0.1702, 01782) [†]
CH ₃	0.3725	0.2238	0.1685

The second set of molecules shows the different fragments of CF_3 and $CClF_2$, where the only difference is the halo-atoms connected to carbon. Unlike the prediction done by the *AFP* model and seen in Figure 2.II, the architecture successfully detects the equal importance of fragments and highlight the more significant contribution from the group CF_3 than $CClF_2$. The raw attention weights of these two are reported in Table 2.

Table 2: Raw attention weights of each fragment to final prediction on halohydrocarbon

First-Order Groups	Chloropentafluoroethane	Hexafluoroethane
CClF ₂	0.3412	-
CF ₃	0.6588	(0.5, 05) [†]

The third and fourth sets are the representations of aromatic amines and carbon-carbon triple bonds respectively. In both cases, the new model manages to most distinguish fragments circling by dashed cycle. Meanwhile, it fixes the predictions of the same atoms at the same location in different molecules. In other words, the common part of various derivatives is expected to be extracted by similar contributions. The last set depicts the ONO_2 groups in nitrate-esters not only have equal contributions in the same molecular but also highlight this fragment as the most significant group.

The hybrid combination with group contribution theory adds a layer of interpretability compared to a purely data-driven approach, establishing the visualization of prediction on basis of chemists' knowledge on property prediction. However, due to the limitation of techniques, such interpretation can only be presented in the form of case studies and not in a general fashion. It is expected to obtain the influence of every single part on the overall prediction, and then it has a chance to reveal the structures of neural networks by adding or masking particular segments.

4. Conclusions

The combination of attention mechanism and prior knowledge adds a level of interpretability to AI-based models compared to a purely data-driven approach, establishing the visualization of prediction on basis of chemists' knowledge on property prediction. In this work, several cases are considered to test the performance and interpretability. By doing this, it is expected to get the effect of each part to the final prediction. Unlike the weights computed by the original *AFP*, *AGC* is able to present results better in accordance with a chemists' intuition. Moreover, the insights provided through the relative importance of the various molecular fragments are consistent with insights gained through the widely-accepted group-contribution models. This work can be beneficial to improve the transparency of models and overcome the possible objection from communities towards these data-driven methods. However, the extent of interpretability must be further tested on other linear and non-linear properties to assess the true extent of interpretability provided by the model developed herein. There will be one day a universal framework instead of isolated case studies can be developed to interpret the inner side of neural networks in the field of chemistry.

References

- Coley, C.W., Barzilay, R., Green, W.H., Jaakkola, T.S. and Jensen, K.F. (2017) 'Convolutional Embedding of Attributed Molecular Graphs for Physical Property Prediction', *Journal of Chemical Information and Modeling*, 57(8), pp. 1757–1772.
- Hukkerikar, A.S., Sarup, B., Ten Kate, A., Abildskov, J., Sin, G. and Gani, R. (2012) 'Group-contribution + (GC +) based estimation of properties of pure components: Improved property estimation and uncertainty analysis', *Fluid Phase Equilibria*, 321, pp. 25–43.
- Jin, W., Barzilay, R. and Jaakkola, T. (2021) 'Chapter 11: Junction Tree Variational Autoencoder for Molecular Graph Generation', *RSC Drug Discovery Series*, 2021-Janua(75), pp. 228–249.
- Vats, D. and Nowak, R.D. (2014) 'A junction tree framework for undirected graphical model selection', *Journal of Machine Learning Research*, 15, pp. 147–191.
- Veličković, P., Casanova, A., Liò, P., Cucurull, G., Romero, A. and Bengio, Y. (2018) 'Graph attention networks', in *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*.
- Wilding, W.V., Knotts, T.A., Giles, N.F. and Rowley, R.L. (2017) 'DIPPR® Data Compilation of Pure Chemical Properties', *Design Institute for Physical Properties, AIChE*.
- Xiong, Z., Wang, D., Liu, X., Zhong, F., Wan, X., Li, X., Li, Z., Luo, X., et al. (2020) 'Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism', *Journal of Medicinal Chemistry*, 63(16), pp. 8749–8760.

Predictive Maintenance in the Digital Era

Aaron S. Yeardley^a, Jude O. Ejeh^a, Louis Allen^a, Solomon F. Brown^a and Joan Cordiner^{a*}

^a*Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom*
j.cordiner@sheffield.ac.uk

Abstract

With the advent of Industry 4.0, predictive maintenance is becoming increasingly popular. It has the potential to improve manufacturing capabilities by identifying equipment faults and preventing the shutdown of the production process. The sheer popularity of predictive maintenance has caused an abundance of research into developing new classification algorithms to successfully predict faults. However, more work is still required in the physical application of predictive maintenance to full industrial plants than in the development of new algorithms using single machines. This work focuses on the application of machine learning for predictive maintenance by prioritising a robust comparison of methods to investigate the accuracy of readily available classification methods. To achieve this, we used a large Fischertechnik (FT) model factory as a realistic and challenging case study to provide data rich with sensors and actuators such as that found in a cyber-physical production plant. We present a rigorous approach to compare and evaluate classifiers by analysing various error metrics. It was found that the Quadratic Discriminant Analysis (QDA) classifier performed the best with respect to the FT model factory dataset. The significance of the results provide validation that readily available machine learning methods are accurate enough to provide predictive maintenance and offer vital analytics. It further proves its suitability for robust maintenance scheduling in many industrial plants. It is expected that, the integration of machine learning and maintenance can significantly improve process safety and save money.

Keywords: predictive maintenance, machine learning, process safety, classification models

1. Introduction

Machine maintenance is of paramount importance to the process industry with regards to both safety and effectiveness. A poor maintenance system can be catastrophic to an organisation in terms of both performance and safety as it has a direct impact on costs, productivity, quality and accidents. It is thus common practise for companies to actively engage in maintenance activities for equipment items within the plant. This is mainly managed through three types of maintenance (Carvalho et al., 2019):

1. Corrective maintenance - maintenance tasks are conducted when a machine fails.
2. Preventative maintenance - maintenance tasks are conducted on a periodic schedule, also known as scheduled maintenance.
3. Predictive maintenance - uses techniques to estimate when maintenance is required.

Predictive maintenance is becoming increasingly popular with the transformation towards Industry 4.0 where process automation and digitisation are becoming the norm (Gilchrist, 2016). One of

the key advantages that Industry 4.0 presents is the abundance of data, which can be used in the control and operation of a plant, improving production efficiency and managing process safety. Predictive maintenance enables the estimation of when maintenance is required on a machine by continuously monitoring sensor information over time (Yan et al., 2017). The implementation of predictive maintenance has the potential to improve manufacturing capabilities by identifying equipment faults and preventing the shutdown of the production process. Currently, Industry 4.0 (Lee et al., 2014) is advancing due to Internet of Things (IoT) technologies providing many sensors that records data enriched by the control commands of actuators. This is transforming manufacturing environments into complex cyber-physical production systems (Gunes et al., 2014). The data measured by these sensors provides an invaluable resource that can be used to train a machine learning method to detect and predict failures. Predictive maintenance improves the decision-making process in manufacturing environments as faults can be predicted before they occur.

Machine learning methods have been increasingly applied to predictive maintenance applications (Carvalho et al., 2019). The most common method uses classification to predict a fault or failure occurring (Susto et al., 2015). Additionally, regression techniques have been used to predict Remaining Useful Life of machines (Van Horenbeek and Pintelon, 2013) and forecast industrial aging processes (Bogojeski et al., 2021). A systematic literature review by Carvalho et al. (2019) showed that the most common machine learning algorithms used are Random Forest, Neural Networks, Support Vector Machine and k-means clustering. Additionally, they found each machine learning method proposed was applied to a specific piece of equipment, for example turbines (Kumar et al., 2018), motors (Dos Santos et al., 2017) and compressors (Prytz et al., 2015). For this reason, it becomes difficult to compare various machine learning algorithms as each study uses vastly different data for validation. Additionally, Industry 4.0 is evolving at such speeds, applying predictive maintenance to just one piece of equipment is not useful when the data collected enables analysis on the whole production plant. As such, further research that compares proposed predictive maintenance strategies instead of developing novel machine learning algorithms is necessary (Carvalho et al., 2019). There is also an abundance of research solely concentrating on developing new machine learning techniques that use historical data to predict when maintenance is required, the integration of predictive maintenance remains briefly addressed in literature.

As noted, more work is required to apply predictive maintenance to full industrial plants and then to integrate it as a method to improve site maintenance. Therefore, this paper focuses on addressing the first issue applying predictive maintenance techniques to a large cyber-physical production system. We present an investigation into readily available classification algorithms for predictive maintenance by using them on a large collection of machine sensor data. We then consider the prospect of implementing the results as an initial tool in a promising application to develop robust maintenance scheduling in an industrial plant. Further, this study provides a rigorous approach to compare and evaluate the classification models, analysing various error metrics and ultimately choosing a machine learning method which can be used for further analysis.

2. Case Study Data

The comparison of predictive maintenance techniques requires the availability of historical maintenance and complex sensor reading data related to an industrial plant. However, the availability of real-data from industry is extremely limited due to confidentiality issues. For this reason, this study uses data provided by a cyber-physical production system developed by Klein and Bergmann (2019). The large Fischertechnik (FT) model factory records sensor readings enriched with control commands providing a realistic and challenging case study for detecting faults. The production plant consists of five workstations as described in an ontological knowledge base (Klein et al., 2019). Altogether, the FT plant consists of 14 machines which use 61 sensor readings indirectly

Table 1: A summary of the data used to train and test predictive maintenance techniques

Machine No.	Train	Test
No Fault	22,763	3,233
1	9	2
2	14	24
3	207	234
4	21	6
5	23	27
6	7	2
7	28	45
8	105	97
9	36	31
10	16	11
11	13	20
12	42	34
13	6	0
14	13	4
Total	23,303	3,770

related to each machine. Classification machine learning models are applied to the data providing an insight into the promising tools readily available to prevent failures in industry.

The large FT plant provides a realistic and challenging case study for detecting faults on 14 machines using the 61 sensor readings as input variables for predictive maintenance. Table 1 shows the raw data which is generated by multiple run-to-failure simulations where the sensor readings and the corresponding classes are recorded. Altogether, the data includes 27073 data points recording 28 different types of faults such as driveshaft slippage in the conveyor. In this work, the type of fault itself is not important as the objective of the predictive maintenance is to predict which machine a fault occurs on.

An investigation of this type requires a consistent methodology to compare each classification models used on the FT simulation data. Therefore, the full data is divided into training and test data based on complete simulations from start to finish. Each individual simulation that leads to a fault is included in either the test data or the training data (Klein and Bergmann, 2019). This means a fault that continues for multiple time steps will not be found in both training and test data. A summary of the classification data is shown in Table 1, where the clear split between training data and test data can be seen for each of the 15 classes.

3. Predictive Maintenance Methodology

This work compares five supervised classification techniques using the Python library, Scikit Learn (Pedregosa et al., 2011). The data described in Section 2 is used as a case study to test the following classification algorithms:

1. Decision Tree (DT) (Breiman et al., 2017)
2. Random Forest (RF) (Breiman, 2001)
3. Neural Network (NN) (He et al., 2015)
4. AdaBoost (AB) (Hastie et al., 2009)

5. Quadratic Discriminant Analysis (QDA) (Friedman, 1989)

The classification models are tested using the case study data and by calculating error metrics. These enable a robust comparison of prediction techniques by comparing the predicted labels to the true observed labels.

The Precision (P) score is defined as a ratio from the number of true positives (T_p) over the number of true positives plus the number of false positives (F_p):

$$P = \frac{T_p}{T_p + F_p} \quad (1)$$

The Recall (R) score is also a ratio. However, for this metric it is the number of true positives over the number of true positives plus the number of false negatives (F_n):

$$R = \frac{T_p}{T_p + F_n} \quad (2)$$

To combine both the Precision and the Recall, the F1-Score weights both Recall and Precision together as a harmonic mean:

$$F1 = 2 \frac{PR}{P+R} \quad (3)$$

Finally, the Accuracy (A) is the total number of true predictions (both true negatives (T_n) and T_p) as a percentage to the total number of predictions:

$$A = 100\% \times \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (4)$$

4. Results

4.1. Comparison of Classifiers

Figure 1 shows the QDA achieving the highest accuracy of 88.3% and all the other classification methods achieved scores above 85%. However, predictive maintenance data commonly shows a significant imbalance between classes and the same is true for the FT model factory. Out of a total of 3,770 data points, 3,233 are in the class "No Fault", therefore, constant predictions of "No Fault" would achieve an accuracy score of 85.8%. Table 2 show the weighted averages obtained for the Precision, Recall and F1-Score. Once again, we found that the QDA performed the best achieving the highest scores all round. Statistically, the significance of the results provide validation that the QDA is the best choice of method for this given dataset.

The results for all five machine learning models are satisfactory, but to get the best from predictive maintenance, trust in the predicted faults is of highest priority. This work has shown a methodology for analysing machine learning models to robustly choose the best method. For the given case study, the QDA was the chosen classification model for predictive maintenance.

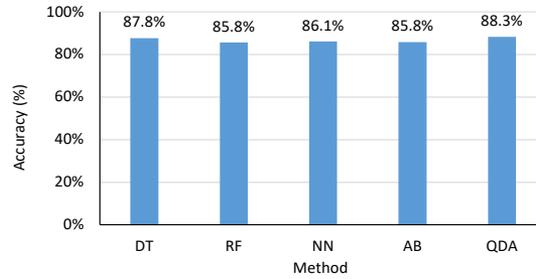


Figure 1: The accuracy score measured for the different classifiers.

Table 2: Resulting diagnostic values from the predictive maintenance

	Precision	Recall	F1 Score
Decision Tree	0.839	0.878	0.838
Random Forest	0.735	0.858	0.792
Neural Network	0.788	0.861	0.799
AdaBoost	0.736	0.858	0.793
Quadratic Discriminant Analysis	0.880	0.883	0.877

Table 3: A confusion matrix showing the condensed results from the QDA.

		True Label		Total
		No Fault	Fault	
QDA Predicted Label	No Fault	3086	95	3181
	Fault	147	442	589
Total		3233	537	3770

4.2. QDA Results

For predictive maintenance, a closer inspection of the label "No Fault" is important because incorrectly predicting a healthy state for all the machines would lead to undetected faults. Conversely, incorrectly predicting a fault (irrelevant to which machine it is on) would result in unnecessary maintenance. Table 3 shows the QDA results in a confusion matrix focusing on the class label "No Fault" by grouping the 14 machine fault classes into just one class, called "Fault". The confusion matrix shows the amount of predicted labels that were in the correct and incorrect label.

Previously, the accuracy of the QDA was given in Figure 1 with the QDA making 441 false predictions out of 3770 data points. Table 3 shows the QDA predicted 95 "No Fault" which were actually "Faults". Additionally the QDA falsely predicted 147 "Faults". This means 45.1% of the original false predictions (199 out of the 441 false predictions) did correctly predict a fault but on the wrong machine. Altogether, the QDA performed well, correctly predicting a fault 442 times out of 537 total faults and correctly predicting no fault 3086 times out of 3181. Therefore, when considering the class "No Fault" only, the QDA achieved an accuracy score of 93.6%. Showing the good performance from the QDA in terms whether a fault has occurred or not.

5. Conclusion

In this paper, we analyse classification models to be used for predictive maintenance in a cyber-physical production system. Confidentiality issues cause a lack of real-data. Hence, we conducted the robust comparison of machine learning classifiers using a case study from a Fischertechnik (FT) simulation model (Klein and Bergmann, 2019). Results showed that the machine learning

methods adopted were accurate in predicting faults, and suitable for predictive maintenance. The QDA classifier was also shown to be the best for this given dataset, achieving an accuracy of 88.3%. Although, when considering the class label "No Fault", the QDA achieved an accuracy of 93.6%, showing good performance from the algorithm for predictive maintenance.

In conclusion, this study investigated the performance of predictive maintenance on a complex cyber-physical production process. Accurate results from all five classifiers have shown that the machine learning algorithms are capable of predictive maintenance in Industry 4.0 when larger amounts of sensors are being used to analyse plant performance.

Future work could use the chosen predictive maintenance classifier to provide a route for robust maintenance scheduling by using it within maintenance workflow. Integrating machine learning techniques such as predictive maintenance, time estimation and schedule optimisation will enable smart maintenance policies to be implemented. This could provide a successful route for many industrial plants to significantly improve process safety and save money.

References

- M. Bogojeski, S. Sauer, F. Horn, K. R. Müller, 2021. Forecasting industrial aging processes with machine learning methods. *Computers and Chemical Engineering* 144, 107123.
URL <https://doi.org/10.1016/j.compchemeng.2020.107123>
- L. Breiman, 2001. Random Forests. *Machine Learning* 45, 5–32.
- L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, 2017. Classification and regression trees.
- T. P. Carvalho, F. A. Soares, R. Vita, R. d. P. Francisco, J. P. Basto, S. G. Alcalá, 2019. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers and Industrial Engineering* 137 (April), 106024.
URL <https://doi.org/10.1016/j.cie.2019.106024>
- T. Dos Santos, F. Ferreira, J. Pires, C. Damasio, 2017. Stator winding short-circuit fault diagnosis in induction motors using random forest. In: 2017 IEEE International Electric Machines and Drives Conference, IEMDC 2017.
- J. H. Friedman, 1989. Regularized discriminant analysis. *Journal of the American Statistical Association* 84 (405), 165–175, cited By :1736.
URL www.scopus.com
- A. Gilchrist, 2016. Industry 4.0. *The Industrial Internet of Things*. Apress.
- V. Gunes, S. Peter, T. Givargis, F. Vahid, 2014. A survey on concepts, applications, and challenges in cyber-physical systems. *KSII Transactions on Internet and Information Systems* 8 (12), 4242–4268.
- T. Hastie, S. Rosset, J. Zhu, H. Zou, 2009. Multi-class AdaBoost. *Statistics and Its Interface* 2 (3), 349–360.
- K. He, X. Zhang, S. Ren, J. Sun, 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification.
- P. Klein, R. Bergmann, 2019. Generation of complex data for AI-based predictive maintenance research with a physical factory model. *ICINCO 2019 - Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics 1* (September), 40–50.
- P. Klein, L. Malburg, R. Bergmann, 2019. FTOnto: A domain ontology for a fischertechnik simulation production factory by reusing existing ontologies. *CEUR Workshop Proceedings* 2454 (October).
- A. Kumar, R. Shankar, L. Thakur, 2018. A big data driven sustainable manufacturing framework for condition-based maintenance prediction. *Journal of Computational Science* 27, 428–439.
- J. Lee, H.-A. Kao, B. Bagheri, 2014. Recent Advances and Trends of Cyber-Physical Systems and Big Data Analytics in Industrial Informatics. *Proceeding of Int. Conference on Industrial Informatics (INDIN)* (October).
URL <https://www.researchgate.net/publication/266375284>
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- R. Prytz, S. Nowaczyk, T. Rönngvaldsson, S. Byttner, 2015. Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. *Engineering Applications of Artificial Intelligence* 41, 139–150.
- G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, A. Beghi, 2015. Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics* 11 (3), 812–820.
- A. Van Horenbeek, L. Pintelon, 2013. A dynamic predictive maintenance policy for complex multi-component systems. *Reliability Engineering and System Safety* 120, 39–50.
URL <http://dx.doi.org/10.1016/j.ress.2013.02.029>
- J. Yan, Y. Meng, L. Lu, L. Li, 2017. Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. *IEEE Access* 5, 23484–23491.

An evolutionary approach for techno-economic assessment

Martin C. De Meio Reggiani,^{a,b} Luciana B. Villar,^{a,c,d} Hernán P. Vigier,^c Nélica B. Brignole^{a,d}

^a *Laboratorio de Investigación y Desarrollo en Computación Científica (LIDeCC), Universidad Nacional del Sur (DCIC-UNS), Av. Alem 1253, Bahía Blanca, Argentina*

^b *Inst. de Invest Económicas y Soc del Sur (IESS), San Andrés 800, B Blanca, Argentina*

^c *Centro de Emprendedorismo y Desarrollo Territorial Sostenible (CEDETS), Univ. Provincial del Sudoeste (UPSO), Ciudad de Cali 320,8000 Bahía Blanca, Argentina*

^d *Planta Piloto de Ingeniería Química (PLAPIQUI), Complejo CCT-UAT, CONICET-UNS, Camino La Carrindanga Km 7, 8000 Bahía Blanca, Argentina*

Abstract

An evolutionary framework has been developed for techno-economic assessment at a decision-making level. A Master-Worker strategy based on Genetic Algorithms is proposed. It was enhanced with an ad-hoc chromosome redefinition, exclusively designed for techno-economic optimization. Specialized crossover and mutation operators were implemented so that the decision-maker places value on the economy of scale and the efficiency in logistics. The methodology enables to tackle systematically changes in the supply of raw materials and market demand altogether. By way of illustration, the eventual natural gas (NG) winter scarcity in Argentina, where industrial plants are forced to work under low-supply conditions, is discussed. For this case, efficient tactical planning was achieved by means of the optimization of the plants' on/off status, together with the assessment of the weekly NG amounts to be individually provided.

Keywords: Techno-economic assessment, Optimization, Genetic algorithm, Planning

1. Introduction

Increasing productivity is a key issue in Industry 4.0. From this perspective, methodologies to achieve operational excellence with further cost reductions have been gaining importance. Decision-making can be facilitated by merging both the production and market scenarios. In an Industry 4.0 environment, techno-economic assessment is crucial in order to achieve a competitive advantage. Its benefits go beyond saving money since it also comprises sustainable actions towards environmental protection.

In this century, interest is focused on the simultaneous improvement of both energy consumption and production. For the operational planning of petroleum supply chains, Neuro and Pinto (2004) applied a large-scale mixed-integer nonlinear programming (MINLP) model. He optimized techno-economically a real-world corporation through the adoption of different strategies. Nowadays, an integrative strategy called enterprise-wide optimization (EWO) is being explored. It deals with the optimal operation of manufacturing facilities, which includes planning, scheduling, real-time optimization and inventory control. In particular, electrical supply-chain management under economic crisis was addressed by Manenti (2013), who proposed a MINLP optimization

methodology that merges enterprise-wide and operational aspects to plan the production of complex networks under a low energy offer for high demand. For EWO, Scholz Dias (2019) recently proposed a combination of data-driven methods and classical optimization techniques, partially incorporating artificial intelligence to the analysis. Techno-economic optimization problems can be challenging to solve, due to their complexity, nonlinearity and potentially high dimensionality. Then, it is worthwhile to devise alternative techniques that deal with complex optimization problems with complicated constraints. Nature-inspired metaheuristic algorithms have demonstrated some promising results in diverse applications where the optimization problems are tough. In particular, genetic algorithms (GAs) tend to be flexible, efficient, highly adaptable and easy to implement. Moreover, GAs can efficaciously be applied in broad practical issues by tailoring both the individuals and the genetic operators to the specific combinatorial problem under study.

2. Problem statement

Throughout history, there have been circumstantial cuts in the supply of non-renewable energy sources, mainly caused by restrictions on the supply side. In this sense, the global recovery in the post-pandemic demand for goods and services was not matched by an equivalent increase in energy supply. In particular, production companies have had their NG supply reduced. Moreover, given the excessive demand in winter seasons (Shuquan et al., 2018) and the lack of investment in infrastructure and production (De Meio Reggiani et al., 2019), both residential and industrial consumption will have to confront higher bills and/ or reductions in NG supply. Therefore, it becomes necessary to optimize the distribution of available NG with a business-planning approach.

It is challenging to expand the scope of evolutionary strategies to be able to exploit them in this complex context. In this paper a metaheuristic methodology is proposed in order to address the particular case of industrial plants that have to face the conditions of scarce NG supply. The adopted test case represents a Gas Supply Problem (GSP) that has recently occurred in Argentina since 2005 (De Meio Reggiani, 2018). In this design, a Central Planner optimizes techno-economic aspects, simultaneously considering both the NG demand by industrial plants and the costs associated with transmission.

3. Fundamental modelling

A metaheuristic approach based on Genetic Algorithms (GA) was designed. It is a parallel implementation of a Master-Worker architecture that performs collaborative optimizations. A module (GA-M) is executed by the Master, while the others (GA-W) are executed by the Workers. Both are standard GA procedures that not only differ in the fitness assessment, but in the genetic operators as well. Given a prospective scenario, GA-W is in charge of evaluating its NPV based on NG availability for each plant's production ($Fitness_W$). In turn, GA-M employs a comprehensive objective function that takes into account both plant-processing NPV and transmission costs ($Fitness_M$).

The Master generates at random the prospective scenario for each Worker, which is defined by some binary values. Then, the Worker executes a GA-W in order to find the best individual. As soon as he has found a candidate solution, he reports it to the Master, who uses GA-M to optimize all the scenarios. In this way, every Worker is simultaneously looking for a good solution, while searching for the most promising places in the space of prospective scenarios. Parallel programming was implemented so that the Workers can simultaneously explore different instances. Thus, many cases are solved as quickly as possible to give effective support to the decision-making task.

As to the representation of the solutions, a dual structure of the phenotype was adopted, where boolean and real genes are combined in a single chromosome. In this way, a wide variety of decision-making variables is included for a rolling horizon comprising H weeks. The chromosome structure is exemplified in Fig. 1. A single chromosome is divided in groups corresponding to the weeks in the rolling horizon H . Each of these groups (W_i) contains the information that corresponds to the i -th week. They consist of a set of tuples $W_i = \{S_i, P_i\}$ with $i = 1 \dots H$, where S_i defines a global scenario through p boolean genes from g_1^i to g_p^i , and P_i contains $(n - p)$ real-valued plant features included in the genes ranging from g_{p+1}^i to g_n^i . A weekly NG provision has to be distributed among p plants. Each plant is characterized by f features. Hence, $n = p + fp$.

The example given in Fig. 1 corresponds to the GSP. A total NG provision of 32 Mm³/day

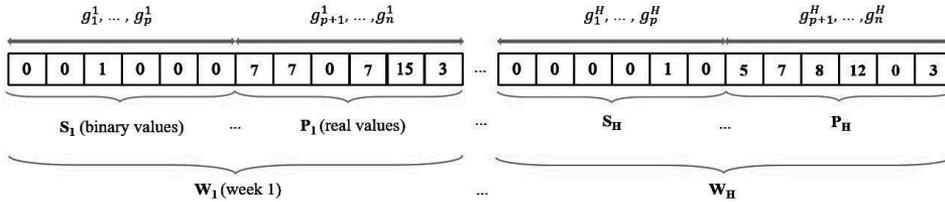


Figure 1. Conceptual representation of a single individual.

for the first week, where this amount has to be distributed among 6 plants with the corresponding NG feed as a single feature ($f = 1$). According to the binary string S_1 , it is proposed that the third plant will be stopped, i.e. $g_3^1 = 1$ and $g_9^1 = 0$.

Both the chromosome choice and the fitness function design are determined largely by the nature of the specific problem that is being tackled. In particular, our goal is to prevent economic losses when the weekly supply of operating industrial gases is forcefully restricted to a lower quantity Q_h^S . Therefore, Eq. 1 must be fulfilled for a time horizon H .

$$Q_h^S \geq \sum_{i=1}^p Q_{i,h} \quad \forall h = 1, H \quad (1)$$

For this problem, the optimization variables comprise the amount to be provided $Q_{i,h}$ for $i = 1, \dots, p$. Let us consider that there are p plants, where $Q_{i,h}^{baud}$ is the business-as-usual demand for the i -th plant and the h -th week. Besides, all plants require a minimum amount $Q_{i,h}^{min}$. Then, the constraints in Eq. 2 have to be incorporated into the formulation, which aims at finding the optima $Q_{i,h}$.

$$\{Q_{i,h}^{min} \leq Q_{i,h} \leq Q_{i,h}^{baud}\} \cup \{Q_{i,h} = 0\} \quad \forall i = 1, p; \forall h = 1, H \quad (2)$$

The Net Present Value (NPV) approach was adopted to assess the objective functions. It is the summation of the difference between the present value of the expected income and costs, which is discounted at a rate r that represents the opportunity cost of capital.

Each Worker calculates his own fitness function $Fitness_W$ related to plant processing (Eq. 3), where $I_{i,h}$, $VC_{i,h}$, and $FC_{i,h}$ are the average processing income, variable and fixed costs, respectively. The summation is discounted at the weekly interest rate r .

$$Fitness_W = \sum_{i=1}^p \sum_{h=1}^H (I_{i,h} \times Q_{i,h} - VC_{i,h} \times Q_{i,h} - FC_{i,h})(1 + r)^{-h} \quad (3)$$

In turn, the Master evaluates the corresponding global NPV by using an aggregation method (Eq. 4) with user-defined weights w_1 and w_2 . In the first term, the scenario is pondered using information about the NG transmission costs $TC_{i,h}$. The boolean

$\delta_{i,h}$ specifies the on/off status of the i -th plant during the h -th week, where $\delta_{i,h} = 0$ means the plant has stopped and $\delta_{i,h} = 1$ indicates that the plant is working. Moreover, it is mandatory to incorporate a Shutdown Constraint. To avoid too short shutdown periods, which would be techno-economically infeasible, a single plant is not allowed to stop more than once during the rolling horizon.

$$Fitness_M = -w_1 \sum_{i=1}^p \sum_{h=1}^H TC_{i,h} (Q_{i,h}^{baud} - Q_{i,h}) \delta_{i,h} + w_2 Fitness_W \quad (4)$$

Fig. 2 summarizes the algorithmic approach. In the first place, the Master’s goal is to intelligently select new starting locations. Then, he assigns them to a single Worker, who starts working from the suggested chromosome to breed the optimal child for the suggested scenario. As soon as possible, all Workers both reject infeasible scenarios and inform the Master about their candidate solution. Notice that there are separate assessment fitness functions: a Central-Planner one (Eq. 4 in GA-M) and a Plant-wide one (Eq. 3 in GA-W). The Master is in charge of judging the global appeal of the results. He ponders the candidate solutions by applying the aggregation method in order to assess the individual’s global fitness. The Master rejects poor solutions and ranks the good ones. The process iterates until the outcome is judged to be good enough. Finally, the Master reports the definitive solution.

Since production usually involves temporal dependency among processes, crossover and mutation operators were properly adjusted for GA-W, where the operations are on real-coded strings. A specialized crossover operator was designed on the basis of a real-coded Simulated Binary Crossover operator (Deb and Agrawal, 1995). It uses specific information about the problem’s constraints to make sure that all ensuing individuals were feasible. Moreover, the use of standard mutation operators was prone to create non-feasible individuals quite frequently. It should be noted that a single chromosome comprises time-ordered information corresponding to a rolling horizon, thus including

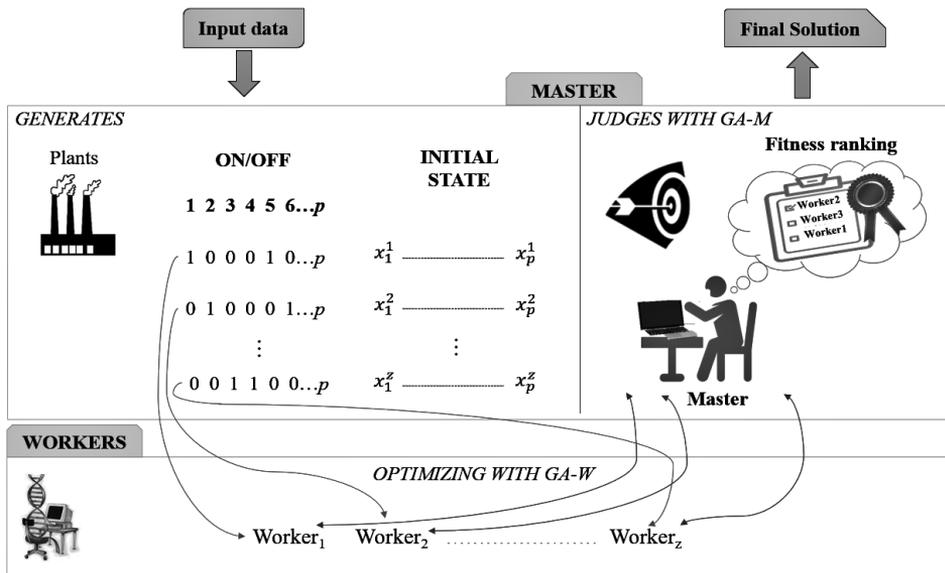


Figure 2. General framework for the techno-economic Master-Worker optimizer.

the state variables for various points in time. Since annual information should not be freely swapped, a mutation operator that always performs permutations within the same year was incorporated, based on a polynomial mutation approach (Deb and Deb, 2014). In contrast, a customized two-point crossover was implemented for the GA-M, where a binary approach with a Shutdown Constraint is necessary. This variation consists in randomly choosing a period of time and exchanging the genes between parents only within that time range. When a portion of each parent is swapped, there is a frequent violation of the constraint that requires that each plant should only stop once in the entire period H . After eliminating repetitions in the non-swapped sections of the children, the Shutdown Constraint is finally fulfilled. This strategy is complemented with a two-stage mutation operator. Firstly, the genes valued 1 are randomly converted to zero. Then, the bits valued 0 are randomly switched to 1, which are carefully assigned to only one plant.

4. Main results

Both technical and economic data for 10 plants of different capacities were considered for this GSP. These plants are located in various regions and the gas shortage lasted for 3 weeks. During this period, the total gas supply had to be reduced with respect to the usual demand in 65%, 55%, and 65% each week, respectively. The hyperparameters for both GAs are reported on Table 1. Since the GA-M model required greater diversity to succeed in finding the optimal pattern, higher population and mutation rates proved to be more effective in the Master's algorithm rather than in the Worker's.

Fig. 3 shows the performance of the optimization method. The parallel GA approach proved to be efficient because function evaluations could be distributed to different processors in order to be completed concurrently. Fig. 3.a depicts the speed-up evolution, showing a satisfactory growth of this ratio with the use of a higher number of cores. A sequential model needed approximately 30 min to find the best solution to this NP-Hard problem, while a 6-core implementation delivered the same result in 8 min. The curve slope implies that there is still room for speed improvement by adding new cores. Moreover, a well-known limitation of the traditional genetic algorithms is that they may fall into local optima quite easily. An alternative to overcome this limiting issue is to opt for methods based on game theory. In this context, parallelism allows the implementation of game operations. Thus, a more accurate and robust optimizer could be achieved.

GA	Population	Crossover Rate	Mutation Rate	Rank Choice	η
Master (GA-M)	100	1	0,2	0,8	-
Worker (GA-W)	50	1	0,05	0,8	15

Table 1. GA parameters.

Fig. 3.b shows the best $Fitness_W$ at each iteration and the corresponding $Fitness_M$, where $w_1 = w_2 = 1$. GA-M stopped when no significant improvements in the best fitness were detected in the last 20 iterations. When the Master stopped, he had found that 3 different plants should shutdown each week, yielding a maximum NPV of US\$ 99.972MM.

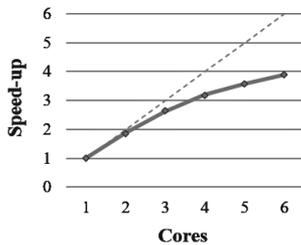


Figure 3.a. Parallel processing speed-up.

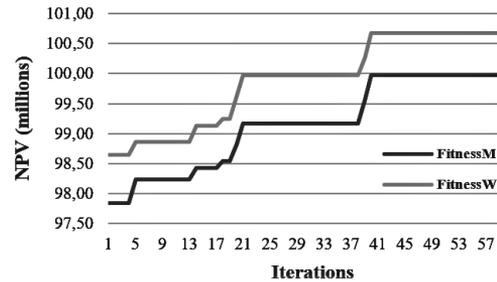


Figure 3.b. NPV variability.

5. Conclusions

Techno-economic assessment can provide energetic and economic incentives and reduce environmental impact when production is carefully optimized. In particular, a metaheuristic methodology based on GA is proposed in order to tackle low-demand conditions. This decision-making approach can anticipate future strategies aiming at NPV maximization. Suitable variants for both mutation and crossover operators were designed to guarantee feasible individuals. Testing on several realistic scenarios has shown that it is useful to make satisfactory comprehensive financial decisions, yielding economically viable scenarios when resource constraints have to be faced. At present, this software product can be considered as a lab-scale prototype, whose basic components have been integrated and computationally tested. The next step is to validate the actual system by testing its operation over a full range of Argentinian environmental conditions.

As to new ways of solving the GSP, it may be profitable to try a game-theory strategy, where population diversity is promoted by letting subgroups compete against each other. Future work should also focus on the incorporation of the metaheuristic approach in a comprehensive EWO framework, not only as a tactical planning methodology, but also at all levels of the decision-making hierarchy. It would also be interesting to determine the extent of model inaccuracies by incorporating uncertainties in the GA strategy.

References

- K. Deb, R. B. Agrawal, 1995, Simulated binary crossover for continuous search space. *Complex systems*, 9, 2, 115-148.
- K. Deb, D. Deb, 2014, Analysing mutation schemes for real-parameter genetic algorithms. *International Journal of Artificial Intelligence and Soft Computing*, 4, 1, 1-28.
- M.C. De Meio Reggiani, 2018, El transporte de gas natural en Argentina: análisis de la ruptura contractual y sus alcances. Doctoral dissertation, Universidad Nacional del Sur, Argentina.
- M.C. De Meio Reggiani, M. Vazquez, M.Hallack, N.B. Brignole, 2019, The role of governmental commitment on regulated utilities. *Energy economics*, 84, C.
- F. Manenti, G. Bozzano, M. D'Isanto, N. M. Lima, L. Z. Linan, 2013, Raising the decision-making level to improve the enterprise-wide production flexibility. *AIChE Journal*, 59, 5, 1588-1598.
- S. M. Neiro, J. M. Pinto, 2004, A general modeling framework for the operational planning of petroleum supply chains. *Computers & Chemical Engineering*, 28, 6-7, 871-896.
- L. M. Scholz Dias, 2019, Enterprise-Wide Optimization: Integrating Planning, Scheduling and Control Problems Using Feasibility Analysis and Surrogate Models. Doctoral dissertation, Rutgers, The State University of New Jersey, School of Graduate Studies.
- Z. Shuquan, S. Zhu, L. Yang, 2018, Reasons and countermeasures for the seasonal gas shortage in 2017. *Natural Gas Industry*, 38, 7, 120-128.

Optimization of an artificial neural network structure for modelling carbon capture in spray columns

Ulderico Di Caprio^a, Emine Kayahan^a, Min Wu^a, Siegfried Mercelis^b, Peter Hellinckx^b, Tom Van Gerven^c, Steffen Waldherr^d and M. Enis Leblebici^{a,*}

^a*Center for Industrial Process Technology, Department of Chemical Engineering, KU Leuven, Agoralaan Building B, 3590 Diepenbeek, Belgium*

^b*IDLab, Faculty of Applied Engineering, University of Antwerp-imec, Sint-Pietersvliet 7, 2000 Antwerp, Belgium*

^c*Process Engineering for Sustainable Systems, Department of Chemical Engineering, KU Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium*

^d*KU Leuven, Department of Chemical Engineering, Celestijnenlaan 200F-bus 2424, Leuven 3001, Belgium*

* muminenis.leblebici@kuleuven.be

Abstract

Anthropogenic CO₂ emissions reinforce global warming. The most mature technology to capture post-combustion CO₂ is the absorption with aqueous monoethanolamine (MEA) in spray or packed columns. However, this process is not considered economically viable. One reason for this hindrance is the low overall mass transfer coefficients ($K_G a$) resulting in large process volumes that are high in capital and operating costs. Spray columns offer some advantages over the packed columns such as the lack of expensive column internals, low pressure drop and higher $K_G a$. Although several spray columns have been investigated experimentally, there are only a few efforts to model this process. Because of the high intercorrelation between the process variables, modelling is a complex task. Machine learning techniques have shown great accuracy in modeling this kind of systems using large amounts of data. Artificial neural network (ANN) is one of the most promising and highly modular techniques in this field. The aim of this work is to find the ANN structure with higher accuracy and generalization capabilities to model the $K_G a$ for CO₂ absorption in spray columns using aqueous MEA. The ANN is trained using the back-propagation algorithm. The structure with higher accuracy and generalization properties is searched by a Bayesian optimizer. The trained model has a coefficient of determination (R^2) of 0.98, and a mean squared error (MSE) of 7.89e-4 on the validation set. It returns prediction errors below 20%. The optimizer found that the autoencoder structure had the higher prediction accuracy and generalization capabilities. This shape has great feature importance extrapolation capabilities, allowing the ANN to have a high prediction accuracy. The proposed procedure can be applied also for other processes where transfer coefficients need to be estimated and optimized.

Keywords: CO₂ capture, artificial intelligence, artificial neural network, absorption, spray columns

1. Introduction

Global warming is one of the main challenges that humanity needs to tackle soon. The anthropogenic CO₂ emissions reinforce it. Several strategies are available to reduce the

CO₂ footprint such as electrochemical separation, oxyfuel technologies, pre-combustion capture and post-combustion capture. Among these technologies, post-combustion capture requires little modification to the currently available processes (Luis, 2016). The most mature technology for the post-combustion CO₂ capture is the chemical absorption using aqueous monoethanolamine (MEA) solutions as absorbents (Luis, 2016; Wang et al., 2017). This process is often run in spray and packed columns. Although this is the most mature technology, it is not economically viable due to the high operating and capital investments required (Li et al., 2016). A strategy towards boosting the applicability of CO₂ capture processes is to increase the overall mass-transfer coefficient ($K_G a$) to enable compact processes which are easier and cheaper to retrofit on existing point carbon sources. $K_G a$ in a gas-liquid column is a combination of liquid-phase and gas-phase mass transfer resistances. Based on the two-flux theory (Afkhampour and Mofarahi, 2017), $K_G a$ can be calculated from the inlet and outlet CO₂ concentration and the inert gas flow rate as shown in (1).

$$K_G a = \frac{G_{inert} \cdot (Y_{CO_2,in} - Y_{CO_2,out})}{Z \cdot P \cdot (y_{CO_2} - y_{CO_2}^*)_{lm}} \quad (1)$$

where G_{inert} is the inert gas velocity (kmol*m⁻²*h⁻¹), P is the total pressure (kPa), Z is the height of the reactor (m), Y_{CO_2} the mole ratio of CO₂ to N₂ in the gas ($Y_{CO_2} = n_{CO_2}/n_{N_2}$), y_{CO_2} is the mole fraction of CO₂ in the gas phase and $y_{CO_2}^*$ is the mole fraction of CO₂ in equilibrium with the liquid phase which is assumed to be zero due to the fast reaction between MEA and CO₂.

An accurate prediction of $K_G a$ is crucial in the design of the absorption columns. Packed columns have been widely used in post-combustion CO₂ capture. For this case, several empirical and semi-empirical correlations for the prediction of $K_G a$ have been proposed (Wang et al., 2005). To the best of our knowledge, there are no empirical equations proposed to predict the $K_G a$ in spray columns for CO₂ capture. It is also not clear if an empirical correlation proposed for a specific column with a specific nozzle applies also to a different column and spray configuration. Therefore, more advanced prediction tools are needed to accurately estimate the $K_G a$.

Machine learning (ML) techniques have shown great capabilities in modelling complex systems with high intercorrelation and interdependency between the process variables (Schweidtmann et al., 2021). They build statistical models starting from a dataset by capturing the correlations between the input and the output variables available in it. Among the various ML techniques, artificial neural networks (ANNs) are very promising and modular (Lee et al., 2018). Their structure is a composition of an elementary unit called neuron. The neurons within the network are organized in layers; each layer takes as input information coming from the layer before, and its outputs are sent to the layers after it. At a given structure, the training of an ANN consists in finding the set of weights for each neuron that return the lower prediction error. Beside training of the neuron weights, number of layers and number of neurons per layer are critical hyper-parameters for the model that require an optimization to increase the performance of the network. The search of best structure is often carried out using the Bayesian optimizer (BO) (Snoek et al., 2012). The optimizer utilizes the previous experiences on the system to map the cost function to optimize it. Thus, the search of the best network is influenced by the past optimization attempt made by the optimizer.

In this work, we applied ML modelling with the aim to predict the $K_G a$ in a spray column for CO₂ capture with MEA solution from process parameters (liquid flowrate, gas

flowrate, CO₂ concentration in the gas, CO₂ concentration in the liquid, MEA concentration in the liquid and nozzle diameter). ANNs were trained on a dataset obtained from literature. ANN structure was optimized and tested. The search was performed using the BO.

2. Dataset description and data preprocessing

The data to train the model were obtained from the paper from (Kuntz and Aroonwilas, 2008) and (Wu et al., 2018). In both the papers, the authors employ single fluid nozzles to atomize the liquid. They reported the value of the $K_G a$ as function of the process variables, namely (1) the flowrate of the absorbent liquid, (2) the flowrate of the flue gas, (3) the concentration of the CO₂ within the gas, (4) the concentration of the sprayed MEA solution, (5) the concentration of CO₂ within the absorbent solution and (6) the diameter of the nozzle. These parameters were used as inputs of the model while the output was the $K_G a$ of the system. The dataset utilized was composed by 180 experimental points. To robustly evaluate the model performances, a cross-validation strategy was applied. From the original dataset, three sets were extracted: the training, the test and the validation sets. The training set contained 100 points and was used to train the neural network. The test set had 48 points. The MSE on this set was as loss function of the BO. The BO aims to minimize the MSE. The validation set contained by 32 points and used to benchmark the performance of the most performant network on the test set proposed by the BO. Prior the model training, the data were scaled in the range [0,1] to improve the performances of the optimizer.

3. ANN optimization description

The weights of the neurons composing the ANN were trained using the back-propagation algorithm. The Adam optimizer was employed (Kingma and Ba, 2014) and the mean squared error (MSE) between the prediction and the experimental value was used as loss function. In addition, Lasso (11) and Ridge (12) regularizations were applied on the training to avoid overfitting. The parameter for the l1 regularization was set to $\lambda_1=1e-5$ and the parameter for the l2 regularization was set to $\lambda_2=1e-4$. Each network structure was trained twice to reduce the impact of the initial guesses on the training performances as more repetition did not change substantially the result. The training returning the lower loss value on the training set was selected.

The BO algorithm implemented in Keras-tuner 1.0.1 was utilized in this study. It was employed to search the network structure with the lower MSE on the test set. To perform the search, we utilized 20 iterations of the algorithm. This value was experimentally observed to be a good balance between computational time and model efficiency. The BO could perform the search in the range [0,12] for the hidden layers number and [5,30] with a step of 5 for numbers of neurons per layer. These ranges were obtained experimentally. We gradually increased the size of the search space until the network found by the optimizer returned prediction with $MSE < 1e-3$ on the test set. The neurons contained in the internal layers of the network utilized LeakyReLU activation function. The output layer employed Sigmoid activation function. The capabilities of the obtained network are a-posteriori evaluated on the validation set. The lower the MSE on the validation set the higher are the generalization capabilities.

4. Results

4.1. Prediction accuracy of the best model

Figure 1 shows the prediction capabilities of the model structure with higher accuracy on the validation set identified by the BO. This figure reports the prediction performances both on the training set and on the validation set. In this figure all the points are clustered around the ideal prediction line with a narrow dispersion. In fact, both the prediction on the training and the validation points report error lower than 20% with only few exceptions. The coefficient of determination (R^2) on the validation set is 0.98 and the MSE is $7.89e-4$. This proves the high prediction accuracy and generalization capabilities of the most performant network identified by the BO. In addition, from Figure 1, one can detect how the dataset used in this study had an unbalanced output. Most of the experimental data points were clustered at middle values of the output with only few data points at the extreme values of $K_G a$. In open literature, it was shown that this data configuration decreased the accuracy of the predictions in the area where the data were rarer (Ribeiro and Moniz, 2020). The accurate prediction of the values located at high output rates was a crucial task when the aim of the model was process intensification. The prediction at high $K_G a$ had errors below 20% without any sensible bias, contributing to the high R^2 of this model. This is an important feature that allows the utilization of the model for design, control and optimization tasks. On the contrary, the model shows poor prediction capabilities at the extreme lower values of $K_G a$.

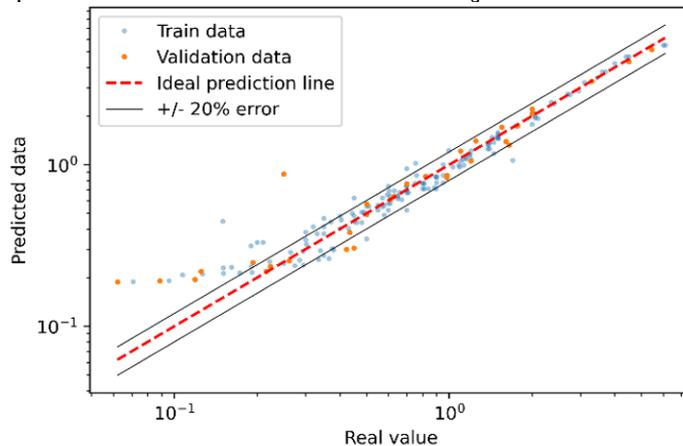


Figure 1: Prediction plot of the model with the higher accuracy on the validation set proposed by the optimizer. Most of the validation points are within the 20% error bend with only few exceptions. The only points with lower accuracy are located at lower output values where the points are rarer.

4.2. Analysis of the best structure

The optimizer identified the autoencoder structure to be the one with the higher accuracy on the test set (Figure 2). The first part of the optimal network is composed by one layer employed for the feature extraction in order to augment the input information (Layer 2 in Figure 2). The features extraction layer is followed by the autoencoder structure (from Layer 3 to Layer 7 in Figure 2). This structure is well known and applied in other fields (Bank et al., 2020). This network identifies the most important combinations of variables for the process and code them into internal variables of the network. This procedure is

done in the encoder of the network. For the case in analysis, the encoder terminates with 10 nodes encoding 10 internal variables (Layer 5 in Figure 2). This is the minimum amount of information that the network needs to perform the computation. Then, the information coded in the encoder are interpret in the decoder. The decoder is composed by two layers (Layer 6 and Layer 7 in Figure2) where the information is expanded and converted to the final output. Using this structure, the network returns low loss function values and very accurate predictions. From Figure 2 one can detect how some of the connections have the weight set to zero or its value is negligible compared to the other connections. This happens especially for the last two layers of the encoder (Layer 3 and Layer 4 in Figure 2). In addition, the optimizer chose 7 hidden layers even if the maximum allowed was 12. This avoided the overfitting on the training data and increased the generalization capabilities of the model. Moreover, the network reported in this study is quite deep and the amount of dataset used for its training could be not enough. However, the accuracy and the generalization capabilities of the model are ensured from the regularization and the cross-validation strategies employed for the training.

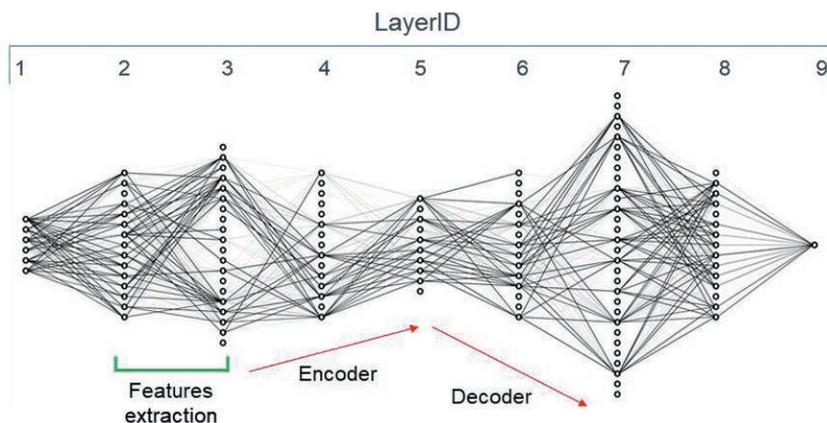


Figure 2: The network structure with the higher accuracy on the validation set proposed by the optimizer. It has the autoencoder structure. In this figure the darker is the color of the connection and the higher is the absolute value of the weight associated to it. The training has a major impact on the encoder. This allows a high selection of the variables used to encode the problem. One can detect how some of the connections have the weight set to zero or its value is negligible compared to the other connections.

5. Conclusions

In this work we proposed a network to predict the overall mass-transfer coefficient in a spray column for CO₂ capture into monoethanolamine. The best model structure returns highly accurate predictions with an R²=0.98 and remarkable generalization capabilities. The network was trained using a Bayesian optimizer to search the best network structure. The Bayesian optimizer identified the most proficient structure in the autoencoder. The network showed noticeable internal features selection to accomplish the prediction task. This methodology is a useful tool to predict internal parameters (such as the $K_G a$ as in this case) of first-principle models. It is a straightforward way to identify the best network structure and it is applicable to predict other process parameters using historical data.

6. Acknowledgements

The authors acknowledge funding from VLAIO Moonshot HBC.2019.0109 'Intensification of CO₂ capture processes (CAPTIN)'.

References

- M. Afkhamipour, M. Mofarahi, 2017, Review on the mass transfer performance of CO₂ absorption by amine-based solvents in low- and high-pressure absorption packed columns, *RSC Advances*, 7, 17857-17872
- D. Bank, N. Koenigstein, R. Giryes, 2020, Autoencoders, arXiv preprint arXiv:2003.05991
- D.P. Kingma, J. Ba, 2014, Adam: A method for stochastic optimization. arXiv Prepr. arXiv1412.6980
- J. Kuntz, A. Aroonwilas, 2008, Performance of Spray Column for CO₂ Capture Application, *Industrial & engineering chemistry research*, 47, 1, 145–153
- J. H. Lee, J. Shin, M. J. Realff, 2018, Machine learning: Overview of the recent progresses and implications for the process system engineering field.
- K. Li, W. Leigh, P. Feron, H. Yu, M. Tade, 2016, Systematic study of aqueous monoethanolamine (MEA)-based CO₂ capture process: Techno-economic assessment of the MEA process and its improvements, *Applied Energy* 165, 648-659
- P. Luis, 2016, Use of monoethanolamine (MEA) for CO₂ capture in a global scenario: Consequences and alternatives, *Desalination*, 380, 93–99
- R.P. Ribeiro, N. Moniz, 2020, Imbalanced regression and extreme value prediction, *Machine Learning*, 109, 9, 1803–1835
- A.M. Schweidtmann, E. Esche, A. Fischer, M. Kloft, J.U. Repke, S. Sager, A. Mitsos, 2021, Machine Learning in Chemical Engineering: A Perspective, *Chemie Ingenieur Technik*, 114, 12, 2029-2039
- J. Snoek, H. Larochelle, R.P. Adams, 2012, Practical bayesian optimization of machine learning algorithms, *Advances in neural information processing systems*, 25
- G. Q. Wang, G. Yuan, K. T. Yu, 2005, Review of Mass-Transfer Correlation for Packed Columns, *Industrial & Engineering Chemistry Research*, 44, 23, 8715–8729
- Y. Wang, L. Zhao, A. Otto, M. Robinius, D. Stolten, 2017, A Review of Post-combustion CO₂ Capture Technologies from Coal-fired Power Plants, *Energy Procedia*, 114, 650–665
- X.M. Wu, Z. Qin, Y.S. Yu, Z.X. Zhang, 2018, Experimental and numerical study on CO₂ absorption mass transfer enhancement for a diameter-varying spray tower, *Applied Energy*, 225, 367–379

Hierarchical Statistical Process Monitoring based on a Functional Decomposition of the Causal Network

Rodrigo Paredes,^a Tiago J. Rato,^a Lino O. Santos,^a Marco S. Reis,^{a,*}

^a *University of Coimbra, CIEPQPF, Department of Chemical Engineering, Rua Sílvio Lima, Pólo II – Pinhal de Marrocos, 3030-790 Coimbra, Portugal*

Abstract

With the continuous growth of data collection systems, even the state-of-the-art multivariate Statistical Process Monitoring (SPM) methods face difficulties in detecting localized faults, whose signatures easily pass unnoticed due to the large normal background noise associated to the many sensors under analysis. These methods are mostly non-causal and do not take into account the inner relationships between the variables or process units. In this work, we propose a new hierarchical monitoring approach based on a functional decomposition of the system's causal network. The methodology consists in finding the natural functional modules of the causal network, by exploring its graph topology and identifying the strongly linked “communities”. Two hierarchical monitoring schemes (aggregating information from the modules and their interactions) are then applied to monitor the overall state of the process. In this way, as the dimensionality of the modules is smaller, the sensitivity of the distributed system to small or localized faults is preserved. Furthermore, the causal nature of the method facilitates fault diagnosis, especially for sensor faults. However, the overall false alarm rate of the methodology must be controlled, which may take away some of the sensitivity of the proposed method. We report results that demonstrate a robust increase in the fault detection sensitivity of the proposed methodologies when compared to methods that monitor the complete causal network. The proposed approach also led to a more effective and unambiguous complementary fault diagnosis activity.

Keywords: Statistical Process Monitoring; Causal Network; Hierarchical Monitoring; Community Detection; Distributed Monitoring

1. Introduction

Causal networks hold a large amount of information on how process variables relate to each other. Nevertheless, most of the current Statistical Process Monitoring (SPM) methodologies, such as those based on Principal Component Analysis (PCA), Partial Least Squares (PLS) and Independent Component Analysis (ICA) (Ge *et al.*, 2013; Qin, 2012), are non-causal and do not account for this inner system structure. This not only reduces their ability to detect faults, but also impacts fault diagnosis due to a smearing-out effect that propagates the fault through non-causal related variables.

As an alternative, causal SPM methodologies have been proposed. One example, also adopted in this work, is the Sensitivity Enhancing Transformation (SET) method (Rato and Reis, 2014). SET uses the causal network to build a whitening filter that decorrelates the variables based on their inner relationships. If the relationships of the inferred causal network are still valid, the filtered variables should remain confined within a Normal

Operating Region (NOC). SET also facilitates fault diagnosis since a deviation on a filtered variable signals a change in the inferred relationships between the original variable and its causal parents.

However, the aforementioned methodologies (including those based on causality, such as SET) lose detection sensitivity with the increase of monitoring dimensionality (*i.e.*, the number of process variables to monitor). To overcome this limitation, decentralized monitoring methodologies have been proposed to segment the process variables into several modules and then perform local monitoring of each module (Ge *et al.*, 2010; Ge and Song, 2013; Tian *et al.*, 2019).

To combine the advantages of causal-based and decentralized monitoring, we propose two hierarchical monitoring schemes (CNET-C: **C**ausal **N**etwork-**C**entralized, and CNET-D: **C**ausal **N**etwork-**D**istributed) based on the functional decomposition of the causal network. In the proposed methodologies, the process variables are divided into modules (defined as *communities*) of closely related variables and then each module is monitored by the SET monitoring methodology. The subdivision into communities can be made based on the process and instrumentation diagram or analysis of the causal network through a community detection algorithm that evaluates the network topology and the density of associations between variables or process units (Masooleh *et al.*, 2021; Javed *et al.*, 2018; Harenberg *et al.*, 2014). As process monitoring is done at the community level, which aggregate subsets of the original variables, the proposed monitoring methodologies become more sensitive to localized fault affecting one or a few variables. Likewise, fault diagnosis also capitalizes from the decorrelation properties of the SET.

This article is organized as follows. In Section 2, the proposed monitoring methodologies are briefly described. Then, in Section 3, we report the results for a simulated case study that compared the proposed methodologies against two monitoring benchmarks. The results are discussed in Section 4. Finally, we present our conclusions in Section 5.

2. Methodologies

2.1. Fault Detection

The proposed centralized (CNET-C) and decentralized (CNET-D) hierarchical monitoring schemes are based on the functional decomposition of the causal network.

The first modelling stage of both methodologies is the same. In this stage, the network is inferred by use of partial correlations (Rato and Reis, 2017) and the dominant causal directions are determined through analysis of the cross-correlation or Granger causality (Yuan and Qin, 2012). Afterwards, the variables are divided into communities using an algorithm that analyses the topology and density of the network. This division into communities effectively reduces the dimensionality of system, and improves the sensitivity to localized faults. To account for inter-community associations, each community is expanded to also include the Markov-blanket of each causal parent in the community. In the next stage, a SET model is built for each community. The SET model consists of a whitening filter that accounts for the causal structure of the data. The SET model is obtained by regressing each variable onto its causal parents (Markovian approach) (Rato and Reis, 2017). The filtering operation of the SET also produces uncorrelated variables, which improves the diagnosis task discussed in Section 2.2. Afterwards, for each community, the SET filtered variables are monitored through a Hotelling's T^2 statistic ($T^2_{SET,i}$, for $i = 1, 2, \dots, c$, where c is the number of communities). The difference between the two proposed methodologies lies in the way the information coming from the communities (*i.e.*, the $T^2_{SET,i}$ statistics) is aggregated.

In the decentralized CNET-D, the $T^2_{SET,i}$ are monitored separately and aggregated through a “OR gate”. In this approach the control limits are established at the $T^2_{SET,i}$ level, with the significance level corrected to control the false alarm rate due to the use of multiple monitoring statistics. The global alarm of the CNET-D is then triggered if at least one community has a monitoring statistic above its control limit.

In turn, for the centralized CNET-C, the $T^2_{SET,i}$ of all communities are concatenated to build a single monitoring statistic (Equation 1):

$$T^2_{SET_{CNET-C}} = (\mathbf{t}_j - \bar{\mathbf{t}}_{NOC})^T \mathbf{S}_{NOC}^{-1} (\mathbf{t}_j - \bar{\mathbf{t}}_{NOC}), \quad (1)$$

where $\mathbf{t}_j = [T^2_{SET,1}, T^2_{SET,2}, \dots, T^2_{SET,c}]$ is vector that concatenates all $T^2_{SET,i}$ at observation j , $\bar{\mathbf{t}}_{NOC}$ is a vector with the sample means of $T^2_{SET,i}$ in normal operation conditions (NOC), and \mathbf{S}_{NOC} is the sample covariance matrix of the $T^2_{SET,i}$ in NOC. For CNET-C, a global alarm is triggered if $T^2_{SET_{CNET-C}}$ is above its control limit.

2.2. Fault Diagnosis

The fault diagnosis is performed in two levels. In the first diagnosis level the communities with abnormal variables are identified. Afterwards, in the second diagnosis level, the abnormal communities are inspected to identify the abnormal variables within them.

In the CNET-D methodology, the first diagnosis level is direct. Since fault detection is carried out at the community level, it suffices to observe which communities triggered an alarm in the logical “OR gate”. In turn, in the CNET-C methodology, the first diagnosis level is performed by assessing the contributions of each community to the $T^2_{SET_{CNET-C}}$ statistic. After identifying the abnormal communities, the second diagnosis level is run to evaluate the contributions of each variable to the $T^2_{SET,i}$ statistics of the abnormal communities. At this point, the contribution of each variable is compared against a control limit, and those above it are deemed to be related with the fault’s root cause.

3. Results

To study the performance of the proposed methodologies, a static linear causal network simulator with 16 variables (Rato and Reis, 2014) was used to generate NOC and faulty data. In this study we considered process, sensor, and correlation faults in several variables, but for the sake of space, only one case is reported here. Each failure magnitude was replicated 100 times to account for variability. The replicated simulations have 5000 observations each. As this work is not focused on network inference, for the proposed monitoring methodologies it was assumed that the causal network was inferred accurately. Furthermore, the causal network was divided into 3 communities by an algorithm based on network density and topology applications. All monitoring methodologies were set to a false alarm rate of $\alpha = 0.01$ (with Bonferroni correction), by taking the $(1 - \alpha/N) \times 100\%$ upper percentile of the monitoring statistics in NOC data (where N is the number of monitoring statistics in the monitoring scheme).

3.1. Fault Detection

Figure 1 shows the fault detection sensitivity study (True Positive Rate (TPR) versus Fault Magnitude) of the proposed methodologies against two benchmark methodologies (SET-Hotelling- T^2 and PCA) for a process fault in variable 1. This figure clearly shows that CNET-C and CNET-D have significantly higher TPR than the benchmarks. This improvement is a result of the functional decomposition into communities as it increases the sensitivity to localized faults.

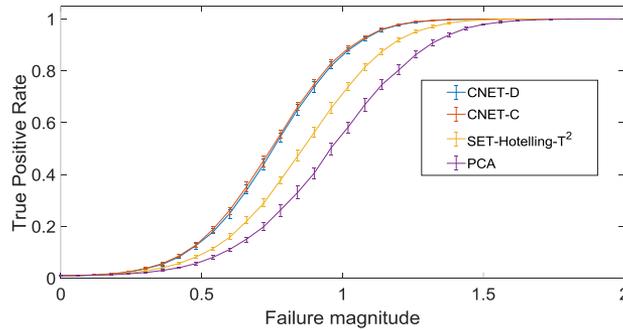


Figure 1: Median true positive rate and interquartile range for a process fault in variable 1. The fault magnitude is defined as k times the standard deviation in NOC.

3.2. Fault Diagnosis

To demonstrate the fault diagnosis capabilities of the proposed methodologies a process fault in variable 1 with a failure magnitude of 1 standard deviation was chosen. Fault diagnosis was performed by use of contribution plots for PCA (Qin *et al.*, 2001) and the contributions to the SET-Hotelling- T^2 (Rato and Reis, 2017). In Figures 2 and 3, the root cause diagnosis for the two benchmark methodologies is graphically represented. Figure 4 shows the two diagnostic levels of the proposed CNET-C methodology. For this case, only SET-Hotelling- T^2 and CNET-C (as well as CNET-D) provide an unambiguous diagnosis, correctly identifying variable 1 in all replicates.

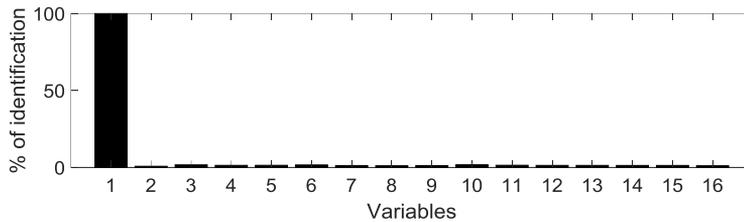


Figure 2: Percentage of times that each variable is significant to the SET-Hotelling- T^2 methodology monitoring statistics. Process fault located in variable 1 with a magnitude of 1 standard deviation.

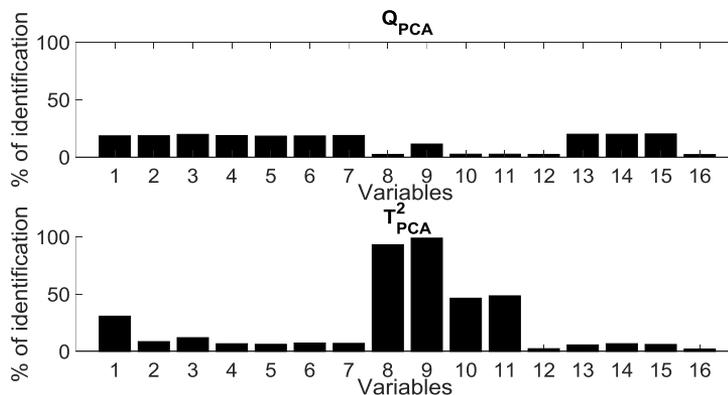


Figure 3: Percentage of times that each variable is significant to the PCA-based methodology monitoring statistics. Process fault in variable 1 with a magnitude of 1 standard deviations.

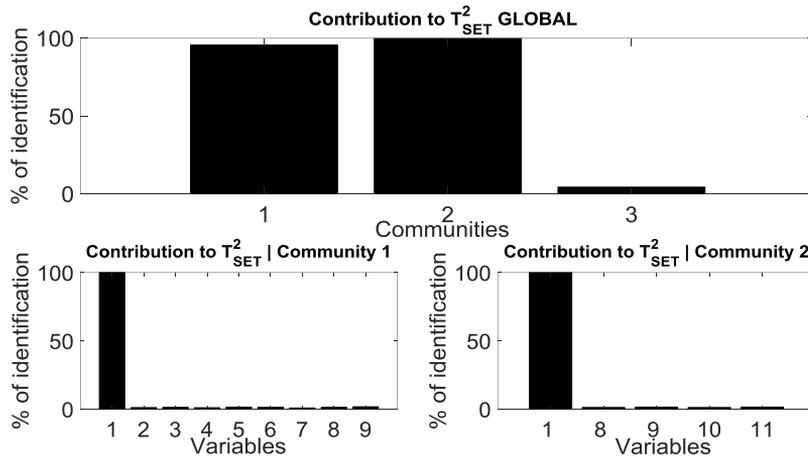


Figure 4: Percentage of times a community is significant for the CNET-C methodology monitoring statistic (first diagnostic level; on top), and the percentage of times a variable is significant for each community's $T^2_{SET,i}$ monitoring statistic (second diagnostic level; on bottom). Process fault in variable 1 with a magnitude of 1 standard deviations.

4. Discussion

Regarding fault detection, the results show that methodologies based on the causal network have greater sensitivity in most scenarios. The causal network functional decomposition into communities reduces the system monitoring dimensionality, which makes the proposed methodologies, CNET-C and CNET-D, to always have greater sensitivity when compared to the benchmark methodology that uses the complete causal network (SET-Hotelling- T^2).

The proposed methodologies have similar performances in terms of fault detection capabilities. Nevertheless, CNET-C tends to perform better in process and correlation failures, while CNET-D is generally better on sensor failures.

CNET-D can lead to sensitivity loss when applied to high dimensional systems due to the correction of the control limits. In this scenario, the high number of communities effectively increases the number of monitoring statistics, and their corrected significance level will tend towards zero. To overcome this limitation, we propose to correct the control limits using methods that control for the false discovery rate (Glickman *et al.*, 2014), instead of the Family Wise Error Rate (such as the Bonferroni correction).

For fault diagnosis, the introduction of the causal network into the monitoring schemes increases the diagnostic capacities compared to the conventional non-causal PCA (Figures 2 to 4). Figure 3 shows that the PCA diagnosis is ambiguous and inconclusive, *i.e.*, it does not diagnose only the variable at fault. In contrast, the SET-Hotelling- T^2 , CNET-C, and CNET-D have a conclusive diagnosis and identified the correct variable in all replicates, while the other variables are identified less than 2% of the times.

The developed methodologies include two hierarchical diagnosis levels. The first diagnosis level identifies the communities with abnormal variables. Afterwards, the second diagnosis level identifies the abnormal variables within the communities diagnosed in the first level. This greatly reduces the monitoring burden requested to process operators.

As the communities were extended to also include the Markov-blanket of the causal parents in each community, some variables may belong to more than one community. If

a fault occurs in these variables, then more than one community may be identified at the first diagnosis level. In these situations, the second diagnosis level is able to identify the correct root cause within the affected communities. Figure 4 represents a scenario where this situation happens: the faulty variable is correctly diagnosed in two communities, proving the good detection performance of the developed methodologies.

5. Conclusions

Causal networks contain information about systems that is typically not used in standard SPM methodologies. The introduction of causal structure into the monitoring methodologies leads to sensitivity gains in fault detection and diagnosis compared to conventional methodologies. Furthermore, the proposed causal network functional decomposition into several communities reduces the system dimensionality, and further increasing the sensitivity to localized faults. The application of the Markov-blanket to the causal parents in each community also allowed for inter-community associations to be included in the models. Fault diagnosis is also made on a reduced set of variables, which improves fault isolating, and leads to a conclusively and unambiguously diagnosis. The CNET-C and CNET-D methodologies demonstrated better performance in fault detection and diagnosis than their PCA and SET counterparts.

References

- Z. Ge, M. Zhang, Z. Song, 2010, Nonlinear process monitoring based on linear subspace and Bayesian inference, *Journal of Process Control*, 20, 5, 676-688.
- Z. Ge, Z. Song, 2013, Distributed PCA Model for Plant-Wide Process Monitoring, *Industrial & Engineering Chemistry Research*, 52, 5, 1947-1957.
- Z. Ge, Z. Song, F. Gao, 2013, Review of Recent Research on Data-Based Process Monitoring, *Industrial & Engineering Chemistry Research*, 52, 3543-3562.
- M.E. Glickman, S.R. Rao, M.R. Schultz, 2014, False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies, *Journal of Clinical Epidemiology*, 67, 8, 850-857.
- S. Harenberg, G. Bello, L. Gjeltema, S. Ranshous, J. Harlalka, R. Seay, K. Padmanabhan, N. Samatova, 2014, Community detection in large-scale networks: a survey and empirical evaluation, *Wiley Interdisciplinary Reviews: Computational Statistics*, 6, 6, 426-439.
- M.A. Javed, M.S. Younis, S. Latif, J.Qadir, A.Baig, 2018, Community detection in networks: A multidisciplinary review, *Journal of Network and Computer Applications*, 107, 87-111.
- L.S. Masooleh, J.E. Arbogast, W.D. Seider, U. Oktem, M. Soroush, 2021, An efficient algorithm for community detection in complex weighted networks, *AIChE Journal*, 67, 7, 1-15.
- S.J. Qin, S. Valle, M.J. Piovoso, 2001, On unifying multiblock analysis with application to decentralized process monitoring, *Journal of Chemometrics*, 15, 9, 715-742.
- S. J. Qin, 2012, Survey on data-driven industrial process monitoring and diagnosis, *Annual Reviews in Control*, 36, 2, 220-234.
- T.J. Rato and M.S. Reis, 2014, Sensitivity enhancing transformations for monitoring the process correlation structure, *Journal of Process Control*, 24, 6, 905-915.
- T.J. Rato, M.S. Reis, 2017, Markovian and Non-Markovian sensitivity enhancing transformations for process monitoring, *Chemical Engineering Science*, 163, 223-233.
- Y. Tian, T. Hu, X. Peng, W. Du, H. Yao, 2019, Decentralized monitoring for large-scale process using copula-correlation analysis and Bayesian inference-based multiblock principal component analysis, 33, 8, e3158.
- T. Yuan, S.J. Qin, 2012, Root cause diagnosis of plant-wide oscillations using granger causality, 8th IFAC International Symposium on Advanced Control of Chemical Processes, 160-165.

A Recurrent Neural Networks-Based Approach for Modeling and Control of a Crystallization Process

Fernando Arrais R. D. Lima^{a,*}, Gabriel F. M. de Miranda^a, Marcellus G. F. de Moraes^b, Bruno D. O. Capron^a, Maurício B. de Souza Jr.^{a,b}

^a*School of Chemistry, EPQB, Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco E, 21941-914, Rio de Janeiro, RJ – Brazil*

^b*PEQ/COPPE – Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030, CT, Bloco G, G115, 21941-914, Rio de Janeiro, RJ – Brazil*

*farrais@eq.ufrj.br

Abstract

Crystallization is a separation and purification process applied in many industrial sectors. The unit operation aims to achieve desired crystal size and shape distribution, making process control a key tool for its success. Despite the importance of controlling batch crystallization processes, there is still a lack of studies applying neural network-based control strategies. Therefore, this work aims to model a crystallization process to predict the moments of the particle-size distribution with neural networks used as the internal model in the predictive controller. Four different neural networks paradigms were considered: a classic single Multilayer Perceptron (MLP) network, a set of four MLP networks in series, and two recurrent networks, the Echo State Network (ESN) and the Long Short-Term Memory (LSTM). The dataset used for training and testing applied a co-teaching learning algorithm, which utilizes simulated and experimental data. The 479 experimental values of concentration, and particle number, length, area, and volume were obtained for several temperatures and ten different batch experiments of potassium sulfate (K_2SO_4) crystallization. The 9000 simulated data were generated using a population balance model for the system. First, the four network structures were trained to predict the moments of the particle size distribution values one step ahead, using the current temperature and moments values as feed. As a result, all strategies were successful, achieving values of R-squared of about 99% for the test samples. Then, the network's predictive performance was studied for larger prediction horizons. The ESN had the best performance, achieving values of R-squared above 90%, for eight out of 10 experiments and up to five steps ahead prediction. In comparison, the other strategies reached values below 90% for more than two experiments. Finally, a Nonlinear Model Predictive Controller (NMPC) based on the selected ESN was successfully applied to the batch crystallization process to maintain crystal size distribution on their desired trajectories by manipulating the operating temperature. The controller behavior was studied for four reference trajectories: constant, 1st order, 2nd order, and adaptive 1st order. As a result, the ESN-based NMPC presented better results, both in terms of performance and computational demand, than an NMPC based on the classic MLP, evidencing the potential of the proposed strategy.

Keywords: Crystallization, Neural Networks, Echo State Network, LSTM, NMPC.

1. Introduction

Crystallization is a purification and separation process applied in many industrial fields, such as the food and pharmaceutical sectors. This unit aims to generate a solid crystalline

product with high purity and desired crystal size and shape distribution, making process control a vital tool for its success [1]. In this way, model predictive control appears as a possible strategy for controlling of a crystallization unit.

Crystallization processes are usually modeled by population balances to predict the size and shape of the particles. One of the reasons for its use is the existence of an established kinetic model [1]. However, neural networks have been applied to model many chemical processes, and there is a lack of application for the modeling and control of crystallization processes.

This work aims to model a crystallization process to predict the moments of the particle-size distribution with neural networks. Four different structures of neural networks were considered: a classic single Multilayer Perceptron (MLP) network, a set of four MLP networks in series, and two recurrent networks, the Echo State Network (ESN) and the Long Short-Term Memory (LSTM). Also, the dataset used for training and testing applied a co-teaching learning algorithm, which utilizes experimental and simulated data [2]. The four networks were compared based on their five steps ahead prediction performance, and finally, the best option was used as an internal model in a Nonlinear Model Predictive Controller (NMPC).

2. Neural Networks

Neural Networks are a paradigm belonging to the class of machine learning techniques; they learn new patterns when presented to new data. The MLP is a feedforward network composed of fully connected neurons. Therefore, each neuron performs a weighted sum of inputs, and this result minus a bias value is sent to an activation function, such as a hyperbolic tangent. The result obtained by the activation function is used as the input for the subsequent neurons. Then, adjusting the weights and biases, the network seeks to learn the general function that relates the inputs and outputs of the given data. MLP is a feedforward network commonly used in the literature.

Recurrent Neural Networks (RNN) are a different kind of neural network created to represent sequential data. The difference between RNN and the feedforward networks is that they contain a feedback loop where data can be fed back to previous layers before it is fed forward again for further processing. This kind of structure gives the network the capability to simulate time-series and understand the dynamic behavior. LSTM and ESN are two different kinds of RNN. The first one presents a cell structure composed of three gates: forgot, input, and output. These gates have the function of controlling the flow of information inside the network [2]. On the other hand, ESN comprises an input layer, a dynamical reservoir, and an output layer. The reservoir layer is the recurrent part, where the outputs of the reservoir and output layers are feedback. Moreover, the weights of the reservoir layer are not trained but randomly fixed, avoiding local minimum problems.

Four different neural networks paradigms were studied to be used as the model for the controller. These networks structures were a single MLP, four MLP in series, ESN, and LSTM. The moments μ and temperature values in a time k were used as the inputs of the networks to predict the moments' values one step (i.e., one sampling time) ahead. Then, the network's capacity to predict the moments' values more steps ahead was studied, and all models were compared. Also, the dataset used for training and testing applied a co-teaching learning algorithm. The 479 experimental values of concentration and particles

number, length, area, and volume were obtained for several temperatures and ten different batch experiments of potassium sulfate (K_2SO_4) crystallization. The 9000 simulated samples were generated from the Population Balance (PB) Model developed by Moraes et al. [1], of which 3000 were used as test samples. The training dataset used data from one experiment and 6000 simulated data. For test samples, the dataset was composed of the other experimental data and three simulations. Each simulation was composed of 1000 data and used different initial conditions.

First, a single MLP network was designed using the temperature and the moments as inputs to predict the future values of the moments. This network structure was composed of a single hidden layer with 400 neurons. Also, a batch size of 350, Relu as the activation function, and the solver Adam were applied.

The second structure is composed of four MLP networks in series. The first network receives temperature and μ_0 values to predict μ_0 one step ahead. The prediction of the first network is used as input of the next one with the temperature and μ_1 at the same time to predict μ_1 one step ahead. The other two networks apply the same methodology, but the third one uses temperature, μ_1 and μ_2 to predict μ_2 one step ahead, and the last one uses temperature, μ_2 and μ_3 to predict μ_3 one step ahead. All four networks used the same specifications, which was one hidden layer with 195 neurons. Moreover, a batch size of 50, Relu as the activation function, and the solver Adam were applied.

The recurrent networks used the same inputs and outputs as the single MLP. For the ESN, a single reservoir and Identity activation function for the output layer and the output recurrence were used. For the LSTM, one LSTM layer with 120 cells and hyperbolic tangent as activation function were employed. Furthermore, sigmoid as a recurrent activation function, a batch size of 250, and the solver Adam were adopted.

3. Model Predictive Controller Design

The optimization problem considered as controlled variables μ_0 , μ_1/μ_0 , μ_2/μ_0 , and μ_3/μ_0 . The manipulated variable u is the process temperature. The MPC goal is to minimize the process performance index J , the objective function of the problem shown in Equation 1. The parameters $\delta(i)$ and γ are the weights of the outputs and the increment input, respectively; $y_i(k + j)$ is the i th predicted output at time $k + j$ and $y_i^r(k + j)$ its reference value. For the ESN, $\delta(i)$ was defined as 32 for μ_0 , and 5 for the other variables. For the MLP, $\delta(i)$ was defined as 45 for μ_0 , 30 for μ_2/μ_0 , and 15 for the other variables. The input increment is defined by $\Delta u(k)$ as a control horizon of 1 was assumed. P is the prediction horizon, which was chosen as 5 sampling intervals.

$$J = \sum_{i=1}^4 \sum_{j=1}^P \delta(i)[y_i(k + j) - y_i^r(k + j)]^2 + \gamma[\Delta u(k)]^2 \quad (1)$$

The optimization constraints are given by Equations 2-4, where C is the concentration of the solute, and C_{eq} is the concentration of the solute in equilibrium.

$$-1 \leq u(k) - u(k - 1) \leq 1 \quad (2)$$

$$C(k)/C_{eq}(k) \geq 1 \quad (3)$$

$$C_{eq}(k) = -686.2686 + 3.579165u(k) - 0.00292874u(k)^2 \quad (4)$$

4. Results and Discussion

First, all four structures were tested to make predictions one step ahead, achieving an R-squared value of 99% for test and training samples. Then, each network's performance was studied for predictions two and five steps ahead, and the results for predictions five steps ahead are illustrated in Figure 1. As a result, the ESN presented the best performance, reaching an R-squared higher than 75% for all experiments.

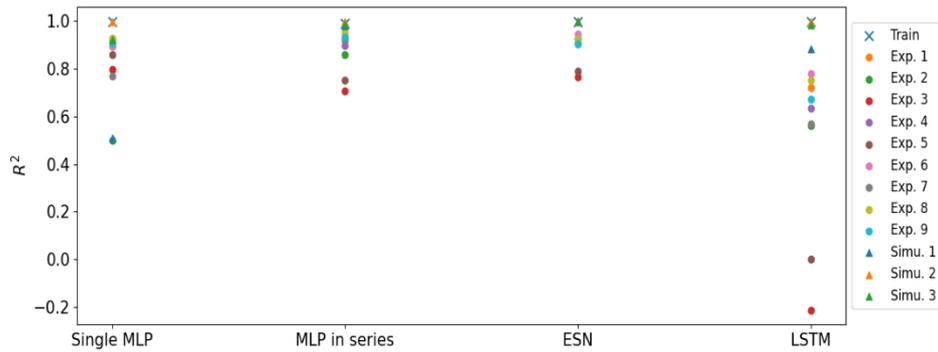


Figure 1: Performance of all neural networks for predictions five steps ahead.

After studying all neural networks behavior, the ESN was chosen to be the model for the predictive controller because it presented the best performance. However, the network's design had to be modified to achieve better control results. Therefore, one ESN was used to predict each moment with the same specifications as the design for four outputs. All four networks presented a performance very close to the other ESN design.

Using the ESN as the model, the controller's performance was compared to a more traditional control approach, using an MLP network. For this second case, four MLP networks were used to predict each moment value with the specifications presented in Table 1, in which T is the temperature. Also, a single hidden layer, Relu as activation function, and the solver Adam were adopted for all four cases. Both strategies predicted the moments five steps ahead and used the same inputs and outputs.

Table 1: Specifications for each MLP used in the MPC.

Inputs	T, μ_0	T, μ_0, μ_1	T, μ_1, μ_2	T, μ_2, μ_3
Outputs	μ_0	μ_1	μ_2	μ_3
Number of neurons in the hidden layer	50	295	300	300
Batch size	20	50	50	50

The proposed scheme for the control loop is illustrated in Figure 2. The process was modeled by the PB Model, and the optimizer used the Successive Quadratic Programming to find the optimum temperature value for each sampling. Also, a sampling time of one minute and the initial conditions of one experiment were chosen.

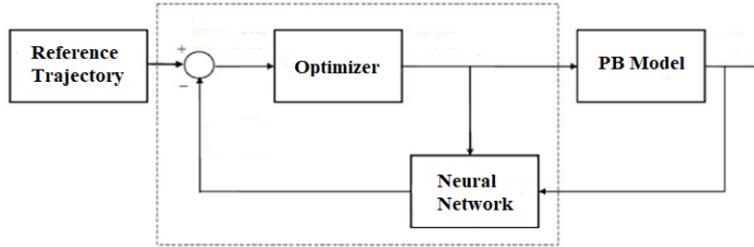


Figure 2: Control loop scheme.

The first reference trajectory tested was a constant trajectory, using the μ_i value at the end of the chosen experiment to define it. Then, a first-order dynamic reference trajectory was studied. This trajectory is described by Equations 5 and 6, which $\mu_i(t_{end})$ is the moment value at the end of the experiment, and γ is a constant equal to 0.9.

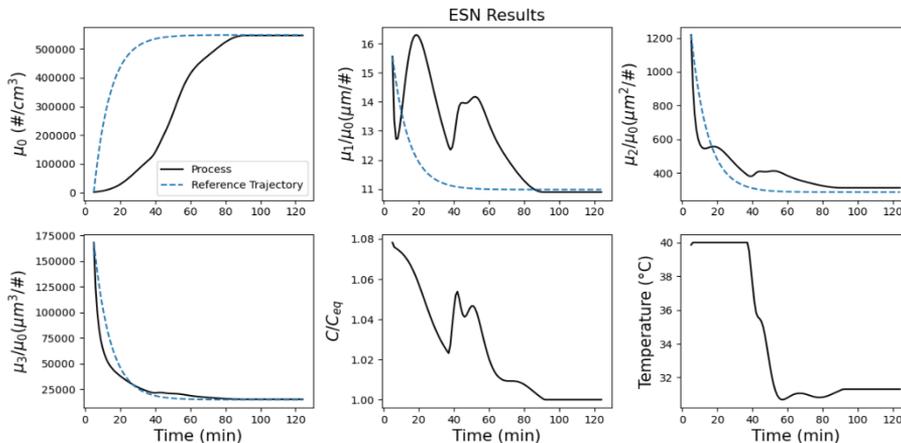
$$\mu_i^{RT}(t_0) = \mu_{i0} \tag{5}$$

$$\mu_i^{RT}(t) = \gamma\mu_i(t - 1) + (1 - \gamma)\mu_i(t_{end}) \tag{6}$$

A 2nd order reference trajectory was modeled as shown in Equation 7 for each μ_i . The parameter τ_i was calculated for each case, reaching the values 21.1, 20.4, 17.5, and 16.4 for μ_0 , μ_1 , μ_2 and μ_3 , respectively.

$$\mu_i = \mu_{i0} + (\mu_i(t_{end}) - \mu_{i0}) \left(1 - \left(1 + \frac{t - t_0}{\tau_i} \right) e^{-(t - t_0)/\tau_i} \right) \tag{7}$$

Finally, another trajectory was defined, which is similar to the 1st order trajectory but is named adaptive, as it departs from the actual system state at each time step.



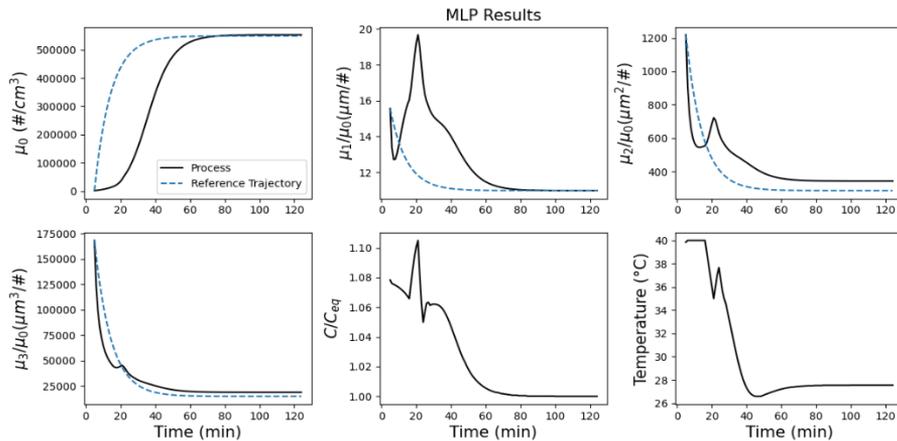


Figure 3: Simulation of the control loop for a 1st order trajectory.

After analyzing the controllers, both structures were efficient for a constant, a 1st order, and an adaptive 1st order reference trajectory, but had some issues for the 2nd order reference trajectory. The ESN had a better performance than the MLP because the calculated overall (considering all the trajectories and all the normalized variables) mean squared error (MSE) was 1.806 for ESN, while this value was 2.018 for the MLP. The results for both networks for a first-order reference trajectory are shown in Figure 3.

5. Conclusion and Prospects

In this work, an MPC controller was developed using a neural network approach to model the process. Four different neural networks designs were studied to predict the moments: a single MLP, a set of four MLP networks in series, LSTM and ESN. The dataset used for training and testing applied a co-teaching learning algorithm, using experimental and simulated data. As a result, the ESN presented the best performance to predict the moments, achieving values of R-squared higher than 75% for predictions five steps ahead.

The ESN was chosen to be the model of the controller, and its behavior was compared to a traditional approach, using MLP as the model. The controller based on the ESN produced better results than the one using the MLP, presenting a lower value of MSE. Four different reference trajectories were analyzed: a constant, a 1st order, a 2nd order, and an adaptive 1st order. Therefore, both structures were efficient for a constant, a 1st order, and an adaptive 1st order reference trajectory but had some issues for the 2nd order reference trajectory.

Finally, it is suggested to try different approaches of neural networks to model this process, such as Gated Recurrent Unit. Also, this same application could be tested on a continuous crystallization process.

References

1. Moraes, M. G. F., Grover, M. A., Souza Jr., M. B., Lage, P. L. C., Secchi, A. R., 2021. Optimal Control of Crystal Size and Shape in Batch Crystallization Using a Bivariate Population Balance Model. *IFAC PapersOnLine*, v. 54, Issue 3, p. 653-660.
2. Wu, Z., Luo, J., Rincon, D., Christofides, P. D., 2021. Machine learning-based predictive control using noisy data: evaluating performance and robustness via a large-scale process simulator. *Chemical Engineering Research and Design*, v. 168, p. 275-287.

Lithium-air battery electrocatalyst identification using Machine Learning and SciBERT word embeddings

Arun Muthukkumaran^a, Arjun Ravichandran^b, Sai Shanbhag^b, Ramprasad Arjun^c, Raghunathan Rengaswamy^{a*}

^a*Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai, 600036, India*

^b*Gyan Data Pvt. Ltd., Indian Institute of Technology Madras Research Park, Chennai, 600113, India*

^c*Department of Metallurgical and Materials Science, Indian Institute of Technology Madras, Chennai, 600036, India*

**raghur@iitm.ac.in*

Abstract

Materials design for functional applications carried out using conventional trial and error methodologies is expensive and time consuming. Machine learning techniques can be used to design materials suitable for functional applications and discover materials for specific applications from existing literature. This study demonstrates a machine learning approach where a combinatorial analysis was used to design perovskite structures and a novel methodology was devised to calculate descriptor values which were used for property prediction of the designed perovskites. Further, natural language processing (NLP) technique was used to obtain word embeddings of perovskite materials. The obtained embeddings were used to identify perovskites that could be used as electrocatalysts. The candidate materials predicted by the NLP algorithm have been investigated for their electrochemical properties and they could be tried as electrocatalysts for the reactions discussed in this study.

Keywords: Machine Learning, Natural Language Processing, perovskite, electrocatalyst

1. Introduction:

Lithium-ion batteries (LIBs) have played a huge role in enabling the digital revolution. The technology has established itself as a reliable energy storage mechanism over the past 20 years. Over the past decade they have become popular for vehicular applications. With increasing energy demand there is a need to look beyond the lithium-ion technology alone to meet the future requirements for energy storage.

Some of the other popular battery chemistries include, lithium-sulphur batteries, lithium-air batteries, zinc-air batteries, etc. Of specific interest is the lithium-air battery. This battery has much higher gravimetric energy storage density compared to all other chemistries (Girishkumar et al. 2010). Because of this it is being researched for use in vehicular applications where weight of the battery system is important. The usable energy density of gasoline for automotive applications is approximately 1700 Wh/kg (Richter et al. 2008). For the present lithium-ion battery technology to achieve this there needs to be a 7-fold increase in the energy density. This warrants a change of approach, which is to look at other battery chemistries. The oxidation of lithium provides 11,680 Wh/kg of energy which is quite close to that of gasoline. This potential of the lithium metal could be made use of by using the lithium-air battery. The working mechanism of the lithium-air battery involves an oxygen reduction reaction (ORR) in the cathode while discharging

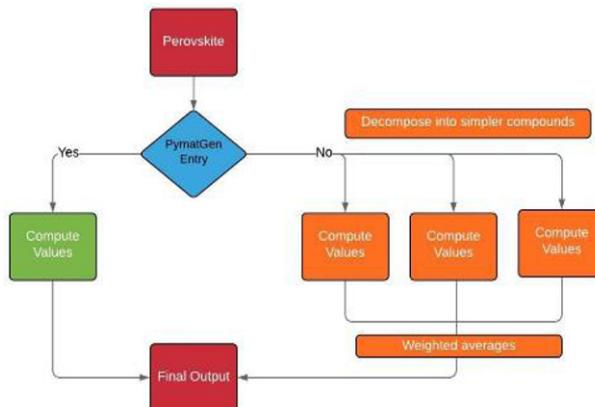


Fig. 2. Descriptor preparation methodology.

The descriptors calculated were then used as features for the machine learning model to predict catalyst activity values which are used to evaluate the suitability of the perovskite to be used as an electrocatalyst. Fig. 1 and Fig. 2 show the methodology for dataset preparation and descriptor calculation. The activity values are reported in different metrics. The details of the metrics are given in the following section.

2.1. Metric transformation and conversion:

Given the widespread scientific reporting of activity results in current as well as voltage units, and the absence of a universal metric to report OER/ORR performance, simple transformations were used to facilitate comparison of activities reported using different metrics. The conversions used for the different metrics are given below.

- Overpotential measurements: The potential values were considered in their native measurements in the case of ORR. For OER, the potential values were converted to negative scale after conversion to overpotentials.
- Current density measurements: For both OER and ORR, the current density values were converted to log scale.
- Experimental data reference compounds: All the activity values were taken as relative percentage changes with respect to LaCoO_3 and LaMnO_3 for OER and ORR respectively. The choice of selecting LaCoO_3 and LaMnO_3 is justified by the fact that they were the most widely reported perovskites for OER and ORR reactions.
- Activity data transformations: To account for the difference in spread resulting from difference in units, each group was normalized by its respective standard deviation.

2.2. Machine learning models and validation techniques:

Different machine learning models such as Ordinary least squares (OLS), Ridge regression (Ridge), Lasso regression (Lasso), Elastic net regression (11 ratio set as 0.25 [EN025], 0.5 [EN05] and 0.75 [EN075]), Random-forest regression (Random_Forest), Gradient boosting regression (gbr), Huber regression, Least angle regression (LARCV), Support vector regression (SVR) and Principal component regression (PCA_CV) were employed for modeling the data and obtaining predictions. Four different validation techniques were used and mean squared error (MSE) was the primary metric used to assess model performance. The best model found using validation was used to give predictions on different unknown A, B combinations of perovskite oxides. The validation techniques are described below.

- K-fold cross validation: 5-fold validation was used to evaluate the model performance on multiple hyperparameters.

- b. Blind A-blend: Model is trained on perovskite data discarding a certain A-element (e.g., La). Whereas, model performance is evaluated using that discarded unseen A-element. Averaging the model performance for different A-elements can be good metric for model validation.
- c. Blind B-blend: Model is trained on perovskite data discarding a certain B-element (e.g., Ni) and model is evaluated using that discarded unseen B-element.
- d. Blind Combination (AB): Combination (combo), here refers to a specific combination of A-site blend and B-site blend. The model was trained on perovskite data discarding a certain AB combination (e.g., SrCaMn) and model is evaluated using that discarded AB combination.

2.3. Materials selection using SciBERT embeddings:

SciBERT (Beltagy, Lo, and Cohan 2020), which is a language model developed based on BERT framework is pretrained on scientific literature across multiple domains. This model was developed to improve performance of BERT framework on downstream scientific NLP tasks. However, the SciBERT model is not trained in the domain of perovskites or electrocatalysts. To get over this hurdle, the model was retrained in these domains using abstracts from Springer and Elsevier journals. 102,260 abstracts were obtained by querying the keywords ‘electrocatalyst’ and ‘perovskite’ directly. The corpus prepared from the abstracts was pre-processed to remove Chinese texts and subsequently the model was trained. Embeddings of perovskites of type ABO_3 were obtained from the trained model. The obtained word embeddings were used to identify candidate electrocatalysts by obtaining the cosine similarities of the perovskites with the keyword ‘*electrocatalyst*’.

3. Results:

3.1. Machine learning model predictions:

	Title	Test_MSE	CV_MSE	Std	Y_test_var	Y_train_var	A_MSE	A_Std	B_MSE	B_Std	AB_MSE	AB_Std
0	Lasso	0.958322	0.4695	0.221096	1.478731	0.871388	2.377066	3.344627	5.844743	14.460464	1.329654	2.520868
1	Ridge	0.761827	0.427494	0.221602	1.478731	0.871388	2.331091	3.414514	5.239869	12.923104	1.277016	2.561755
2	EN025	0.732273	0.414262	0.215897	1.478731	0.871388	2.383850	3.421740	3.775924	8.830914	1.298490	2.582024
3	EN050	0.724438	0.410489	0.215859	1.478731	0.871388	2.420791	3.394287	4.738948	11.716463	1.325761	2.567973
4	EN075	0.637202	0.40411	0.216404	1.478731	0.871388	2.402587	3.372402	5.362065	13.498782	1.327730	2.545984
5	Lars	1.016232	0.432297	0.223419	1.478731	0.871388	1.132952	1.376610	6.380603	17.161459	0.770053	1.077042
6	OLS	0.761827	0.42561	0.222703	1.478731	0.871388	2.346757	3.366826	5.486369	13.563721	1.293076	2.529997
7	PCA_CV	0.761827	0.420595	0.222033	1.478731	0.871388	2.350896	3.362774	3.417171	7.239046	1.275236	2.534316
8	Forward_CV	0.625114	0.41256	0.22299	1.478731	0.871388	2.346757	3.366826	5.486369	13.563721	1.293076	2.529997
9	Backward_CV	0.780785	0.40019	0.218933	1.478731	0.871388	2.346757	3.366826	5.486369	13.563721	1.293076	2.529997
10	Huber_regression	0.672265	0.39727	0.21784	1.478731	0.871388	1.345892	1.975930	6.259976	16.251491	0.778015	1.389715
11	Random_Forest	0.585536	0.389611	0.213907	1.478731	0.871388	0.834821	0.872661	0.949613	6.604615	0.698617	0.768084
12	svm	0.638256	0.381731	0.212913	1.478731	0.871388	0.823278	0.817979	0.910278	6.641432	0.687174	0.767762
13	gbr	0.608289	0.377698	0.212765	1.478731	0.871388	0.756461	0.794638	0.840252	6.486634	0.640579	0.658752

Fig. 3. Training and validation results for OER dataset.

Fig. 3 shows the results obtained from training and validating OER dataset using various machine learning models. The train test split was chosen as 80:20. Test_MSE is test data mean squared error, CV_MSE is 5-fold cross validation mean squared error. A_MSE, B_MSE, AB_MSE are blind A, blind B and blind AB validation mean squared errors respectively. From the results it can be seen that the Gradient boosting regression has outperformed all the other models. gbr is an ensemble learning method. In gbr, while the decision trees are built, each new tree corrects the errors made by the previously trained

trees. This feature enhances the accuracy of the model. The validation of activities predicted using gbr is shown in Fig. 4.

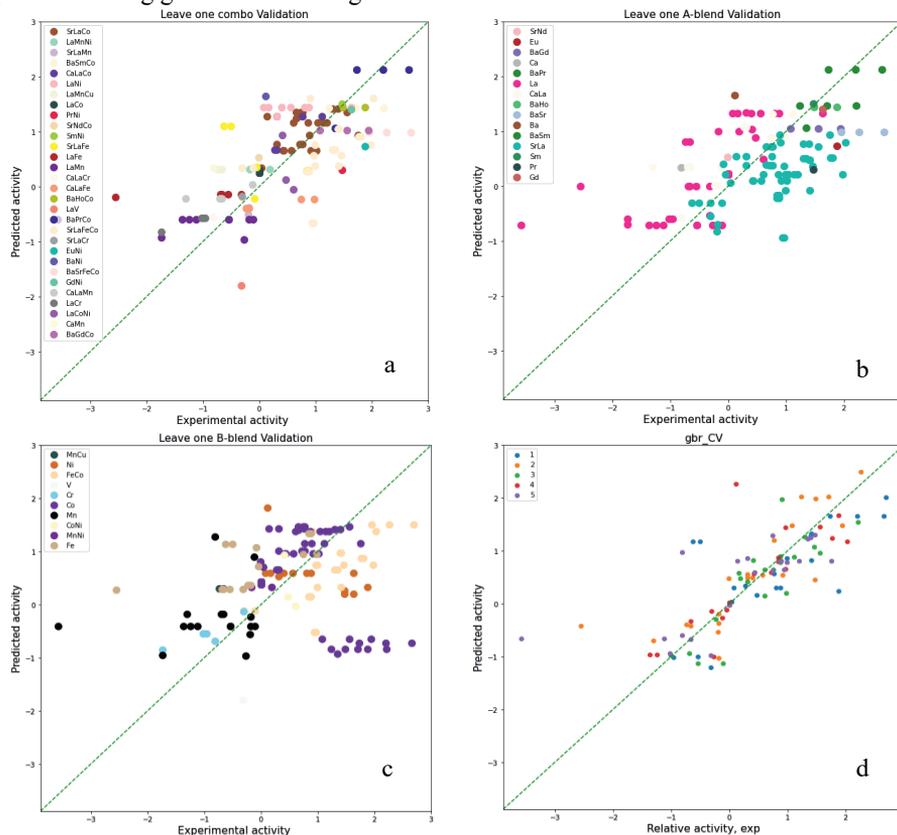


Fig. 4. Validation for OER activities predicted using gbr model, a. Validation using blind combination, b. Validation using blind A-blend, c. Validation using blind B-blend, d. Validation using 5-fold cross validation.

Scatter plot of blind A-blend validation has lesser MSE compared to blind B-blend validation. This suggests that blinding A element did not affect the model performance whereas blinding B element did. According to the results obtained from the study it can be said that B elements are more important in finding the OER activity of a perovskite. If we blind a random B element and train the model the model fails to predict the activity accurately (for most of the perovskites considered in the plot) for the B element which is blinded. This demonstrates the importance of B element. Table 1 shows the top 4 perovskites ranked based on scaled activities. These perovskites can be seen as candidates for non-precious metal based electrocatalysts. Similar analysis was carried out with ORR dataset and the candidate materials obtained are shown in Table 1.

3.2. SciBERT embedding predictions:

The SciBERT language model, unlike other language models (Skip-gram, CBOW, etc.) is a context-dependent model and is trained on scientific text. The word embeddings are computed taking into consideration all the words in a sentence along with their position in the sentence hence are better than the embeddings obtained from other models.

The top 8 perovskites predicted by computing cosine similarity with the keyword ‘electrocatalyst’ are shown in Table 2. Cosine similarity measures the cosine of the angle between two vectors projected in a multi-dimensional space. Vectors oriented in similar direction will have a high cosine similarity. A literature review of the predicted materials revealed that some of these materials have already been studied for their electrochemical properties (e.g., LaCoO_3 , SrCoO_3 , $\text{SrCo}_{0.7}\text{Ni}_{0.3}\text{O}_3$, etc.). One of the most popular electrocatalyst for OER and ORR is platinum. It was observed that the cosine similarities of ‘Pt’ with ‘*electrocatalyst*’ and ‘*LaCoO₃*’ with ‘*electrocatalyst*’ were alike (0.095 for ‘Pt’ and 0.102 for ‘*LaCoO₃*’). These prove that the predictions are consistent with existing scientific knowledge and are capable of possibly identifying newer materials.

Table 1. Top 4 perovskite-based electrocatalysts for OER and ORR predicted from ML analysis.

OER		ORR	
Perovskite	Scaled activity	Perovskite	Scaled activity
$\text{Sr}_{0.5}\text{Ba}_{0.5}\text{Fe}_{0.5}\text{Co}_{0.5}\text{O}_3$	2.113	$\text{Sr}_{0.5}\text{Sm}_{0.5}\text{Cu}_{0.5}\text{Ni}_{0.5}\text{O}_3$	0.200
$\text{Sr}_{0.5}\text{Ba}_{0.5}\text{Fe}_{0.5}\text{Mn}_{0.5}\text{O}_3$	2.105	$\text{Ca}_{0.5}\text{Sm}_{0.5}\text{Cu}_{0.5}\text{Ni}_{0.5}\text{O}_3$	0.197
$\text{Ce}_{0.5}\text{Ag}_{0.5}\text{MnO}_3$	2.091	$\text{Ca}_{0.5}\text{La}_{0.5}\text{Cu}_{0.5}\text{Ni}_{0.5}\text{O}_3$	0.197
$\text{SrFe}_{0.5}\text{Mn}_{0.5}\text{O}_3$	2.085	$\text{Sr}_{0.5}\text{Y}_{0.5}\text{Cu}_{0.5}\text{Ni}_{0.5}\text{O}_3$	0.192

Table 2. Top 8 perovskite-based electrocatalysts predicted using word embeddings

Rank	Perovskite	Rank	Perovskite
1	LaCoO_3	5	$\text{Nd}_{0.5}\text{Ba}_{0.5}\text{CoO}_3$
2	$\text{Ca}_{0.1}\text{Ba}_{0.9}\text{SnO}_3$	6	$\text{Ce}_{0.8}\text{Sr}_{0.2}\text{Co}_{0.6}\text{Ni}_{0.2}\text{Fe}_{0.2}\text{O}_3$
3	$\text{SrCo}_{0.7}\text{Ni}_{0.3}\text{O}_3$	7	$\text{Sm}_{0.8}\text{Sr}_{0.2}\text{Fe}_{0.8}\text{Co}_{0.2}\text{O}_3$
4	AuWO_3	8	SrCoO_3

4. Conclusion:

This study demonstrated a ML based approach for materials design and an NLP based approach for materials discovery. Through the ML based methodology, perovskites of a certain structure were generated and their electrochemical activities were predicted. The designed perovskites could be seen as novel materials that could be synthesized. The NLP results clearly demonstrate the usefulness of the methodology for materials discovery. This technique shows how existing knowledge can be analyzed in a better way. Since the volume of scientific literature is growing by the day, this approach can be seen as a paradigm shift in the way literature has been analyzed and assimilated.

5. Future work:

Following the extraction of the embeddings for the perovskites, they will be subsequently used as features for training the ML model along with the already present descriptors. Research is in progress to develop newer descriptors and metrics.

References:

- Beltagy, Iz, Kyle Lo, and Arman Cohan. 2020. “SCIBERT: A Pretrained Language Model for Scientific Text.” *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*: 3615–20.
- Girishkumar, G. et al. 2010. “Lithium-Air Battery: Promise and Challenges.” *Journal of Physical Chemistry Letters* 1(14): 2193–2203.
- Richter, Burton et al. 2008. “Energy Future: Think Efficiency.” *American Physical Society*: 1–109. <http://www.aps.org/energyefficiencyreport>.
- Rong, Ziqin et al. 2016. “An Efficient Algorithm for Finding the Minimum Energy Path for Cation Migration in Ionic Materials.” *Journal of Chemical Physics* 145(7). <http://dx.doi.org/10.1063/1.4960790>.
- Wang, Hao Fan et al. 2020. “MOF-Derived Electrocatalysts for Oxygen Reduction, Oxygen Evolution and Hydrogen Evolution Reactions.” *Chemical Society Reviews* 49(5): 1414–48.

A digital twin-concept for smart process equipment assemblies supporting process validation in modular plants

Jonathan Mädler,^a Julian Rahm,^a Isabell Viedt,^a Leon Urbas^a

^a*Technische Universität Dresden, 01062 Dresden*
jonathan.maedler@tu-dresden.de

Abstract

Modular plants provide the opportunity to increase flexibility and reduce time-to-market in the process industry. Process validation is time-consuming and limits the realization of this potential. Therefore, it should be supported efficiently. Smart process equipment assemblies (sPEAs) built from the real module, a digital twin (DT), and suitable methods and algorithms provide high potential to do so. In this paper, we investigate how a suitable DT should look like to support the process design phase in process validation. Therefore, the semantics and information demand of different relevant simulation and optimization problems are analyzed. We reason that the DT should combine structural information like engineering data, e.g. through DEXPI, and behavioral models in form of simulation models. Furthermore, it should provide the descriptive capability to capture the semantics of different application cases like e.g. design of experiments. We suggest a linked-data-based architecture to meet these requirements. The simulation models are semantically lifted into linked data through an information model describing its purpose, quality and variables. The approach provides the potential to reduce manual effort of the user since information is interconnected, accessible and processible automatically.

Keywords: Modular plants, smart PEA, digital twin, linked data, process validation

1. Introduction

Modular plants (MPs) built from process equipment assemblies (PEAs) strive to reduce time-to-market and increase production flexibility in the process industry. Therefore, this technology is attractive for industry sectors with small product charges and short product life cycles like the pharma industry. This industry is highly regulated and requires a rigid process validation (Katz and Campbell, 2012). The MP concept accelerates construction and automation of process plants. To develop MPs into highly flexible and smart production systems, process validation should be efficiently supported as well.

Process design is the first stage in process validation (Katz and Campbell, 2012). In recent years, the quality by design (QbD) approach is adopted in the pharma industry to set a sound scientific foundation in this stage. The approach emphasizes knowledge management, risk-based approaches, and design of experiments. We are convinced, that smart PEAs built from the real PEA, its digital twin (DT) and tailor-made methods and algorithms can reduce manual efforts of the owner/operators (O/Os) and therefore accelerate process validation. Thus, this paper aims to answer the question how a DT of a sPEA should be structured to support the process design phase in process validation and the solution of related simulation and optimization problems as efficiently as possible. Simulation and optimization methods (e.g. black-box optimization methods) for systems

with distributed knowledge are not considered in this publication. The remainder of the paper is structured as follows: section 2 introduces an illustrative example; section 3 reviews the DT and key simulation and optimization problems to derive requirements to the DT for process validation; section 4 presents a concept for a linked-data-based DT of sPEAs to integrate structural and behavioral information for process validation; section 5 provides a conclusion and further research potential.

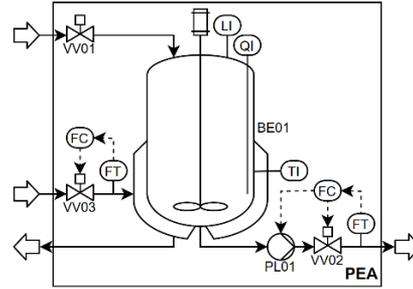


Figure 1: P&ID of a simple STR-PEA

2. Illustrative example

Let us assume that an O/O aims to carry out an irreversible, exothermic reaction $A \rightarrow B$. The quality target product profile (QTPP) requires the concentration of component A c_A to sink below 1 mol/L in the product stream F , while the process must not reach a temperature T above 80 °C in the vessel. The educt stream is expected to vary in volume flow F_{in} , concentration $c_{A,in}$ and temperature T_{in} in the production. The O/O chooses the stirred tank reactor (STR)-PEA in figure 1 to carry out the process. The PEA can control the product volume flow (VV02, PL01) and the cooling water flow (VV03). Furthermore, it provides a service to open the valve VV01 to receive any process stream provided to it. The PEA has sensors to measure the temperature T , the level in the tank L , the concentration c_A , and the flows F_{in} and F . Assuming constant physical property parameters, the following model equation system can be derived:

$$\begin{aligned} \frac{dV}{dt} &= F_{in} - F; \quad L = \frac{4V}{\pi D^2}; & \frac{dT}{dt} &= \frac{F_{in}}{V}(T_{in} - T) - \frac{\lambda}{\rho c_p} k c_A - \frac{UA_J}{\rho V c_p}(T - T_J) \quad (1) \\ \frac{dc_A}{dt} &= \frac{F_{in}}{V}(c_{A,in} - c_A) - k_0 \exp\left(-\frac{E_A}{RT}\right) c_A; & \frac{dT_J}{dt} &= \frac{F_J}{V_J}(T_{J,in} - T_J) + \frac{UA_J}{\rho_J V_J c_{p,J}}(T - T_J) \end{aligned}$$

We assume that the O/O already knows about the density ρ and specific heat capacity c_p and did choose a suitable kinetic model. Yet, the O/O has only rough estimations of the kinetic parameter values k_0 and E_A as well as the reaction enthalpy λ . The sPEA now shall provide a DT, and methods and algorithms to support the O/O to identify the model more precise and derive the design space (DS) of the process.

3. Requirements analysis

3.1. A digital twin for smart PEAs

Smart PEAs are envisioned to support their user in application beyond the capabilities of current PEAs. Therefore, it is necessary to extend their descriptive capabilities. The DT concept provides a far superior form of description to sPEAs. Currently, descriptions to PEAs focus on single aspects. For example, the module type package (MTP) describes the automation interface of a PEA (VDI/VDE/NAMUR, 2019), or DEXPI which describes piping and instrumentation (Wiedau et al., 2019). Boschert et al. (2018) define the DT as a *'linked collection of [...] engineering data, operation data and behavior descriptions via several simulation models'*. These simulation models are use case specific and therefore of varying, suitable fidelity (Boschert et al., 2018). Networking structural models from the planning phase, such as the MTP and DEXPI, was considered in preliminary work (Rahm et al., 2021). For the DT of our illustrative example, the MTP would e.g. describe the services controlling the control loops and the input valve. To

support process validation, the DT of a sPEA must additionally integrate behavioral models to capture process behavior. Furthermore, it must integrate with simulation and optimization methods, which accelerate process validation through an increased degree of automation and reduced manual efforts. In addition, an integration of structural models describing process validation data like e.g. the QTPP would provide an added value.

3.2. *Quality by Design methods for smart PEAs*

Design space identification is an important task in process validation. Von Stosch et al. (2020) argue that a dynamic DS can provide increased process flexibility. Thus, a dynamic model should be used to capture the process behavior. Since, smart PEAs will be reused in different application scenarios, they can be expected to become pre-characterized. Therefore, model-based design of experiments (MBDoE) and system identification approaches should be foreseen to identify the dynamic model. Finally, Ochoa et al. (2021) did consider flexibility analysis (FA) for DS identification from static simulation models. This approach can be extended applying the dynamic flexibility analysis developed by Dimitriadis and Pistikopoulos (1995).

3.2.1. *Dynamic modeling for behavioral models*

As derived above, the DT of a sPEA should provide dynamic behavioral models of the PEA and the process happening within the PEA. Differential algebraic equation systems (DAEs) of the following form provide a well-known form of dynamic model:

$$\begin{aligned} \mathbf{0} &= \mathbf{f}(t, \dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}); \quad \mathbf{x}(t_0) = \mathbf{x}_0 \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}(t), \boldsymbol{\theta}) \end{aligned} \quad (2)$$

Herein, t represents the time, \mathbf{x} are the differential and algebraic states, \mathbf{x}_0 are the initial states of the DAE-system, $\dot{\mathbf{x}}$ denote the derivatives of the differential states, \mathbf{u} are external inputs, \mathbf{y} are the outputs, and $\boldsymbol{\theta}$ are equipment or physical property data related parameters. These variables are defined in the mathematical domain. In our illustrative example e.g. the external inputs \mathbf{u} have to be mapped to the input mass flow F_{in} , T_{in} , $c_{A,in}$ and the input information flows F and F_j . Thus, PSE software defines a semantic layer upon the mathematic model, which allows to differentiate between these input types. Meanwhile, the semantic meaning of a variable can often be related to information in other structural models. E.g., equipment properties like the volume of the jacket V_j are described in DEXPI. In a DT, potentially conflicting descriptions of the same quantity should be prevented. Therefore, the behavioral model should be integrated into the structural model. But since the semantic is usually proprietary to a software, a digital connection between the behavioral model and the structural model is often not established yet. Our new DT concept must overcome this obstacle. Furthermore, the DT must provide meta information on the purpose and quality of a simulation model and describe the relationship between different simulation models which are part of the DT.

3.2.2. *Simulation and optimization problems applying the digital twin*

In the following, the optimization problems **MBDoE** and **FA** are analyzed as examples. **MBDoE** is an optimal control problem. It is applied to design highly informative experiments. An example are D-optimal experiments. The related optimization problem can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{u}(t), \mathbf{x}_0} & \left[\det \left(\left[\sum_{iExp=1}^{NExp} \left[\sum_{iy=1}^{Ny} \sum_{jy=1}^{Ny} \sigma_{iy,jy} \mathbf{S}_{iy}^T \mathbf{S}_{jy} \right]_{iExp} + \mathbf{FIM}^0 \right]^{-1} \right) \right] \\ s.t. & \quad \mathbf{0} = \mathbf{f}(t, \dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}_{CP}, \boldsymbol{\theta}_{TP}); \quad \mathbf{x}(t_0) = \mathbf{x}_0; \\ & \quad \mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t), \boldsymbol{\theta}_{CP}, \boldsymbol{\theta}_{TP}); \quad \mathbf{u}_L(t) \leq \mathbf{u}(t) \leq \mathbf{u}_U(t); \quad \mathbf{x}_L(t) \leq \mathbf{x}(t) \leq \mathbf{x}_U(t) \end{aligned} \quad (3)$$

Herein, **FIM** denotes the Fisher information matrix, θ_{CP} the capability property parameters (cf. Bamberg et al., 2021), θ_{TP} the transformation property parameters, σ are the variances of the different measurement signals and **S** denotes the parameter sensitivities of the measurements. Firstly, it is important to recognize the ‘factors’ of the experiments are formed by the information flow and mass flow related inputs **u** and the initial state \mathbf{x}_0 . In our illustrative example, these are the information flows F, F_j along with the mass flow related variables $F_{in}, c_{A,in}, T_{in}$ and the initial states $V_R, c_{A,0}, T_0$ which might be adjustable under experimentation conditions. The parameter vector θ must be split into capability property parameter θ_{CP} , which characterize the sPEA capabilities and are known beforehand, and the transformation property parameters θ_{TP} , which are related to varying substance systems processed by the PEA. In the illustrative example, θ_{CP} the variables V_j, UA_j, ρ_j and $c_{p,j}$ are capability parameters, while ρ, c_p, k_0, E_A and λ are transformation parameters. The parameter sensitivities **S** are not provided by the model according to equation (2). For this purpose, numerical differentiation might be applied or a related model including the sensitivity equations might be included in the DT. The simulation model must be considered for nonlinear equality constraints. Furthermore, inequality constraints might be provided through information from a structural model like e.g. a maximal operation temperature of an equipment.

FA can be used to calculate the design space (Dimitriadis and Pistokopoulos, 1995):

$$\chi = \max_{\mathbf{cma}(t)} \left[\min_{\mathbf{cpp}(t)} \left[\max_{ih \in \{1, N_h\}} \left[h_{ih}(\mathbf{x}(t), \mathbf{cqa}(t), \mathbf{cma}(t), \mathbf{cpp}(t), \theta) \right] \right] \right] \quad (4)$$

s.t. $\mathbf{0} = \mathbf{f}(t, \dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{cma}(t), \mathbf{cpp}(t), \theta); \quad \mathbf{x}(t_0) = \mathbf{x}_0;$

$$\mathbf{cqa}(t) = \mathbf{g}(\mathbf{x}(t), \theta); \quad \mathbf{cma}_L(t) \leq \mathbf{cma}(t) \leq \mathbf{cma}_U(t); \quad \mathbf{cpp}_L(t) \leq \mathbf{cpp}(t) \leq \mathbf{cpp}_U(t)$$

In process validation, critical material attributes (CMAs), critical process parameters (CPPs) and critical quality attributes are differentiated (CQAs). FA can be used to check if the CQAs can be met under varying CMAs adjusting the CPPs. CMAs and CPPs are both related to the input variables **u**. Thus, the optimization problem requires an adjusted semantic. In our illustrative example, CMAs could be related to the variables describing the feed mass flow F_{in}, c_A, T_{in} , which are dictated by a prior process step in the MP, while potential CPPs are F_j and F . In addition, a connection to structural models describing process validation quantities like requirements to the CQAs could provide added value.

In summary, (1) the DT of a sPEA should be built from integrated structural and behavioral models to provide networked information spaces and hence make information accessible more easily or even automatically. Furthermore, (2) the DT should provide the capability to describe and connect the different semantic contexts to make it understandable. In addition, (3) the DT should be extendable easily, allow the integration of new information models, and be adjustable to different PEA configurations easily.

4. A linked-data-based digital twin concept for smart PEAs

Over the life cycle of a PEA, different domain-specific models emerge that have specific views of a PEA (Wiedau et al., 2019). The systematic networking of models is an essential prerequisite for a smart PEA. Existing information relationships can thus be used directly. In addition, the networking is a basic prerequisite for the synchronization of the models, so that all models are always free of contradictions regarding defined rules (Rahm et al., 2021). Graube et al. (2012) developed a concept using linked data that meet the requirements for a shared and distributed information space in industry. The basis are graph-based structures that are described using the Resources Description Framework (RDF) (Graube et al., 2012). The SPARQL query language can be used to retrieve the

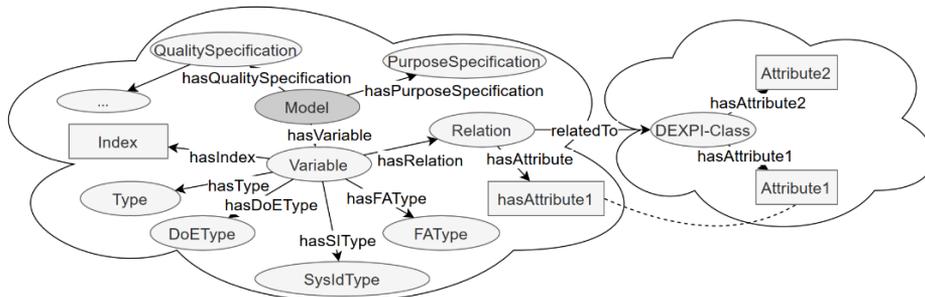


Figure 2: Information model for simulation models in linked data information. Behavioral models like simulation models must be semantically lifted into the graph-based structure to unlock the descriptive power of linked data for the DT. We suggest defining an additional structural model in linked data which describes meta information of simulation models (see figure 2). This structural model relates simulation models to variables, a purpose specification, and a quality specification (cf. Mädler et al., 2021). This paper does focus on the variables. The most important attributes of a variable are the type and the index. These are used to relate the variable to the characterizing vectors of a simulation model according to equation (2). The type might therefore be ‘input’, ‘output’, ‘state’, ‘initial state’, or ‘parameter’. The index attribute provides the index of a variable in the vector. The further attributes provide information on the semantic meaning of the variable e.g. in the DoE context. Therefore, the DoEType might be ‘dynamic factor’, ‘constant factor’, ‘measurement’ or ‘none’. These types do not force the user to use a variable as factor, but rather indicate its semantic meaning. In addition, the variable can have relationships into other information models like DEXPI. In this case, an additional attribute to a variable must indicate the relation, point to the node in the foreign information model and name the related attribute. In this way, the semantic lifting of simulation models into linked data permits the connection of the considered quantities in a simulation model with several other information spaces (see figure 3). Firstly, e.g. variables representing quantities related to a sensor might point to the particular sensor in DEXPI. Secondly, the same variables could become directly or indirectly connected to digital representations of equipment data sheets. Thirdly, information models describing the knowledge space of process validation including process design, process qualification and continued process verification might be integrated. Finally, meta information models for experiments in linked data can be envisioned. This would allow to relate time series data to digital data sheets of the sensors and to the outputs of the simulation model. The combination of a linked data information space for structural models, a repository for behavioral models and a time series database for experimental data does provide a well-suited containment for each type of data and its semantic interconnection. SPARQL queries can be used to set up simulation and

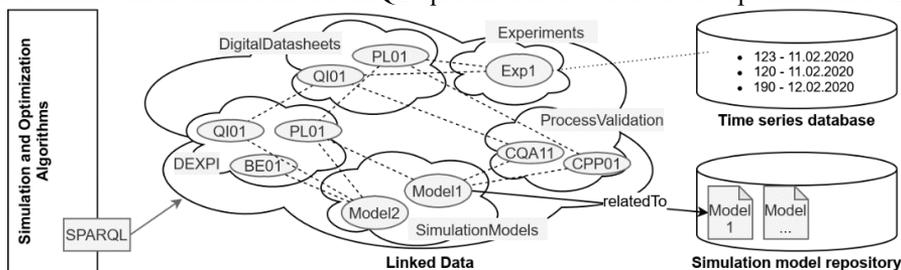


Figure 3: Information architecture for the digital twin of smart PEAs

optimization problems by asking the DT for simulation models of suitable purpose specification and quality specification. Furthermore, the information model provides knowledge on the semantic structure of the simulation model.

5. Conclusion

This paper aimed to answer the question how a DT for the support of process design in process validation in sPEAs should be constructed. It was reasoned that the DT should include and interconnect structural and behavioral models to provide information for different simulation and optimization tasks. We suggested a linked-data-based architecture to build a digital twin integrating structural models like engineering data, behavioral models like simulation models and time series data. The simulation models are semantically lifted into linked data using a descriptive information model. This reduces manual efforts for the O/O since information becomes digitally interconnected, accessible, and automatically processible. In future work, this approach must be extended to include the purpose specification and quality specification of simulation models. Furthermore, the relationship between different simulation models for the same PEA should be added to the information model. Finally, the information models related to digital datasheet, experimental data and process validation must be integrated.

Acknowledgement

The authors thank the P2O-Lab of the TU Dresden for the discussions and application examples.

References

- A. Bamberg, L. Urbas, S. Bröcker, M. Bortz, and N. Kockmann, 2021, The Digital Twin—Your Ingenious Companion for Process Engineering and Smart Production, *Chem Eng Technol*, 44, 6, 954–961
- S. Boschert, C. Heinrich, R. Rosen, 2018, Next generation digital twin. In *Proc. tmce (Vol. 2018, pp. 7-11)*, Las Palmas de Gran Canaria, Spain
- V. D. Dimitriadis, E. N. Pistikopoulos, 1995, Flexibility analysis of dynamic systems, *Ind. Eng. Chem. Res.*, 34, 12, 4451-4462
- M. Graube, J. Pfeffer, J. Ziegler and L. Urbas, 2012, Linked Data as integrating technology for industrial data, *Int. J. Distrib. Syst. Technol*, 3, 3, 40–52
- P. Katz and C. Campbell, 2012, FDA 2011 process validation guidance: Process validation revisited, *J. GXP compliance*, 16, 4, 18
- J. Mädler, I. Viedt, L. Urbas, L., 2021, Applying quality assurance concepts from software development to simulation model assessment in smart equipment, In *Computer Aided Chemical Engineering (Vol. 50, pp. 813-818)*. Elsevier
- M.-P. Ochoa, S. García-Muñoz, S. Stamatidis, and I. E. Grossmann, 2021, Novel flexibility index formulations for the selection of the operating range within a design space, *Comput Chem Eng*, 149, 107284
- J. Rahm, M. Theißen, A. Klose, M. Soemers, H. Temmen, C. Schäfer, and L. Urbas, 2021, Efficient Automation Engineering of Modular Process Equipment Assemblies Using the Digital Twin, *Chem Ing Tech*, 93, 2081-2091
- VDI/VDE/NAMUR, 2019, VDI/VDE/NAMUR 2658-Automation engineering of modular systems in the process industry-General concept and interfaces – Part 1, Beuth Verlag GmbH.
- M. von Stosch, R. Schenkendorf, G. Geldhof, C. Varsakelis, M. Mariti, S. Dessoy, A. Vandercammen, A. Pysik, and M. Sanders, 2020, Working within the Design Space: Do Our Static Process Characterization Methods Suffice? *Pharmaceutics*, 12, 6, 562
- M. Wiedau, L. von Wedel, H. Temmen, R. Welke and N. Papakonstantinou, 2019, ENPRO Data Integration: Extending DEXPI Towards the Asset Lifecycle, *Chem Ing Tech*, 91, 240-255

Bayesian Optimization for techno-economic analysis of pressure swing adsorption processes

Leif Erik Andersson^{a,*}, Johannes Schilling^b, Luca Riboldi^a, André Bardow^{b,c} and Rahul Anantharaman^a

^a*SINTEF Energy Research, Sem Sælands vei 11, 7034 Trondheim, Norway*

^b*Energy and Process Systems Engineering, ETH Zurich, 8092 Zurich, Switzerland*

^c*Institute of Energy and Climate Research, Energy Systems Engineering (IEK-10), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*
leif.andersson@sintef.no

Abstract

Pressure swing adsorption (PSA) can remove CO₂ from flue gases. The full potential of the technology can only be exploited if the optimal combination of adsorbent material and process is identified. This identification requires screening the large database of adsorbent materials by performing computationally intensive process optimizations. To ensure a suitable compromise between accuracy and computational lightness, the PSA process can be described by reduced-order models. However, those models might involve several discrete states making the objective function discontinuous and not continuously differentiable, challenging gradient-based methods. The selection of suitable optimization methods is therefore an open issue. This study compares three optimization algorithms, Bayesian optimization, NOMAD and KNITRO, for two cases and adsorbent materials.

For the tested cases and adsorbent materials, none of the three algorithms is clearly superior to the other. Bayesian optimization (BO) needs fewest function evaluations and computational time to converge and outperforms the other two for one case. However, BO is less reliable than NOMAD and KNITRO for the other case tested.

Keywords: pressure swing adsorption, CO₂ capture, techno-economic optimization, Bayesian optimization

1. Introduction

Rigorous pressure swing adsorption (PSA) models consist of a set of partial differential and algebraic equations and are thus computationally intensive to solve. In addition, the performance of a PSA cycle must be assessed at cyclic steady state, which requires to simulate several cycles. Rigorous PSA models are discontinuous and are not continuously differentiable. An overview of the related challenges, with a focus on efficient simulation and optimization strategies for the design of PSA systems, is provided by Biegler et al. (2004).

A key for the design of an efficient PSA process is the choice of the adsorbent (Farmahini et al., 2021). To maximize the economic efficiency of a PSA process, adsorbents have to be assessed using a thermo-economic process metric. However, the impact of the adsorbent properties on the economic performance is challenging to predict (Leperi et al., 2019). To address this challenge, a techno-economic framework is developed to screen thousands of adsorbent structures for PSA processes and rank them for a given set of process metrics (Riboldi et al., 2020). For efficient screening, fast and robust process optimization is essential. In this study, a reduced-order

PSA model is used (Riboldi et al., 2020) and linked to a Bayesian Optimization algorithm (BO). BO uses Gaussian Process (GP) regression to identify a surrogate model of the original objective function and optimizes the surrogate to find the next evaluation point in each iteration. GP is a probabilistic, non-parametric modelling technique well known in the machine learning community (Rasmussen and Williams, 2006). The advantage of using GP regression is that it is not bounded by specific model structures as e.g. parametric models. Consequently, it can identify the surrogate model in a flexible manner (de Avila Ferreira et al., 2018). We compare the BO algorithm to two optimization algorithms from the literature, namely an algorithm implemented in the software library KNITRO (Byrd et al., 2006) and the blackbox optimisation software NOMAD (Abramson et al., 2021), and discuss their performance in terms of objective function value, computational time and how reliable they converge to the same optimum.

2. Optimization algorithms

2.1. Bayesian Optimization

BO minimizes a function $f(\mathbf{x})$. For this purpose, a probabilistic model for the function $f(\mathbf{x})$ is constructed using Gaussian process regression, which is used to decide where to evaluate the function next. Thus, BO is a sequential approach in which first a regression model is trained as surrogate model, which is then minimized instead of the original function $f(\mathbf{x})$. The optimum of the surrogate model is evaluated using the original function $f(\mathbf{x})$, the surrogate model updated, and the optimization repeated. BO is especially powerful when function evaluations are computationally expensive (Snoek et al., 2012) as it is the case for the investigated PSA model.

Gaussian Processes (GP) are non-parametric, probabilistic kernel methods (Rasmussen and Williams, 2006) and aim to identify a function f , which describes a data set. The noisy observations of $f(\mathbf{x})$ are given by

$$y = f(\mathbf{x}) + v, \quad (1)$$

where the noise v is Gaussian with zero mean and variance σ_m and x is the input vector. Smoothness properties of the underlying function $f(\mathbf{x})$ are enforced by choice of mean and covariance function without relying on parametric assumptions (Snelson and Ghahramani, 2006). In this article, a zero-mean function and the automatic relevance squared-exponential covariance are chosen. The performance of the GP depends on the hyperparameters Φ , which are commonly unknown and need to be inferred from data. The hyperparameters are the noise variance, signal variance and the values in the scaling matrix of the covariance function. Here, the marginal likelihood is used to estimate the hyperparameters.

An important choice for BO is the formulation of the acquisition function. The acquisition function is used to determine the next point to evaluate. In this work, we test two acquisition functions: First, the GP lower confidence bound (GP-LCB) is chosen

$$\mathbf{x}_{i+1} = \arg \min_{\mathbf{x}} (\mu(x; \{\mathbf{x}_i, y_i\}, \Phi) - \kappa \sigma(x; \{\mathbf{x}_i, y_i\}, \Phi)), \quad (2)$$

where μ and σ mean and variance of the GP regression model. The observations are in the form $\{\mathbf{x}_i, y_i\}_{n=1}^N$ and κ is a parameter balancing exploitation against exploration (Snoek et al., 2012). The second acquisition function tested is maximizing the expected improvement (GP-EI) over the current best.

$$\mathbf{x}_{i+1} = \arg \max_{\mathbf{x}} (\sigma(x; \{\mathbf{x}_i, y_i\}, \Phi)(\gamma(\mathbf{x})\Psi(\gamma(\mathbf{x})) + \mathcal{N}(\gamma(\mathbf{x}); 0, 1))), \quad (3)$$

where Ψ is the cumulative distribution function of the standard normal and $\gamma(\mathbf{x})$ is defined as:

$$\gamma(\mathbf{x}) = \frac{f(\mathbf{x}_{best}) - \mu(x; \{\mathbf{x}_i, y_i\}, \Phi)}{\sigma(x; \{\mathbf{x}_i, y_i\}, \Phi)}. \quad (4)$$

2.2. NOMAD

NOMAD is a blackbox optimization tool suited for difficult problems with a small number of variables. NOMAD uses Mesh-Adaptive-Direct-Search to perform the optimization, which is a gradient-free optimization method (Le Digabel, 2011). In this work, NOMAD version 4 is used, which is available from the website (Abramson et al., 2021).

2.3. KNITRO

KNITRO is a commercial optimization software library for solving nonlinear mathematical optimization problems (Byrd et al., 2006). The library contains several local optimization solvers and algorithms for continuous and discrete problems with and without constraints. KNITRO allows for gradient-based or derivative-free and black-box optimization. In this work, we use KNITRO version 12.3. For a description of the used settings, please see Section 4.1.

3. PSA model

The PSA model builds on an approach presented in the literature (Maring and Webley, 2013), which assumes local equilibrium, i.e., CO₂ gets instantaneously adsorbed on the adsorbent materials. The PSA model is customized to simulate a 4-step cycle process (adsorption, blowdown, evacuation, light product pressurization) (Riboldi et al., 2020). The model includes a realistic representation of the operation of vacuum pumps, based on the work from Maruyama et al. (2020). The model has four optimization variables: the minimum, intermediate and maximum pressure and the adsorption temperature and two constraints, purity and recovery. A techno-economic analysis (TEA) framework is integrated into the PSA model enabling the consideration of the CO₂ avoidance cost (CAC) as objective function of the optimization. The TEA model is based on rigorous simulation of PSA (Subraveti et al., 2021) and adapted for utilization in the reduced-order model. The model embeds external models and functions, so the optimization algorithm must be able to handle blackboxes. Two adsorbent materials are investigated in this study: a zeolite 13X (Haghpanah et al., 2013) and a metal-organic framework (MOF) UTSA-16 (Agueda et al., 2015). The competitive adsorption isotherms are described by the dual-site Langmuir model. The CO₂ emission intensity of energy demands and reference processes is taken fromecoinvent version 3.7.1, using the impact assessment method "Environmental Footprint 2.0" and allocation at the point of substitution (Wernet et al., 2016).

4. Case study

4.1. Setup of the optimizations

Each algorithm is tested 20 times considering randomly generated initial values. For NOMAD and KNITRO, the same 20 initially generated initial values are considered. In contrast, the BO needs an initial set of several function evaluations to create the first GP regression model. The size of the initial set is set to ten. Each initial set contains the initial value of NOMAD and KNITRO of the corresponding test run and additionally nine randomly generated evaluations. The BO does not evaluate gradients of the original function, but only of the surrogate. Therefore, BO is terminated if the maximum amount of function evaluations is reached, which is set to 56. In addition, BO is terminated if the constraints are satisfied, and the last iterations is within 1% of the best previously obtained objective function value and the scaled inputs (between 0 and 1) of these two points have a distance of less than 0.015. A large exploration factor prevents quick convergence towards the same region and premature termination of the optimization. A multi-start gradient-based optimization is used to maximize the marginal likelihood to estimate the hyperparameters and optimize the acquisition function.

The NOMAD algorithm terminates when the maximum amount of function evaluations are reached. We test a maximum amount of 100 and 150 function evaluations. The constraints are handled with a progressive barrier approach. Therefore, constraints must just be satisfied at the solution. For KNITRO, we performed a parameter study to identify the solver setting resulting in a good compromise between accurate gradients and reliable convergence due to discontinuities of the model. From this study, we chose the Active Set algorithm to solve the problem. The gradients of the objective and constraints are approximated using forward finite differences and the Hessian of the Lagrangian by (dense) quasi-Newton BFGS Hessian. To account for the discontinuities in the objective function, constraints and gradients, we select a comparatively high relative step size of 0.05 for finite-difference gradients.

4.2. Cases

The performance of the two adsorbent materials is tested on the adsorption of CO₂ from flue gas for two relevant industries, namely a cement kiln and a natural gas (NG) offshore combined cycle power plant. The two cases are representative of rather different CO₂ concentrations (25% and 4% of CO₂ in N₂, respectively). The power required by the PSA process is drawn from the grid for the cement case while internally generated in the case of the NG offshore combined cycle.

5. Result

The optimization algorithms are compared in Table 1. For the cement case, BO achieves the smallest average CAC with the smallest computational time. In addition, it reaches the same optimum with high reliability. KNITRO achieves a similar objective function value in case of the zeolite 13X adsorbent but requires more function evaluations and thus, computational time. For UTSA-16, the mean objective function value differs by 10% for both of KNITRO and NOMAD 100/150 from the mean CAC obtained with BO. Moreover, the maximal difference in the objective function value is 39% and 24% for NOMAD 150 and KNITRO, respectively, while it is only 14% and 15% for BO EI and BO LCB, respectively. For the cement case, BO is also about twice as fast compared to the other algorithms. The optimal values of two of the four process degrees of freedom are at the bounds. This result is identified consistently by the BO, while KNITRO and

Table 1: Comparison of optimization algorithms based on 20 tests. The CAC, the elapsed time, the number of PSA function evaluations and the maximal difference in the objective function over the 20 tests are displayed. The best values are indicated by bold font.

Case	Adsorbent	Method	CAC			Elapsed in s		# Evals		Max diff. CAC in %
			Mean	Std	Min	Mean	Std	Mean	Std	
Cement	zeolite 13X	NOMAD 100	33.4	2.1	31.2	62.2	1.2	100	0	23.6
		NOMAD 150	32.2	0.9	31.2	93.2	1.5	150	0	12.4
		KNITRO	31.4	0.4	31.1	48.1	10.5	86.0	18.6	5.4
		BO LCB	31.3	0.2	31.1	27.1	6.6	36.2	8.0	3.3
		BO EI	31.4	0.3	31.2	28.8	8.6	40.9	12.4	2.9
	UTSA-16	NOMAD 100	37.5	3.1	33.1	62.4	1.6	100	0	39.4
		NOMAD 150	36.6	3.1	33.1	93.6	1.2	150	0	39.1
		KNITRO	37.0	2.6	33.2	61.1	12.1	108.5	20.5	24.4
		BO LCB	33.7	1.1	33.3	33.9	7.1	45.0	8.6	15.4
		BO EI	33.8	1.0	33.2	26.8	8.2	35.8	10.4	13.6
NG offshore	zeolite 13X	NOMAD 100	566.5	51.2	539.3	63.0	1.5	100	0	42.6
		NOMAD 150	549.3	25.8	536.8	93.71	1.1	150	0	22.9
		KNITRO	542.3	5.3	535.4	63.6	14.1	109.8	22.3	3.3
		BO LCB	567.8	30.1	538.5	39.2	7.1	53.2	8.6	18.4
		BO EI	547.8	11.5	537.0	45.9	4.04	56	0	7.3
	UTSA-16	NOMAD 100	622.4	6.3	612.4	62.9	1.0	100	0	3.0
		NOMAD 150	617.9	4.8	609.2	93.7	1.5	150	0	3.8
		KNITRO	626.5	9.7	612.3	61.9	9.1	109.3	15.8	6.7
		BO LCB	642.5	21.7	612.2	40.6	5.4	54.3	6.1	13.1
		BO EI	629.9	28.8	610.0	40.9	8.1	50.8	9.7	17.8

NOMAD converge to sub-optimal solutions in some test runs, where these variables are not at the bounds. Nevertheless, all algorithms converge to the same optimum in at least one run. Therefore, the best achieved objective function values are almost identical.

For the NG offshore case, KNITRO achieves the best mean objective function value for the zeolite 13X adsorbent and NOMAD 150 for the UTSA-16 adsorbent. BO still needs less computational time than the other algorithms but is less reliable for this case. In one to two of the 20 test runs, the maximal difference in the objective function is 18% for BO. For these cases, at least one optimization variable gets stuck close to or at a bound. The BO LCB is more prone to this behavior than the BO EI. The BO EI has a stronger focus on exploring the optimization region, which improves converging to the global optimum when many local solutions exist. Nevertheless, the best function value achieved by every algorithm over the 20 tests is very similar.

The optimization time mainly results from evaluating the PSA model. BO needs fewest function evaluations but needs more overall optimization time per function evaluation than the other algorithms. The time difference per function evaluation between BO and KNITRO is about 0.2s. BO optimizes hyperparameters of three surrogate models (objective function plus two constraints) and minimizes the surrogate model in each iteration, which causes the larger time consumption per function evaluation.

6. Discussion

For the considered cases and adsorbents, none of the optimization algorithms can be recommended unconditionally. The discontinuities in the PSA model result in many local optimal solutions making it difficult for the optimization algorithm to converge consistently to the global optimum. The BO is the fastest algorithm. For the BO algorithm, the user has to choose the covariance function, the maximum allowed function evaluations and the termination criteria. More function evaluations improve the reliability but lead to unnecessary function evaluations around the optimum. The maximum number of function evaluations chosen in this work balances reliability of finding a solution close to the global optimum and computational time. The chosen criterion to terminate the BO before reaching the maximum allowed function evaluations decreases the number of iterations by about 15 in the cement case, while it almost never triggered in the NG offshore case. Defining a suitable termination criterion for BO can be challenging. However, without additional termination criterion, the computational time of the BO for the cement case would just increase by about 10 s. For the BO, a squared exponential covariance function is chosen, which expects a certain smoothness of the system. The convergence might be improved by a less smooth covariance function, for which a suitable optimization method has to be chosen to optimize the Gaussian Process regression model, e.g., a gradient-free optimization method. For GP-LCB, we choose $\kappa = 16$ to focus on exploration since the BO-LCB tends to converge to local optima if κ is too small. The main reason for choosing the LCB acquisition function is to avoid exploring likely uninteresting regions, which is counteracted by a large κ . The EI acquisition function is easier to apply for this case, which does not contain a tuning parameter.

For NOMAD, a maximum of 100 and 150 function evaluations are tested. For these investigated cases and adsorbents, fewer than 100 function evaluations are not recommended without the risk of significantly degrading performance. 150 function evaluations increase the reliability but also the computational time. The advantage of the NOMAD is the easy applicability and the convergence to the global optimum giving enough function evaluation. Similar to BO, KNITRO achieves a good balance between reliability and computational time. However, a detailed study of the user options has to be performed in advance (e.g., algorithm, methods for gradients and Hessian or step size of finite differences) to tackle the challenges of the discontinuities of the model. A combination of BO and KNITRO, where BO provides the initial values for KNITRO, could reduce the computational time of KNITRO.

7. Conclusions

A Bayesian optimization algorithm is tested for optimizing the carbon-avoidance cost of a PSA process for two adsorbent materials and two industry cases and compared to two optimization algorithms from the literature. For the investigated cases, BO shows a good balance between computational time and reliability and is thus a promising option for adsorbent screenings. However, no algorithm outperforms the others in all investigated cases. Each optimization algorithm is at least for one case/adsorbent combination less reliable than the other algorithms. Consequently, in future work, a detailed comparison of the algorithms has to be performed for more cases and adsorbent materials.

8. Acknowledgement

The authors are supported by the PrISMa project (No 299659) that is funded through the ACT Programme (Accelerating CCS Technologies, Horizon 2020 Project No 294766), and it receives financial contributions from BEIS, NERC and EPSRC (UK), RCN (Norway), SFOE (Switzerland) and US-DOE (USA).

References

- M. Abramson, C. Audet, G. Couture, J. Dennis, Jr., S. Le Digabel, C. Tribes, 2021. The NOMAD project. Software available at <https://www.gerad.ca/nomad/>.
- V. I. Agueda, J. A. Delgado, M. A. Uguina, P. Brea, A. I. Spjelkavik, R. Blom, C. Grande, 2015. Adsorption and diffusion of H₂, N₂, CO, CH₄ and CO₂ in UTSA-16 metal-organic framework extrudates. *Chemical Engineering Science* 124, 159–169.
- L. T. Biegler, L. Jiang, G. Fox, 2004. Recent Advances in Simulation and Optimal Design of Pressure Swing Adsorption Systems. *Separation and Purification Reviews* 33, 1–39.
- R. H. Byrd, J. Nocedal, R. A. Waltz, 2006. Knitro: An integrated package for nonlinear optimization. In: *Large-scale nonlinear optimization*. Springer, pp. 35–59.
- T. de Avila Ferreira, H. A. Shukla, T. Faulwasser, C. N. Jones, D. Bonvin, 2018. Real-time optimization of uncertain process systems via modifier adaptation and gaussian processes. In: *2018 European Control Conference (ECC)*. IEEE, pp. 465–470.
- A. H. Farmahini, S. Krishnamurthy, D. Friedrich, S. Brandani, L. Sarkisov, 2021. Performance-based screening of porous materials for carbon capture. *Chemical Reviews* 121 (17), 10666–10741.
- R. Haghpanah, A. Majumder, R. Nilam, A. Rajendran, S. Farooq, I. A. Karimi, M. Amanullah, 2013. Multiobjective optimization of a four-step adsorption process for postcombustion CO₂ capture via finite volume simulation. *Industrial and Engineering Chemistry Research* 52, 4249–4265.
- S. Le Digabel, 2011. Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm. *ACM Transactions on Mathematical Software* 37 (4), 1–15.
- K. T. Leperi, Y. G. Chung, F. You, R. Q. Snurr, 2019. Development of a general evaluation metric for rapid screening of adsorbent materials for postcombustion CO₂ capture. *ACS Sustainable Chemistry & Engineering* 7 (13), 11529–11539.
- B. Maring, P. Webley, 2013. A new simplified pressure/vacuum swing adsorption model for rapid adsorbent screening for CO₂ capture applications. *International Journal of Greenhouse Gas Control* 15, 16–31.
- R. Maruyama, K. Pai, S. Subraveti, A. Rajendran, 2020. Improving the performance of vacuum swing adsorption based CO₂ capture under reduced recovery requirements. *International Journal of Greenhouse Gas Control* 93, 102902.
- C. E. Rasmussen, C. K. Williams, 2006. *Gaussian processes for machine learning*. the MIT Press.
- L. Riboldi, C. Charalambous, E. Moubarak, R. Anantharaman, S. Roussanaly, C. Fu, B. Smit, J. Young, M. van der Spek, E. Sanchez-Fernandez, D. Ongari, S. Majumdar, E. García-Díez, V. Kulakova, S. Garcia, 2020. Advanced methodology for screening of novel adsorption materials for cost-efficient CO₂ capture. *SSRN Electronic Journal*.
- E. Snelson, Z. Ghahramani, 2006. Sparse Gaussian processes using pseudo-inputs. In: *Advances in neural information processing systems*. pp. 1257–1264.
- J. Snoek, H. Larochelle, R. P. Adams, 2012. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* 25.
- S. G. Subraveti, S. Roussanaly, R. Anantharaman, L. Riboldi, A. Rajendran, 2021. Techno-economic assessment of optimised vacuum swing adsorption for post-combustion CO₂ capture from steam-methane reformer flue gas. *Separation and Purification Technology* 256, 117832.
- G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, B. Weidema, 2016. The ecoinvent database version 3 (part I): overview and methodology. *The International Journal of Life Cycle Assessment* 21 (9), 1218–1230.

Online state of charge estimation of lithium-ion battery using surrogate model based on electrochemical model

Seunghyeon Oh^a, Kyojin Jang^a, Jiyong Kim^b, Il Moon^a

^a*Department of Chemical and Biomolecular Engineering, Yonsei University, Seoul 03722, Republic of Korea*

^b*Department of Chemical Engineering, Sungkyunkwan University, Suwon 16419, South Korea*

Abstract

As electric vehicles (EVs) become popular, the use of lithium-ion batteries is increasing. However, it is difficult to ensure the reliable use of the battery because the state of charge (SoC) of a lithium-ion battery cannot be measured directly. SoC estimation is necessary to ensure the safety of battery systems by preventing over-charge and discharge of the battery and to extend battery life through efficient use. The optimal battery balancing strategy, estimation of the remaining driving range of EV, and Vehicle to Grid strategy can be developed through accurate SoC estimation. Therefore, it is necessary to study how to estimate the SoC of the battery in real-time. As methods for estimating SoC, equivalent circuit model, electrochemical model, and recently artificial neural network-based model are being studied. The electrochemical model is not suitable for real-time use due to its high computational complexity. Artificial neural network-based models require a large amount of data for learning, but most of the data is collected in the lab, which can lead to a lack of data and is difficult to ensure accuracy in such cases.

In this study, a surrogate model based on electrochemical models for SoC estimation is developed to solve computational complexity problems of electrochemical models and the accuracy problems of data-driven models due to data dependence. Based on the 1C discharge experimental data, parameter identification of the electrochemical model was performed using a Genetic Algorithm. Output variables such as Li-ion amount in the negative electrode and voltage for various drive cycle loads are derived from this model to create a sufficient amount of data for model training. The surrogate model is trained using these output variables and compared with data-driven models in terms of accuracy and computational complexity. Based on the Long Short Term Memory (LSTM) architecture, the artificial neural network-based model was trained gradually with driving cycle data and then compared with the surrogate model at each step. As a result, the surrogate model based on an electrochemical model using fewer data showed feasible computational complexity and high accuracy. Through the proposed method, it is expected to accurately estimate battery SoC in real-time and make the battery use efficiently.

Keywords: Lithium-ion battery, Battery modelling, State-of-Charge

1. Introduction

As electric vehicles (EVs) become popular, the importance of lithium-ion battery management is increasing. However, it is impossible to accurately monitor the electrochemical behavior in a lithium-ion battery, so it is very difficult to estimate the State of Charge (SoC) required for optimizing the use of the battery. Representative methods used to estimate SoC include the Coulomb Counting Method, the Open Circuit Voltage Method, the electrochemical model based method, and the data-driven model based method. The Coulomb Counting Method measures the current of the battery and integrates the current over time. The limitation of this method is that when the battery is used for a long time without being fully charged or discharged, errors accumulate, making it difficult to accurately estimate SoC. The open circuit voltage method uses the SOC versus OCV curve. However, to measure the relaxed OCV, the battery has to reach equilibrium, so it is difficult to use it in real-time because it takes several hours. The electrochemical model based method uses the electrochemical model to simulate the electrochemical behavior inside a battery. This method shows very accurate performance, but the model is so complex that it is difficult to use in real-time. The data-driven model based method is training and using artificial neural networks with sufficient data. If the data is sufficient, it shows high performance, but if the data is insufficient, the reliability of the model decreases.

Chemali et al. introduced a deep neural network to map battery signals directly to SoC. Chemali et al. used the Long Short Term Memory (LSTM) based artificial neural network model to estimate SoC. Jiao et al. introduced a GRU-based artificial neural network to estimate. Dawson-Elli et al. developed the surrogate model of P2D by using machine learning techniques such as a random forest, decision tree, and gradient boosted machine. Tian et al. employed an adaptive cubature Kalman filter to reduce the fluctuation of LSTM network estimation. In this study, the LSTM-based data-driven model, the P2D model, and the P2D model-based surrogate model are compared and analyzed in terms of accuracy and calculation time.

2. Datasets

This research uses a dataset obtained from a Panasonic 18650PF cell. The battery is described in Table 1. The dataset involves 1C discharge data, HPPC (Hybrid Pulse Power Characterization) data, drive cycle data at ambient temperatures 25 °C

Table. 1. Panasonic 18650 Cell Parameters

Nominal Open Circuit Voltage	3.6V
Capacity	Min. 2.75 Ah / Typ. 2.9 Ah
Min / Max Voltage	2.5V / 4.2V
Mass / Energy Storage	48g / 9.9Wh
Minimum Charging Temperature	10
Cycles to 80% Capacity	500 (100% DOD, 25)

3. Method

As methods for estimating SoC, the performance of the data-driven model, electrochemical model, and surrogate model is compared in terms of accuracy and computation time in this study. A flowchart representing the methodology is shown in Figure 1.

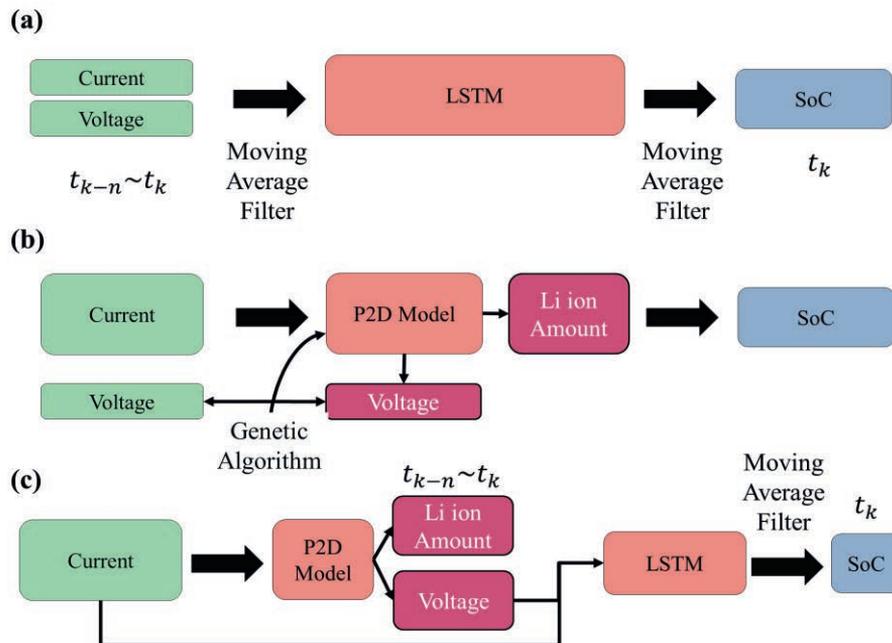


Fig. 1. Process Flowchart for Developing Models (a) Data-driven model (b) P2D model (c) Surrogate model based on electrochemical model

3.1. Data-driven model

LSTM is introduced to solve the performance degradation of RNNs in long-term sequences. Fig. 2 presents the structure of LSTM, in which x , h are the input of the network at time step k ; σ is sigmoid function and \tanh is hyperbolic tangent function. There are three gates in LSTM to capture long-term dependencies. The input gate, output gate, and forget gate allow the LSTM to forget or memorize newly acquired information to the memory cell. The LSTM model is trained on 9 drive cycle data. Of the ten drive cycle datasets, 80% of 9 datasets are used for training and 20% for validation. The model was tested with the remaining dataset. The exponential weighted moving average filter is used for smoothing input data.

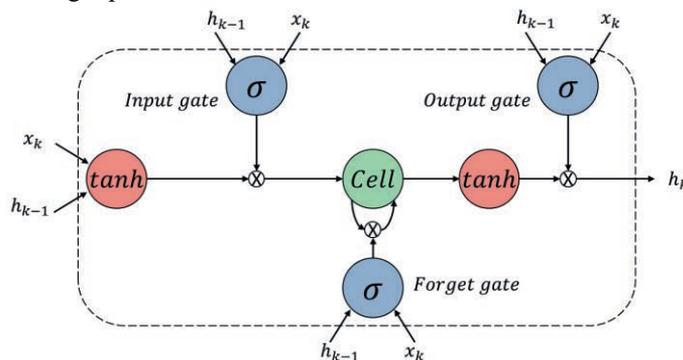


Fig. 2 Structure of the LSTM

3.2. Electrochemical model

The pseudo-two-dimensional(P2D) model, also called the Doyle-Fuller-Newman(DFN) model, is based on the porous electrode theory and the concentrated solution theory. Table. 2 shows nonlinear partial differential equations that represent mass balances, charge balances, electrochemical reaction kinetics. The 1C Discharge data is used to identify the parameters for the P2D model by Genetic algorithm.

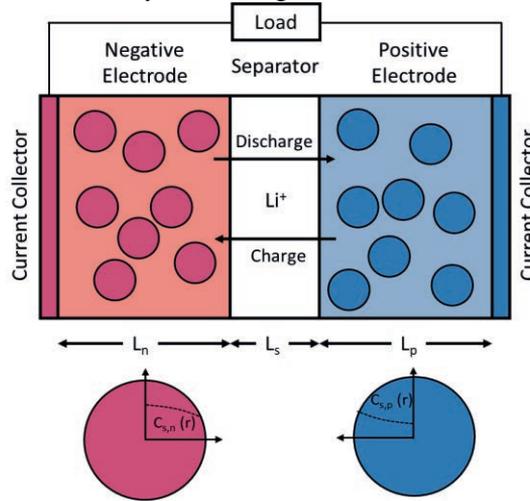


Fig. 3 Schematic of P2D Model

Table. 2 Governing Equations of P2D Model

Region	Governing equations
Charge Conservation in the solid particles	$\sigma_{eff,s,i} \frac{\partial^2 \phi_{s,i}}{\partial t} = a_{sep,i} F j$
Mass Conservation in the solid particles	$\frac{\partial c_{s,i}}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(D_{s,i} r^2 \frac{\partial c_{s,i}}{\partial r} \right)$
Charge Conservation in the electrolyte	$-\sigma_{eff,s,i} \frac{\partial \phi_s}{\partial x} - \kappa_{eff,i} \frac{\partial \phi_l}{\partial x} + \frac{2\kappa_{eff,i} RT}{F} (1 + t_+) \frac{\partial \ln c}{\partial x} = \frac{I_{app}}{A}$
Mass Conservation in the electrolyte	$-\varepsilon_{l,i} \frac{\partial c_l}{\partial t} = D_2 \frac{\partial^2 c_2}{\partial x^2} + a_{s,i} (1 - t_+) j$
Intercalation at the particle surface	$j = 2k_i (c_{s,i,max} - c_{s,i r=r_i})^{0.5} (c_{s,i r=r_i})^{0.5} (c_l)^{0.5} \sinh \left(\frac{0.5F}{RT} (\phi_s - \phi_l - U_i - FR_{SEI}) \right)$

3.3. Surrogate Model

The Surrogate model was trained in the same way as the Data-drive model using simulated data. The current datasets of nine drive cycles are used for the P2D model. It is different from the data-driven method in that it is trained using simulation data, not actual measurements. Based on the P2D model in which parameters are identified using 1C discharge data, the voltage dataset for training can be generated using any current dataset, which can be used if the measurement dataset is insufficient.

4. Result

The SoC estimation results of each method are shown in the Table. 3. The algorithms are executed on Core i7 4.0 GHz processor with 16 GB RAM. The surrogate model showed competitive performance without experimental data. Figure. 4 presents the graphs of the

estimated and actual values. In Fig. 5 below, the Mean Absolute Error(MAE) of the data-driven model becomes lower than the MAE of the surrogate model.

Table. 3. SoC estimation results

Model	MAE(%)	Computation time(s)
Data-drive model	1.21	1.02
P2D model	0.72	103.74
Surrogate model	2.01	0.77

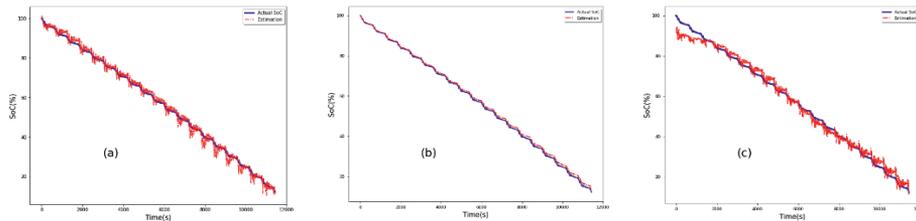


Fig. 4. SoC estimation results of models (a) data-driven model; (b) electrochemical model; (c) Surrogate model based on electrochemical model

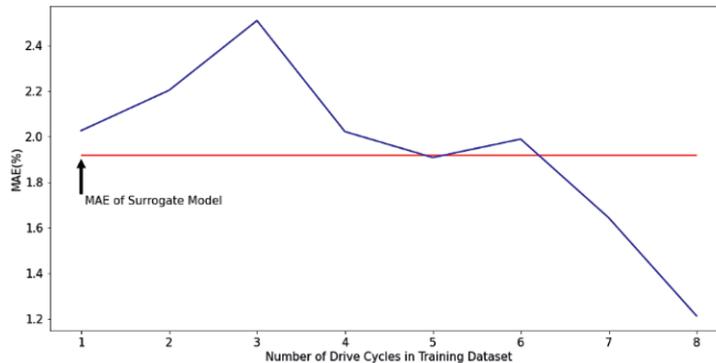


Fig. 5. MAE of the data-driven model

5. Conclusion

The P2D model estimated the SoC most accurately, but the computation time took the longest. The surrogate model and the data-driven model show similar computation times. The MAE of the data-driven model is lower than the surrogate model after training 5 drive cycles datasets. If data is insufficient, the Surrogate model could be an alternative to the data-driven model. The surrogate model also can be trained using more simulated data from the P2D model. If the battery condition changes by aging, it is difficult to gather sufficient data to train the data-driven model. However, simulated data can be gathered after the parameter identification of the P2D model. It is considered to be an advantage of the surrogate model for real-time SoC estimation rather than the data-driven model, which requires a lot of data when the battery condition changes.

Nomenclature

$i = (p, sep, n)$	T = ambient temperature [K]
p = positive electrode	A = cross section [m ²]
sep = separator	c = lithium concentration [mol/m ³]
n = negative electrode	D = diffusivity of lithium ion [m ² /s]
s = solid phase	j = the rate of lithium ion flowing out of a particle across a
l = liquid phase	boundary between the solid and the electrolyte [mol/m ² ·s]
κ_{eff} = effective conductivity of the electrolyte [S/m]	k = intercalation/deintercalation reaction-rate constant
ϕ = electrical potential [V]	[m ^{2.5} /mol ^{0.5} ·s]
σ = electronic conductivity [S/m]	r = radius of particle [m]
σ_{eff} = effective electronic conductivity [S/m]	R_{SEI} = resistance of SEI layer [Ω /m ²]
ϵ = volume fraction [-]	t_+ = transference number in the electrolyte [-]
F = Faraday constant [C/mol]	U = Open circuit potential [V]
R = ideal gas constant [J/mol·K]	
I_{app} = current applied to the lithium-ion battery [A]	

References

- P. J. Kollmeyer, 2018, Panasonic 18650PF Li-ion Battery Data, Mendeley Data, V1
- E. Chemali, et al., 2018, "State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach", Journal of Power Sources, vol. 400, 242-255
- E. Chemali, et al., 2018, "Long Short-Term Memory Networks for Accurate State-of-Charge Estimation of Li-ion Batteries", IEEE Transactions on Industrial Electronics, vol. 65, no. 8
- Jiao, et al., 2020, "A GRU-RNN based momentum optimized algorithm for SOC estimation", Journal of Power Sources, vol. 459, 228051
- N. Dawson-Elli et al., 2018, "Data Science Approaches for Electrochemical Engineers: An Introduction through Surrogate Model Development for Lithium-Ion Batteries", J. Electrochem. Soc., 165, A1
- Zhu et al., 2020, "Co-estimation of model parameters and state-of-charge for lithium-ion batteries with recursive restricted total least squares and unscented Kalman filter", Applied Energy, vol. 277, 115494

Machine-learning based prediction of infinite-dilution activity coefficients of ionic liquids using physicochemical properties

Seongju Lee,^a Changsu Kim,^b Thai Ngan Do,^b and Jiyong Kim^{b,*}

^a*Department of Energy and Chemical Engineering, Incheon National University, 22012, Republic of Korea*

^b*School of Chemical Engineering, Sungkyungwan University, 16419, Republic of Korea*

**Corresponding Author's E-mail: jiyongkim@skku.edu*

Abstract

Ionic liquids (ILs) is an organic salt in the liquid state, which have an excellent property as a favorable solvation for a range of polar and non-polar compounds. Recently, a number of ILs have been widely used for a chemical separation, e.g., extraction of aromatic and aliphatic compounds from hydrocarbon mixtures, due to its high-energy efficiency compared to conventional separation processes that requires a considerable thermal energy or electricity. One of critical properties forward a wide application of ILs is the infinite dilution activity coefficient (IDACs) of solutes and the equation of state (EoS) parameters. However, the discovery of a perplexing principle of ILs behaviors requires highly sophisticated experiments on a number of solutes.

Moreover, the validity of conventional property models such as UNIFAC, Abraham and COSMO-RS for predicting IDACs show some limits in that only few solutes are applicable within a certain range of operating temperature. Therefore, in this work, we aimed to propose the machine learning (ML) based IDAC prediction model to rapidly provide precise properties and parameters of ILs separation process. First, ILs experimental data and their physicochemical features are collected from the literature. We then developed an IDACs prediction model using artificial neural networks (ANNs) and validated the model with accuracy metrics such as coefficient of determination (R^2) and root mean square error (RMSE). Finally, the capability (accuracy and quantity) of the ML-based prediction model is discussed by comparing conventional models. As a result, it was revealed that the ML-based approach shows higher accuracy (approximately 10~50%) compared to conventional property models. A database of predicted properties based on data-driven strategy could lead to wide applications of ionic liquids in separation technique. From proposing a strategy for ILs utilization, the ML-based approach was able to provide the essential information on the ILs design and promote the future environmental-friendly separation processes.

Keywords: Ionic Liquids; Infinite-dilution activity coefficients; Machine-Learning.

1. Introduction

Ionic liquids (ILs) are the green solvents, which receive the great attention as replacement for conventional separation media. Since ILs could bring the significant economic and environmental benefits, the potential of ILs in industrial deployment has been proven (Urszula et al., 2010, Pablo et al., 2019, Siliang et al., 2020).

While ILs are indeed a promising separation technology, there are obstacles delaying the real-world deployment. The precise modelling of ILs process is the one of the most difficult problems. Conventionally, the ILs behavior is largely understood by the thermodynamics such as the equation of state (EoS). Even though such empirical approaches have described the principle of solute-solvent interactions in ILs, identifying critical parameter for EoS such as infinite dilution activity coefficient (IDACs) are extremely difficult. Because the experiments take time and trial-and-errors, which eventually became a bottleneck of ILs application.

Machine learning (ML) has been a practical tool for discovering and understanding the perplexing patterns in chemical domain. Especially, artificial neural networks (ANNs) have successfully assisted wide range of chemical engineering R&Ds including novel catalyst discovery, battery lifecycle prediction, and metal-organic-framework candidate material screening.

Fabian et al. utilized the matrix completion method (MCM) to develop ML-based activity coefficients prediction model (Fabian et al, 2020). Although such an approach successfully analyzes the mixture behavior of the 240 solutes and 250 solvents, the accuracy of the proposed model is a limitation for generating reliable prediction data. Therefore, we propose the ML-based IDAC prediction model for the precise and rapid modelling of ILs process. To implement such data-driven method, the ILs experiment datasets and solute-solvent property descriptors are collected. Then, ANNs were exploited to develop the ML-prediction model. Finally, from comparing to the conventional IDAC prediction model, the reliability of ML-based IDAC prediction model is evaluated.

2. Data and method

2.1. Overview

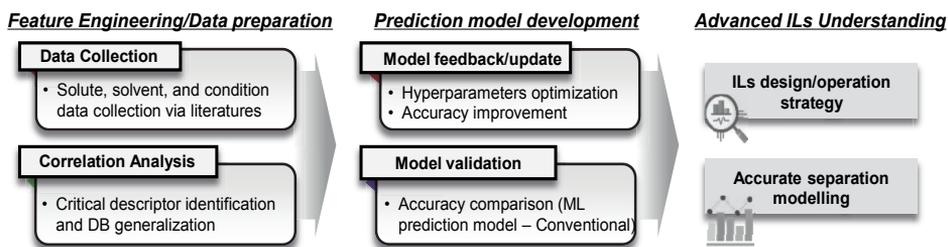


Figure 1: ML-based IDAC prediction model development overview

The main step of ML-based model development is described in Figure 1. It is composed of three steps: *Feature engineering and data preparation*, *Prediction model development*, and *Advanced ILs understanding*. Firstly, at *Feature engineering and data preparation* step, ILs synthesis and operation data are collected via literature survey (Urszula et al., 2009, Li et al., 2011). Since defining the chemical dimension of ILs process is important to develop highly-reliable prediction model, the most meaningful descriptor is extracted by using correlation analysis. As generalized ILs database is constructed, the datasets are randomly divided into two different datasets, which are for the model train and model validation. In this case, 70% of total datasets are used for model train and 30% are used for model test. Then, by using model train data, artificial neural networks (ANNs) develop the IDAC prediction model. To develop the accurate prediction model, the model

hyperparameters are optimized through a manual search technique. Also, the accuracy of the prediction model is evaluated by Pearson coefficient of determination value (R^2). Finally, by comparing the accuracy of ML-based prediction model and conventional prediction model (UNIFAC), the merits of ML-based prediction model are argued.

2.2. ILs separation data acquisition

As shown in Table 1, there are 9 types of solute and 7 types of solvent are used. Also, in Table 2, the properties and conditions of solvent and solute are described. Property descriptors include the thermodynamical-, physical-, electrochemical-, and interaction property. Firstly, thermodynamic properties at critical point (critical temperature, pressure, volume) and acentric factor are collected since it is important for navigating the phase change. Physical properties include the molecular weight, density, and viscosity. Electrochemical properties (dipole moment, Kamlet-Taft parameters) are also collected since the separation principle using ILs rely on different electric interaction property of substances in the mixture (Lee et al., 2008). Interaction property indicates the Hildebrand's solubility parameter, which determines the solubility and structural affinity of a solute-solvent couple (Marciniak, A., 2010). As a result, 1679 datasets with 14 descriptors are comprised.

Table 1: Solvent type and solute type of ILs separation data

Solute type		
n-Pentane	n-Hexane	n-Heptane
n-Octane	n-Nonane	n-Decane
Benzene	Toluene	Ethylbenzene
Solvent type		
[BMPYR][CF ₃ SO ₃]	[EMIM][TCB]	[4BMPy][NTf ₂]
[EMIM][TFA]	[BMIM][CF ₃ SO ₃]	[BMIM][SCN]
[3BMPy][SF ₃ SO ₃]		

Table 2: ILs separation process descriptor

Data description		
Descriptor type	Descriptor Name	Unit
Solute property	Critical temperature	K
	Critical pressure	bar
	Critical volume	cm ³ /mol
	Dipole moment	D
	Molecular weight	g/mol
	Acentric factor	-
Solvent property	Molecular weight	g/mol
	Density	kg/m ³
	Viscosity	m·Pas
	Hildebrand's solubility parameter	Mpa ^{0.5}
	Kamlet-Taft parameters, α	-
	Kamlet-Taft parameters, β	-
	Kamlet-Taft parameters, π^*	-
Operating conditions	Temperature	K
Outcome	Natural logarithm of infinite-dilution of Activity Coefficients	-

2.3. Prediction model development

Figure 2 (a) and (b) show the topology of ANNs and activation function used in this model. The optimal topology (50-100-150-200-250-150-100-50) is found by the manual search technique. Datasets were normalized for preventing the over-fitness of model. The activation function of ANNs is ‘swish’ function (in figure 2 (b)), which is well-known to outperform other activation functions in training deep neural networks.

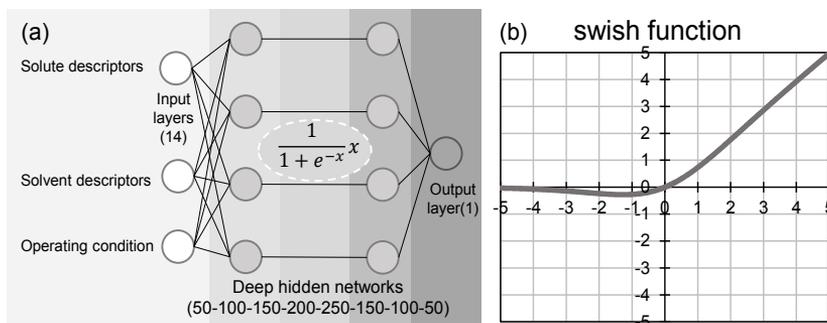


Figure 2: ANNs prediction model development details

Validation of the model was performed in the leave-one-group-out cross-validation (LOGOCV) method. LOGOCV is a cross-validation scheme which holds out the samples according to third-party provided array of integer groups. In order to numerically represent the reliability of models, the Pearson correlation coefficient (R^2) is employed.

2.4. Results

From developing more than 100 models while diversifying topologies, the final model is finally gathered. The train/test set validation of IDAC prediction model is described in Figure 3. For the accuracy evaluation, experimental IDACs and predicted IDACs are measured by R^2 value and compared to each other. Note that R^2 value above 0.95 means high accuracy. In Figure 3 (a) and (b), the prediction accuracies of training data sets and test data sets are almost perfect ($R^2 = 0.984$ and 0.974 , respectively). Also, small difference between two accuracy means that the model overfitting is quite unlikely. Since the ML-based prediction model shows the high accuracy and secure from the overfitting, the capability of this model is further investigated.

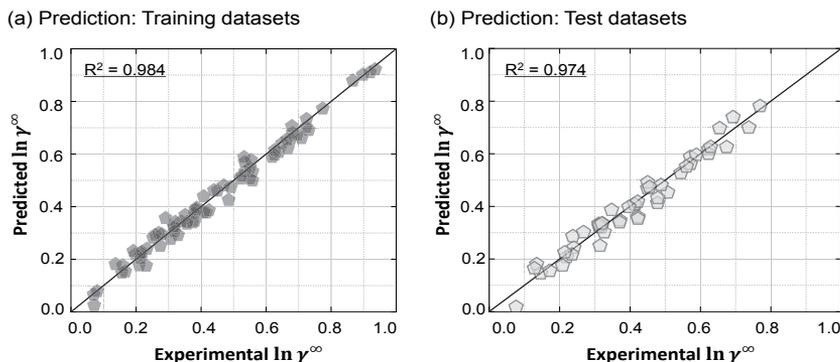


Figure 3: Predicted result and accuracy of training datasets and test datasets

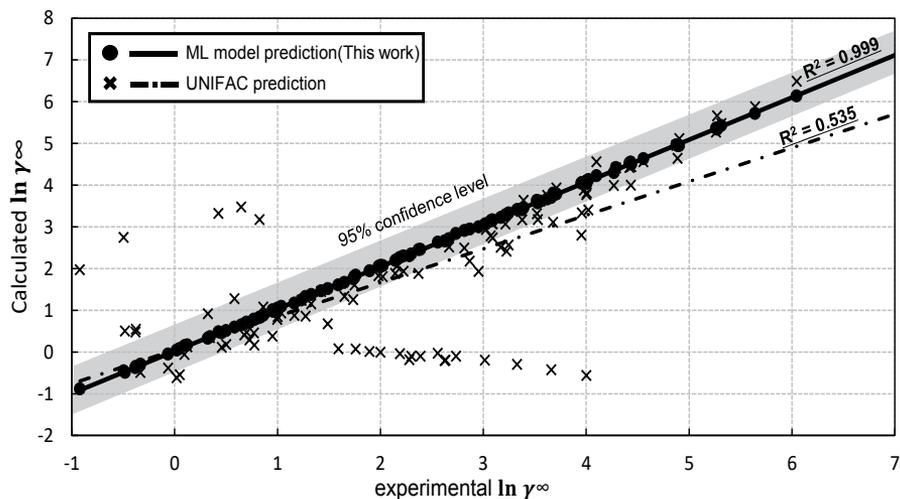


Figure 4. Accuracy comparison of the ML-based prediction model and UNIFAC

The accuracy of the ML-based prediction model and the conventional prediction model (UNIFAC, Thomas et al., 2019) is compared in Figure 4. Each model predicted the IDAC of test data sets where random solute and solvent combinations are given. Note that two types of point are the predicted results of each model (ML model – O, UNIFAC – X). In result, the ML-based model shows better prediction accuracy with higher R^2 value (0.999). It is interesting that the ML-based prediction model produces a relative confident result over 95% level in all prediction cases. In such event, it could be concluded that the ML-based model successfully found the IDAC patterns of ILs process. On the other hand, UNIFAC produced a result of relative high variance. Even though many predictions show a high accuracy, there are unignorable numbers which show incorrect prediction.

3. Conclusion

In this study, we developed and introduced the ML-based IDAC prediction model for ILs process applications. By collecting and defining the critical descriptors, the preliminary chemical dimension of ILs process is constructed. As descriptors, the thermodynamical,

physical, electrochemical, and interaction properties of solute and solvent are utilized. In result, the ML-based IDAC prediction model successfully achieved the high accuracy ($R^2 > 0.95$) and secured from overfitting problem. Also, from comparing the ML-based IDAC prediction model to the conventional IDAC prediction model, it is concluded that the ML-based model is more precise and understood the solute-solvent interaction trend in ILs process. Based on this work, it is able to i) suggest the proper ILs design and operation strategy, ii) provide the efficient and clean separation method from promoting and facilitating the real-world deployment of ILs.

References

- Urszula Domanska, Marta Krolikowska, 2010, Measurements of activity coefficients at infinite dilution in solvent mixtures with thiocyanate-based ionic liquids using GLC technique, *The Journal of Physical Chemistry B*, 114, 8460-8466.
- Pablo Navarro, Ignacio de Dios-Garcia, Marcos Larriba, Noemi Delgado-Mellado, Miguel Ayuso, Daniel Moreno, Jose Palomar, Julian Garcia, Francisco Rodriguez, 2019, Dearomatization of pyrolysis gasoline by extractive distillation with 1-ethyl-3-methylimidazolium tricyanomethanide. *Fuel Processing Technology*, 195, 106156.
- Siliang Gao, Wencheng Tang, Ming Zhao, Siyuan Qie, Weiwei Pang, Longsheng Tian, 2020, Extractive distillation of benzene, toluene, xylenes from pyrolysis gasoline using methylsulfonylthane as a cosolvent, *Asia-Pacific Journal of Chemical Engineering*, 16, 3.
- Fabian Jirasek, Rodrigo A. S. Alves, Julie Damay, 2020, Machine Learning in Thermodynamics: Prediction of Activity Coefficients by Matrix Completion, *Journal of Physical Chemistry Letters*, 11, 3, 981–985.
- Urszula Domanska, Gan G. Redhi, Andrzej Marciniak, 2009, Activity coefficients at infinite dilution measurements for organic solutes and water in the ionic liquid 1-butyl-1-methylpyrrolidinium trifluoromethanesulfonate using GLC, *Fluid Phase Equilibria*, 278, 1-2, 97-102.
- Yi Li, Li-Sheng Wang, Yun-Xia Feng, Chun-Yuan Zhang, 2011, activity coefficients of organic solutes at infinite dilution in ionic liquids. 1. 1-hexyl-3-methylimidazolium hexafluorophosphate and 1-octyl-3-methylimidazolium hexafluorophosphate and their application to alkane/aromatic and aromatic/aromatic hydrocarbon separation, *Ind. Eng. Chem. Res.*, 50, 10755-10764.
- Jong-Min Lee, Sebastian Ruckes, John M. Prausnitz, 2008, Solvent Polarities and Kamlet-Taft Parameters for Ionic Liquids Containing a Pyridinium Cation, *J. Phys. Chem. B*, 112, 1473-1476
- Andrzej Marciniak, 2010, The solubility parameters of ionic liquids, *International Journal of Molecular Sciences*, 11, 1973-1990.
- Thomas Brouwer, Boelo Schuur, 2019, Model Performances Evaluated for Infinite Dilution Activity Coefficients Prediction at 298.15 K, *Ind. Eng. Chem. Res.*, 58, 8903-8914.

The Impact of Reward Shaping in Reinforcement Learning for Agent-based Microgrid Control

Valentin Pèrè^a, Fabien Baillon^a, Mathieu Milhe^a and Jean-Louis Dirion^a

^a*Université de Toulouse, IMT Mines Albi, UMR CNRS 5302, Centre RAPSODEE, Campus Jarlard, F-81013 Albi Cedex 09, France*
valentin.pere@mines-albi.fr

Abstract

In order to reduce CO₂ emissions, electricity networks must increasingly integrate renewable energies. Microgrids are distributed electrical networks with their own generation and load, often supported by an electrical storage system. It can be connected to the external electrical network or isolated. Since electricity consumption, price and renewable production are stochastic phenomena, the control of microgrids must adapt to uncertainties. Data-driven models and in particular reinforcement learning (RL) have become efficient algorithms in high-level microgrid control. RL are agent-based algorithms, which interact with their environment and learn with a numerical reward signal. A certain behavior can implicitly be expected when the reward system is formulated. For example, a reward system that encourages the agent to interact as little as possible with the external network will explicitly increase the autonomy of the microgrid. Implicitly, it can be expected to schedule the battery to maximize the ratio of renewable energy used to the amount producible. Q-learning algorithm has been used due to its performance in discrete action space, which simplified the benchmark complexity. An agent is trained with different reward functions commonly found in the literature related to data-driven microgrid control algorithms. The agent parameters do not vary from one case study to another. Indicators are set up to evaluate the agent behavior. They are based on implicit behavioral criteria in the definition of the reward system such as the ratio of renewable energy used, the amount of energy stored during peak hours, etc. This study enables to find a way to rationalize the choice of a reward system to control in a near-optimal way microgrid while meeting implicit secondary objectives. It could lead to a choice on weighting coefficient in a combination of reward functions.

Keywords: Microgrid, Reinforcement Learning, Control, Reward

1. Introduction

1.1. Electricity storage scheduling with reinforcement learning

To remain reliable despite uncertainties in renewable electricity production and consumption, microgrid control must be efficient. Bidram and Davoudi (2012) has distinguished 3 categories of control. The first 2 categories are frequency and voltage regulation with very low time scale. The third category is what is called high-level control in this study. It concerns power flow long-term planning with higher time granularity (minimum 15 minutes). This high level planning can be done with several methods (Abdelhedi et al. (2018)): rule-based, optimization-based or learning-based methods are efficient. However, with uncertainties of electricity consumption, price and renewable generation, complete physical model based approaches are inappropriate. Thus, optimization-based methods will have difficulties to achieve optimal planning without predicting the future values of the stochastic variables. Data-driven methods are proven to be efficient in this context.

Especially, reinforcement learning (RL) algorithms (Sutton and Barto (1995)) learn policies given an environment and objectives. A RL agent makes decisions in its environment using Markov decision processes, the environment responds and the agent receives a reward signal, indicating if the reached environment state is suitable or not in respect to the objective. With this signal, it learns to value states or actions taken in specific states and thus can build a control policy according to the defined reward system.

High-level control can focus on every microgrid unit, including consumption (demand side management), controllable production (unit commitment) and electricity storage scheduling. This study aims to provide a view of the impact of the choice of reward signal in storage scheduling with respect to implicit indicators.

1.2. Case Study

The system studied is a simulation of a microgrid composed by photovoltaic (PV) panels for electricity production, a point of consumption, an electrochemical battery for short-term storage system and hydrogen storage for long-term electricity storage. This simulation has data-driven units (electricity consumption and PV production) and analytical models (storage). Both data and short-term storage characteristics are taken from François-Lavet et al. (2016). The microgrid has two operation modes: connected to the main grid and isolated.

Some simplifications were made: The maximum power of the battery is not taken into account, the power to be supplied is multiplied by its efficiency and it automatically balances the network provided that its energy capacity is high enough. The RL agent controls the hydrogen storage system. Its maximum power is 1.1kW, its electrolyser efficiency is 0.65 and its fuel cell efficiency is 0.5. No maximum storage capacity is considered. The data are two years of PV production data in Belgium and consumption data respectively. When the net demand for electricity cannot be supplied, the short-term storage is discharged and charged when there is a surplus of energy. The main goal here is to test different reward functions to observe their effects on in behaviors that are implicitly expected from the agent. These behaviors are tracked with indicators.

The RL algorithm used is Deep Q-learning (Mnih et al. (2013)). The agent receives continuous state values that are the electricity consumption, the PV electricity production and the short-term battery state of charge (SOC). All these values are normalized between 0 and 1. The actions that are available for the agent at each timestep are the operating mode of the hydrogen storage system. These actions are discrete, the first action available uses electricity to charge the long-term storage with electrolysis at maximum power if possible. The second action discharges the hydrogen storage with fuel cell at its maximum nominal power (if enough energy is stored). The agent can also choose to do neither. Thus, its action space is composed of three actions.

2. An introduction to Q-learning and deep Q-learning

2.1. Q-learning

The objective of a RL algorithm is to find policies (i.e. probability to take one action from a given state) that maximize the rewards received in an episode (i.e. a series of interaction within the environment). To decide between several actions, the agent values state and action pairs through the rewards following the choice. These pairs are called Q-values and denoted $Q(s, a)$, a stands for action and s for state. The idea behind building an efficient policy is to select the action that maximizes this value from the state in which the agent is located. However, in order to value these pairs, the agent has to explore states and values to sample rewards. A Q-learning agent uses a behavioral policy to sample actions and a learned policy to update the different pairs value. With this algorithm, only the immediate reward perceived by the agent after his action and the following

action (from the next state) that brings to the maximum Q-value are taking into account to update the learned policy.

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t) \right] \tag{1}$$

This update rule for Q-value mapping is shown in Equation 1, with $\gamma \in [0; 1]$ the discount factor to level out the extent to which future actions are considered in the estimation of a Q-value. Once the agent is trained, an estimate of every Q-values is stored in a table. The agent can then choose every action that maximizes immediate and future reward directly according to this table.

2.2. Deep Q-learning

With Q-learning, the main drawback is from the use of a table, which implicitly requires countable and therefore discrete spaces. Also, wide spaces lead to long computations before the agent is trained.

With Deep Q-learning, the table is replaced by a neural network (NN). In deep learning, the target of a NN has to be stationary. Here, as showed in Figure 1, the target value the NN must predict Q-values, the second part of equation 2 (which the same equation as 1 but factorized in an different way).

$$Q(S_t, A_t) = (1 - \alpha)Q(S_t, A_t) + \alpha \times (R_t + \gamma \times \max_a Q(S_{t+1}, a)) \tag{2}$$

With these Q-values, the problem is that a part of the target, $\max_a Q(S_{t+1}, a)$ (with S_{t+1} the state at time $t + 1$ and a the action that can be taken from state S_{t+1}) depends on the NN that is updated. To solve this problem, an other NN with the same parameters is used to estimate $Q(S_{t+1}, A_{t+1})$. These parameters are frozen and actualized slowly. Another reinforcement learning problem is the influence of the output on the next input. This problem is solved with a *ReplayMemory* that injects randomly a sample of state, action, reward, next state and next action into the NN. In this way, the impact of previous prediction is negligible for incoming output and the idea of transition in the system is kept.

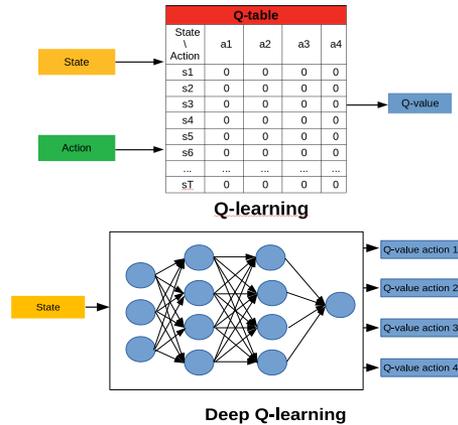


Figure 1: Visual comparison between Q-learning and deep Q-learning

3. Analysis on the impact of reward shaping on implicit indicator

Each and every microgrid system is built with a particular objective. Isolated microgrids follow the objective to be autonomous for example. Some have diesel generator as support electricity production (Kofinas et al. (2018)), and the objective can be to supply users autonomously without this support. Grid-connected microgrids can minimize operation cost or emissions for example. Even though objectives and therefore reward systems can be different, whoever made them implicitly expects awaited behaviors. In the case study microgrid presented in 1.2, whether the reward system depends on autonomy or operating cost, the system is expected to buy less energy from the main grid. Of course, in an isolated grid that aims to be autonomous, the equivalent of buying electricity from the main grid to compensate the grid imbalance is poor quality electricity or even blackout periods. Thus, this need of extra energy can be penalized the same way for these very

different systems. Also, systems aims to be energy efficient. Electricity production can not much exceed consumption, or else equilibrium is lost. However, it is important to get as much electricity as possible from the PV panels, whatever the objective. When the system does not need electricity (consumption is supplied, battery is fully loaded and long-term storage is supplied at maximum power), the remain production is extra-energy and lost. Another implicit objective in every case is to minimize excess of electricity.

3.1. Methodology

Every hour, the agent makes a choice in its microgrid. It can charge or discharge the long-term storage, or do nothing. Its action is perceived as an extra electricity demand or production, if the long term storage is charged or discharged respectively. In order to underline the impact of reward shaping, deep Q-learning agents are trained with different reward systems and indicators are identified. First, in an objective of operating cost minimization, different microgrid configurations are tested. The microgrid buys electricity from the main grid at a price of 2€/kWh. As the time granularity of the simulation is one hour, the agent perceives a -2 reward in this case. It occurs when the microgrid net demand (electricity demand action minus electricity production) including additionnal production or consumption from agent action (hydrogen storage charge or discharge perceived as consumption or production) is positive and superior to the electrochemical battery capacity. On the opposite, when this net demand is negative and the extra energy exceeds electrochemical battery capacity, the remaining produced energy is wasted.

Whatever which mode (isolated or grid-connected) is selected, this system is adopted, and the negative reward for buying electricity is applied to penalize isolated microgrid instability (in this case, no electricity is bought, but the simulation is similar). In the grid-connected mode, three configurations are tested, with the objective of reducing operating cost:

- Case 1 The system can not sell energy to the main grid. With the operating cost reward system, the only perceived rewards are negative and correspond to the electricity purchased from the main grid.
- Case 2 The system can sell energy to the main grid, without power constraint at the common coupling point (the electricity exchange point between the micorgrid and the main grid). In this configuration, agent can receive rewards from selling extra-energy to the main grid. In this specific case, it is impossible to waste energy and PV panels produce 100% of what they should. The electricity sold is four times cheaper than the bought electricity.
- Case 3 The system can sell energy to the main grid, with a power constraint at the common coupling point. The agent sells its extra energy and receives positive rewards in this case. However, this amount of sold energy is limited by a power constraint and so is the reward.

In isolated microgrid, as the main objective is to be autonomous, a negative reward for system instability is applied. It is exactly the same system configuration as the first case listed above.

What if the reward system is explicitly giving the agent penalties for wasting excess energy ? To analyse this situation, the first and last configuration are adopted with an extra negative reward equal to the wasted electricity. Again, it has no sense to use this reward system on the second case because no energy waste is allowed. At last, the energy excess negative reward will be applied without the energy bought negative reward to know how the system behave with only this objective.

A comparison of the quantity of excess energy and bought energy from the main grid will be done in the next section.

3.2. Results and analysis

Agents were trained in every case, with three reward systems for each except the second case in which extra energy does not exist. It means seven agent were trained. The convergence of obtained rewards converged at approximately 30 episodes for every agent. Once the training is over, data of their control behavior in the last training episode (1 year) are collected to analyse results.

3.2.1. Excess energy and purchased energy

Surprisingly, there is very little variation in excess energy between the reward systems for each case.

The same pattern is observed with some small variations. Obviously, this amount is smaller when only excess energy governs the agent’s reward system. It is larger when reward system only considers operating cost. This has many common features with the purchased energy curves. As shown in Figure 2, more energy is brought from the main grid in winter. The agent tends to buy more energy when its reward system penalize it.

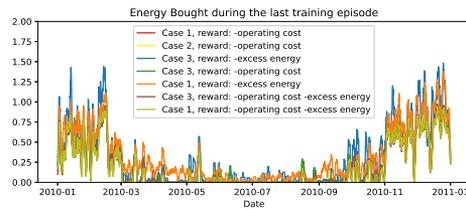
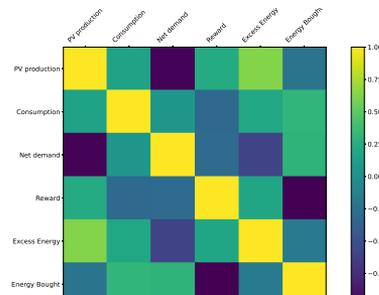


Figure 2: Smooth curve of hourly bought energy in the 3 reward systems

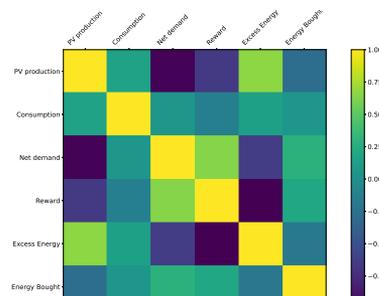
3.2.2. Correlation between variables

Correlation matrices between exogenous variable have been made. It gives important information on how the agent behaves.

Figure 3a shows that rewards are negatively correlated with net demand when operating cost defines reward system. Of course, purchased energy is negatively correlated with rewards. Net demand is the only parameter affected by the agent decision. To increase net demand, the agent has to discharge the long-term storage. In order to do that, the hydrogen storage can not be empty. The agent’s game would therefore be to charge and discharge the hydrogen storage at the right time, so that it can discharge when the purchase of electricity is necessary, to alleviate the negative reward. On the contrary, Figure 3b shows that rewards are positively correlated with net demand when excess energy penalizes the agent. The agent’s action must therefore charge the hydrogen storage when the PV production exceeds the consumption. He can buy energy without being penalized and emptying the storage (infinite in capacity) does not increase his reward. The graph of stored energy (Figure 4) is interesting, it underlines the fact that long-term



(a) Correlation matrix of exogenous variables in Case 1 with only operating cost considered in reward system



(b) Correlation matrix of exogenous variables in Case 1 with only excess energy considered in reward system

storage is used as short-term storage to increase the rewards when buying and selling energy define the reward system. What was observed in the correlation matrix is confirmed in the agent behavior, giving penalties only for excess energy tends to make the agent store bigger hydrogen quantity when it is possible. When both excess energy and operating cost are considered, the agent also tends to store hydrogen the way expected in Case 3. Net demand is highly correlated with PV production and very little correlated with consumption. This is because PV energy production has a range of values that can go very high when non-zero, compared to the consumption which is more constant and low.

When PV panels produce a lot of electricity (between July and September), it may be more attractive to store energy to avoid big excess energy penalties rather negative reward for energy bought. In Case 3, the stored energy can be sold to the main grid in winter. This explains why the agent prefers to store energy in summer in Case 3 rather than in Case 1 (where energy can not be sold) with multi-objective reward system.

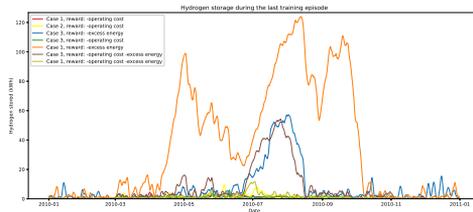


Figure 4: Energy stored in long term storage during the last training episode in every cases

4. Conclusion

The reward functions defines the behavior of a RL based control algorithm in microgrid. Making explicit certain implicitly expected behaviors changed totally the decisions. The multi-objective reward system seems more interesting for the microgrid control in this case study. However, the way to modelize electrochemical battery and hydrogen storage was too simplistic. The sizing of the storage systems and PV panels was arbitrary, as was the price of energy both when purchased and when sold to the main grid. In these simulation conditions, the sensibility study of the defined reward system can not be unbiased. However, it highlighted a behavioral difference of the agent on how to store hydrogen. With a robust sizing of the system and a good storage systems modeling, optimal weights for reward systems can be found through sensibilization analysis on different criteria.

References

- R. Abdelhedi, A. Lahyani, A. C. Ammari, A. Sari, P. Venet, Feb. 2018. Reinforcement learning-based power sharing between batteries and supercapacitors in electric vehicles. In: 2018 IEEE International Conference on Industrial Technology (ICIT). IEEE, Lyon, pp. 2072–2077.
- A. Bidram, A. Davoudi, Dec. 2012. Hierarchical Structure of Microgrids Control System. IEEE Transactions on Smart Grid 3 (4), 1963–1976.
- V. François-Lavet, D. Taralla, D. Ernst, R. Fonteneau, 2016. Deep Reinforcement Learning Solutions for Energy Microgrids Management. European Workshop on Reinforcement Learning (EWRL 2016), 7.
- P. Kofinas, A. Dounis, G. Vouros, Jun. 2018. Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids. Applied Energy 219, 53–67.
- V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M. Riedmiller, Dec. 2013. Playing Atari with Deep Reinforcement Learning. arXiv:1312.5602 [cs].
- R. S. Sutton, A. G. Barto, 1995. Reinforcement Learning: An Introduction, 352.

Knowledge mining from scientific literature for acute aquatic toxicity: classification for hybrid predictive modelling

Gulnara Shavaliyeva,^a Stavros Papadokonstantakis,^{a,b} Greg Peters^c

^a*Chalmers University of Technology, Department of Space, Earth and Environment, Hörsalsvägen 7b, Gothenburg, SE-41296, Sweden*

^b*TU Wien, Institute of Chemical, Environmental and Bioscience Engineering, Getreidemarkt 9, Vienna, 1060, Austria*

^c*Chalmers University of Technology, Department of Technology, Management and Economics, Vera Sandbergs Allé 8, Gothenburg, SE-41133, Sweden*

Abstract

This work proposes a systematic method consisting of state-of-the-art text processing approaches and human-machine interaction for the extraction of useful sentences and data in tabular, graphical, and numerical form, containing information particularly relevant for hybrid modelling. It is applied to the domain of acute aquatic toxicity of chemicals, which is particularly relevant for the safety, health, and environmental hazard assessment of chemicals. Nearly 400 papers from 2000-2021 were identified and processed with the proposed method. The results indicate that the vast amount of knowledge can be efficiently processed in orders of magnitude faster than conventional methods without loss of detail and interpretation depth. The information is in a form that can be useful in hybrid modelling with respect to model and predictor selection, prioritization, and constraints, addressing data gaps, and validating and interpreting model performance.

Keywords: machine learning, text mining, sustainability

1. Introduction

Properties of molecules lie in the heart of any sustainability assessment method. Predictive models, developed via typical statistical regression or advanced machine learning (ML) approaches, reduce the amount of experimental testing but strongly depend on availability of pre-existing data. Limitations on data often hinder the development of robust models, leading often to overfitting of model parameters and decreasing their applicability domain (Galushka et al., 2021). This can be, for instance, of great importance during computer-aided design of novel materials. An approach combining ML methods with knowledge existing in the field, typically called hybrid modelling, might be one of the solutions to this problem (Willard et al., 2020). However, the amount of work required to process the enormous volumes of information might hinder the utilization of the existing knowledge. In different disciplines the expertise has been accumulated for years in form of scientific publications. The enormous volumes of information exist in almost every domain and researchers have been striving to reduce the amount of manual work required to process this information (Shahid et al., 2020). Nevertheless, systematic and to the extent possible fast methods of knowledge extraction from the vast amount of scientific literature are still lacking.

This step resulted in the identification of around 400 publications. Analysis of bibliometric information of the collected articles is presented in Figure 1 using a term map for words with minimum co-occurrence equal to ten. The map is based on text data of titles and abstracts of the articles. The larger the circle the more frequent the term is. More correlated terms have a shorter distance between the circles. The lines between the terms represent co-occurrence links. The earlier studies seem to be dealing with modes and mechanisms of toxic action, analysis of the relationships between structure, molecular descriptors, and toxicity. The later publications are addressing prediction models and their performance. Overall, model-related topics also seem to dominate the research area of the collection of articles.

2.2. Text mining

A python-based package was developed containing several modules that automate the text mining step consisting of three main parts: extraction of article texts and single sentences, key phrase extraction, and extraction of the relevant sentences. Single sentences were identified and checked for completeness using the library for natural language processing spaCy (<https://explosion.ai/blog/introducing-spacy>). From the complete sentences relevant sentences were extracted on the basis of reader-provided input. First, main terms were used to reduce the number of sentences to those that include any main terms. Then, key phrases were extracted from this reduced set of sentences using the open source python-based “pke” package (Boudin, 2016) with the implemented graph-based key-phrase extraction model SingleRank (Wan and Xiao, 2008). The sentences whose key phrases contain the main words and the connection words were extracted as relevant sentences. The main terms included words such as “toxicity”, “acute”, “LC50”, “EC50”, while the connection words such as “increase”, “decreas”, “relat”, “correlate”, “structure”, “fragment”, “class”, “significant”, “high”, “affect” etc.

2.3. Analysis of results and secondary text screening

The extracted set of the relevant sentences was then inspected based on human expertise to identify useful sentences. The potential use of the extracted knowledge defines the usefulness of the sentence. In this study, a sentence was considered useful if it contained information that could be used in hybrid predictive modeling. The useful sentences were collected either as directly extracted knowledge or used to identify articles and parts of the text for additional manual screening. The article screening was also performed to extract tables, figures, and equations since it was not yet possible to retrieve all of them automatically in a readable format. The analysis of the extracted sentences and/or article screening can be iterated based on altered input parameters (e.g., main terms, connection words, addition of articles). The information retrieved in this step in the form of useful sentences, models, figures, and tables was used for structuring knowledge in the next step.

2.4. Knowledge collection

The information extracted from the articles published in 2000-2014 (225 articles) was used to develop the initial knowledge structuring. The information extracted from the rest of the articles (165 articles) was used to demonstrate an update procedure (Figure 2). The newly extracted information is considered to compete with the previously collected if it provides the same type of information (e.g., a QSAR for the same species and class of chemicals using the same molecular descriptors as predictor variables). When this kind of information improves the results of the previous studies (e.g., a QSAR model with better performance based on a larger dataset), then the newly acquired knowledge replaces the previously collected one, otherwise it can be discarded or stored depending on the purpose of the classification. The new information could also contradict the

previously extracted one, for instance when the descriptors of the QSAR model are reported to have a positive correlation with the toxicity endpoint instead of a negative one reported by the information analyzed earlier. In that case, the decision can be to either replace the previously collected knowledge (e.g., if the new information is supported by more extensive experimental work), not to replace it but store the new information (e.g., a new and sufficiently diverse argument is provided that still lacks extensive experimental evidence) or completely discard it. If the new information complements the previously extracted information by providing additional depth (e.g., QSAR for the same class of chemicals but based on a different set of molecular descriptors or species), the new information is added to the previously collected information. If the new information does not satisfy the criteria to be considered competing or complementary knowledge, then the information is collected under a new classification category.

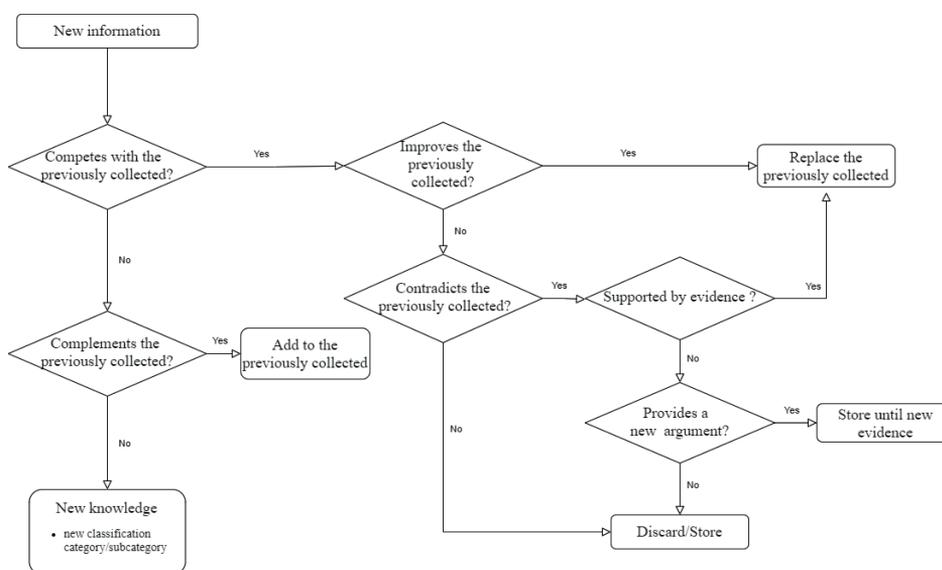


Figure 2: Update mechanism for knowledge collection and classification.

3. Results and Discussion

The method resulted in a significant reduction of text for initial reading (>85%). Most of the sentences extracted by the proposed method are directly useful for hybrid modeling or indicate to certain parts of the initial article for the manual screening. The collected knowledge was first classified to quantitative and qualitative information. The quantitative knowledge was further classified to QSAR models and experimental data and the qualitative knowledge in general statements (20) and key aspects, trends and patterns (650), single feature and structural alerts and multi-feature and combinatorial alerts (240). The QSAR models were further classified depending on the functional form (430 linear, 180 non-linear), the descriptors (600), the interspecies relationships (50), and the endpoint metric for specific species (50). Then, with respect to the applicability domain, quantitative and qualitative classes of the collected knowledge were further associated with the mode of action (i.e., narcosis, reacting, specifically acting, unidentified) and the chemical classes (40, i.e., aliphatic, aromatic, cyclic aliphatic, and application based such as pharmaceuticals, pesticides, surfactants). Most of the QSARs applied linear modeling due to its simplicity and interpretability. The nonlinear models often showed an increased

accuracy. The nonlinear models included most frequently support vector machine (SVM), k-nearest neighbors, neural networks, tree-based methods or gradient boosting. Examples of the extracted quantitative knowledge are presented in Table 1.

Table 1: Examples of quantitative knowledge extracted with respect to mode of action.

	Model type	Endpoint	Descriptors	Performance
Nonpolar narcosis	Interspecies relationship	log(1/LC50) <i>Pimephales promelas</i>	log(1/LC50) <i>Oncorhynchus mykiss</i>	n=19, R ² =0.96
Reactive chemicals	Linear modeling	log(1/EC50) <i>Pseudokirchneriella subcapitata</i>	nN>0: GATS1e, \bar{V}^+_s , SpMin1_Bh(p)	n=10, R ² =0.89
			nN=0, n(C=O)=0: logK _{ow}	n=9, R ² =0.80
Specifically acting chemicals	Linear modeling	log(1/LC50) <i>Poecilia reticulata</i>	E _a (max), $\sum C_a$, Nv1	n=31, R ² =0.77

It is generally accepted that acute toxicity is defined by the mode of toxicological action and chemical characteristics. Higher molecular toxicity values were often associated with increased lipophilicity. Most toxic compounds are related to hydrophobicity and their behavior as hydrogen-bonding or electron-acceptors. Specific functional groups like cyano, isothiocyanate, halogens were reported to enhance the toxicity of molecules, depending also on the molecular structure and position of the group in the molecule. Halogens are reported to increase toxicity. However, the contribution is more significant if the halogen is activated. Examples of qualitative knowledge are presented in Table 2.

Table 2: Examples of qualitative knowledge extracted with respect to molecular features reported to increase or decrease toxicity.

	Toxicity increase	Toxicity decrease
Property	Lipophilicity, hydrophobicity, increase	electrophilicity
		LogPo/w < 2 and ΔE (LUMO-HOMO) > 9 eV
Structure	Cyano, isothiocyanate, halogens (enhanced by activation (e. g., adjacent to an ester or other unsaturation)), amino group, nitro group, nitrile, disulfide, phosphoric acid derivatives, pyrazolyl group, formamide groups, ring aromaticity, sulfur, aromatic esters vinyl moiety, double and triple bonds, acrylate, carbamate groups	Higher polarity substitution of a hydrophobic group Presence of nitrogen in sp ² state Simultaneous presence of Sulfur together with double bond

The extracted knowledge can guide predictor variables selection, prioritization, or even identify which new predictor variables may add valuable information. Experimental data collected from various studies can serve for training or external validation of the developed models. Interspecies correlations can be used for closing the data gaps in a toxicity dataset of certain species. Data on the species' sensitivity, outliers could be used to explain the results of the prediction or deviations in the predictions. QSAR equations

or alerts could be directly integrated into the training phase of the data science models. The previously observed aquatic toxicity trends and patterns could be helpful for the evaluation of the results obtained by the developed models. Table 3 conceptually summarizes examples of the potential use of the extracted knowledge.

Table 3: Examples of potential use of the extracted knowledge.

	Model parameter selection	Predictor selection	Model constraints	Addressing data gaps	Model validation	Interpretation of results
QSAR models	√		√		√	√
QSAR descriptors		√				√
QSAR statements	√	√			√	√
Interspecies relationships				√	√	√
Experimental data	√			√	√	
Alerts		√	√			√
Applicability domain						√

4. Conclusions

The work performed the extraction and collection of knowledge existing in the field of acute aquatic toxicity. The procedure reduces the amount of manual work required to process a high number of scientific articles while extracting both generic and case-specific models, statements and alerts. The overall collected knowledge and its classification could be useful in hybrid modelling studies with respect to model and predictor selection, prioritization, and constraints, addressing data gaps, and validating and interpreting model performance. Thus, this work demonstrates one more aspect of digitalisation in PSE, in particular regarding knowledge about sustainability related properties, which is indispensable for multicriteria evaluation of process design.

References

- F. Boudin, 206, pke: An Open Source Python-Based Keyphrase Extraction Toolkit, COLING - 26th International Conference on Computational Linguistics, Proceedings, 69–73.
- M. Galushka, C. Swain, F. Browne, M.D. Mulvenna, R. Bond, D. Gray, 2021, Prediction of Chemical Compounds Properties Using a Deep Learning Model, Neural Computing and Applications, 33, 13345-13366.
- A.I. Papadopoulos, G. Shavaliieva, S. Papadokostantakis, P. Seferlis, F.A. Pedromo, A. Galindo, G. Jackson, C.S. Adjiman, 2020, An Approach for Simultaneous Computer-Aided Molecular Design with Holistic Sustainability Assessment: Application to Phase-Change CO₂ Capture Solvents, Computers and Chemical Engineering, 135, 106769.
- A. Shahid, M.T. Afzal, M. Abdar, M.E. Basiri, X. Zhou, N.Y. Yen, J.W. Chang, 2020, Insights into Relevant Knowledge Extraction Techniques: A Comprehensive Review, The Journal of Supercomputing, 76, 9, 1695-1733.
- X. Wan, J. Xiao, 2008, CollabRank: Towards a Collaborative Approach to Single-Document Keyphrase Extraction, COLING 2008 - 22nd International Conference on Computational Linguistics, Proceedings, 969–976.
- J. Willard, X. Jia, S. Xu, M. Steinbach, V. Kumar, 2020, Integrating Physics-Based Modeling with Machine Learning: A Survey, ArXiv, 2003.04919v4.

Machine Learning for the prediction of the thermochemical properties (enthalpy and entropy of formation) of a molecule from its molecular descriptors

Cindy Trinh^a, Dimitrios Meimaroglou^{a,*}, Silvia Lasala^a and Olivier Herbinet^a

^a*Université de Lorraine, CNRS, LRGP, 1 rue Grandville, 54000 Nancy, France
dimitrios.meimaroglou@univ-lorraine.fr*

Abstract

In this work, the use of ML methods for the prediction of the enthalpy and entropy of formation of molecules is investigated on the basis of different molecular and structural properties of these molecules, also known as "descriptors". Accordingly, different families, containing between 19 and 247 molecules, are extracted from the DIPPR (Design Institute for Physical Properties) database and used to train Support Vector Regression (SVR) models. 5666 molecular descriptors are calculated using dedicated software and a preliminary data preprocessing work is also carried out. The first results are very promising for the prediction of the enthalpy of formation of almost all the considered families of molecules, while the trained models for the prediction of the entropy of formation display lower accuracy and higher variations among the different families. This work, which is still under development, is carried out within the framework of a larger project aiming at the design and discovery of new molecules that meet specific requirements for use in thermodynamic cycles.

Keywords: machine learning, enthalpy of formation, entropy of formation, molecular descriptors, feature selection

1. Introduction

In recent years, machine learning (ML) methods have shown very good performance in tackling particularly complex problems, extending beyond the domains of computer science and automatic recognition. In this sense, they represent an interesting alternative to commonly employed quantum chemistry (QC) or group contribution (GC) methods, for the prediction of the thermochemical properties of molecules. Indeed, although QC and GC methods have been largely employed to calculate these properties due to their accuracy and low computational cost respectively, their applicability is often limited to small or simple molecules. QC methods see their computational cost increase drastically with the size and/or complexity of the molecules. With GC methods, on the other hand, the main difficulty lies in the decomposition of large or complex molecules into known (i.e., already identified) groups as well as in the necessity to account for important interactions (Li et al. (2021)). Inversely, data-driven ML methods have already demonstrated their ability to tackle complex problems in other fields, when classical approaches fail or are inefficient (Trinh et al. (2021)). Additionally, the use of diverse approaches (QC, GC, ML...) can also be proven extremely valuable in confirming the validity and reliability of the predicted values in the absence of relevant experimental data. Several works have already been published on the use of ML methods for the prediction of thermochemical properties of molecules (Yalamanchi et al. (2019, 2020); Dobbelaere et al. (2021); Plehiers et al. (2021)). However, these works are often

focused to specific families and/or properties of interest. The present work, investigates the use of ML methods for the prediction of both the enthalpy and entropy of formation of molecules from a large number of different families. In this respect, an exhaustive analysis of different eventual dimensionality reduction approaches, for the optimal selection of the most relevant molecular characteristics for each property, is carried out and several regression techniques are tested. The work is part of a larger project aiming at designing reactive working fluids for thermodynamic cycles. Since the work is still ongoing, this paper describes the implemented methodology along with some preliminary results, as well as the envisioned future steps for model improvement.

2. Methods

2.1. Data set

Public data from the Design Institute for Physical Properties database (DIPPR), containing 2230 molecules in SMILES notation and classified in different families and subfamilies, were used to train different ML prediction models. In particular, this work was focused on alcohols, alkanes, alkenes, alkynes, aromatics, esters, ethers, halogen compounds and sulfur compounds, as these families present a specific interest with respect to the objective of discovering molecules for implementation in thermodynamic cycles. The number of molecules per family varied significantly (i.e., between 19 and 247 molecules), as shown in Table 1, which can significantly affect the choice and performance of the ML technique. As for the enthalpy and entropy of formation, they have an average uncertainty of respectively 3 kJ/mol and 7 J/mol/K, for the considered molecules.

Table 1: Classification of DIPPR molecules per family

Family	Alcohols	Alkanes	Alkenes	Alkynes	Aromatics	Esters	Ethers	Halogen compounds	Sulfur compounds
Molecules	168	171	144	19	153	213	67	247	99

2.2. Descriptors calculation

The molecular descriptors are different properties, characteristic of the molecular and topological characteristics of the different molecules, that are commonly employed in similar regression studies. They can be calculated by means of different software tools, such as PaDEL (Yap (2010)), RDKit (Landrum (2021)), CDK (Steinbeck et al. (2003)) or AlvaDesc (Mauri (2020)), on the basis of a standardised description of the molecules, such as their SMILES (Simplified Molecular Input Line Entry System) notation. In this work, two free-source (PaDEL and RDKit) and one closed-source (AlvaDesc) tools were tested. Among them, AlvaDesc was finally retained on the basis of different factors, such as the number of calculated molecular descriptors (i.e., 5666 descriptors were provided by AlvaDesc), the robustness in the provided results, the execution speed and the proposed documentation and support. Note that, some geometrical descriptors required information that cannot be provided via the SMILES notation (i.e., related to the 3D conformation of the molecules). It was therefore necessary to convert the molecules' SMILES notation to a "MDL MOL" standard, prior to importing them into AlvaDesc, which was realized using the RDKit ETKDG method for conformer generation (Riniker and Landrum (2015)). The ETKDG method, which is a stochastic method using knowledge-based and distance geometry algorithms, is considered to be an accurate fast conformer generation method, especially for small molecules (Hawkins (2017)).

2.3. Data preprocessing and feature selection

The use of well-curated data sets is of profound importance to the implementation of any data-driven technique. In this sense, an important step of this study consisted in preparing the data

(e.g., removal of molecules or descriptors with missing values) and selecting relevant molecular descriptors (i.e., features influencing the targeted thermochemical properties) for the different molecules. This step is crucial as using data with missing, constant, redundant or irrelevant values can affect significantly the performance of the ML algorithm. For example, constant descriptors do not display any effect on the target property and, thus, should not be considered by the algorithm. In addition, the selection of only a subset of the most relevant descriptors will help to decrease the computational effort and prevent overfitting phenomena (Bommert et al. (2020)).

Accordingly, the first step of the performed data preparation procedure was to remove molecules with missing property values as well as descriptors with missing, constant, or quasi-constant values (i.e., when at least 95% of their values were identical). Finally, Pearson correlation coefficients were calculated for all descriptors and the ones that were correlated to more than 98% to another descriptor were removed. For the family of alkanes, the final number of descriptors obtained after the implementation of the aforementioned steps was reduced from 5666 to 586.

Following these initial data treatment steps, the implementation of feature selection methods is necessary to further reduce the number of descriptors and assess their effect on the targeted properties. Feature selection refers to a class of dimensionality reduction techniques, consisting in identifying a subset of the initial features that best represent the property of interest. These methods can be classified into different groups (i.e. filter methods, wrapper methods and embedded methods) (Danishuddin and Khan (2016); Saeys et al. (2007)), according to the adopted statistical tests and to whether the correlation between the features and the performance of the algorithm is taken into account. For example, filter methods rely only on the intrinsic properties of the data and use feature relevance score to identify the most significant features while wrapper methods select the relevant features by calculating the feature impact on the performance of a learning algorithm (Danishuddin and Khan (2016)). In the context of this work, the implementation of several feature selection techniques is investigated. This work is currently underway.

2.4. Training, hyper-parameters optimization and performance evaluation

An initial test of the implementation of a regression technique was performed on the data after the pretreatment steps described in the previous paragraph. Support Vector Regression (SVR) was selected as it can perform with high precision and generalization with a small number of training samples, high dimensional and noisy data (Trinh et al. (2021)). Prior to the implementation of SVR, the data within each family of molecules was randomly shuffled (i.e., due to the fact that the original data from DIPPR were grouped into sub-families, within each family). The data was then partitioned into distinct groups, following a k -fold cross-validation principle (with $k=5$), where for each fold k , the k^{th} data subset served for testing the performance of the algorithm, this latter trained on the basis of the remaining $k-1$ groups. This procedure ensures a representative use of the data and allows assessing the average overall regression performance of the trained model. The value for k was selected so that 20% of the available data, for each family, is implemented for testing within each fold. Generally, $k=5$ or $k=10$ ensures a good trade-off between low computational cost, low bias and low variance for the evaluation of the model performance (Kohavi (1995)).

In parallel to the training of the SVR model, a set of hyperparameters (i.e. kernel function, box constraint, kernel scale, epsilon, kernel polynomial order, etc.) of the algorithm were optimized within each fold (i.e., using the *OptimizeHyperparameters* option of the *fitrsvm* function of the Matlab toolbox). A more detailed description of the hyperparameters can be found in Vapnik (1995). In this sense, within the $k-1$ data sets, used for model training, a secondary j -fold partition took place (with $j=5$) for further identifying and separating a validation data set, serving for the identification procedure of the algorithm hyperparameters. The performance of the algorithm was evaluated in terms of the typically employed performance indexes (i.e., RMSE, MAE and R^2). Matlab was used for all the above developments.

3. Results and discussion

The obtained performance results and the corresponding optimized hyperparameters at each fold for the SVR algorithm, as implemented for the prediction of the enthalpy of alkanes, is given in Table 2. As can be seen, the mean value of the coefficient of determination R^2 for training-validation and test data sets are both quite satisfactory as they are respectively equal to 0.999 and 0.984.

Table 2: Performances for enthalpy prediction for alkanes for each fold k and with the associated optimized hyperparameters.

Fold	RMSE (kJ/mol)		MAE (kJ/mol)		R ²		Kernel function	Box constraint	Kernel scale	Epsilon
	Training-Validation	Test	Training-Validation	Test	Training-Validation	Test				
1	7.866	12.386	7.065	9.290	0.997	0.982	'linear'	987.219	1	10.228
2	4.085	20.640	3.830	13.102	0.999	0.979	'linear'	409.623	1	4.680
3	1.405	31.040	1.370	16.993	1.000	0.977	'linear'	280.166	1	1.476
4	2.813	16.515	2.724	8.098	1.000	0.988	'polynomial order 2'	3.032	1	3.062
5	8.191	13.393	2.288	9.770	0.997	0.995	'gaussian'	982.251	12.928	0.411
Mean	4.872	18.795	3.455	11.451	0.999	0.984				

Table 3: Performances for enthalpy prediction for different families of molecules and their combination (color code for R^2 : green: > 0.96; red: < 0.90; orange: in between).

Family of molecules	Number of molecules	Number of descriptors	ENTHALPY					
			RMSE (kJ/mol)		MAE (kJ/mol)		R ²	
			Training-Validation	Test	Training-Validation	Test	Training-Validation	Test
Alcohols	166	940	2.130	33.695	1.782	21.309	1.000	0.976
Alkanes	171	586	4.872	18.795	3.455	11.451	0.999	0.984
Alkenes	144	837	5.039	21.346	2.454	11.948	0.996	0.980
Alkynes	19	647	0.601	20.571	0.578	18.105	1.000	0.964
Aromatics	153	946	8.050	25.114	3.851	14.958	0.997	0.974
Esters	213	554	192.176	278.421	44.170	92.658	0.737	0.557
Ethers	67	1250	0.200	30.160	0.200	20.284	1.000	0.981
Halogen Compounds	244	1496	61.155	104.240	13.120	60.027	0.988	0.982
Sulfur Compounds	99	1344	2.031	23.401	1.958	16.461	1.000	0.980
All	1276	533	95.297	121.130	13.792	29.496	0.963	0.926

Table 4: Performances for entropy prediction for different families of molecules and their combination (color code for R^2 : green: > 0.96; red: < 0.90; orange: in between).

Family of molecules	Number of molecules	Number of descriptors	ENTROPY					
			RMSE (kJ/mol)		MAE (kJ/mol)		R ²	
			Training-Validation	Test	Training-Validation	Test	Training-Validation	Test
Alcohols	166	940	14.874	46.674	4.924	22.745	0.984	0.905
Alkanes	171	586	15.052	25.749	3.531	11.601	0.995	0.981
Alkenes	143	831	1.526	20.125	1.435	12.422	1.000	0.994
Alkynes	19	647	1.446	17.997	1.331	13.038	0.999	0.978
Aromatics	153	946	36.472	39.118	11.592	16.401	0.927	0.912
Esters	212	552	275.855	311.044	67.420	103.181	0.640	0.514
Ethers	67	1250	2.853	34.440	0.641	21.757	0.998	0.880
Halogen Compounds	244	1497	11.030	17.676	3.577	9.946	0.991	0.968
Sulfur Compounds	99	1344	16.312	34.100	7.217	22.145	0.968	0.897
All	1274	533	124.520	132.311	22.054	29.676	0.780	0.751

The overall performance of the models trained for enthalpy and entropy of formation prediction, for the different families of molecules as well as for the combination of all of them, is shown in

Tables 3 and 4. For the prediction of the enthalpy of formation, the performance of the models on all the molecule families, except that of esters, displayed a value for the determination coefficient superior to 0.96. On the other hand, the predictive capabilities of the trained SVR models for the entropy of formation was satisfactory only for the families of alkanes, alkenes, alkynes and halogen compounds, while some families (i.e., esters) displayed a significantly low model predictive capacity. However, it was more generally observed that considering the molecules per family resulted in better prediction performances. Note that, these are only preliminary results obtained using a single regression ML technique and after only a preliminary data pretreatment step. In addition, some families (e.g., esters) contain some very large molecules (e.g., glycerides such as tripalmitin or tristearin) for which the conformer generation is quite complex, introducing additional uncertainty in the descriptors calculation step, which needs to be further analyzed (Wang et al. (2020); Cleves and Jain (2017)).

Finally, some preliminary tests on the implementation of some feature selection methods were implemented to identify the most influencing descriptors for the family of alkanes. The tested methods are: univariate feature ranking for regression using F-tests, importance ranking of predictors using RReliefF algorithm, Pearson correlation coefficient (between a descriptor and the target) and low-variance. The two first methods were available as Matlab functions and were chosen according to the supported problem and data type while the two other methods are commonly encountered feature selection techniques (Pedregosa et al. (2011); Mangal and Holm (2018)). The top 5 ranking descriptors with highest scores, resulting from these different filter techniques, are presented in Table 5. It can be observed that very different descriptors have been identified by the different techniques. Although filter techniques are simple and fast to implement, their drawback is that they rank features independently from the ML algorithm. Conversely, the comparison with wrapper methods (such as forward selection or backward elimination) could be very interesting.

Table 5: Top 5 descriptors with highest scores using different filter techniques (*the bold italic indications correspond to the descriptors category as defined by AlvaDesc*).

	Univariate feature ranking for regression using F-tests	Importance ranking of predictors using RReliefF algorithm
1	<i>Information indices</i> Structural Information Content index (neighborhood symmetry of 0-order)	<i>Drug-like indices</i> Modified drug-like score from Zheng et al. (2 rules)
2	<i>Charge descriptors</i> Relative positive charge	<i>Ring descriptors</i> Normalized number of ring systems
3	<i>P_VSA-like descriptors</i> P_VSA-like on Molar Refractivity, bin 5	<i>Walk and path counts</i> Self-returning walk count of order 5
4	<i>Information indices</i> Complementary Information Content index (neighborhood symmetry of 1-order)	<i>Charge descriptors</i> Submolecular polarity parameter
5	<i>Information indices</i> Complementary Information Content index (neighborhood symmetry of 0-order)	<i>Charge descriptors</i> Maximum positive charge
	Pearson correlation coefficient	Low-variance
1	<i>Information indices</i> Complementary Information Content index (neighborhood symmetry of 1-order)	<i>Chirality descriptors</i> Number of neighbouring atoms of the chiral centre (level 1)
2	<i>P_VSA-like descriptors</i> P_VSA-like on Molar Refractivity, bin 5	<i>Ring descriptors</i> Normalized number of ring systems
3	<i>3D-MoRSE descriptors</i> Signal 21 / unweighted	<i>3D autocorrelations</i> 3D Topological distance based descriptors – lag 7 unweighted
4	<i>RDF descriptors</i> Radial Distribution Function – 030 / unweighted	<i>3D autocorrelations</i> 3D Topological distance based descriptors – lag 8 unweighted
5	<i>RDF descriptors</i> Radial Distribution Function – 020 / unweighted	<i>2D Atom Pairs</i> Presence/absence of C – C at topological distance 9

4. Conclusions

In this work, SVR data-driven models were developed to predict the enthalpy and entropy of formation for different families of molecules, each one containing between 19 and 247 molecules extracted from the DIPPR database. 5666 molecular descriptors were calculated using dedicated software and some preliminary work was done on the data preprocessing and feature selection. The obtained results are especially promising for the prediction of the enthalpy of formation for

the majority of families, whilst the prediction of the entropy of formation displays higher variations between the different molecular families. Further work is currently underway for the improvement of the model performance, especially in terms of a more thorough analysis of feature selection, model selection and hyperparameter optimization. More generally, this work is carried out within the framework of a long-term project, whose ultimate objective is to employ the developed ML models for the design and discovery of new molecules that meet specific requirements for use in thermodynamic cycles.

References

- A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, M. Lang, 2020. Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics and Data Analysis* 143, 106839.
- A. E. Cleves, A. N. Jain, 2017. ForceGen 3D structure and conformer generation: from small lead-like molecules to macrocyclic drugs. *Journal of Computer-Aided Molecular Design* 31 (5), 419–439.
- Danishuddin, A. U. Khan, 2016. Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug Discovery Today* 21 (8), 1291–1302.
- M. R. Dobbelaere, P. P. Plehiers, R. Van de Vijver, C. V. Stevens, K. M. Van Geem, 2021. Learning Molecular Representations for Thermochemistry Prediction of Cyclic Hydrocarbons and Oxygenates. *The Journal of Physical Chemistry A*.
- P. C. Hawkins, 2017. Conformation Generation: The State of the Art. *Journal of Chemical Information and Modeling* 57 (8), 1747–1756.
- R. Kohavi, 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference of Artificial Intelligence* (June).
- G. Landrum, 2021. Rdkit: Open-source cheminformatics software. <http://www.rdkit.org>.
- Q. Li, G. Wittreich, Y. Wang, H. Bhattacharjee, U. Gupta, D. G. Vlachos, 2021. Accurate Thermochemistry of Complex Lignin Structures via Density Functional Theory, Group Additivity, and Machine Learning. *ACS Sustainable Chemistry and Engineering* 9 (8), 3043–3049.
- A. Mangal, E. A. Holm, 2018. A Comparative Study of Feature Selection Methods for Stress Hotspot Classification in Materials. *Integrating Materials and Manufacturing Innovation* 7 (3), 87–95.
- A. Mauri, 2020. alvaDesc: A tool to calculate and analyze molecular descriptors and fingerprints. *Methods in Pharmacology and Toxicology* (January 2020), 801–820.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- P. P. Plehiers, I. Lengyel, D. H. West, G. B. Marin, C. V. Stevens, K. M. Van Geem, 2021. Fast estimation of standard enthalpy of formation with chemical accuracy by artificial neural network correction of low-level-of-theory ab initio calculations. *Chemical Engineering Journal* 426 (July), 131304.
- S. Riniker, G. A. Landrum, 2015. Better Informed Distance Geometry: Using What We Know to Improve Conformation Generation. *Journal of Chemical Information and Modeling* 55 (12), 2562–2574.
- Y. Saeys, I. Inza, P. Larrañaga, 2007. A review of feature selection techniques in bioinformatics. *Bioinformatics* 23 (19), 2507–2517.
- C. Steinbeck, Y. Han, S. Kuhn, O. Horlacher, E. Luttmann, E. Willighagen, 2003. The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics. *Journal of Chemical Information and Computer Sciences* 43 (2), 493–500.
- C. Trinh, D. Meimaroglou, S. Hoppe, 2021. Machine learning in chemical product engineering: The state of the art and a guide for newcomers. *Processes* 9 (8).
- V. N. Vapnik, 1995. *The Nature of Statistical Learning*.
- S. Wang, J. Witek, G. A. Landrum, S. Riniker, 2020. Improving Conformer Generation for Small Rings and Macrocycles Based on Distance Geometry and Experimental Torsional-Angle Preferences. *Journal of Chemical Information and Modeling* 60 (4), 2044–2058.
- K. K. Yalamanchi, M. Monge-Palacios, V. C. Van Oudenhoven, X. Gao, S. M. Sarathy, 2020. Data Science Approach to Estimate Enthalpy of Formation of Cyclic Hydrocarbons. *Journal of Physical Chemistry A* 124 (31), 6270–6276.
- K. K. Yalamanchi, V. C. Van Oudenhoven, F. Tutino, M. Monge-Palacios, A. Alshehri, X. Gao, S. M. Sarathy, 2019. Machine Learning to Predict Standard Enthalpy of Formation of Hydrocarbons. *Journal of Physical Chemistry A* 123 (38), 8305–8313.
- C. W. Yap, 2010. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *Journal of computational chemistry* 32, 174–182.

A Novel Machine Learning-Based Optimization Approach for the Molecular Design of Solvents

Zihao Wang^a, Teng Zhou^{a,b,*}, Kai Sundmacher^{a,b}

^a *Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, D-39106 Magdeburg, Germany*

^b *Process Systems Engineering, Otto-von-Guericke University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany*

zhout@mpi-magdeburg.mpg.de

Abstract

In this contribution, we propose a new CAMD approach for solvent design by combining machine learning with deterministic optimization. Variational autoencoder (VAE), a powerful generative machine learning method, is used to transfer a molecular structure into a continuous latent vector with an encoder and to convert the latent vector back to the molecule with a decoder. Solvent properties of interest are estimated by a feedforward neural network (FNN) using the latent vector as input. Through the joint training of the VAE and FNN-based property model with collected data, a continuous latent design space can be constructed for the optimization-based solvent molecular design. Specifically, nonlinear optimization is first performed to identify optimal latent variables featuring desirable solvent properties based on the established property prediction model. Knowing the optimal latent vector, the corresponding target molecule is then generated by the pre-trained VAE decoder. The separation of 1-butene and butadiene is used as an example to investigate the reliability of the proposed design method. A set of solvent candidates showing superior separation performance are generated. To further filter suitable solvents for industrial applications, additional analyses on molecular synthesizability are conducted.

Keywords: molecular design, solvent discovery, machine learning, variational autoencoder, nonlinear optimization

1. Introduction

The selection of suitable solvents is of central importance in separation processes, such as gas purification and extractive distillation, for achieving a high product purity and reducing energy consumption. High-throughput computational screening can help discover solvent candidates showing desirable properties. Unfortunately, the huge number of possible organic molecules (Bohacek et al., 1996) makes the brute force solvent screening extremely costly and time-consuming. By contrast, computer-aided molecular design (CAMD) of solvents based on available property estimation models (Song et al., 2020; Wang et al., 2020) is much more efficient for direct targeting of optimal solvents possessing certain desirable properties.

In the past decade, the CAMD method has been developed and widely used for solvent design from a given set of molecular building blocks under the constraints of molecular structural feasibility and property specifications (Song et al., 2018; Scheffczyk et al., 2017; Papadopoulos et al., 2020; Zhou et al., 2021; Zhang et al., 2021). In most of the

previous CAMD studies, group contribution (GC) based property models are typically used for solvent property prediction (Adjiman et al., 2021). However, one should note that standard GC models cannot accurately predict certain properties, and more importantly, they cannot well distinguish between structural isomers. These limitations can be potentially tackled by combining advanced machine learning (ML) methods with special molecular representations such as chemical text (e.g., SMILES string) and molecular graph.

Variational autoencoder (VAE), a powerful generative ML method, can generate simple but informative latent variables from complex inputs (such as image, text, and graph) and reconstruct these inputs from the latent variables (Sanchez-Lengeling and Aspuru-Guzik, 2018; Gómez-Bombarelli et al., 2018). In this contribution, we propose a new CAMD approach for solvent design by combining VAE with deterministic optimization. Using SMILES strings as input, we first build a VAE to transform molecules into latent variables. Using these variables as input, a feedforward neural network (FNN) model is trained to predict molecular properties of interest. Nonlinear optimization is finally performed to identify optimal latent variables and the corresponding molecules featuring desirable properties. The proposed CAMD approach is applied to 1-butene and butadiene separation. A set of solvent candidates with high separation performance and molecular synthesizability are identified.

2. Method

The CAMD approach consists of two stages, namely generative modeling and molecular design, as shown in Figure 1. In the first stage, the VAE model is built to freely transform between molecules and continuous latent variables. Using the latent variables as input, an FNN model is trained to predict molecular properties of interest. In the next molecular design stage, based on the property prediction model, a nonlinear optimization problem is formulated and solved to identify the optimal latent variables featuring desirable properties. By decoding the optimal latent variables with the VAE decoder, the corresponding best molecular structure can be finally generated.

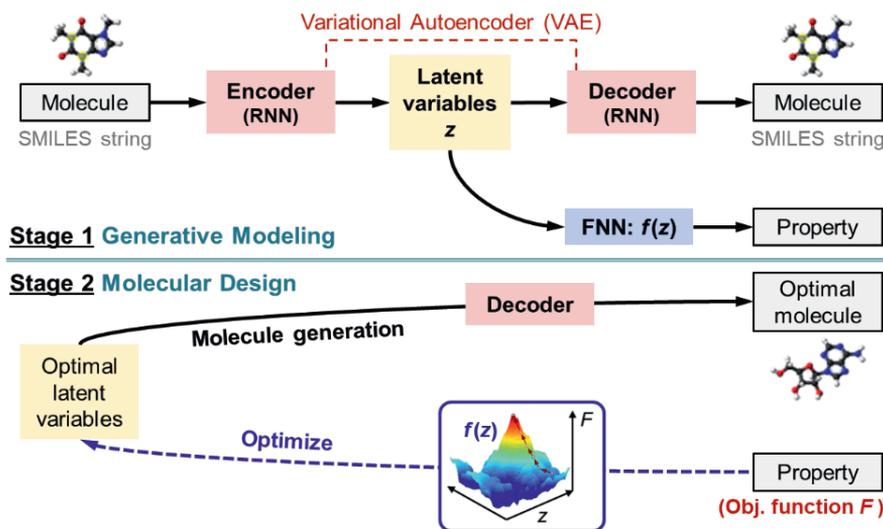


Figure 1. The proposed ML-based optimization approach for molecular design

2.1. Generative modeling

The VAE model is composed of two components, an encoder and a decoder. Starting from the molecular structure (represented by the SMILES string), the encoder converts it into a continuous latent vector while the decoder enables the reproduction of the original molecule from the latent vector. In this work, both encoding and decoding are achieved by recurrent neural networks (RNNs). Using the latent vector as input, molecular properties of interest are predicted by another FNN model. Notably, the VAE and FNN models are trained simultaneously so that the molecules close to each other in the latent space exhibit similar properties.

2.2. Molecular design

Relying on the FNN model, the mathematical relation between latent variables and molecular properties is obtained. With an objective function defined on molecular properties of interest, a nonlinear optimization problem is formulated and solved to identify the optimal latent variables. The corresponding SMILES string is generated by decoding the latent variables and the molecular structure is reconstructed using RDKit (2021).

3. Results and Discussion

C4 hydrocarbons (mostly consisting of 1-butene, 2-butene, isobutene, and butadiene) are mainly co-produced by the thermal cracking of various feedstock such as naphtha and gasoil (Streich et al., 2016). These unsaturated components are important intermediates in the chemical industry. For example, butadiene is a monomer in the manufacture of synthetic rubbers, and therefore it needs to be purified from the C4 mixture. However, similar physicochemical properties of these components make them difficult to separate by conventional methods. Adding a suitable solvent into such a mixture to change the relative volatility, extractive distillation can be conducted to achieve efficient separation.

In order to illustrate our proposed method for solvent design, the separation of simplified C4 mixtures (butadiene and 1-butene) is carried out as a case study. Because butadiene generally shows strong interaction with organic solvents in contrast to 1-butene, a suitable solvent can be designed to efficiently extract butadiene from the mixture. Two important solvent thermodynamic properties (i.e., selectivity and capacity at the infinite dilution condition), indicating the separation performance of the solvent, are considered (Zhou et al., 2019). The infinite dilution selectivity is defined by

$$S_{C_4H_6/C_4H_8}^\infty = \gamma_{C_4H_8}^\infty / \gamma_{C_4H_6}^\infty$$

The infinite dilution capacity of the solvent toward butadiene is determined by

$$C_{C_4H_6}^\infty = 1 / \gamma_{C_4H_6}^\infty$$

As represented above, both the selectivity and capacity are calculated with activity coefficients at infinite dilution. Notably, an increase in the selectivity usually leads to a decrease in capacity (Krummen et al., 2000). Thus, the product of selectivity and capacity is used as the objective function for solvent design.

3.1. Model Training

A dataset consisting of 2,496 solvents is adopted, with infinite dilution activity coefficients of 1-butene and butadiene calculated using COSMO-RS (Klamt and Eckert, 2000). The training and test sets are randomly selected and they account for 90% and 10% of the dataset respectively. ML models (i.e., VAE and FNN) are constructed using the

open-source ML framework TensorFlow (2021), and they are optimized on the training set and evaluated using the test set. Table 1 summarizes the accuracy of the VAE model for molecular reconstruction and the performance (i.e., mean absolute error (MAE) and coefficient of determination (R^2)) of the FNN model for property prediction. As indicated, both ML models present high performance on both training and test sets. Thus, for an unknown solvent represented by a latent vector, the decoder of the VAE model can reliably reconstruct its molecular structure and the FNN model can accurately predict its infinite dilution activity coefficients toward 1-butene and butadiene. This guarantees the reliability of the solvent design results.

Table 1. Model performance in the molecular reconstruction and property prediction

Dataset	Molecular reconstruction accuracy	$\ln(\gamma_{C_4H_8}^\infty)$		$\ln(\gamma_{C_4H_6}^\infty)$	
		MAE	R^2	MAE	R^2
Training	90.91%	0.0833	0.9504	0.0769	0.9221
Test	89.24%	0.0896	0.9625	0.0886	0.9323

3.2. Solvent design

To achieve energy-efficient extractive distillation of butadiene and 1-butene, the solvent should be optimally designed by maximizing its separation performance (i.e., the product of selectivity and capacity). Based on the FNN model, we first identify the optimal latent vector by solving a nonlinear optimization problem where the solvent performance is maximized. Later, the VAE decoder is used to convert the optimal latent vector into solvent molecular structure. In this way, a new solvent (the red star) showing superior separation performance is designed with the objective function of 7.930, as shown in Figure 2. It is significantly better than the best solvent in the employed dataset (blue star, objective function 5.413). The grey dots indicate the locations of all the 2,496 solvents in the latent space visualized in two dimensions using principal component analysis.

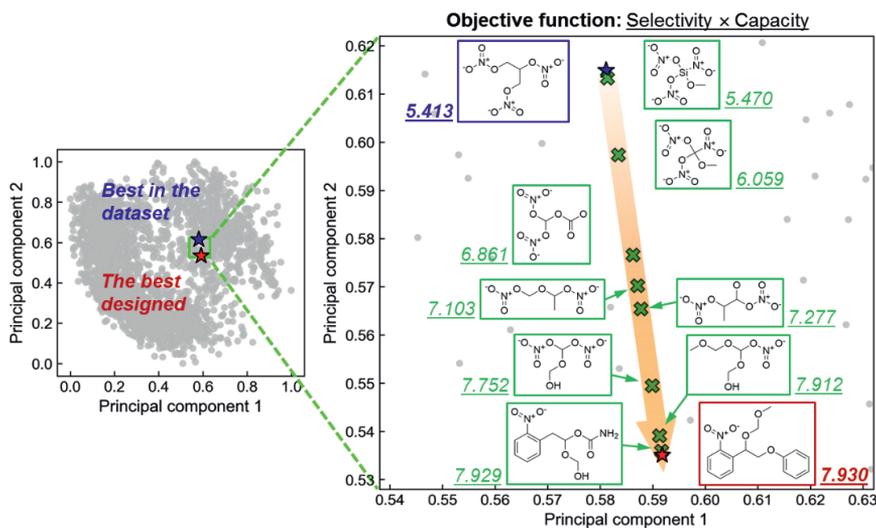


Figure 2. Interpolation-based solvent design for the separation of butadiene and 1-butene

When training the VAE model, the latent space tends to cluster molecules with similar structures. Meanwhile, because the VAE model is jointly trained with the FNN model,

the distribution of molecules in the latent space is indicated by the molecular property. In this case, molecules locating closely in the latent space should also share similar properties. Taking these into account, interpolation is performed between the best solvent in the original database (blue star) and the optimally designed solvent (red star) to generate more solvent candidates (green crosses), as indicated in Figure 2. Following the route, an increase in the solvent separation performance is identified, and structural similarity is also observed between neighboring solvents. This demonstrates that the latent space allows smooth navigation of solvent performance because of its local similarity on both solvent structure and property.

To further filter suitable solvents from the above-designed candidates, additional analyses on molecular synthesizability are carried out. The synthetic accessibility score (SAscore) (Ertl and Schuffenhauer, 2009) can estimate the ease of synthesis of compounds based on available synthetic knowledge and molecular structural complexity, and it varies between 1 (easy to synthesize) and 10 (very difficult to synthesize). Table 2 lists the SAscore for the top five solvent candidates presented in Figure 2. In addition to SAscore, another two synthesizability indicators, Retro score (RScore) and Retro step (RStep), are estimated from the Spaya platform (2021) based on literature reaction templates. RScore indicates the probability of the synthetic route matching with existing reaction templates, and RStep is the number of reaction steps of the considered route. Obviously, a large RScore and a small RStep are preferred. As shown in Table 2, the optimal solvent with the best separation performance also displays the highest synthesizability, as revealed by the three indicators.

Table 2. Evaluation on the synthesizability of designed solvent candidates

Rank	Molecular SMILES string	Performance	SAscore	RScore	RStep
1	<chem>COCOC(COclcccc1)clcccc1[N+](=O)[O-]</chem>	7.930	2.67	0.8	2
2	<chem>NC(=O)OC(Cc1cccc1[N+](=O)[O-])OCO</chem>	7.929	3.19	0.7	3
3	<chem>COCOC(OCO)O[N+](=O)[O-]</chem>	7.912	4.18	0.6	4
4	<chem>O=[N+](O)OC(OCO)O[N+](=O)[O-]</chem>	7.752	4.00	0.7	4
5	<chem>CC(O[N+](=O)[O-])C(O)O[N+](=O)[O-]</chem>	7.277	4.32	0.8	2

4. Conclusion

Combining ML and deterministic optimization, we propose a new CAMD approach for the molecular design of solvents. After developing accurate ML models, deterministic nonlinear optimization is first performed to identify the optimal latent variables, and the corresponding optimal solvent structure is then obtained by decoding the latent variables. Besides the best solvent, a few more candidates are identified by interpolating in the latent space and the practical applicability of all the solvent candidates is evaluated from the perspective of molecular synthesizability. The proposed CAMD approach has been successfully applied to solvent design for butadiene and 1-butene separation.

The generated molecules represented by SMILES strings sometimes are proven to be structurally infeasible by RDKit. Imposing grammar constraints in encoding and decoding steps or replacing SMILES strings with molecular graphs can potentially guarantee the feasibility of the designed molecules. Additionally, density functional theory calculations can provide further evaluations on molecular structural stability, and process design can be incorporated into the CAMD framework in order to generate solvents better serving industrial processes.

References

- C.S. Adjiman, N.V. Sahinidis, D.G. Vlachos, B. Bakshi, C.T. Maravelias, C. Georgakis, 2021, Process systems engineering perspective on the design of materials and molecules, *Ind. Eng. Chem. Res.*, 60, 14, 5194-5206.
- R.S. Bohacek, C. McMartin, W.C. Guida, 1996, The art and practice of structure-based drug design: a molecular modeling perspective, *Med. Res. Rev.*, 16, 1, 3-50.
- P. Ertl, A. Schuffenhauer, 2009, Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions, *J. Cheminf.*, 1, 1, 1-11.
- R. Gómez-Bombarelli, J.N. Wei, D. Duvenaud, J.M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T.D. Hirzel, R.P. Adams, A. Aspuru-Guzik, 2018, Automatic chemical design using a data-driven continuous representation of molecules, *ACS Cent. Sci.*, 4, 2, 268-276.
- A. Klamt, F. Eckert, 2000, COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids, *Fluid Phase Equilib.*, 172, 1, 43-72.
- M. Krummen, D. Gruber, J. Gmehling, 2000, Measurement of activity coefficients at infinite dilution in solvent mixtures using the dilutor technique, *Ind. Eng. Chem. Res.*, 39, 6, 2114-2123.
- A.I. Papadopoulos, G. Shavaliyeva, S. Papadokostantakis, P. Seferlis, F.A. Perdomo, A. Galindo, G. Jackson, C.S. Adjiman, 2020, An approach for simultaneous computer-aided molecular design with holistic sustainability assessment: Application to phase-change CO₂ capture solvents, *Comput. Chem. Eng.*, 135, 106769.
- RDKit: Open-Source Cheminformatics Software, 2021, <https://www.rdkit.org/>.
- B. Sanchez-Lengeling, A. Aspuru-Guzik, 2018, Inverse molecular design using machine learning: Generative models for matter engineering, *Science*, 361, 6400, 360-365.
- J. Scheffczyk, L. Fleitmann, A. Schwarz, M. Lampe, A. Bardow, K. Leonhard, 2017, COSMO-CAMD: A framework for optimization-based computer-aided molecular design using COSMO-RS, *Chem. Eng. Sci.*, 159, 84-92.
- Z. Song, H. Shi, X. Zhang, T. Zhou, 2020, Prediction of CO₂ solubility in ionic liquids using machine learning methods, *Chem. Eng. Sci.*, 223, 115752.
- Z. Song, C. Zhang, Z. Qi, T. Zhou, K. Sundmacher, 2018, Computer-aided design of ionic liquids as solvents for extractive desulfurization, *AIChE J.*, 64, 3, 1013-1025.
- Spaya, 2021, <https://spaya.ai/>.
- T. Streich, H. Kömpel, J. Geng, M. Renger, 2016, Secure the best benefit from C₄ hydrocarbon processing-Part 1: Separation sequences, *Hydrocarbon Process.*, 73-78.
- TensorFlow, 2021, <https://www.tensorflow.org/>.
- Z. Wang, Y. Su, S. Jin, W. Shen, J. Ren, X. Zhang, J.H. Clark, 2020, A novel unambiguous strategy of molecular feature extraction in machine learning assisted predictive models for environmental properties, *Green Chem.*, 22, 12, 3867-3876.
- X. Zhang, J. Wang, Z. Song, T. Zhou, 2021, Data-driven ionic liquid design for CO₂ capture: molecular structure optimization and DFT verification, *Ind. Eng. Chem. Res.*, 60, 27, 9992-10000.
- T. Zhou, H. Shi, X. Ding, Y. Zhou, 2021, Thermodynamic modeling and rational design of ionic liquids for pre-combustion carbon capture, *Chem. Eng. Sci.*, 229, 116076.
- T. Zhou, Z. Song, X. Zhang, R. Gani, K. Sundmacher, 2019, Optimal solvent design for extractive distillation processes: A multiobjective optimization-based hierarchical framework, *Ind. Eng. Chem. Res.*, 58, 15, 5777-5786.

A methodology for gray-box modeling of nonlinear ODE systems

Joschka Winz^{a,*} and Sebastian Engell^a

^aProcess Dynamics and Operations Group, Department of Biochemical and Chemical Engineering, TU Dortmund, Emil-Figge Str. 70, 44227 Dortmund, Germany

joschka.winz@tu-dortmund.de

Abstract

For the application of advanced process control and optimization methods, a dynamic process model is necessary. Developing a purely mechanistic white-box process model is time consuming and challenging when the underlying physical, chemical or biochemical phenomena are not fully understood. Black-box models can be used in such cases when sufficient process data is available. These models on the other hand suffer from limited accuracy for inputs that are not well represented in the collected data. For extended extrapolation capabilities and improved interpretability, gray-box models can be used where white-box and black-box parts are combined. If the black-box models describe embedded variables of the white box model, the determination of a suitable black-box model structure is a challenging task. In this work we present a methodology to systematically decompose the parameter estimation and model structure selection tasks. The methodology is applied to a fermentation use case with promising results.

Keywords: machine-learning, gray-box modeling, dynamic modeling, fermentation

1. Introduction

The problem of finding a suitable process model for a chemical or biochemical process is challenging. Oftentimes the underlying phenomena are not fully understood and only basic equations of the system dynamics that result from heat and mass balances are available where not all terms are fully specified. If by simple correlations or kinetics, the experimental data at hand cannot be matched sufficiently accurately, one way to improve the model accuracy is to embed data-based or machine learning (ML) models into the fundamental equations. This combination of traditional modelling efforts with ML approaches is commonly known as gray-box modelling.

There exists a multitude of ways to realize combinations of different types of models, an overview can be found in [1]. Here, we consider the case that a set of ordinary differential equations derived from process knowledge is available, where some variables in these equations are described by ML models. These kinds of systems have been studied by different authors already. One important distinction between different situations is whether there exists data specifically for the input/output relation that the ML model is supposed to represent, as investigated in [2], or not.

In the case that such data is not available, there are two ways to approach the issue. First, an artificial data-set can be estimated from the available dynamic data. Several techniques have been applied in the literature, for example the use of a state observer that estimates the unknown parameter [3]. [4] and [5] replace the embedded variables by functions of

time which are fitted to data to generate a training set for the parameterization of the ML model. Second, one can learn the parameters of the machine learning model by performing a full dynamic parameter estimation for the overall model, including the parameters of the embedded ML models and other parameters of the balance equations, which involves simulating the full model and minimizing the errors with respect to the experimental data.

Also a combination of the two approaches is possible, as proposed in this work.

2. Simulation model use case: Fermentation of a sporulating bacterium

We are considering the use case of a fermentation of a microorganism that undergoes sporulation. Fermentation processes are a typical example of processes where deriving a fully mechanistic model is complicated as a complete mathematical description of the cell internal metabolism would be necessary. The simulation use case involves three state variables: the dry basis concentration of vegetative cells X_v , the dry basis concentration of sporulated cells X_s and the concentration of substrate S . These state variables undergo two main reactions resulting in the following dynamics. In these equations it is assumed for simplification that the cell weight does not change and is the same for both vegetative and sporulated cells.

$$\begin{aligned}\dot{X}_v &= r_g(T, S)X_v - r_s(T, S)X_v \\ \dot{X}_s &= r_s(T, S)X_v \\ \dot{S} &= -r_g(T, S)X_v Y_{X/S}^{-1}\end{aligned}\quad (1)$$

Here, r_g and r_s denote growth and sporulation reaction that both nonlinearly depend on temperature and substrate concentration with relations taken from [6–9]. $Y_{X/S}^{-1}$ is the yield coefficient. This system is modelled using the following gray-box model structure.

$$\begin{aligned}\hat{\dot{X}}_v &= \varphi_1(T, \hat{S}) \hat{S} \hat{X}_v - \varphi_2(T, \hat{S}) \hat{X}_v \\ \hat{\dot{X}}_s &= \varphi_2(T, \hat{S}) \hat{X}_v \\ \hat{\dot{S}} &= -\varphi_1(T, \hat{S}) \hat{S} \hat{X}_v \hat{Y}_{X/S}^{-1}\end{aligned}\quad (2)$$

The main issue of fitting this model is finding an appropriate machine learning (ML) model structure for the two embedded variables φ_1 and φ_2 . To this end a methodology is developed, which is described in the next section.

3. Novel methodology for gray-box modelling of nonlinear ODE systems

In this work we discuss a methodology to find a suitable structure and parameters for dynamic gray-box models of the general structure shown in Eq. 3:

$$\dot{x} = f_{\Theta} \left(x, u, \varphi_{\Theta_{ML}}(d) \right) \quad (3)$$

Here, x denotes the state vector and u denotes a vector of input variables. Additionally, there are embedded variables φ that are described by ML models. They depend on a set of ML-model parameters Θ_{ML} . The inputs to these ML-models are denoted by the vector of descriptors d . The elements of this vector are a subset of the elements of the other variables of the system x and u . The differential equations are described by the possibly nonlinear functions f_{Θ} . These itself depend on a set of parameters Θ which occur in the mechanistic relations.

To find a suitable ML model structure for φ and values for the sets of parameters Θ and Θ_{ML} , a full dynamic parameter estimation problem can be set up as shown below.

$$\begin{aligned}
 & \min_{\Theta, \Theta_{ML}} J(\mathbf{Y}^{exp}, \mathcal{P}) \\
 & \text{s. t.} \quad \dot{x} = f_{\Theta}(x, u, \varphi_{\Theta_{ML}}(d)) \\
 & \quad x_i(t_{j,0}) = x_{i,j}^{0,exp} \quad \forall i = 1, \dots, n_x, j = 1, \dots, n_{exp} \\
 & \quad \hat{y} = h(x)
 \end{aligned} \tag{4}$$

In this optimization problem, the cost function J depends on the set of measured output values \mathbf{Y}^{exp} and the set of simulated output values $\hat{\mathbf{Y}}$ as shown in Eq. 5.

$$J(\mathbf{Y}^{exp}, \mathcal{P}) = \frac{1}{n_y} \sum_{i=1}^{n_y} \frac{1}{n_{exp}} \sum_{j=1}^{n_{exp}} \frac{1}{n_{samp}} \sum_{k=1}^{n_{samp}} J_r(r_{i,j,k}) \tag{5}$$

Here, n_y denotes the number of outputs and n_{exp} the number of experiments. n_{samp} denotes the number of samples taken, which is assumed to be the same in each experiment for ease of notation. J_r describes the kind of cost function used. Commonly, the mean squared error is used with $J_r(x) = x^2$. $r_{i,j,k}$ refers to a residual, which is defined as

$$r_{i,j,k} = y_{i,j,k}^{exp} - \hat{y}_i(t_{j,k}^{exp}). \tag{6}$$

Solving the full parameter estimation problem in Eq. 4 is a difficult task as the evaluation of the fitness, described by the cost function J , requires the simulation of the dynamics of the system. The iteration over all parameters and hyperparameters is computationally costly especially because good initialization of Θ_{ML} are not known and therefore many simulations will diverge.

To overcome this issue, we propose to decompose the model selection and parameterization problem into several steps as shown in Fig. 1.

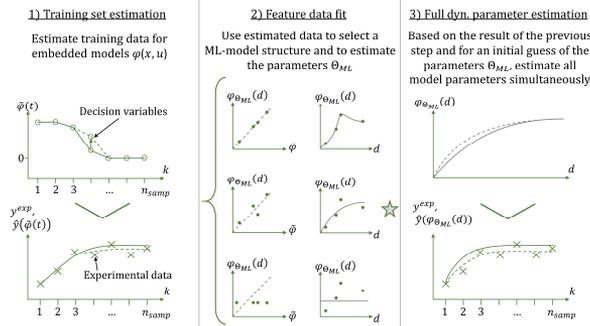


Fig. 1: Steps of applying the methodology for dynamic gray-box modelling

Step 1) involves the estimation of training data of the embedded functions by replacing the embedded variables φ by placeholders $\tilde{\varphi}(t)$, which are computed as continuous piecewise linear functions of time as shown in Eq. 7, similar to the approaches in [4] and [5].

$$\tilde{\varphi}_i(t) = \frac{\tilde{\varphi}_{i,j,k+1} - \tilde{\varphi}_{i,j,k}}{t_{j,k+1} - t_{j,k}}(t - t_{j,k}) + \tilde{\varphi}_{i,j,k} \quad t \in [t_{j,k}, t_{j,k+1}] \tag{7}$$

Here, $\tilde{\varphi}_{i,j,k}$ denote the values of the piece-wise linear function at the sampling time points $t_{j,k}$. With this replacement, determining the set of descriptors d and estimating the parameters Θ_{ML} is simplified. To estimate training data for the embedded functions, the

knot values of the piece-wise linear fit are adjusted as degrees of freedom in the parameter estimation problem shown in Eq. 8.

$$\begin{aligned} \min_{\Theta, \tilde{\Phi}} \quad & J(\mathbf{Y}^{exp}, \hat{\mathbf{Y}}) + \lambda \text{Reg}(\tilde{\Phi}) \\ \text{s. t.} \quad & \dot{x} = f_{\Theta}(x, u, \tilde{\varphi}(t)) \\ & x_i(t_{j,0}) = x_{i,j}^{0,exp} \quad \forall i = 1, \dots, n_x, j = 1, \dots, n_{exp} \\ & \hat{y} = h(x) \end{aligned} \quad (8)$$

In this optimization problem, the set of all values $\tilde{\varphi}_{i,j,k}$ is denoted as $\tilde{\Phi}$. As there is a large number of degrees of freedom, regularization is applied to prevent overfitting. This can be realized by adding the mean absolute slope $\text{Reg}(\tilde{\Phi})$ to the cost function, weighted by a regularization constant λ , see Eq. 9.

$$\text{Reg}(\tilde{\Phi}) = \frac{1}{n_{\varphi}} \sum_{i=1}^{n_{\varphi}} \frac{1}{n_{exp}} \sum_{j=1}^{n_{exp}} \frac{1}{n_{samp} - 1} \sum_{k=2}^{n_{samp}} \left| \frac{\tilde{\varphi}_{i,j,k} - \tilde{\varphi}_{i,j,k-1}}{t_{j,k} - t_{j,k-1}} \right| \quad (9)$$

The results of solving the optimization problem in Eq. 8 are the values of the predicted state \mathbf{X}^* and input variables \mathbf{U}^{exp} and the corresponding values $\tilde{\varphi}_{i,j,k}$.

In nonlinear dynamic systems, it can occur that the deviation from the experimental data is insensitive to changes of the values $\tilde{\varphi}_{i,j,k}$. One such example is when the embedded variable is multiplied with a state variable that vanishes during an experiment. To ensure that the estimated data is reliable, the collected optimal values $\tilde{\Phi}^*$ are filtered using the sensitivity of the residuals $r_{i,j,k}$, see [10] for further details.

Using the resulting data $\tilde{\Phi}^*$, the descriptors d , a suitable ML-model structure and a set of parameters Θ_{ML} can be found efficiently during step 2), the feature data fit.

This model structure and the estimated set of parameters are initial values for performing step 3), the full dynamic parameter estimation, described by Eq. 4.

4. Results

The methodology presented in section 3 is applied to the simulated case study from section 2. Experimental data was generated in silico by generating random initial conditions and simulating the ODEs presented in section 2. Afterwards, Gaussian noise was added to the measured outputs. X_t is the total cell concentration as $X_t = X_v + X_s$.

In step 1), the optimization problem in Eq. 8 is solved using the simulated data and a regularization constant with a value of $\lambda = 2$, the results are shown in Fig. 2.

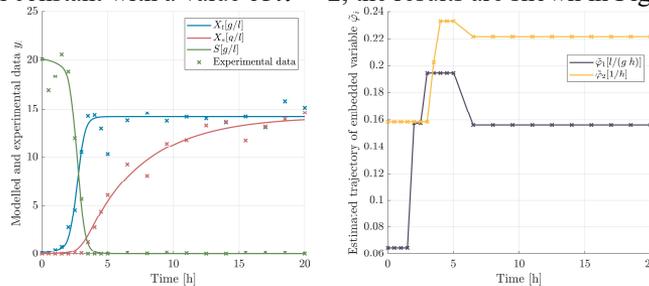


Fig. 2: Results of the training data estimation step. Left: predicted and measured output values; Right: Corresponding continuous piece-wise linear trajectories of the embedded variables

In Fig. 2 on the left hand side, the simulated and experimental values of the measured variables are shown. It can be seen that even though the experimental values are subject

to significant noise, the overall behavior is represented well. The substrate concentration S is initially high but decreases monotonously as the cells grow. This growth leads to an increase in the total cell concentration. After about 4-5 h the growth process is concluded and the sporulation begins. The sporulated cell concentration increases until about 20 h. On the right hand side, the corresponding estimated values of the embedded variables are shown. The fact that a penalization term on the slope is added to the cost function is reflected by the fact that intervals of constant $\tilde{\varphi}$ emerge.

Conducting step 2) of the methodology, different surrogate models were trained on the estimated data. Due to the training data estimation, any ML-toolbox can be used. In this case a constant value and a linear model were applied. Additionally two symbolic regression routines were applied, a Lasso regression [11] and an ALAMO [12] model were trained based on the same set of basis functions. Lastly, an artificial neural network model with 1 hidden layer and 5 nodes was fitted to the data. All models use the temperature T and substrate concentration \hat{S} as descriptors d . The regression plots are shown in Fig. 3.

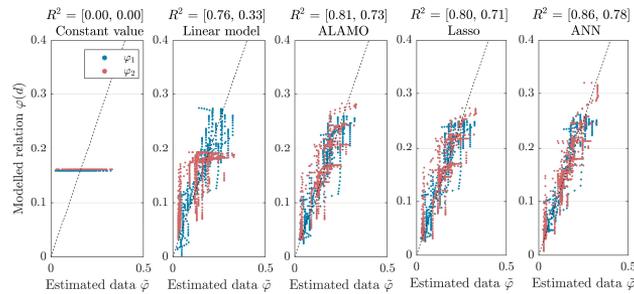


Fig. 3: Regression plot of all considered ML-model structures

From Fig. 3, it can be seen that both the constant value as well as the linear model fail to accurately describe especially the embedded variable φ_2 , thus nonlinear models have to be used. The highest coefficient of determination can be achieved using ANN models. Initialized with these parameters, the full dynamic parameter estimation problem in Eq. 4 (step 3)) was solved. The results of this optimization are displayed in Fig. 4.

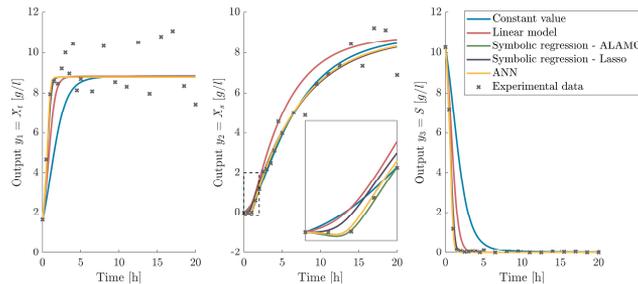


Fig. 4: Results of the full parameter estimation

From this figure, it can be seen that the sharp initial increase of the number of vegetative cells is not accurately represented using a constant value or a linear model for the embedded variable φ_2 . While there is a slight undershoot in the dynamic prediction of the ANN and of the ALAMO model, these model structures describe the dynamic behavior well. The Lasso model can be considered the best model as no negative concentrations of sporulated cells X_s are predicted.

5. Conclusions and outlook

The proposed methodology of systematically determining a dynamic gray-box model of nonlinear dynamic systems was successfully applied to a simulation use case of a fermentation process. It was shown that using a training data estimation step enables the training of data-based models using common ML-toolboxes. The model selection can be performed efficiently without the need of dynamic simulations. Finally, the validated model structures are fine-tuned by solving the full dynamic parameter estimation problem.

In further work, we will include constraints into the parameter estimation to ensure that the resulting models are meaningful.

6. Acknowledgment

This research has been supported by the project "KI-Inkubator-Labore in der Prozessindustrie - KEEN", funded by the Bundesministerium für Wirtschaft und Klimaschutz (BMWK) under grant number 01MK20014T. This support is gratefully acknowledged.

7. References

- [1] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, 2020, "Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems,". arXiv preprint arXiv:2003.04919
- [2] A. Azarpour, S. R. Wan Alwi, G. Zahedi, M. Madooli, and G. J. Millar, 2015, "Catalytic activity evaluation of industrial Pd/C catalyst via gray-box dynamic modeling and simulation of hydropurification reactor," *Applied Catalysis A: General*, vol. 489, pp. 262–271
- [3] L. Scheffold, T. Finkler, and U. Piechottka, 2021, "Gray-Box system modeling using symbolic regression and nonlinear model predictive control of a semibatch polymerization," *Computers & Chemical Engineering*, vol. 146, p. 107204
- [4] L. Hebing, T. Neymann, and S. Engell, 2020, "Application of dynamic metabolic flux analysis for process modeling: Robust flux estimation with regularization, confidence bounds, and selection of elementary modes," *Biotechnology and Bioengineering*, vol. 117, no. 7, pp. 2058–2073
- [5] C. de Prada, D. Hose, G. Gutierrez, and J. L. Pitarch, 2018, "Developing Grey-Box Dynamic Process Models," *IFAC-PapersOnLine*, vol. 51, no. 2, pp. 523–528
- [6] G. Bastin and D. Dochain, 1990, *On-line estimation and adaptive control of bioreactors*. Amsterdam: Elsevier
- [7] S. Das and R. Sen, 2011, "Kinetic modeling of sporulation and product formation in stationary phase by *Bacillus coagulans* RK-02 vis-à-vis other Bacilli," *Bioresource Technology*, vol. 102, no. 20, pp. 9659–9667.
- [8] E. Baril, L. Coroller, O. Couvert, M. El Jabri, I. Leguerinel, F. Postollec, C. Boulais, F. Carlin, and P. Mafart, 2012, "Sporulation boundaries and spore formation kinetics of *Bacillus* spp. as a function of temperature, pH and $a(w)$," *Food microbiology*, vol. 32, no. 1, pp. 79–86
- [9] P. Atehortúa, H. Alvarez, and S. Orduz, 2007, "Modeling of growth and sporulation of *Bacillus thuringiensis* in an intermittent fed batch culture with total cell retention," *Bioprocess Biosyst Eng*, vol. 30, no. 6, pp. 447–456
- [10] J. Winz, S. Engell, 2021, "A methodology for reliable dynamic nonlinear gray-box modeling," submitted to 13th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems (DYCOPS)
- [11] R. Tibshirani, 1996, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996
- [12] A. Cozad, N. V. Sahinidis, and D. C. Miller, 2014, "Learning surrogate models for simulation-based optimization," *AIChE J.*, vol. 60, no. 6, pp. 2211–2227

Optimal aeration of wastewater treatment plants based on recurrent neural networks and real plant data

Christian Hoffmann,^a Joris Weigert,^a Gerardo Brand Rihm,^a Erik Esche,^a and Jens-Uwe Repke^a

^a*Technische Universität Berlin, Process Dynamics and Operations Group, Str. des 17. Juni 135, 10623 Berlin, Germany*

Abstract

A significant share of the energy consumption by wastewater treatment plants stems from the air compressors in the aerobic processing steps. To minimize their energy consumption, dynamic optimization is a viable option. Modeling the complex (bio)chemical processes within wastewater treatment plants is a complex task and the simulation based on those models requires significant computational effort and is hard to initialize. An alternative to this approach is data-driven modeling of these plants. Therefore, this work features recurrent neural networks. These networks are trained and tested with real data from a German wastewater treatment plant. The identified models are used to obtain optimal control trajectories. The obtained trajectories are compared to the control actions applied in reality to assess the economic benefit of this approach.

Keywords: Machine learning, Recurrent neural networks, Wastewater treatment, Optimal control, Dynamic real-time optimization.

1. Main Text

In wastewater treatment plants (WWTPs), sewage is recycled in several anaerobic and aerobic processing steps. The aeration for these aerobic steps is done via compressors, which account for 50-60 % of the consumed energy of these plants (Viholainen et al., 2015). Therefore, minimizing energy consumption of the compressors would have major economic and environmental benefits. Consequently, rigorous, dynamic models were proposed to describe the transient behavior of WWTPs. Many of these models were critically analyzed and compared (Hauduc et al., 2010; 2013). However, while these models are certainly helpful in assessing the process dynamics, they may not be suited for online optimization because many of the required inputs for these models, e.g., some of the balanced pseudo-components, cannot be measured online.

Contrarily, data-driven models offer many advantages as they do not require the complex (bio)chemical phenomena to be modeled rigorously and can thus be evaluated much faster. This approach works if the available data amount and quality suffices for identifying a data-driven model with acceptable prediction error. As a result, machine learning methods have been proposed for applications in WWTPs. For example, Cheng et al. (2020) and Arismendy et al. (2020) studied different types of neural networks to forecast various process variables. An extensive review of machine learning methods for wastewater plants is given by Newhart et al. (2019). As of now, however, machine learning methods have hardly been used for optimal control in WWTPs. Only recently, Icke et

al. (2020) used a self-learning algorithm for control and reduced energy consumption by 15 %, thus showing the large economic potential of such approaches. In this contribution, recurrent neural networks (RNNs) are used to model the process dynamics of WWTPs. In RNNs, the feedback of outputs allows previous regressors to have an impact on the current prediction. This is advantageous for dynamic systems and gives these models much flexibility. The RNNs are trained with real data from a plant in Germany within (1) TensorFlow 2 (Abadi et al., 2015) with simple RNN architecture and with a long short-term memory (LSTM) layer and (2) a self-implemented RNN based on a multi-layer perceptron from Scikit-learn (Pedregosa et al., 2011). Instead of choosing the hyperparameters of the RNNs by trial-and-error, a hyperparameter tuning is carried out for optimal performance (Akiba et al., 2019).

2. Wastewater treatment plant

A simplified flowsheet of the WWTP that provided the data for this contribution is shown in Figure 1. The influent enters the plant and may be split into a stream that fills the rain basins (RB1–RB4). These basins usually fill in the case of heavy rain. Eventually, the sewage enters the sewer route (B1–B6) in which the (bio)chemical decomposition processes take place. Basin B5 and B6 are aerated. This aeration is realized by the compressors COM1–COM4. Furthermore, oxygen can be measured online in B5 and B6. Additionally, pH value and the concentrations of NH_4^+ and PO_4^- ions, and total suspended solids (TSS) are measured online in B6. The influent flow, the liquid volumes in the rain basins, and the power consumption P of the four compressors will be treated as inputs to the model whereas the measured concentrations will be outputs. This means that none of the basins are modeled in detail. Instead, only the input-output relation between the named variables must be found via training of the RNN. In general, the compressor duties depend on the required volume flow of air to achieve the desired oxygen concentration in the aerated basins.

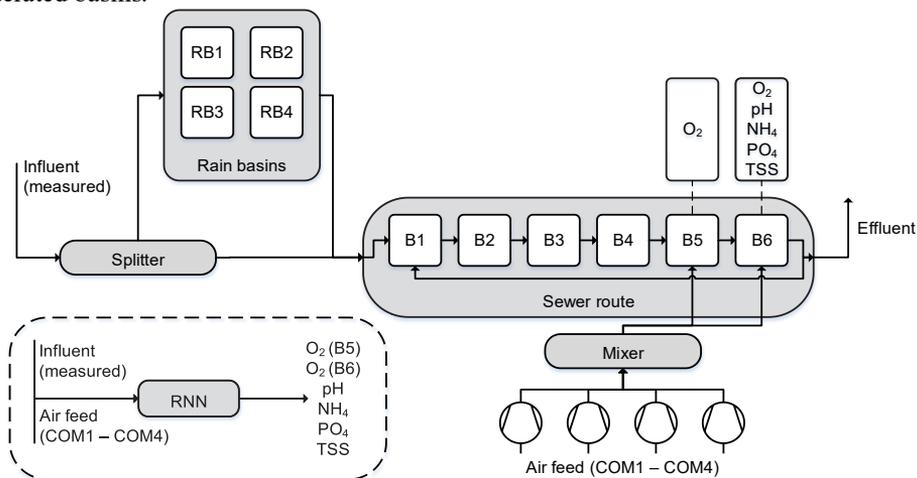


Figure 1: Simplified flowsheet of the WWTP in Weinheim (Germany). A representation of the RNN, including inputs and outputs, is given in the bottom left corner.

3. Training framework

The Python framework used in this contribution is illustrated in Figure 2 and based on the preliminary work by Weigert et al. (2020). In the data processing step, the plant data is read from Excel sheets and stored as a pickle file for efficient storage. Afterwards, the data is split into training and testing data. Throughout this work, the training-to-testing ratio is 9, i.e., the first 90 % of the data are used for testing while the remaining 10 % are used for cross validation and hyperparameter tuning. Once the training and testing data have been set and the input data have been windowed, the training data is also shuffled as part of the training algorithm.

In the model generation step, the data are separated into inputs and outputs as described in the previous section. Before the training begins, the training data are normalized so that all inputs have mean 0 and standard deviation of 1 whereas outputs lie between 0 and 1. In an external loop, the hyperparameter tuning via Optuna (Akiba et al., 2019) is initiated. Starting from an initial set of parameters, Optuna varies the hyperparameters of the respective model in the intervals defined in Table 1. Within the inner loop, the RNN is trained using the algorithms provided by TensorFlow 2 (Abadi et al., 2015) and Scikit-learn (Pedregosa et al., 2011), respectively. The objective function of the inner tuning is the mean squared error (MSE) of measurement data and prediction. The hyperparameter tuning minimizes the MSE for the testing data. Here, it is possible to use the MSE of the one-step prediction, the recurrent prediction – the prediction of the entire time horizon using the original data only in the first time point – or the MSE of an arbitrary prediction horizon. Within this contribution, the hyperparameter tuning was always minimized subject to a prediction horizon of 4 prediction steps. The inner tuning is repeated until the number of pre-defined outside iterations (here: 400) is reached. This number of iterations must be sufficiently large. In this study, the best solution was usually found after approximately 200 iterations, further increasing the number of iterations revealed no additional benefit. Finally, the hyperparameter set with the lowest MSE is chosen and the model is stored. In the optimal control step, the horizon is initialized by starting the optimization at time step $t_{\text{start}} = t_0 + n \Delta t_k$ where $n = \max(n_c; n_s)$ (see Table 1 for definitions). This ensures that enough past regressors are available to solve the RNN at the initial point. During the optimization, the plant influent remains an input to the RNN, but it can obviously not be used as degree of freedom. Contrarily, the liquid volumes in the rain basins and the energy consumption of the compressors are decision variables. The length of the prediction horizon can be

Table 1: Possible parameter range for hyperparameter tuning via Optuna.

Parameter	Range
Number of neurons in hidden layer (NN)	[10 ... 200]
Activation function (AF)	tanh or relu
L2 regularization weight (α)	[1e-5 ... 10]
Number of past states (n_s)	[3 ... 20]
Number of past controls (n_c)	[3 ... 20]

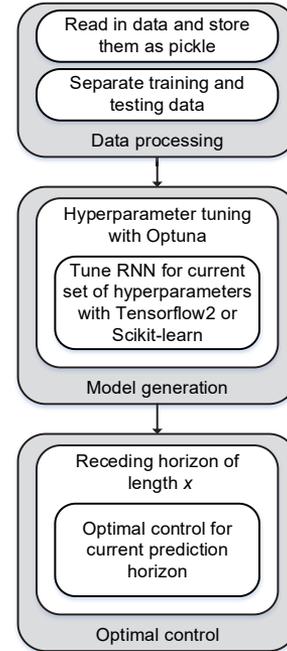


Figure 2: Framework for model generation and optimal control within Python.

repeated until the number of pre-defined outside iterations (here: 400) is reached. This number of iterations must be sufficiently large. In this study, the best solution was usually found after approximately 200 iterations, further increasing the number of iterations revealed no additional benefit. Finally, the hyperparameter set with the lowest MSE is chosen and the model is stored. In the optimal control step, the horizon is initialized by starting the optimization at time step $t_{\text{start}} = t_0 + n \Delta t_k$ where $n = \max(n_c; n_s)$ (see Table 1 for definitions). This ensures that enough past regressors are available to solve the RNN at the initial point. During the optimization, the plant influent remains an input to the RNN, but it can obviously not be used as degree of freedom. Contrarily, the liquid volumes in the rain basins and the energy consumption of the compressors are decision variables. The length of the prediction horizon can be

chosen arbitrarily: on the one hand, the horizon should be long enough to ensure a stable solution of the dynamic optimization; on the other hand, the prediction error usually increases with increasing time horizon. Therefore, a suitable horizon length must be determined. In this contribution, a prediction error of 1 led to profiles that occasionally violated inequality constraints. This did not occur for a prediction horizon of 2 for which the following results are shown.

4. Results and discussion

In the following, we compare the results obtained with different RNNs using the hyperparameter tuning described above. Results were obtained for three multiple-input-single-output (MISO) RNNs and one multiple-input-multiple-output (MIMO) RNN. For MISO RNNs, one RNN is trained per output variable. In the case of the MIMO RNN, one RNN is trained that predicts all outputs simultaneously. Figure 3 shows the results obtained with the MIMO RNN, which is based on a multi-layer perceptron in which the outputs are fed back as inputs (pseudo-RNN). Surprisingly, both the available training and testing data can be predicted very well given the small number of inputs. Table 2 shows a comparison of the obtained results for all trained RNNs. The first three columns contain ranges as every output is described by a separate model with different parameterizations.

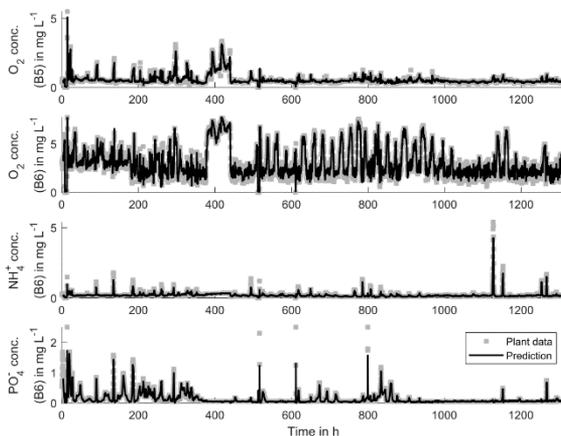


Figure 3: Comparison of plant data and prediction with MIMO RNN.

Table 2: Training results for all investigated RNNs in TensorFlow (TF) and Scikit-learn (SL) after hyperparameter tuning. Models in TensorFlow do not use external past states since the amount of recycled information is determined internally.

Parameter	Trained RNNs			
	MISO RNN	MISO RNN	MISO RNN	MIMO RNN
RNN Type	simple (TF)	LSTM (TF)	pseudo (SL)	pseudo (SL)
NN	35 to 143	61 to 196	46 to 165	164
AF	relu and tanh	relu and tanh	relu	relu
α	0.002 to 4.37	$6.5 \cdot 10^{-5}$ to 3.39	0.017 to 0.33	0.036
n_s	n. a.	n. a.	9 to 19	15
n_c	6 to 15	3 to 20	3 to 8	3
MSE $\cdot 10^3$	11.39	10.1	2.57	2.40

The differences in the obtained hyperparameters are significant and include a large variation in the number of neurons and the weighting parameter, and different activation functions. Moreover, the number of past states is much higher for the MISO RNNs compared to the MIMO RNN. Overall, the models trained in Scikit-learn perform better than those trained in TensorFlow 2. In addition, the MIMO RNN is slightly better than the MISO version. A possible explanation is that the MIMO model also considers the interdependency between the various outputs. Therefore, the MIMO architecture is used in the following. The influent profile for this case study is shown in Figure 4 (bottom). The objective function is the MSE between measured concentrations and pre-defined setpoints. As Figure 4 shows over a period of about one week, the effluent concentrations show notably less fluctuations for both ammonia and phosphorous components. As in the original data, compressors 3 and 4 are only used irregularly. However, this applies to compressor 2 as well in the optimization. In the case of lower plant load (between hour 30 and hour 75), the overall compressor power is also reduced. As a result, a reduction of the energy consumption of about 50 % is predicted for this week.

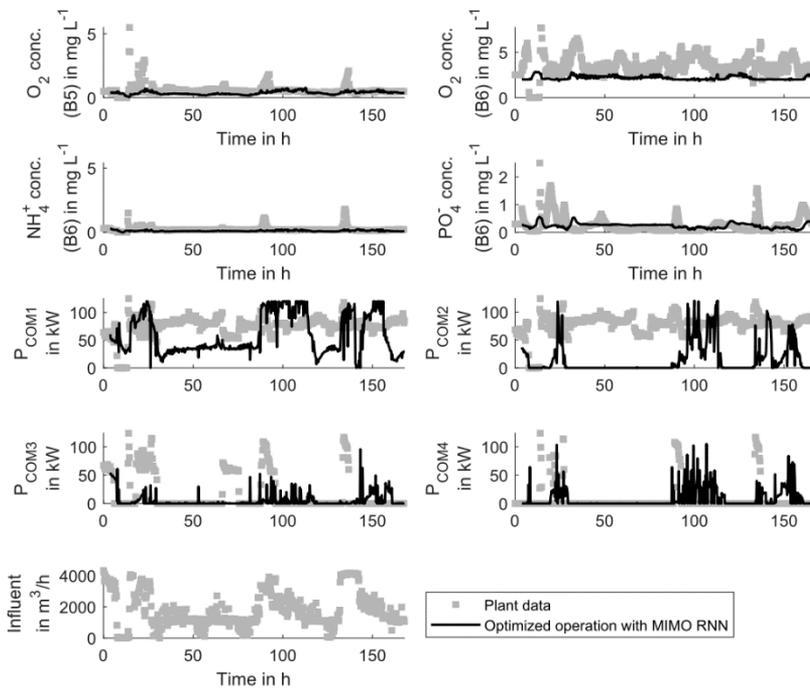


Figure 4: Optimized trajectory for 1 week. Levels of the rain basins are not shown here for brevity.

5. Conclusion and outlook

This contribution showed that a MIMO RNN can reliably predict the concentrations of components that are decisive for the effluent quality of a WWTP. The MIMO RNN also performed better than MISO RNNs where a single RNN per variable is trained at a time. The obtained MIMO RNN was used for optimal control in a receding horizon. Both the energy consumption and the liquid volumes in the rain basins were selected as decision

variables. This approach resulted in significantly lower energy consumption compared to the original plant data: a reduction of 50 % was predicted for the investigated week. In future work, the number of data points for the RNN must be extended and should include other seasons as well. In particular, the accurate description of outputs for small concentrations of dissolved oxygen should be checked to verify that the determined optimal solution is indeed physically feasible and not only a pseudo-solution of the trained RNN.

Acknowledgements

The authors gratefully acknowledge the financial support of the German Federal Ministry of Economic Affairs and Energy and the AiF Projekt GmbH (project number KK5118001RH0).

References

- M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Jozefowicz, Rafal, Y. Jia, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, M. Schuster, R. Monga, S. Moore, D. Murray, C. Olah, J. Shlens, Steiner, Benoit, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, 2015. Tensorflow: Large-scale machine learning on heterogeneous systems (white paper).
- T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, 2019. Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2623–2631.
- L. Arismendy, C. Cárdenas, D. Gómez, A. Maturana, R. Mejía, C. G. Quintero M., 2020. Intelligent system for the predictive analysis of an industrial wastewater treatment process. *Sustainability* 12 (16), 6348.
- T. Cheng, F. Harrou, F. Kadri, Y. Sun, T. Leiknes, 2020. Forecasting of wastewater treatment plant key features using deep learning-based models: A case study. *IEEE Access* 8, 184475–184485.
- H. Hauduc, L. Rieger, A. Oehmen, M. van Loosdrecht, Y. Comeau, A. Héduit, P. A. Vanrolleghem, S. Gillot, 2013. Critical review of activated sludge modeling: State of process knowledge, modeling concepts, and limitations. *Biotechnology and Bioengineering* 110 (1), 24–46.
- H. Hauduc, L. Rieger, I. Takács, A. Héduit, P. A. Vanrolleghem, S. Gillot, 2010. A systematic approach for model verification: application on seven published activated sludge models. *Water Science and Technology* 61 (4), 825–839.
- O. Icke, D. M. van Es, M. F. de Koning, J. J. G. Wuister, J. Ng, K. M. Phua, Y. K. K. Koh, W. J. Chan, G. Tao, 2020. Performance improvement of wastewater treatment processes by application of machine learning. *Water Science and Technology* 82 (12), 2671–2680.
- K. B. Newhart, R. W. Holloway, A. S. Hering, T. Y. Cath, 2019. Data-driven performance analyses of wastewater treatment plants: A review. *Water Research* 157, 498–513.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, 2011. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research* 12 (85), 2825–2830.
- J. Viholainen, K. Grönman, A. Jaatinen-Värri, A. Grönman, P. Ukkonen, M. Luoranen, 2015. Centrifugal compressor efficiency improvement and its environmental impact in waste water treatment. *Energy Conversion and Management* 101, 336–342.
- J. Weigert, C. Hoffmann, E. Esche, J.-U. Repke, 2020. Enabling dynamic real-time optimization under uncertainty using data-driven chance constraints. In: *Computer Aided Chemical Engineering* 48. Elsevier, pp. 1189–1194.

Soft Sensor of Key Components in Recirculating Aquaculture Systems, using Feedforward Neural Networks

Allyne M. dos Santos^{a,*}, Espen Karlsen^a, Sigurd Skogestad^a and Kari J.K. Attramadal^{b,c}

^a*Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway*

^b*Department of Biotechnology and Food Science, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway*

^c*Nofitech AS, 7042 Trondheim, Norway*

allyne.dos.santos@ntnu.no

Abstract

Recirculating Aquaculture Systems are known to have its water quality conditions controlled, despite being common to have a simple control structure and a lot of human interaction to achieve that. To avoid long-term exposure to toxic levels of carbon dioxide and ammonia, its concentration needs to be monitored more often than manual measurements are available. In this work, we analyze the multilayer perceptron's ability to monitor water quality components that are important for the development of the fish. This alternative method for monitoring has the potential to complement the current sensor structure and laboratory procedures for manual measurement collection, but more studies need to be done on the type of machine learning model.

Keywords: Soft sensor, Recirculating aquaculture systems (RAS), Feedforward neural networks

1. Introduction

Recirculating Aquaculture Systems (RAS) have two dynamic sub-systems: fish metabolism and water treatment. The water treatment system is responsible by keeping the water quality at high standards, reducing water consumption, and reducing contact with external pathogens (European Market Observatory for Fisheries and Aquaculture Products, 2021).

Regarding the water quality, some components are important to be monitored and controlled due to toxicity, but are hard or not able to be measured continuously, such as ammonia. In addition, it requires either several sensors or a central sensor station with sampling system to get information about the levels of dissolved carbon dioxide and ammonia in all fish tanks and water treatment system. One way of using information of the process to estimate the concentration of these toxic components is by using soft sensors (Fortuna et al., 2007).

The soft sensor technique is a combination of data, for parameter estimation, and process knowledge, for feature selection. The development of the data-driven models can be done, for example, using machine learning models, such as feedforward neural networks (FNN). In this work, the main objective is to apply multilayer perceptron (MLP) (Werbos, 1974), which is a classic type of FNN and a universal approximator (Hornik et al., 1989), as a soft sensor for recirculating aquaculture of Atlantic salmon (*Salmo salar*).

In RAS, even if fish, feed and waste production increase in an exponential way, the goal is that the water quality should be kept stable and good. Measurements of ammonia are performed manually

daily, whereas CO_2 is measured continuously at least in one point in the system. However, their levels might vary during the day depending on feeding and other factors, and data about this may be missed. Optimizing the water quality and fish growth would be easier with higher resolution of information. In addition, manual collection and laboratory analyzes are time consuming and therefore costly. Therefore, it is useful to develop an alternative method to monitor these water quality variables.

As the system is assumed to reach a steady state after each day, the water treatment system is approximated to a steady-state model developed in previous work (Dos Santos et al., 2021). The training and validation data are acquired from this model, where some parameters are considered as uncertainties. To improve the soft sensor identification, the uncertainties are changed using latin hypercube sampling (LHS) (Jin et al., 2005), so it contains the operating region, which provides condition for fish optimal growth. After addition of white noise, the data is used to train different MLP configurations for predicting carbon dioxide, ammonia and ammonium concentrations.

2. Process Description

Figure 1 shows a diagram of the RAS this work is focused on. The process consists of a fish tank, a biofilter, a stripper and an oxygen cone. The model is a simplified version of the process, as it does not consider the effect of the water quality on the fish metabolism, if the conditions are kept within bounds. This assumption is reasonable for each phase of the fish life, which can last from weeks to years depending on a lot of factors. Therefore, this work is only valid for the phase the model represents, which is the smolt phase. This could be easily extended to other phases by changing some parameters in the model, such as the amount of product generated by the fish metabolism per kg of fish feed.

The measurements that are available from sensors or human addition include recirculating volumetric flow rate, q ; pH ; fish feed rate, F ; buffer additions, \dot{m}_{buffer} ; base additions, \dot{m}_{base} ; makeup water, q^m ; air inlet flow rate, \dot{m}_{air} ; makeup oxygen, \dot{m}_{O_2} ; average salinity, S ; average temperature, T . More details about the process can be found in Dos Santos et al. (2021).

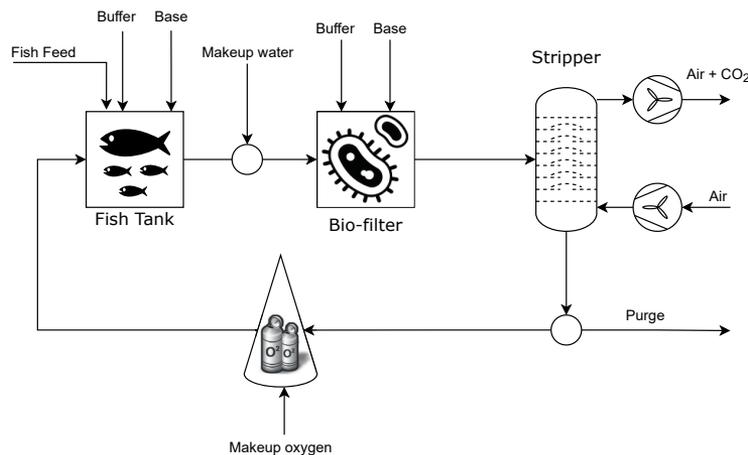


Figure 1: Process diagram of a recirculating aquaculture system

3. Methodology

In order to develop a soft sensor, a standard procedure was followed. Gather data, preprocess it and divide it into training and validation data. After that, fit the model and choose the best model of validation phase. And finally, test it with industrial data.

3.1. Training and Validation Data Acquisition

For the training and validation data acquisition, 6967 steady-state data points were generated within the region described by Table 1, using CasADi v3.5.5 (Andersson et al., 2019) in Python v3.8.8. 5967 of these data points were generated using latin hypercube sampling (LHS) built in pyDOE package in Python. LHS is a popular algorithm for planning computer experiments covering the entire range of uncertainties and disturbance in an optimally and distributed way for training. The rest was generated randomly within the same region for validating the soft sensor. After that, 1% white noise was added to both input and output data of both training and validation data.

Table 1: Region of operation

Parameter	Mean	Unit	Range	Description
G/L	2.7	-	$\pm 50\%$	Gas-liquid ratio in equilibrium over the stripper
ξ_B	0.8	-	$\pm 25\%$	Biofilter efficiency
T	14	$^{\circ}\text{C}$	$\pm 30\%$	Average temperature of the system
pH ^m	7.0	-	$\pm 10\%$	pH of the makeup water
$y_{\text{CO}_2}^{\text{in}}$	4.15e-04	-	$\pm 10\%$	CO ₂ composition in the air inlet
S	15.95	ppt	$\pm 30\%$	Average salinity of the system
pH _{des} ^B	7.2	-	$\pm 1\%$	Desired pH for the biofilter
q	20	m ³ /min	$\pm 50\%$	Recirculating volumetric flow rate
F	580.6	g/min	+ 50%	Fish feed rate

3.2. Data Preprocessing

Some concentrations are really low in RAS, when comparing with other concentrations. Therefore, it is essential that all data is submitted to preprocessing. As the soft sensor model applied in this work is a deep learning neural network, the normalization of the datasets uses a minmax calculation of the training dataset. After that, the datasets are ready to be used to train and validate the soft sensor.

3.3. Soft Sensor

The feedforward neural network (FNN) architectures were created and optimized using Auto-keras package (Jin et al., 2018), which is a package in Python that automatizes the neural architecture search (NAS) of models supported by another python package named Keras.

In this work, the FNNs are multilayer perceptron (MLP), trained using backpropagation method with a batch size of 500 samples. The objective function of the NAS was the validation loss, the loss function was the mean squared error, and the maximum number of trials is 50. The features were chosen based on knowledge of the process and the available measurements in a real RAS.

After choosing the model with the lowest validation loss, the MLP performance was tested predicting the targets concentration from real data. To avoid breaking the non-disclosure agreement, the collected industrial data was normalized. The performances were compared using the root mean squared error between scaled predicted and scaled data.

4. Results

The choice of inputs was based on the available measurements from sensors or manual insertion of inlet streams described on the previous section. Three types of models were tested: multiple-input, single-output MLP (MISO-MLP), multiple-input, multiple-output MLP (MIMO-MLP), and a hybrid model, which consists of a MIMO-MLP predicting ammonium and dissolved CO₂ concentrations with ammonia concentration being calculated using the equilibrium equation, see Eq. 1. The inputs of the models were the same for MISO-MLP_{NH₄⁺} and MISO-MLP_{NH₃} models: fish feed rate, F , recirculating volumetric flow rate, q , pH in the tank, pH^T . The MIMO-MLP, MISO-MLP_{H₂CO₃} and the hybrid models' features included the same as the previous with addition of a new feature: air inlet flow rate, \dot{m}_{air} .

$$c_{NH_3}^T = \frac{K_3(S, T) c_{NH_4^+}^T}{c_{H^+}^T} \quad (1)$$

where the equilibrium constant, K_3 , is dependent on salinity and temperature, and the concentrations unit is mmol/L.

Figure 2 shows the prediction of the validation data using the MISO-MLP models, Figure 3 shows the results using the hybrid model, and Figure 4 shows the results using the MIMO-MLP model. Comparing Figures 2 and 3, prediction of ammonium and dissolved carbon dioxide were similar, but ammonia predictions are worse using the hybrid model. Comparing Figures 2 and 4, prediction of ammonia is slightly worse on the extremes using MIMO-MLP, and the other predictions were similar.

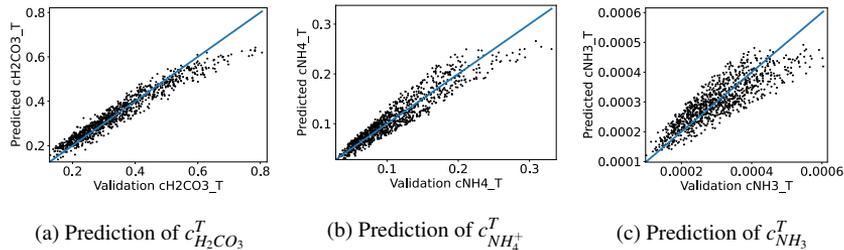


Figure 2: Prediction of validation data using the MISO-MLP models separately

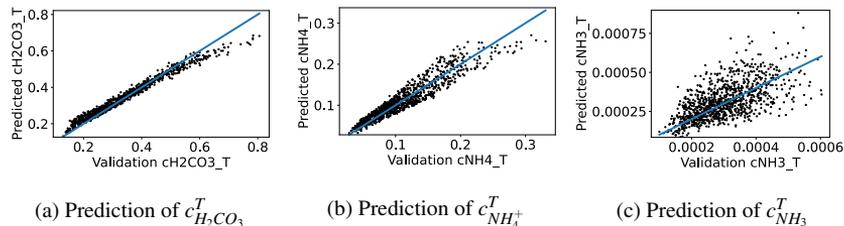


Figure 3: Prediction of validation data using the hybrid model

In Table 2, we summarize the performance of the models at the validation phase using the RMSE index. The best MLP architecture for this case study was the MISO-MLP models put together, which gave the lowest final RMSE at the validation phase, and MIMO-NLP model was the second best giving similar, but slightly higher, RMSE.

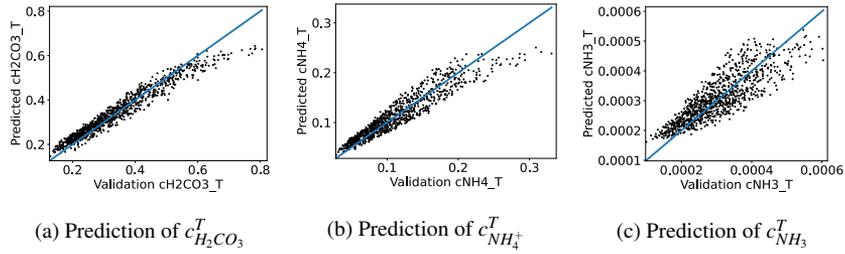


Figure 4: Prediction of validation data using the MIMO-MLP model

Table 2: Summary of the models performance at the validation phase - RMSE index

Output	MISO-MLPs	Hybrid	MIMO-MLP
$c_{H_2CO_3}^T$	0.0645	0.0787	0.0694
$c_{NH_4^+}^T$	0.1204	0.1201	0.1230
$c_{NH_3}^T$	0.1322	0.2611	0.1351
Final	0.1097	0.1720	0.1129

Their architecture of the MLPs are described in Table 3. MISO-MLP $_{H_2CO_3}$ model has two dense hidden layers with 32 nodes each using the rectified linear activation function (ReLU), while the others have one dense hidden layer, but with 64 and 128 nodes using the same activation function on MISO-MLP $_{NH_4^+}$ and MISO-MLP $_{NH_3}$ models, respectively. Note that ammonia concentration turned out to be harder to estimate when compared with ammonium, and possibly the NAS found a flat optimum, which means that probably a MLP with lesser nodes would not improve but would not make the prediction much worse also.

Table 3: Number of nodes in each layer of each MISO-MLP model

Layers	MISO-MLP $_{H_2CO_3}$	MISO-MLP $_{NH_4^+}$	MISO-MLP $_{NH_3}$
Input	4	3	3
Normalization	3	3	3
Dense ₁	32	64	128
Dense ₂	32	0	0
Dropout	0	64	128
Output	1	1	1

The prediction of the industrial data using the MISO-MLP models is shown in Figure 5. The predicted values of ammonia concentration seem to be closer to the real data compared with ammonium predictions, which is unexpected. The predicted $c_{H_2CO_3}$ were the same for the first few samples, which means that the model could not extract enough information from the features on those points.

5. Discussion

The prediction of ammonia was revealed to be much harder than of ammonium, resulting in a poor prediction of that variable. This might be due to the noise of the input variables that affects a lot the ammonia concentration, as its magnitude is much smaller. This effect does not disappear after

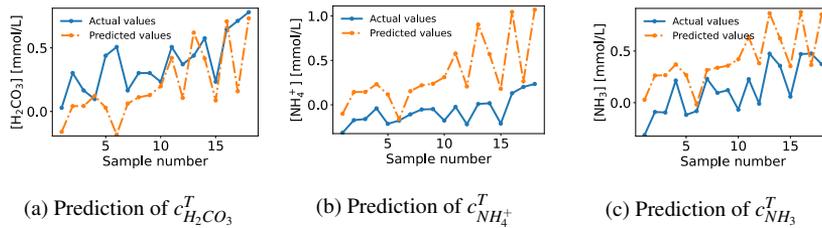


Figure 5: Prediction of industrial data using the MISO-MLP models

preprocessing the data; it expands instead, as the magnitude after normalization is approximately 1000 times higher.

The MISO-MLP models gave the best performance due to different input features to each model, as adding not so important features can make NAS more complex and add a lot of useless cases, as happened with the MIMO-MLP model. This could be solved by increasing the maximum number of trials at the NAS step, but it would take longer, and would still have the possibility of finding different local minimum, for better or worse.

6. Conclusion

The measurement of key waste products are not always easy to collect in real-time or at a required frequency, which reduces the possibility to stabilize and optimize the water quality for the fish. This can be improved by using machine learning models, such as multilayer perceptron.

The MLP models trained in this work are deep neural networks, and its architectures were optimized using an automated neural architecture search and tuning of hyperparameters. Three configurations were compared: MISO-MLPs; hybrid model; and MIMO-MLP. The best configuration was the MISO-MLP models together, although their performance was not so good. This might be due to the possibility of NAS reaching a local minimum or the models could not capture the information it needed, so a different type of model could perform better. A soft sensor using these models for monitoring would perform better than the hybrid model, and the MIMO-MLP model, but would also complement the manual measurements, when estimating dissolved CO_2 and NH_3 concentrations.

References

- J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, M. Diehl, 2019. CasADi: a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* 11 (1), 1–36.
- A. M. Dos Santos, L. F. Bernardino, S. Skogestad, K. J. Attramadal, 2021. Steady-state and Dynamic Modelling of Water Quality in Recirculating Aquaculture Systems. In Progress.
- European Market Observatory for Fisheries and Aquaculture Products, 2021. Freshwater aquaculture in the EU. No. March. European Union.
- L. Fortuna, S. Graziani, A. Rizzo, M. G. Xibilia, 2007. *Soft Sensors for Monitoring and Control of Industrial Processes* (Advances in Industrial Control). Springer-Verlag London.
- K. Hornik, M. Stinchcombe, H. White, 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2 (5), 359–366.
- H. Jin, Q. Song, X. Hu, 2018. Auto-Keras: An Efficient Neural Architecture Search System. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1946–1956. URL <https://arxiv.org/abs/1806.10282v3>
- R. Jin, W. Chen, A. Sudjianto, 2005. An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference* 134 (1), 268–287.
- P. J. Werbos, 1974. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph.D. thesis, Harvard University.

Interpretability of neural networks predictions using Accumulated Local Effects as a model-agnostic method

Tina Danesh^{a,*}, Rachid Ouaret^{a,*}, Pascal Floquet^a and Stéphane Negny^a

^a*Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse 31030, France*

tinasadat.daneshalaghehband@toulouse-inp.fr, rachid.ouaret@toulouse-inp.fr

Abstract

Recently, machine learning methods such as neural networks have been applied in various applications thanks to their accuracy and flexibility. However, the main drawback of these methods is the lack of interpretability, which is the reason for being uncommon in chemical engineering applications. At the same time, the recent rise of interpretability research has led to some confusion in various communities. In order to deal with this issue for the CAPE community, we propose to discuss the notion of interpretability under the prism of neural network predictions using an example from chemical engineering. To do this, we first set out the framework for defining interpretability for machine learning methods. Then a post-hoc method for model evaluation, explicitly named "model-agnostic methods", will be presented. In this study, we try to enhance the interpretability of the neural network predictions and visualize the effect of features on the model's output by a model-agnostic method named Accumulated Local Effects. As a case study, we work on predicting electrical power output and prioritizing the parameters of a combined cycle power plant. We could conclude that the most influential input parameter among Ambient Temperature (AT), Atmospheric Pressure (AP), Vacuum (V), and Relative Humidity (RH) is AT, and the most interaction is between AT and V.

Keywords: Interpretability, Neural network, Model-agnostic method, Accumulated local effects

1. Introduction

Recently, Machine Learning (ML) techniques, especially deep neural networks models, have become a crucial tool for various engineering applications such as chemical engineering (Zhang et al., 2020). These techniques have achieved high predictive accuracy primarily due to their ability to model complex nonlinear interactions between predictors and model outcomes. However, most accurate predictions remain in the black box frameworks, meaning that interpretation of the model's parameters interaction and explanation is partially (or entirely) hidden to the experts. Understanding the reasons behind predictions is very important to one plan to act based on these predictions.

In this paper, we will restrict the definition of the concept of *interpretability* to the ease of understanding the relationships established by ML models, which are able to work with non-linearly related variables. More precisely, for prediction problems, interpretability can be defined as the process of extracting relevant knowledge from a model about the learned relationships between features and model outputs. The knowledge referred here concerns information with a given meaning, allowing a particular audience to understand a chosen problem better (Murdoch et al., 2019).

To this end, we have two options, namely model-specific tools (easily interpretable models) and model-agnostic methods. Model-specific methods have their own tools to interpret the results, such as weights (statistical significance) in linear or logistic regression and the hierarchical graphical tool for decision trees. Nevertheless, they cannot be generalized to other ML models, particularly regarding the nonlinear impacts between variables in the Neural Networks learning. Thus, model-specific interpretation tools have a significant disadvantage: the limitation to one type of model and the difficulty of switching to other ML regression methods. The second approach, which is the model-agnostic interpretation method, is designed to remedy these shortcomings. Model-agnostic tools can be used on any machine learning model and are applied after the model has been trained (post-hoc) (Molnar, 2019).

In this paper, in order to enhance the interpretability in machine learning predictions, Model-agnostic methods are used and applied to real complex data and prediction problems. The prominent advantages of model-agnostic methods are their flexibility in model, explanation, and representation (Molnar, 2019). These methods separate the explanations from the machine learning model and could be applied to any ML model (regression and classification) (Ribeiro et al., 2016). The visualization provided by these tools is an important model diagnostic technique.

Model-agnostic methods can be distinct into local and global methods. Global methods concentrate on how features affect the prediction on average. In comparison, local methods explain the individual effect on predictions. Global methods are mainly represented as expected values based on the data distribution. Global methods include Partial Dependence Plots (PDP), Accumulated Local Effect plots (ALE), global surrogate models. Local methods include Individual Conditional Expectation curves (ICE), Local Interpretable Model-agnostic Explanations (LIME), and counterfactual explanations (Molnar, 2019).

One of the interesting model-agnostic methods available is ALE Plots. Our previous work dived into PDP and ICE plots as model-agnostic methods and discussed the main effect of each predictor x_1, x_2, \dots, x_n on $f(x_1, x_2, \dots, x_n)$ (Danesh et al., 2021). This study will work on ALE Plot to detect the lower order interaction effects among different predictors and compare it with other visualization tools for interpretability purposes.

As a real-world application, data from Combined Cycle Power Plant (CCPP) proposed by Tüfekci (2014) are used. The dataset is collected over six years and comprises 9568 data points. The paper focused on testing and comparing some machine learning regression methods. It aimed to extend an accurate predictive model and evaluate the prediction accuracy by Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

This paper's main contribution falls within the framework of the interpretability methods of neural network predictions (the Multi Layer Perceptron (MLP)) from model-agnostic points of view. The primary purpose of this study is to present to the CAPE community a simple way to interpret, evaluate and validate a machine learning prediction model. For this purpose, we used ALE plots as model-agnostic methods to visualize the behavior and effect of inputs on the output.

2. Accumulated Local Effects

2.1. Theory

Accumulated local effects explain the average impact of features on the prediction of a machine learning model (Apley and Zhu, 2020). ALE methods could work while the features are dependent, although the assumption of features independence is the biggest problem with PDPs.

In what follows, upper case X is used to identify random variables, and lower case is used to identify specific values of the random variables. The total feature space x contains two different sets. The first one, x_i represents the features that their effect on the prediction is studied. The

second one named x_j represents the other features that exist in the machine learning model except for x_i .

The ALE main effect of predictor $x_i, i \in \{1, \dots, d\}$ is defined by the following equation:

$$\hat{g}_{i,ALE}(x_i) = \int_{z_{0,i}}^{x_i} E[\hat{g}^i(X_i, X_j) | X_i = z_i] dz_i - C$$

Where, $\hat{g}^i(X_i, X_j) = \frac{\partial \hat{g}(X_1, \dots, X_d)}{\partial X_i}$. Here, $g(x)$ is assumed a black box supervised learning model that is a neural network in our study. $\hat{g}(x)$ represents the fitted model. $z_{0,i}$ refers to an approximate lower bound of X_i , and it affects the vertical translation of the ALE plot. C is considered as a constant that aims to make the mean of $\hat{g}_{i,ALE}(x_i)$ equal to zero concerning the marginal distribution of X_i or to center the plot vertically.

The ALE second-order effect of predictors $\{x_i, x_l\}, \{i, l\} \subseteq \{1, \dots, d\}$ is defined by the following equation:

$$\hat{g}_{\{i,l\},ALE}(x_i, x_l) = \int_{z_{0,i}}^{x_i} \int_{z_{0,l}}^{x_l} E\left[\frac{\partial^2 \hat{g}(X_1, \dots, X_d)}{\partial X_i \partial X_l} | X_i = z_i, X_l = z_l\right] dz_i dz_l - f_i(x_i) - f_l(x_l) - C$$

Where, $z_{0,i}$ and $z_{0,l}$ refer to approximate lower bounds of X_i and X_l , respectively. The functions $f_i(x_i)$ and $f_l(x_l)$ are the function of single variables X_i and X_l , respectively. The functions and the constant aims to make the mean of $\hat{g}_{\{i,l\},ALE}(x_i, x_l)$ equal to zero concerning the marginal distribution of X_i and X_l .

2.2. Estimation

For estimation, the features are divided into many intervals, and then the differences in the predictions are computed. We could approximate the derivatives by this procedure. The advantage of this procedure is that it is able to work for models without derivatives. The equation of estimation, as proposed by Apley and Zhu (2020), is as follows:

$$\hat{g}_{i,ALE}(x) \approx \sum_{k=1}^{k_i(x)} \frac{1}{n_i(k)} \sum_{j: x_{j,i} \in N_i(k)} [\hat{g}(z_{k,i}, x_{j,\setminus i}) - \hat{g}(z_{k-1,i}, x_{j,\setminus i})] - C$$

Where the notations are as follows: the constant is chosen to $\frac{1}{n} \sum_{j=1}^n \hat{g}_{i,ALE}(x_{j,i}) = 0$. For each $k \in \{1, 2, \dots, K\}$, $n_i(k)$ refers to the number of training observation that falls into k th interval $N_i(k)$. Let $x_{j,\setminus i} = (x_{j,m} : m = 1, \dots, d; m \neq i)$ where the subscript $\setminus i$ refers all variables except the i th. For each $i \in \{1, 2, \dots, d\}$, $N_i(k) = (z_{k-1,i}, z_{k,i}] ; k = 1, 2, \dots, K$ indicates an enough good partition of the sample rang of $x_{j,i} : j = 1, 2, \dots, n$ into k intervals (K is an input argument in the ALEPlot function, and we typically use K around 100, with larger values often give better result.). $z_{k,i}$ is chosen as the $\frac{k}{K}$ quantile of the empirical distribution of $\{x_{j,i} : j = 1, 2, \dots, n\}$ that $z_{0,i}$ is chosen below the smallest observation, and $z_{K,i}$ is chosen as the largest observation.

3. Case study: Combined Cycle Power Plant

A combined cycle power plant was selected as the real-world application to carry out the study. The components of a CCPP are Gas Turbines (GT), Steam Turbines (ST), and Heat Recovery Steam Generators (HRSG). One of the advantages of CCPP is using waste heat to produce supplementary steam in order to generate additional electricity (Niu and Liu, 2008). One of the efficient machines to generate mechanical and electrical power from gas fuels is the gas turbine. Consequently, gas turbines are more common for power generation. They are utilized especially in combined cycle mode to recover the waste heat to produce extra electricity. In addition to the high

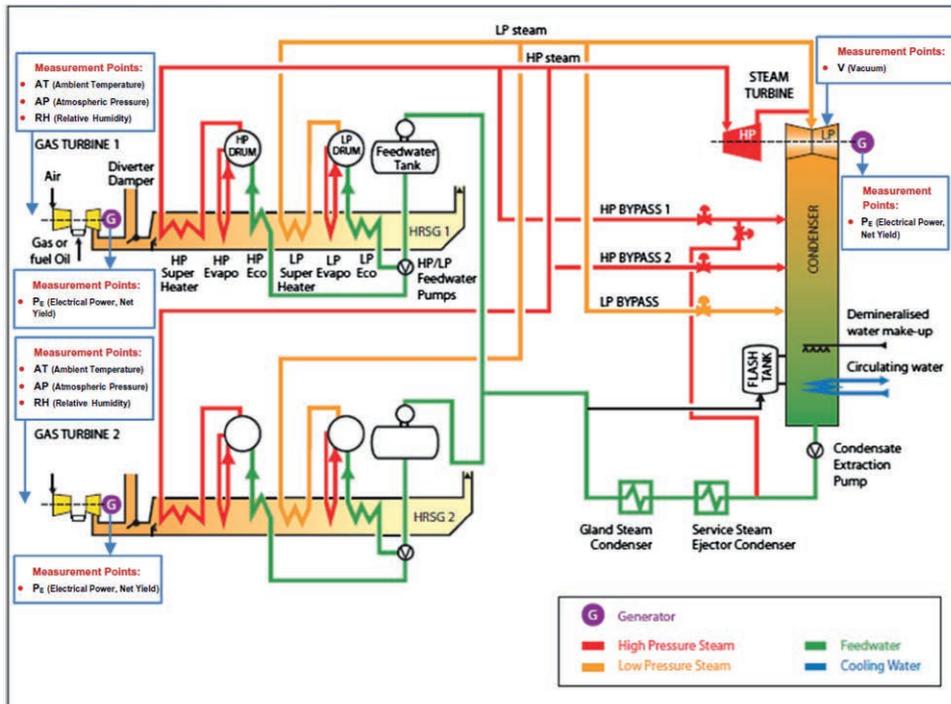


Figure 1: The combined cycle power plant layout (Tüfekci, 2014). It contains two gas turbines, a steam turbine, and heat recovery steam generators. The figure shows the measurement points of the input and output variables.

power outputs of CCPPs, they release partly low exhaust gases. CCPP generates 68% of electricity and the remaining 32% waste. In contrast, other kinds of power plants can generate only 33% electricity.

Given the advantages listed, it is clear that the use of CCPP increases, and predicting and interpreting the prediction model of a power plant became a trending topic for researchers and a critical problem. We need to have information on the influential factors and interactions between them to accurately predict a power plant's power output. Furthermore, it is favorable to have a power plant with high efficiency, reliable and sustainable and to maximize the income from the available megawatt-hours (MW/h). Figure 1 shows the CCPP layout and the sensors location (Tüfekci, 2014). The parameters that affect the CCPP are the ambient conditions such as ambient temperature (AT), atmospheric pressure (AP), and relative humidity (RH); and the exhaust steam pressure (or vacuum, V) effect on the steam turbine. These parameters are the input variables of the system, and the electrical power from both gas and steam turbines is the target variable.

The power plant was designed with a nominal generating capacity of 480 MW, consisting of two 160MW ABB 13E2 Gas Turbines, two dual pressure heat recovery steam generators, and one 160MW ABB Steam Turbine. The dataset contains data over six years. The dataset from the plant and data preprocessing is described in detail in the original paper (Tüfekci, 2014). The dataset consists of 9568 data points collected when the plant worked with a full load over 674 different days. This study is performed on the dataset of the mentioned paper with the same CCPP. The significant contribution of the paper is that we bring interpretability and explainability point of view in the framework of supervised machine learning approaches to realize and visualize the

effects and interaction of the predictor variables on the predicted response.

4. Results and discussion

Here we present the ALE plots used to visualize the effects of the predictors and interaction between them. All the results are attained from R software (R Core Team, 2020).

Figure 2 shows the ALE main-effect plot of input variables for an MLP prediction with one layer and 50 neurons. The AT main effect has sigmoidal behavior. Increasing the AT makes PE decrease. The RH main effect behaves quadratically. The V main effect has almost sigmoidal behavior, and the AP and RH main effects have quadratic behavior.

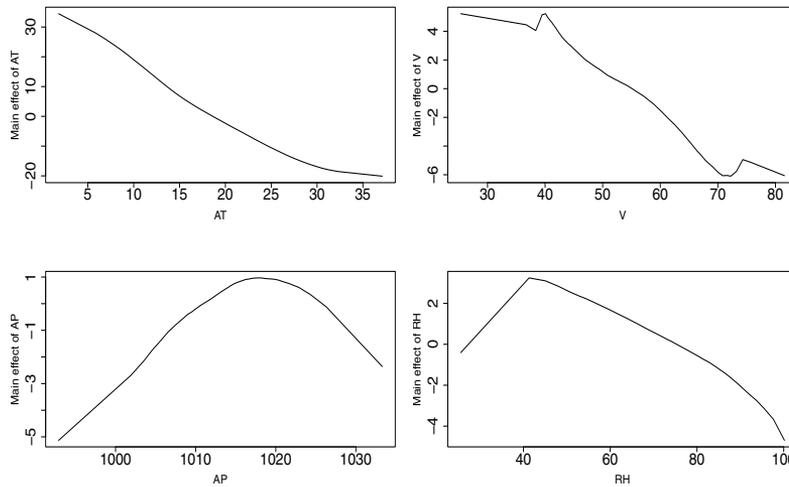


Figure 2: ALE main-effect plots for neural network with 50 neurons.

Figure 3 is the ALE second-order effect plot without the main effect of each input variable. They reveal the interaction between input variables for a neural network with 50 neurons. The numbers on the contours show the function values. The darker the chart color shows, the higher the function value. Figure 3 reveals notable interactions between AT and V since the contour values change over a range of 90 units (from -90 to 0), which is almost as large as the range for the main effect of AT in figure 2. Figure 3 shows almost moderate to negligible interaction in other subplots.

We can conclude that AT and V have the most interaction and AT-RH and AP-RH has the most negligible interaction based on figure 3.

5. Conclusion

This study aimed to improve the interpretability and explainability of a prediction model for the output of a CCPP as a case study by applying model-agnostic methods. A clarification on the concept of interpretability of machine learning predictions was provided. The importance of this study comes from the fact that most of the prediction models are like a black box. However, they are more accurate and faster than physical approaches with many nonlinear equations that might give undesirable and invalid results.

This study was performed on the same data set as Tüfekci's paper (Tüfekci, 2014). She studied different regression methods to present a model to predict an electrical power output. We worked

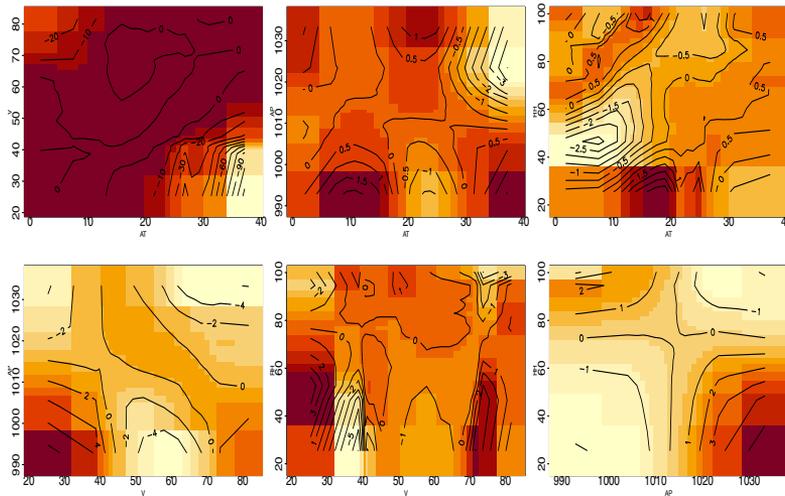


Figure 3: ALE second-order effect plots for for neural network with 50 neurons.

on the interpretability point of view of neural networks as machine learning methods. More precisely, our study focuses on analyzing the interactions between different variables and their individual effect on the output of the system. It is based on the accumulated local effects tool as a model-agnostic method for visualization to make more effortless interpreting. The model-agnostic flexibility is their remarkable advantage because it can be applied to any ML model. ALE would show the most important input variable, the relation between the input and output variables, and the second-order interaction between the input variables and the output variable.

To sum up, we can use ALE when the features of a machine learning model are correlated. While other model-agnostic methods, such as PDPs, are not trustable because they may average predictions of unrealistic artificial data instances in the mentioned case. A further research objective will include the comparison of ALE for different neural network architectures. Moreover, we could apply other model-agnostic methods such as partial dependence plots, global surrogate models, individual conditional expectation curves, and local interpretable model-agnostic explanations.

References

- D. W. Apley, J. Zhu, 2020. Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (4), 1059–1086.
- T. Danesh, R. Ouaret, P. Floquet, 2021. Towards robust process design. the sensitivity analysis using machine learning methods. *ENBIS 2021 Spring Meeting on Data Science in Process Industries*.
- C. Molnar, 2019. *Interpretable Machine Learning*.
- W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, B. Yu, 2019. Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences* 116 (44), 22071–22080.
- L. Niu, X. Liu, 2008. Multivariable generalized predictive scheme for gas turbine control in combined cycle power plant. In: 2008 IEEE Conference on Cybernetics and Intelligent Systems. IEEE, pp. 791–796.
- R Core Team, 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL <https://www.R-project.org/>
- M. T. Ribeiro, S. Singh, C. Guestrin, 2016. Model-agnostic interpretability of machine learning. arXiv preprint arXiv:1606.05386.
- P. Tüfekci, 2014. Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *International Journal of Electrical Power & Energy Systems* 60, 126–140.
- T. Zhang, Y. Li, Y. Li, S. Sun, X. Gao, 2020. A self-adaptive deep learning algorithm for accelerating multi-component flash calculation. *Computer Methods in Applied Mechanics and Engineering* 369, 113207.

Hybrid AI Models in Chemical Engineering – A Purpose-driven Perspective

Arijit Chakraborty^a, Sven Serneels^{b,c}, Heiko Claussen^b and Venkat Venkatasubramanian^a

^a*Department of Chemical Engineering, Columbia University, New York, NY 10027, United States of America*

^b*Aspen Technology Inc., 20 Crosby Dr, Bedford, MA 01730, United States of America*

^c*Department of Mathematics, University of Antwerp, B-2020 Antwerp, Belgium*

Abstract

Recent successes of machine learning in applications such as gaming, computer vision, and natural language processing, have generated considerable excitement for the application of purely *black-box* data driven techniques in other areas. However, unlike such applications, chemical engineering systems are governed by fundamental principles comprising of conservation laws and constitutive equations. Incorporating such natural constraints is valuable in many applications. However, as the complexity of systems increases, obtaining these first-principles models becomes exceedingly difficult. Hence the appeal of black-box models, which manage to perform well in some practical applications. This, however, comes at the cost of not being able to interpret and explain such a model's performance, which might limit its acceptance. As a result, hybrid AI models that combine first-principles with data driven techniques have been proposed in the literature. These attempt to eliminate the drawbacks of both approaches and provide insights into the system. Hybrid AI models can be developed for different end-user purposes – interpretability, interoperability, meeting desired performance targets and constraints, etc. This review article describes these disparate but related approaches, and provides a summary of recent progress in this field. Further, it provides a perspective for potential future research in this domain.

Keywords: mechanistic models, hybrid models, data driven model, reduced-order model

1. Introduction

In this emergent era of artificial intelligence (AI), particularly machine learning (ML), purely data driven *black-box* models are gaining great importance in many applications. Driven by their successes in game playing, natural language processing, and computer vision, they are seen as a panacea in all other domains as well.

However, unlike such applications, many science and engineering based systems, such as the ones in chemical engineering and materials science, are governed by fundamental principles comprising of, but not limited to, conservation laws and constitutive equations. Incorporating such natural constraints is valuable, even critical, in many applications. The complexity of many practical chemical engineering systems, however, poses challenges in formulating such *mechanism-based* transparent models. Often, in many practical settings, such models cannot be developed as not enough fundamental information is available to formulate them. Further, even if developed, they tend to be a large set of differential and algebraic equations, often nonlinear, with many variables and parameters that are hard to measure or compute and hence solve satisfactorily. Such drawbacks

have motivated people to develop data-driven "black-box" models with their own limitations, such as their inexplicability.

The recent trend on 'explainable AI', or XAI, focuses on making existing black-box systems explainable and interpretable to the end-user – a system that can explain its recommendation and reasoning, preferably in a natural-language-like fashion. While such an explanation might not be crucial for movies recommendations (*e.g.*, by Netflix), or for dinner choices (*e.g.*, by Yelp), they are critical for many engineering applications such as in control and safety (Venkatasubramanian, 2019). As the recent Boeing 737 Max-8 control system failures have demonstrated, the cost of a mistake is potentially high. Thus, the need to explain and justify the recommendations of an AI agent is imperative in such applications.

Combining first-principles knowledge with data driven techniques seems like a reasonable compromise to address this challenge. Hybrid models can be developed for different end-user purposes, such as to achieve interpretability, interoperability, meeting desired performance targets and constraints, etc. One of the earliest hybrid-AI model was developed by Sundaram et al. (Sundaram et al., 2001) for the inverse design of fuel additives. They combined a population balance-based, equations-driven, fundamental model that exploited physicochemical mechanisms of fuel additive function with a neural network-based model. The same group further developed a different hybrid AI model, but with a similar structure, for the inverse design of rubber compounds (Ghosh et al., 2003). Yet another example was the Reaction Modeling Suite system for catalyst design (Caruthers et al., 2003, Katare et al., 2001, 2004). All three approaches leveraged the underlying physics and chemistry effectively, even though they were all very different owing to the diverse nature of the applications. Such exploitation resulted in more explainable and reliable predictive models that did not require great amounts of data that were hard to acquire. These hybrid AI models combined *symbolic AI* (*i.e.*, a priori knowledge in the form of rules, schemas, and equations) with *numeric AI* (*i.e.*, data-driven techniques such as neural networks) and were implemented and used in industrial settings. Despite such early successes, the development of hybrid AI models remains a major challenge requiring a lot of time, effort, and considerable modeling skills. Hence, there is an urgent need to make such development more systematic, routine, relatively easy, and quick in a variety of applications.

Consequently, we feel there is a need to provide a coherent narrative to the various hybrid modeling approaches, specifically from the perspective of the end-user's purpose. We depict the wide range of models that one can obtain, and that hybrid AI models attempt to combine the best

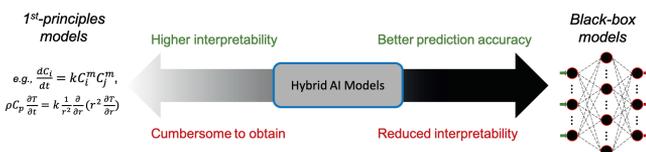


Figure 1: Combining 1st principles with black-box models yields hybrid AI models.

of the two extremes, in fig. 1. The various levels of granularity that go into developing a hybrid model often depend on the problem it aims to solve. In particular, this ranges from achieving interpretability, incorporating fundamental physicochemical mechanisms, targeted performance (execution speed and accuracy), adherence to constraints, interoperability (including extrapolation), etc. In this paper, we discuss the opportunities and challenges underlying such approaches.

2. Objectives of Hybrid Models

2.1. Interpretability

As noted, many safety-critical applications in chemical engineering and materials science require the models to be interpretable. This involves, among other things, being able to describe the underlying phenomena (*e.g.*, type and/or order of reaction kinetics) in fundamental mechanisms-based terms, explaining the quantitative dependence of state variables, demonstrating dimensionally consistent models, and producing causal models.

Recent work in attempting to model process systems considering the above guiding principles, has resulted in the combination of symbolic knowledge with domain dependent (intelligent) feature engineering into a machine-learning system. Symbolic knowledge (*i.e.*, symbolic AI) is accounted for through symbolic regression, where the task is to identify the model form that could have generated the data, and subsequently obtain the parameters. Genetic algorithms have been used as a feasible method for searching through the large function space, under user specified first-principles-based mechanistic constraints (Chakraborty et al., 2020). These models can be linear or nonlinear (Chakraborty et al., 2021). Constraints can be applied on the dimensionality of features generated, as has been successfully demonstrated in the domain of computational physics (Udrescu and Tegmark, 2020).

Such symbolic AI based models are inherently more simpler than black-box models. For a plant operator, one can easily perform a sensitivity analysis and obtain the dependence on such a model. Further, statistical techniques that provide uncertainty estimates of the parameters allow insight into the importance of features of the model. Parametric models such as these require significantly lower number of parameters than black-box models. One can even account for a priori knowledge (*e.g.*, in the form of priors, for Bayesian regression). Such parametric models can be beneficial in surrogate-based modeling.

Domain knowledge can also be injected into such models by penalizing intractable function transformations. This drawback is a recurring theme in purely black-box approaches, where overly complex function transformations are used to represent the model trajectory, resulting in poor generalizability. Additionally, such symbolic models tend to require much less data than their purely black-box counterparts. This is a major differentiator for hybrid models in many engineering applications that are in the 'precious data' domain, as noted, in contrast with standard machine learning techniques for 'big data' domains (such as computer vision, natural language processing, gaming, etc).

2.2. Adherence to Constraints

To deploy predictive models in many engineering applications, it is of great importance that the models respect constraints imposed by the first principles. Such constraints can be of varying nature, such as univariate limit constraints to more complex linear and nonlinear constraints that involve multiple input and output variables simultaneously. Examples of constraints derived from first principles that commonly occur in chemical engineering are mass, atom, and energy balance, or constraints on concentrations prescribed by the chemical reactions taking place in the equipment. Each of these constraints is typically linear in all or a subset of the variables considered, whereas some of these can be nonlinear as well. Atom balance constraints, for instance, are almost always nonlinear constraints.

In many applications, model predictions need to meet such constraints for the model to be deployable. In fact, one could consider machine learning models that respect a set of constraints, to be a first subset of hybrid models that combine black-box machine learning techniques with domain knowledge, with the latter being represented by the constraints. Several approaches can be taken to build machine learning models that respect such constraints. These approaches differ along

the guiding principle: either first-principles models generate the eventual estimate, adjuvated by inputs from machine learning, or the models are mainly data-driven machine learning models, but learning and predictions are adjusted such that they respect constraints. The former category is an active field of research, referred to as *Physics-Inspired Neural Networks* (PINNs) (Raissi et al., 2019) or *Neural Differential Equations* (NDE) (Chen et al., 2018). While such models are mathematically elegant, they can still lack predictive accuracy as they represent an approximation to the true first principles, since typically more phenomena take place than those represented by the differential equation layer in them. For these reasons, one can also consider the opposite approach, where the model is mainly data-driven but just slightly adjusted to the constraints. This can be done post-hoc, e.g., by reconciling predictions such that they meet constraints (Narasimhan and Jordache, 1999), but the best predictive performance can be obtained when the constraints are embedded into model estimation. For models that are defined according to an explicit optimization criterion, one can train models by numerically optimizing in the presence of first principle constraints. An example of such a model could be a sparse linear regression model that meets mass balance. In the case where all flows are defined, that becomes a mixed-integer linear optimization problem that can be solved numerically. On the other hand, more complex models such as (deep) neural networks are not defined according to single optimization problem. To build neural network models that inherently respect first principle constraints, one has to ascertain that the constraints are embedded into the training framework. In such case, common neural network optimization techniques such as ADAM can still be applied.

2.3. Scientific Machine Learning

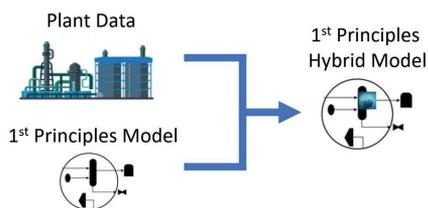


Figure 2: Combining 1st principles model knowledge and plant data yields 1st principles-driven hybrid model.

There has been a resurgence in the recent years in the field of incorporating domain knowledge into hybrid models through the use of known a priori model equations; however, the reader will find it interesting that this approach is not recent, as noted, with early contributions in 1990s and 2000s, e.g., (Psychogios and Ungar, 1992; Sundaram et al., 2001; Ghosh et al., 2003). Similar remarks apply to building more general hybrid models as they do to the approaches to account for constraints. Both first-principles-driven models (fig. 2), and predominantly data-driven models can be built (fig. 3), in pure form or as a nuanced mix of both approaches.

While first-principles-driven hybrid models such as neural differential equations do respect the constraints established through the differential equation, they obviously embed more first-principles information than just the constraints. However, as there are often multiple governing equations to the phenomena that describe a certain piece of equipment, data driven approaches' flexibility can be enticing.

Again there are several ways in which first-principles knowledge that extends beyond constraints can be embedded into models that are essentially data-driven. A first, and often successful, approach consists of domain-based feature engineering. Chemical engineering knowledge can both complement and simplify common statistical transformations, such as expanding the data with all cross products of variables and/or their squares. Governing physicochemical first-principles can inspire transformations specific to only a few variables that greatly enhance predictive power, taking advantage of the sparse nature of the dynamic system (Brunton et al., 2016). Examples of such variable transformations are common dimensionless numbers, such as the Reynolds or Damköhler numbers. In many cases, such transformations can already successfully describe the nonlinearity present within a certain operating window in (chemical) manufacturing processes. Therefore, to build predictive models in such a context, it can be sufficient to combine these transformations

with common regularized regression estimators, *e.g.*, the LASSO (Tibshirani, 1996) or Partial Least Squares (PLS) (Wold, 1975). Note that this approach can still be adopted when the data can be expected to contain outliers and uninformative variables, in which case one can substitute PLS for Sparse Partial Robust M-regression (SPRM) (Hoffmann et al., 2015), which is a sparse version of the older Partial Robust M-regression (PRM) (Serneels et al., 2005).

Of course, there are applications where the non-linearity provided by domain inspired feature engineering is insufficient to describe a real world process or piece of equipment, even when constrained to a narrow operating window. In such case, more involved nonlinear models should be estimated. Variable transformations such as the ones mentioned above can also be applied in combination with decision trees, where the interpretability of a certain split in the tree being based on a given threshold in a dimensionless number, can be valuable to a process engineer. Likewise, less involved transformations can be of great practical importance, such as being able to interpret the model's actions in terms of ratios of additives dosed to the bulk mass instead of the plain mass flows of those additives. Note that a lot of progress has recently been made to calculate decision trees optimally (Bertsimas and Dunn, 2019), which in many cases obliterates the necessity to combine decision trees into (random) forests. Such trees can be estimated from original and transformed variables and the required set of constraints can be passed on into the numerical mixed-integer optimization.

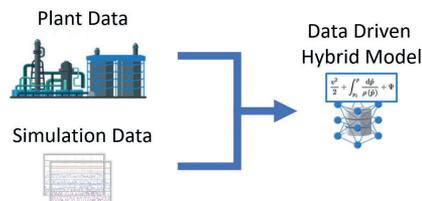


Figure 3: Combining simulation data and plant data yields a data-driven hybrid model.

It is well known that deep neural networks are efficient universal approximators for nonlinear functions (Kreinovich, 1991) and therefore, do not per se require nonlinear variable transformations to attain a satisfactory model fit. However, neural networks can also benefit from having input based on domain specific transformations, which can lead to simpler models (fewer variables, fewer hidden layers), which then on its turn will facilitate interpretability when model agnostic tools are applied to them for such purpose, such as LIME (Ribeiro et al., 2016). Moreover, the resulting simpler neural network models require lower amounts of training data and are less prone to over-fitting. Beyond variable transformations, it is possible to design custom layers that embed first principles knowledge. The effect of these first principles can then be analyzed by inspection of the corresponding layer estimates. Note that in a similar way, one can design the final layer and/or activation function in such a way that constraints are respected (Beucler et al., 2021).

3. Conclusion & Future Perspectives

This article has highlighted that different end-user requirements result in the need for different hybrid modeling approaches. Mechanistic insight can be obtained through symbolic models, where knowledge is captured through the relationships of state variables and their transformations. They facilitate interpretable models. On the other hand, enhanced predictive capabilities with the injection of domain knowledge, are the forte of constraints-based models. These account for mechanistic knowledge about the system by virtue of formulating the problem as a parameter estimation exercise, within the permissible regions in the parameter space. Consequently, we obtain models that are capable of predicting the system with considerable accuracy, but at the cost of providing mechanistic insight. Additionally, there has been a resurgence in the domain of scientific machine learning, where the strengths of function approximating neural networks are combined with first-principles knowledge, to yield mechanistically feasible neural networks. Such models are able to extrapolate outside the domain of training data.

As the future is aimed towards the goal of Industry 4.0 (Sansana et al., 2021), hybrid models will play a major role. The increasing availability of data opens a multitude of possibilities. It is imperative that we take advantage of the same – by augmenting the rapid progress of AI in data driven modeling with process/system knowledge. Further research is needed to develop systematic frameworks that quickly and easily help develop hybrid AI models that incorporate first-principles knowledge of physicochemical mechanisms, cause-and-effect relationships, constraints, performance targets, and more.

References

- D. Bertsimas, J. Dunn, 2019. *Machine Learning Under a Modern Optimization Lens*. Dynamic Ideas, LLC, Belmont, MA, USA.
- T. Beucler, M. Pritchard, S. Rasp, J. Ott, P. Baldi, P. Gentine, 2021. Enforcing analytic constraints in neural networks emulating physical systems. *Physical Review Letters* 126 (9), 098302.
- S. L. Brunton, J. L. Proctor, J. N. Kutz, 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences* 113 (15), 3932–3937.
- J. Caruthers, J. A. Lauterbach, K. Thomson, V. Venkatasubramanian, C. Snively, A. Bhan, S. Katare, G. Oskarsdottir, 2003. Catalyst design: knowledge extraction from high-throughput experimentation. *Journal of Catalysis* 216 (1-2), 98–109.
- A. Chakraborty, A. Sivaram, L. Samavedham, V. Venkatasubramanian, 2020. Mechanism discovery and model identification using genetic feature extraction and statistical testing. *Computers & Chemical Engineering* 140, 106900.
- A. Chakraborty, A. Sivaram, V. Venkatasubramanian, 2021. Ai-darwin: A first principles-based model discovery engine using machine learning. *Computers & Chemical Engineering* 154, 107470.
- R. T. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, 2018. Neural ordinary differential equations. *arXiv preprint arXiv:1806.07366*.
- P. Ghosh, S. Katare, P. Patkar, J. M. Caruthers, V. Venkatasubramanian, K. A. Walker, 2003. Sulfur vulcanization of natural rubber for benzothiazole accelerated formulations: from reaction mechanisms to a rational kinetic model. *Rubber Chemistry and technology* 76 (3), 592–693.
- I. Hoffmann, S. Serneels, P. Filzmoser, C. Croux, 2015. Sparse partial robust M regression. *Chemometrics and Intelligent Laboratory Systems* 149, 50–59.
- S. Katare, J. M. Caruthers, W. N. Delgass, V. Venkatasubramanian, 2004. An intelligent system for reaction kinetic modeling and catalyst design. *Industrial & engineering chemistry research* 43 (14), 3484–3512.
- S. Katare, P. Patkar, P. Ghosh, J. Caruthers, W. Delgass, V. Venkatasubramanian, 2001. A systematic framework for rational materials formulation and design. In: *Proceedings of the Symposium on Materials Design Approaches and Experiences*. pp. 321–332.
- V. Y. Kreinovich, 1991. Arbitrary nonlinearity is sufficient to represent all functions by neural networks: a theorem. *Neural networks* 4 (3), 381–383.
- S. Narasimhan, C. Jordache, 1999. *Data reconciliation and gross error detection: An intelligent use of process data*. Elsevier.
- D. C. Psychogios, L. H. Ungar, 1992. A hybrid neural network-first principles approach to process modeling. *AIChE Journal* 38 (10), 1499–1511.
- M. Raissi, P. Perdikaris, G. Karniadakis, 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378, 686–707.
- M. T. Ribeiro, S. Singh, C. Guestrin, 2016. Why should i trust you?: Explaining the predictions of any classifier.
- J. Sansana, M. N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L. H. Chiang, M. S. Reis, 2021. Recent trends on hybrid modeling for industry 4.0. *Computers & Chemical Engineering*, 107365.
- S. Serneels, C. Croux, P. Filzmoser, P. Van Espen, 2005. Partial robust M-regression. *Chemometrics and Intelligent Laboratory Systems* 79 (1-2), 55–64.
- A. Sundaram, P. Ghosh, J. M. Caruthers, V. Venkatasubramanian, 2001. Design of fuel additives using neural networks and evolutionary algorithms. *AIChE journal* 47 (6), 1387–1406.
- R. Tibshirani, 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B* 58, 267–288.
- S.-M. Udrescu, M. Tegmark, 2020. Ai feynman: A physics-inspired method for symbolic regression. *Science Advances* 6 (16), eaay2631.
- V. Venkatasubramanian, 2019. The promise of artificial intelligence in chemical engineering: Is it here, finally. *AIChE J* 65 (2), 466–478.
- H. Wold, 1975. Soft modeling by latent variables: the non-linear iterative partial least squares approach. In: J. Gani (Ed.), *Perspectives in probability and statistics, papers in honor of M.S. Bartlett*. Academic Press, London, UK, pp. 117–142.

Determination of Adsorption Energies from DFT databases using Machine Learning techniques

José I. Arsuaga^a and Ana I. Torres^a

^a*Instituto de Ingeniería Química, Facultad de Ingeniería, Universidad de la República, J. Herrera y Reissig 565, Montevideo 11300, Uruguay*
aitorres@fing.edu.uy, +59827142714 ext. 18102

Abstract

This paper discusses the estimation of adsorption energies for reaction intermediates for a given metallic surface and molecule. Regression models are learned from DFT data available in the literature in a two step approach. First, metallic surfaces are characterized by a principal component analysis (PCA) followed by a suitable orthonormal rotation to find a set of species that can be used as descriptors for the metallic surface. Then, different machine learning techniques are considered for the regression using the previous descriptors for the metallic surface and molecular descriptors such as the number and type of bonds for the adsorbate. With the available data, CH_3 , CO_2 and CH_2 were found to explain 93% of the total variance, thus were used as surface descriptors. Three of the tested models were found to adjust similarly well to validation data.

Keywords: Adsorption energies, Machine Learning, Electrocatalysis

1. Introduction

Electrocatalytic reactions have recently gained a lot of attention as they can be powered by electric energy from renewable non-programmable sources to achieve zero or even negative carbon processes. In particular, reduction of CO_2 on metallic surfaces is a promising process to transform industrial CO_2 emissions into valuable fuels and products. However, electrocatalytic processes are characterized by their very low selectivity; for example in the case of CO_2 many C1 and C2 products including acids (e.g formic acid, acetic acid), alcohols (e.g methanol and ethanol), and light hydrocarbons (for example ethylene) have been experimentally reported.

It is clear that the properties of the metallic surface play a role on the selectivity. As examples, copper and copper alloys or copper oxides seem to favor the production of ethylene and methanol from CO_2 (Dinh et al. (2018), Wang et al. (2018)), and platinum the production of CH_4 (Umeda et al. (2020)). Yet, the reaction mechanisms are not completely understood, and the lack of understanding hinders reactor and process design. Reaction network generators can be used to build the possible reaction pathways; still, thermodynamical properties of the proposed adsorbates need to be known to infer which of the pathways are feasible. One option to compute the required thermodynamical properties is the use of density functional theory (DFT). However, this approach requires very specialized knowledge, is time consuming and resource intensive in terms of computational power. Thus, it may not be suitable if a large amount of DFT-derived data is needed for calculations in a particular application.

In this work, we have taken another approach to estimate the required thermodynamic properties (adsorption energies) which is based on machine learning using data from DFT that is already available in the literature. Figure 1 schematizes the procedure. The rationale is that adsorption

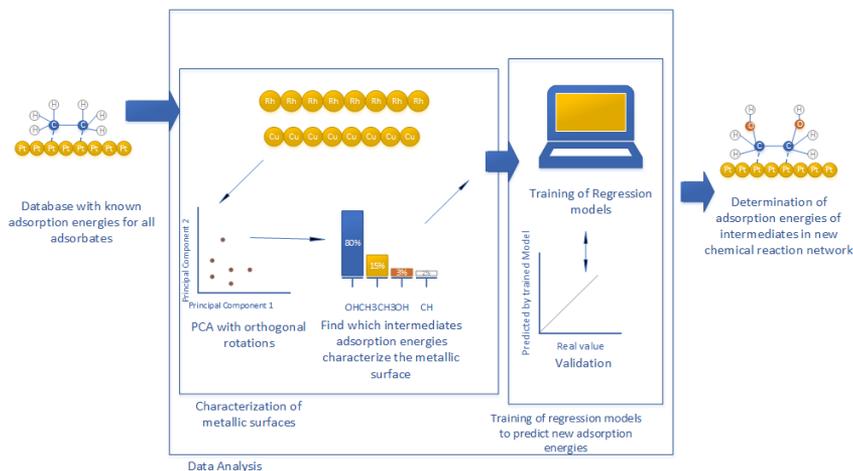


Figure 1: Schematic of the machine learning approach to find a model for estimating adsorption energies of adsorbates in metallic surfaces.

energies depend on properties of the adsorbates (molecules) and properties of the metallic surfaces. While it is fairly clear which properties should be considered as descriptors of the molecules (e.g. atoms, type and amount of bonds, etc.), it is not that clear which ones should be considered for the metallic surfaces. Therefore, the main hypothesis is that given a dataset containing the adsorption energies for several adsorbates in different metallic surfaces, it is possible to find a set of descriptors for the surfaces that are based on a subset of these energies. This idea is not new, Chowdhury et al. (2018) has proposed a similar approach using a dataset of 29 molecules and 8 metallic surfaces. In here, we have largely expanded the dataset including new surfaces and adsorbates, as a result, a different set of descriptors is obtained. On the basis of these new descriptors for the surfaces, and those of the molecules, a regression model was trained to predict adsorption energies of other molecules on the previous surfaces. The longer term objective is to use this regression model together with network generators to predict preferred reaction pathways for electrocatalytic reactions of different species on different surfaces.

2. Selection of descriptors of the metallic surface

As mentioned in the introduction, in order to find a correlation able to estimate the adsorption energies of any pair of adsorbate/metallic surface, a way of characterizing the surfaces needs to be found. A traditional DFT approach to describe a surface would need for example information on the atoms that compose the surface and their geometrical arrangement. These characteristics will affect the adsorption energies of all the adsorbates, although the effect on the possible adsorbates is different for each one, and depends mainly on the adsorbate itself. The idea then is to find those adsorbates whose energy of adsorption changes the most when the metallic surface changes. If this set of adsorbates is small, then using them as descriptors of the surface is a very practical and efficient way to characterize the surface, as instead of performing DFT calculations for all the components, we need DFT calculations for just a few.

2.1. PCA analysis of the energy of adsorption data

Principal component analysis (PCA) is a technique that given a data matrix $M_{N \times P}$, finds a new space of reduced dimensions which conserves a maximum amount of the variance of the original

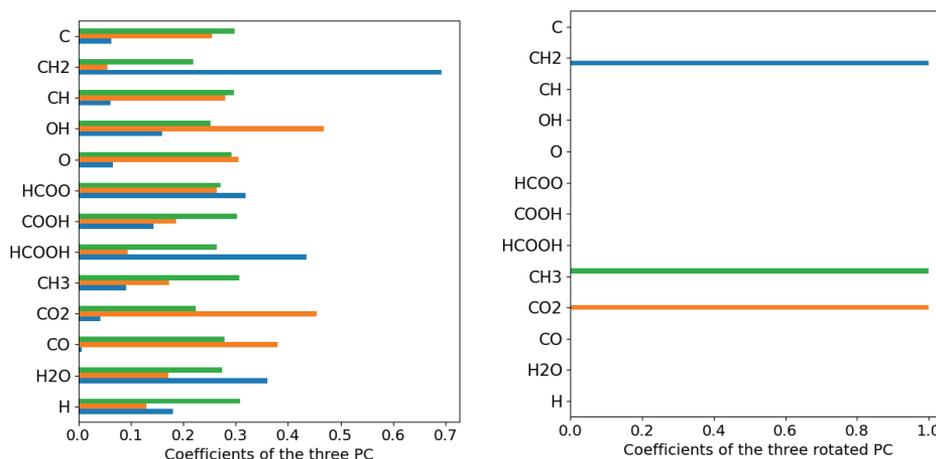


Figure 2: PCA analysis: left plain PCA; right PCA+varimax rotation. In both, green, blue and red are first, second and third component respectively.

data. This new space can be built by finding the eigenvectors of the covariance matrix (Σ_M) of the data in M . To establish which are the principal components of the data in M we look for the eigenvectors of Σ_M associated with the largest eigenvalues. In practice, those that added together make up for a certain threshold of the variance (usually 90-95%).

In our case, M is the matrix of adsorption energies (taken from DFT databases); each M_{np} entry is the adsorption energy for the p -th adsorbate on the n -th metallic surface. Figure 2 shows the results of applying PCA to a dataset built using adsorption energies reported in Plauack et al. (2016); Herron et al. (2012, 2013, 2014); Xu et al. (2018a); Bai et al. (2019); Ford et al. (2010); Mavrikakis et al. (2002); Ojeda et al. (2010); Scaranto and Mavrikakis (2016a); Singh et al. (2014); Scaranto and Mavrikakis (2016b); Ferrin et al. (2012); Greeley and Mavrikakis (2002); Ford et al. (2005); Chen et al. (2019); Xu et al. (2018b); Kregelberg et al. (2004); Hahn and Mavrikakis (2014); Grabow and Mavrikakis (2011); Gokhale et al. (2008); Li et al. (2016); Herron et al. (2014); Saliccioli et al. (2010, 2012); Lu et al. (2015, 2012); Schmidt and Thygesen (2018); Wellendorff et al. (2015). In here, it is important to mention that to apply PCA-techniques, M has to be complete, which means that only those adsorbates for which we found DFT data for *all* the surfaces of interest (Cu,Pt,Pd,Rh,Re,Ru,Ag,Au,Fe,Ir,Os,Co,Ni) were included.

Figure 2-left shows the three principal components as green (first component), orange (second component) and blue (third component) bars. Together these three are able to explain 93% of the variance of the original dataset (results obtained using the `scikit-learn` package in Python Pedregosa et al. (2011)). The adsorption energy for each adsorbate can then be expressed in terms of these three principal components; the absolute value of the weights that need to be applied are represented in the figure by the length of each bar.

2.2. PCA with varimax rotation

Unfortunately, the results in Fig.2-left are of little use as they are, as the principal components lack of physical interpretation. It would be desirable to have results where each component is clearly dominated by one or a few adsorbates. This is accomplished by finding a new set of orthogonal axis that represent a basis of the same space as the principal components, but in which the axis align better with some of the adsorbates. In this way, the coefficients of many of the adsorbates

Table 1: Summary of the performance of the regression

Method	KRR-poly	KRR-rbf	KSVR-poly	KSVR-rbf	CART	RF
RMSE	0.22	0.25	0.23	0.27	0.2	0.15
RMSE wo/H	0.12	0.13	0.12	0.13	0.18	0.11
sign mismatch	CO_2	-	CO_2	-	O	-

become zero, and those that are non-zero can be interpreted as the descriptors.

What was discussed was solved by Kaiser (1958) who proposed to find the new axes by solving the optimization problem in Eq. 1.

$$\max_R \sum_{j=1}^{j=K} \left(\frac{1}{N} \sum_{i=1}^{i=N} ((A'R_{ij}^2)^2) - \alpha \frac{1}{N^2} (A'\bar{R}_{ij}^2)^2 \right) \quad (1)$$

$$s.t : R^T R = I$$

In here, A is the original $P \times K$ eigenvector matrix ($K = 3$ as there are three principal components in our case study) and R the $K \times K$ rotation matrix. α is a parameter of the problem, if $\alpha = 1$ Eq. 1 is the Varimax rotation.

Fig.2-right shows the results when applying the varimax rotation. These results indicates that CH_3 , CO_2 and CH_2 , as first (79% of the variance), second (9% of the variance) and third component (5% of the variance) respectively, can be used as descriptors of the metallic surface (results also obtained with `scikit-learn`). Notice that this is different from the results in Chowdhury et al. (2018) who obtained OH and $CHCHCO$ as descriptors. The difference lies in the expansion of the dataset to include data from different sources, as the exactly same results as in the Chowdhury et al. (2018) are obtained when considering their database.

3. Learning a model to predict adsorption energies

After a suitable set of descriptors for the metallic surfaces is found, a regression problem that uses them and those of the adsorbates, can be formulated to learn a model for the energies of adsorption from data. As descriptors of the molecules we have considered the number and type of bonds in the adsorbate; we have also added facet and coverage as additional descriptors for the metallic surface when available. Notice now that completeness of the data is not required for this step, thus all available data can be used.

The following techniques were considered for learning the model: Kernel Ridge Regression (KRR, with polynomial and radial basis functions as kernels), Kernel Support Vector Regression (KSVR, with polynomial and radial basis functions as kernels), Classification and regression trees (CART) and Random Forest (RF). For the sake of space, we will not describe these methods, they are well explained in several references including `scikit-learn` documentation Pedregosa et al. (2011). In all cases, an 8-fold cross validation scheme was performed to define the set of possible values for the hyperparameters for each technique. Learning/ validation division of the dataset was 85/15% respectively. A summary of the results is in Table 1. As seen KRR-rbf KSVR-rbf and RF provided similar results in terms of RMSE over the validation set. Most importantly, they did not predict positive energies as negative nor the other way around. This was a problem that we observed when using polynomial kernels (wrong sign prediction for CO_2) and CART (wrong sign prediction for O). From a physical viewpoint, this is troublesome because it would imply that certain adsorbates cannot be adsorbed when in reality they can. In here, it has to be commented that we identified some outliers for the energy of adsorption of H (2 out of 90 datapoints for H) from the reported DFT data; the table includes the RMSE results with and without considering these outliers.

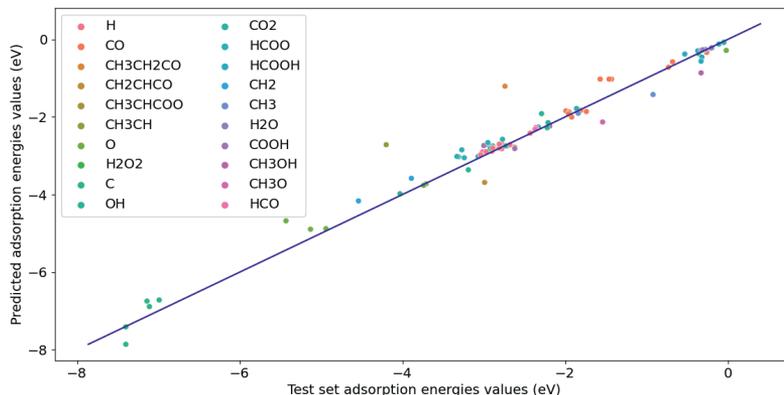


Figure 3: Results of the regression using Random Forest. Adjustment for $y = x$: $R^2 = 0.96$

Finally, Fig.3 presents the predicted vs DFT adsorption energies for data in the validation set. This type of plot showing a good regression is typical for all the techniques we have tested. This stresses the need of verifying that the chosen model correctly assigns the sign of the energy of adsorption for all data in the validation set, a point that may be overseen by just looking for the best RMSE and predicted vs real data fitting.

4. Conclusion

A large data set of DFT-based adsorption energies for different metallic surfaces and adsorbates was used to train a regression model. In a first step, PCA followed by Varimax rotation was found to be able to characterize the metallic surfaces using CH_3 , CO_2 and CH_2 as principal components, with a loss of information less than 10% in terms of variance of the data. Six regression models based on either Kernel Ridge, Support Vector, CART or Random Forest were considered. All models were found to provide good estimations in terms of RMSE, but some had trouble in assigning a correct sign to those adsorbates whose energy of adsorption was close to zero. Those that can correctly estimate the sign could be used together with reaction network generators to predict thermodynamically feasible pathways.

References

- Y. Bai, D. Kirvassilis, L. Xu, M. Mavrikakis, 2019. Atomic and molecular adsorption on ni(111). *Surface Science* 679, 240–253.
- B. W. J. Chen, D. Kirvassilis, Y. Bai, M. Mavrikakis, Apr 2019. Atomic and molecular adsorption on ag(111). *The Journal of Physical Chemistry C* 123 (13), 7551–7566.
- A. J. Chowdhury, W. Yang, E. Walker, O. Mamun, A. Heyden, G. A. Terejanu, 2018. Prediction of adsorption energies for chemical species on metal catalyst surfaces using machine learning. *The Journal of Physical Chemistry C* 122 (49), 28142–28150.
- C. T. Dinh, T. Burdyny, M. Kibria, A. Seifitokaldani, C. Gabardo, F. P. García de Arquer, A. Kiani, J. Edwards, P. Luna, O. Bushuyev, C. Zou, R. Quintero-Bermudez, Y. Pang, D. Sinton, E. Sargent, 05 2018. Co 2 electroreduction to ethylene via hydroxide-mediated copper catalysis at an abrupt interface. *Science* 360, 783–787.
- P. Ferrin, S. Kandoi, A. U. Nilekar, M. Mavrikakis, 2012. Hydrogen adsorption, absorption and diffusion on and in transition metal surfaces: A dft study. *Surface Science* 606 (7), 679–689.
- D. C. Ford, A. U. Nilekar, Y. Xu, M. Mavrikakis, 2010. Partial and complete reduction of o2 by hydrogen on transition metal surfaces. *Surface Science* 604 (19), 1565–1575.
- D. C. Ford, Y. Xu, M. Mavrikakis, 2005. Atomic and molecular adsorption on pt(111). *Surface Science* 587 (3), 159–174.

- A. A. Gokhale, J. A. Dumesic, M. Mavrikakis, Jan 2008. On the mechanism of low-temperature water gas shift reaction on copper. *Journal of the American Chemical Society* 130 (4), 1402–1414.
- L. C. Grabow, M. Mavrikakis, Apr 2011. Mechanism of methanol synthesis on cu through co₂ and co hydrogenation. *ACS Catalysis* 1 (4), 365–384.
- J. Greeley, M. Mavrikakis, Jun 2002. A first-principles study of methanol decomposition on pt(111). *Journal of the American Chemical Society* 124 (24), 7193–7201.
- K. Hahn, M. Mavrikakis, Feb 2014. Atomic and molecular adsorption on re(0001). *Topics in Catalysis* 57 (1), 54–68.
- J. A. Herron, J. Scaranto, P. Ferrin, S. Li, M. Mavrikakis, Dec 2014. Trends in formic acid decomposition on model transition metal surfaces: A density functional theory study. *ACS Catalysis* 4 (12), 4434–4445.
- J. A. Herron, S. Tonelli, M. Mavrikakis, 2012. Atomic and molecular adsorption on pd(111). *Surface Science* 606 (21), 1670–1679.
- J. A. Herron, S. Tonelli, M. Mavrikakis, 2013. Atomic and molecular adsorption on ru(0001). *Surface Science* 614, 64–74.
- H. F. Kaiser, 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23 (3), 187–200.
- W. P. Krekelberg, J. Greeley, M. Mavrikakis, Jan 2004. Atomic and molecular adsorption on ir(111). *The Journal of Physical Chemistry B* 108 (3), 987–994.
- S. Li, J. Scaranto, M. Mavrikakis, Oct 2016. On the structure sensitivity of formic acid decomposition on cu catalysts. *Topics in Catalysis* 59 (17), 1580–1588.
- J. Lu, S. Behtash, A. Heyden, 2012. Theoretical investigation of the reaction mechanism of the decarboxylation and decarbonylation of propanoic acid on pd(111) model surfaces. *The Journal of Physical Chemistry C* 116 (27), 14328–14341.
- J. Lu, M. Faheem, S. Behtash, A. Heyden, 2015. Theoretical investigation of the decarboxylation and decarbonylation mechanism of propanoic acid over a ru(0001) model surface. *Journal of Catalysis* 324, 14–24.
- M. Mavrikakis, J. Rempel, J. Greeley, L. B. Hansen, J. K. Nørskov, 2002. Atomic and molecular adsorption on rh(111). *The Journal of Chemical Physics* 117 (14), 6737–6744.
- M. Ojeda, R. Nabar, A. U. Nilekar, A. Ishikawa, M. Mavrikakis, E. Iglesia, 2010. Co activation pathways and the mechanism of fischer–tropsch synthesis. *Journal of Catalysis* 272 (2), 287–297.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- A. Plauck, E. E. Stangland, J. A. Dumesic, M. Mavrikakis, 2016. Active sites and mechanisms for h₂o₂ decomposition over pd catalysts. *Proceedings of the National Academy of Sciences* 113 (14), E1973–E1982.
- M. Saliccioli, Y. Chen, D. G. Vlachos, 2010. Density functional theory-derived group additivity and linear scaling methods for prediction of oxygenate stability on metal catalysts: Adsorption of open-ring alcohol and polyol dehydrogenation intermediates on pt-based metals. *The Journal of Physical Chemistry C* 114 (47), 20155–20166.
- M. Saliccioli, S. M. Edie, D. G. Vlachos, 2012. Adsorption of acid, ester, and ether functional groups on pt: Fast prediction of thermochemical properties of adsorbed oxygenates via dft-based group additivity methods. *The Journal of Physical Chemistry C* 116 (2), 1873–1886.
- J. Scaranto, M. Mavrikakis, 2016a. Density functional theory studies of hcooh decomposition on pd(111). *Surface Science* 650, 111–120, the surface science of heterogeneous catalysis: In honor of Robert J. Madix.
- J. Scaranto, M. Mavrikakis, 2016b. Hcooh decomposition on pt(111): A dft study. *Surface Science* 648, 201–211, special issue dedicated to Gabor Somorjai’s 80th birthday.
- P. S. Schmidt, K. S. Thygesen, 2018. Benchmark database of transition metal surface and adsorption energies from many-body perturbation theory. *The Journal of Physical Chemistry C* 122 (8), 4381–4390.
- S. Singh, S. Li, R. Carrasquillo-Flores, A. C. Alba-Rubio, J. A. Dumesic, M. Mavrikakis, 2014. Formic acid decomposition on au catalysts: Dft, microkinetic modeling, and reaction kinetics experiments. *AIChE Journal* 60 (4), 1303–1319.
- M. Umeda, Y. Niitsuma, T. Horikawa, S. Matsuda, M. Osawa, 2020. Electrochemical reduction of co₂ to methane on platinum catalysts without overpotentials: Strategies for improving conversion efficiency. *ACS Applied Energy Materials* 3 (1), 1119–1127.
- L. Wang, S. A. Nitopi, E. Bertheussen, M. Orazov, C. G. Morales-Guio, X. Liu, D. C. Higgins, K. Chan, J. K. Nørskov, C. Hahn, T. F. Jaramillo, Jul 2018. Electrochemical carbon monoxide reduction on polycrystalline copper: Effects of potential, pressure, and pH on selectivity toward multicarbon and oxygenated products. *ACS Catalysis* 8 (8), 7445–7454.
- J. Wellendorff, T. L. Silbaugh, D. Garcia-Pintos, J. K. Nørskov, T. Bligaard, F. Studt, C. T. Campbell, 2015. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected dft functionals. *Surface Science* 640, 36–44, reactivity Concepts at Surfaces: Coupling Theory with Experiment.
- L. Xu, D. Kirvassilis, Y. Bai, M. Mavrikakis, 2018a. Atomic and molecular adsorption on fe(110). *Surface Science* 667, 54–65.
- L. Xu, J. Lin, Y. Bai, M. Mavrikakis, Jun 2018b. Atomic and molecular adsorption on cu(111). *Topics in Catalysis* 61 (9), 736–750.

Probabilistic graphical models for the identification and analysis of reaction pathways in water treatment

Rachid Ouaret^{a*}, Ali Badara Minta^{a,b}, Claire Albasi^a, Jean-Marc Choubert^b and Antonin Azaïs^{b*}

^a*Laboratoire de Genie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France*

^b*INRAE, UR REVERSAAL, 5 Rue de La Doua, CS 20244, F-69625, Villeurbanne Cedex, France*
rachid.ouaret@toulouse-inp.fr; antonin.azais@inrae.fr

Abstract

In this paper, we proposed a new pathway enrichment analysis for reaction schemes based on Bayesian Network (BN). Several reaction schemes are proposed in the waster waters treatment literature, however, these data diverge greatly and, to date, are unexploited. As an innovative alternative, probabilistic graphical models could allow establishing parentage between the most relevant molecules and Transformation Products (TPs). Based on the analysis of an extensive bibliography gathering more than 45 articles for more than 140 molecules, this study proposed an innovative methodology based on knowledge graphs, pathway data enrichment, and Bayesian Networks (BN) analyses. The proposed methodology has been applied to elucidate the degradation of the sulfamethoxazole. Probabilistic approaches and graphical models bring a new light on the identification of the TP parentage.

Keywords: Bayesian Networks, Graphical Models, Pathway Enrichment Analysis, Knowledge Engineering,

1. Introduction

Water is a scarce resource, the ultimate receptacle of anthropogenic pollution. Its preservation and treatment are the subject of new paradigms in terms of analytical development, processes and even data management. In the last decades, the scientific community has pointed out the presence of new synthetic molecules (manufactured products) in all aquatic compartments (domestic, surface and groundwater). Several studies have proposed various breakdown pathways of a targeted micropollutant, which could be different depending on study objectives, analytical methodology, etc. Often these studies are oriented on the formation of Transformation Products (TPs) in controlled conditions and with a singular process. Most of TPs can retain or sometimes amplify toxic effects on aquatic life. The formation and the evolution of TPs remains a subject in full expansion in wastewater treatment. Their prediction is difficult as there is no consensus on their pathways even if a same couple "parent compound – process" is applied. The available data on the occurrence of TPs remain fragmentary, but sufficient to prevent any attempt at manual synthesis. Indeed, the reaction schemes proposed in the literature diverge greatly, as much in their structure as in the number and nature of the molecules (elements) represented. It is assumed that these differences are the result of a combination of causes including the type of process(es) used, their operating conditions, the analytical power (resolution) available and finally the expertise intrinsic to the research consortium.

However, even considering recent studies Chen et al. (2019); Yazdanbakhsh et al. (2020) con-

ducted on the same contaminant (sulfamethoxazole, an antibiotic) and the same type of process, here oxidation, it appears that only one of the produced molecules is common between the two examples considered. To date and to the authors' knowledge, the pool of data generated by these mechanistic (physicochemical) studies is unexploited from statistical methodology point of view. Also, no attempt has been made to consolidate these data. Therefore, data science appears to be an innovative alternative to establish the most probable filiation between molecules and to identify the preferential reaction pathways. To address these issues, the ANR TRANSPRO project (2019-2022), led by a consortium of public partners (EPOC, LGC and INRAE), aims to improve knowledge on the nature, origin and dynamics of TPs. In view of the set of molecules breakdown pathways proposed in the literature, graph data mining techniques would provide new insights into the most common degradation paths of a targeted micropollutant. It is in this perspective that probabilistic graphical models have been chosen to conduct this study. From the table of all the molecules observed on the analyzed literature. The analysis of data would allow disentangling the relevant relations between the molecules, the dependencies or independences between several groups of molecules. The use of Bayesian networks will allow the extraction of a compact representation of pathways of degradation without loss of information. In addition to the reaction schemes identification, a methodology of data enrichment of the pathways' degradation database and specific data analysis have been proposed. From a preferential degradation scheme, it is possible to prioritize the most recurring Products of Transformation (TP) and thus to privilege them for all environmental and toxicological studies. The methodology was applied to a largely described micropollutant, sulfamethoxazole (SMX).

2. Methodology

2.1. Literature analysis of degradation Pathways of micropollutants

This study is articulated around the three complementary parts (Figure 1): **(A)** deals with pathway enrichment analysis step and data mining pre-processing from dedicated search engines such as ScienceDirect. Part **(B)** consists of the identification of molecules breakdown pathways of a target pollutant (here, SMX) and matrix representation of degradation levels, and finally part **(C)** represents the use of information from different reaction schemes for Bayesian Networks (BN) analysis and probabilistic graph based analysis.

Part (A) : Data pathway enrichment analysis. In this step, 45 scientific papers have been selected and analyzed for the SMX degradation pathways identification. To date, no automatic method to extract pathway information is proposed in the literature. For this reason, part **(B)** is designed to feed a database, which is created for this purpose. **The part (B)** is dedicated to the extraction of degradation pathways, requiring a manual transcription of all the molecules as well as their breakdown level, or position, in the degradation scheme. To date, this part of the data enrichment is done manually. Two different types of processes have been distinguished in these papers, 30 scientific papers based on Advanced Oxidation Processes (AOP) and 15 present the degradation of SMX through the implementation of various BIOlogical processes (BIO). From these 45 papers, 141 molecules (SMX + 140 TPs) and 177 reaction pathways were considered. Then, an adjacency matrix weighted by the number of arcs observed between the different by-products is deduced.

Figure 2 shows an example of the transcription procedure of two degradation patterns into an easily exploitable matrix in data processing step. Assuming that there are two papers in which two SMX degradation patterns leading to the formation of the following products: A, B, C, D and E. In the matrix, we consider the molecules in the column as the incoming molecules from the degradation of the starting molecules that are in the row. The values in the matrix correspond to the number of times the degradation of the parent molecule into the child molecule is observed in all the articles for all the proposed reaction pathways. This example allows to better understand, from two articles, the challenge and the complexity of the data enrichment part, especially when

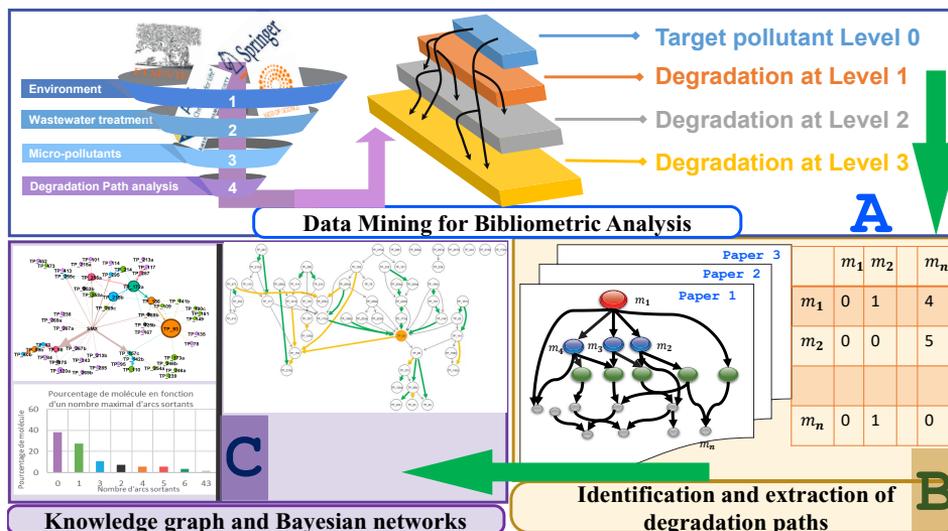


Figure 1: General scheme of the methodology followed for the analysis of the degradation pathways of molecules observed in the literature in water treatment. Two complementary blocks (A-B) for the enrichment of the database and a block (C) for the analysis by probabilistic graphical models and knowledge graph based analysis.

Table 1: Data matrix used for Bayesian networks analysis. For each paper, the process (AOP or BIO) of degradation of the SMX molecule is identified. The transformation products (TP) are identified at their respective level of degradation.

Paper	Path	SMX	TP(1)	TP(2)	...	TP($n-1$)	TP(n)
Paper 1	path1	AOP	level 5	level 2	...	level 5	level 7
	path2	AOP	level 2	level 1	...	level 2	level 3
	path3	BIO	level 1	level 4	...	level 1	level 8
Paper 2	path4	AOP	level 6	level 3	...	level 2	level 4
	path5	BIO			...		
...
Paper p	path m	AOP	level 6	level 3	...	level 7	level 8

the number of papers to analyze becomes high.

Part (C) consists in representing the graphs of interactions between the target molecule (SMX) and a set of its transformation products. Two approaches have been used. First, by integrating information on the number of direct links between molecules (arc weight) and the number of arcs coming out from the different molecules (node weight). In this case, the data is extracted from and the model is assumed to be faithful to the physical reality. Then, from the table 1, Bayesian networks modeling based approach is applied to find similarities with the reference model. In the latter case, very little information is provided: a table with 177 rows (pathways) and 141 variables (molecules). The molecules obtained from the degradation of SMX were noted by TP followed by their m/z ratio: the mass of the molecules relative to their charge number. For example, TP_{93} corresponds to the transformation product with an m/z ratio of 93. In some cases, two molecules with a different structure are detected with the same molar mass. It was therefore chosen to add a letter to their molar masses to differentiate them, as in the case of 283a and 283b. For each molecule, a binary coding has been applied: observed molecule (Yes) for a given pathway and unobserved molecule (Not Observed).

2.2. Bayesian Networks

The main topic of this work can be summarized in one question: *can we conceive of graph mining methods that can extract the most probable pathways from several reaction schemes?* Probabilistic graphical models, and more precisely Bayesian networks, seem to us appropriate to shed new light on this question. Moreover, a graphical model has the advantage of being easy to interpret, and according to the modeled properties (independence for example) can be more or less powerful to detect relationships of parentage between molecules. From the table obtained in part B (see table 1), it is possible to use a Bayesian Network on nominal variables. In the absence of an a priori structure, the use of this type of model requires learning of the graphical structure (nodes and arcs) as well as the estimation of the conditional probability distribution associated to each molecule (random variables). In our case, an a priori structure can be attributed to the graph obtained from the adjacency matrix. We will first test a model with an a priori structure, then a series of simulations is performed for graph learning. The probabilistic relationships in a Bayesian network are represented by a qualitative description- a graph (\mathcal{G}), and a quantitative description- an underlying joint probability distribution. Formally, a Bayesian network $B(\mathcal{G}, \theta)$ is defined by (i) a directed acyclic (without circuit) graph $\mathcal{G}(X, E)$ whose nodes are associated with a tuple of categorical random variables $X = \{X_1, X_2, \dots, X_n\}$ and E represents a set of arcs, and (ii) a set of probabilities $\theta = \{\mathbb{P}(X_i | Pa(X_i))\}$ of each node conditional on the state of its ancestors (parents) $Pa(X_i)$ in \mathcal{G} . In the case where all variables are observed, the simplest and most used method for estimating probabilities is the Maximum Likelihood (ML) method which gives $\widehat{\mathbb{P}}(X_i = x_k | Pa(X_i) = x_j) = \widehat{\theta}_{ijk}^{ML} = \frac{N_{ijk}}{\sum_k N_{ijk}}$, where N_{ijk} is the number of instances in the data table where X_i takes its k^{th} value x_{ik} and the variables in $Pa(X_i)$ takes their j^{th} configuration for all k . Regarding the learning of the graph, there are two types of techniques to build the structure of the Bayesian network: Score Based Algorithm (SBA) and Constraint Based Algorithm (CBA). For a review of the graph learning methods, we recommend the very complete books by Koller and Friedman (2009) and Murphy (2012). The use of CBA based algorithms on SMX physico-chemical reaction data often leads to a construction of graph that are not very faithful to the structures observed in the literature. For this reason, this study focuses on the score methods and more particularly on the Hill-climbing (HC) algorithms (Chickering, 2002). Score-based approaches range into the search space by examining only possible local changes in the neighborhood of the current solution and applies the modification (add arrows, delete,...) that maximizes the score function. Very briefly, these algorithms structure the network from an empty, randomly generated or predefined graph (a priori structure). Then, a score is computed on the

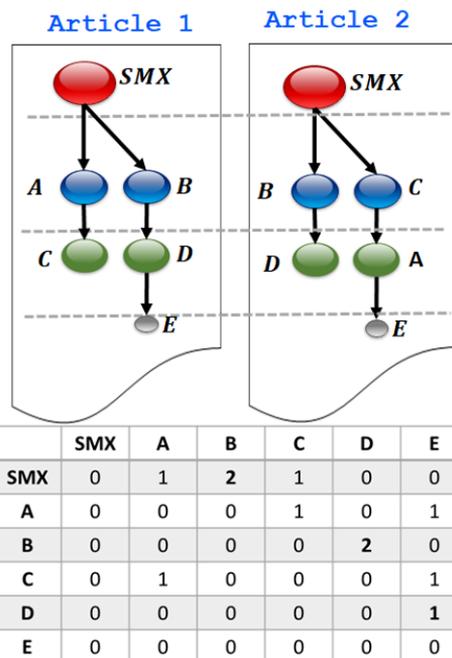


Figure 2: Simplified example of two degradation paths with the associated adjacency matrix. A value greater than or equal to 1 shows the number of times a direct link between two molecules was observed in all papers. The value 0 codes the absence of link detected between the molecules.

basis of this initialization and the modifications in the graph (addition, inversion or deletion of an arrow) are performed. If the modification does not increase the score, the algorithm returns to the previous state and another modification is performed; if on the contrary the score increases, the new state is kept and a new modification is performed. The algorithm stops when no modification increases the score. The score-based methods are easily trapped in the numerous local minima and the final graph obtained depends strongly on the initial conditions. Thus, an empty graph represents the most frequent choice in the absence of a priori knowledge. Regarding the choice of the score, the Bayesian Dirichlet Equivalent (BDe) score proposed in (Heckerman et al., 1995) provides a graphical structure most similar to the graph obtained via the literature database we implemented (*cf.* data from the table 1).

Let the values $\{x_{i1}, \dots, x_{ir_i}\}$, $r_i \geq 1$, $i = 1, \dots, n$ of all the set that each X_i can take, r_i is the number of states of the finite random variable, D is the database, and G the structure of the network on X (set of random variables). The number of possible configurations for the parents of X_i is determined by $q_i = \prod_{X \in Pa(X_i)} r_i$. From these elements, BDe score and is defined by:

$$\mathbb{P}(G, D) = \mathbb{P}(G) \mathbb{P}(D | G) \iff \mathbb{P}(D | G) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + N_{ijk})}{\Gamma(\alpha_{ijk})} \quad (1)$$

where $\mathbb{P}(G)$ represents the a priori probability assigned to the structure G . The parameters α_{ijk} defined by $\alpha_{ijk} = \eta \times \widehat{\mathbb{P}}(X_i = x_k, Pa(X_i) = x_j | G_c)$ with G_c the complete connected graph and the equivalent sample size η expresses the strength of our belief in the prior distribution.

3. Results and discussion

Developing Bayesian networks was also one of the main objectives of our study. These networks allow the observation of dependency or independence relationships between the different variables (here, molecules). Figure 3-(A) shows the representation of the Bayesian network (HC algorithm with a BIC score) obtained, as well as the common parentage to the knowledge graph from the literature review. Three different colors are proposed, green if the observed parentage is identical and orange if one of the two molecules (parent or child) is reversed. The black arcs illustrate false detections or the absence of parentage between the molecules. First, we notice that there are 40% of links (green arcs) that coincide with the knowledge graph from the literature (green color code). Furthermore, note that nodes with no direct relatives are considered the first level of degradation of the SMX, so arcs (not shown in the graphs) from the SMX to these molecules are considered "valid". The molecule *TP_93* was also detected as a very frequent transformation product, i.e. several incoming arcs from different nodes. In view of the results obtained, we can confirm our hypothesis on the usefulness of Bayesian networks for the identification of preferential reaction patterns. For all the scores tested, the BDe gives a good compromise between graph structure and interpretation.

Figure 3-(B) presents a Bayesian network built with an a priori structure initialization: the graph is extracted from the literature review. The advantage of this approach compared to the knowledge network is to be able to (i) simulate the preferential paths, (ii) identify and control the monitoring molecules, and (iii) analyze the behavior of some degradation pathways. It is also possible to perform any kind of inference based on the conditional probabilities obtained from this graph and the collected data. For example, the joint probability of the most recurrent pattern ($\mathbb{P}(SMX, TP_{98}, TP_{83}, TP_{98b}, TP_{60b})$) can be obtained through a table of calculation of all the modalities taken by the TPs. The probability of simultaneously observing TPs of preferential path and SMX is about 7.5%. It is also easy to retrieve all the tables of conditional or marginal probabilities associated with the different transformation products. As a simple example, $\mathbb{P}(TP_{98} | SMX) = 19.15\%$, $\mathbb{P}(TP_{98}, TP_{83} | SMX) \approx 8\%$, or as presented in (Figure 3):

$$\mathbb{P}(TP_{93} = \text{Observed}, TP_{172a} = \text{Observed} | \{SMX, TP_{255a} = \text{Observed}\}) = 31.9\%.$$

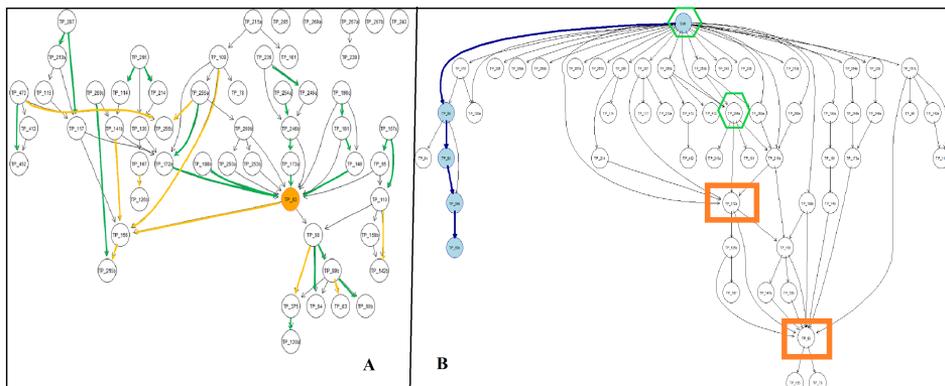


Figure 3: Bayesian networks obtained for Biological processes. Left (A), without “a priori” network, the structure was learned from data (BIO processes) and compared to the knowledge graph from the literature review. Right (B) Bayesian network with structure initialization and highlighted (blue path) the most recurrent degradation path. A query such as searching probability of orange molecules knowing green ones : $\mathbb{P}(\text{orange rectangles} \mid \text{green hexagon}) = 31.9\%$ is possible from the BN inference (Right B)

Specifying the arguments in the inference procedures requires some care, but the result is an extremely flexible framework to compute the probability of arbitrary combinations of events. As an example of a more complex query, we can compute:

$$\mathbb{P}(TP_{93} = \text{Observed}, TP_{172a} = \text{Observed} \mid \{SMX, TP_{255a} = \text{Observed}\} \cup \{TP_{89b} = \text{Observed}\}) = 22.3\%$$

4. Conclusions and future work

To day, no methodology for consolidating and processing data from molecule degradation pathways are considered in the literature. The proposed methodology explores a new way and attempts to respond to this lack, to create consensus on the basis of studies where data are widely dispersed, from heterogeneous or even contradictory sources. Regarding the Bayesian networks, the graphical models shed new light on the data enrichment pathways of molecule degradation between the transformation products when passing through wastewater treatment plants. To date, the part of the database enrichment is done manually. This database is often fed by updating the articles and identifying new molecules and reaction pathways. The most ambitious perspective is the automation of the database enrichment phase. In future work we intend to develop a framework for the data retrieve automation of degradation pathways. In these conditions, an identification of molecules and their properties would be possible by extending our field of investigation in graph mining.

References

- M. Chen, C. Guo, S. Hou, L. Wu, J. Lv, C. Hu, Y. Zhang, J. Xu, 2019. In-situ fabrication of ag/pg-c3n4 composites with enhanced photocatalytic activity for sulfamethoxazole degradation. *Journal of hazardous materials* 366, 219–228.
- D. M. Chickering, 2002. Optimal structure identification with greedy search. *Journal of machine learning research* 3 (Nov), 507–554.
- D. Heckerman, D. Geiger, D. M. Chickering, 1995. Learning bayesian networks: The combination of knowledge and statistical data. *Machine learning* 20 (3), 197–243.
- D. Koller, N. Friedman, 2009. *Probabilistic graphical models: principles and techniques*. MIT press.
- K. P. Murphy, 2012. *Machine learning: a probabilistic perspective*. MIT press.
- A. Yazdanbakhsh, A. Eslami, M. Massoudinejad, M. Avazpour, 2020. Enhanced degradation of sulfamethoxazole antibiotic from aqueous solution using mn-wo3/led photocatalytic process: Kinetic, mechanism, degradation pathway and toxicity reduction. *Chemical Engineering Journal* 380, 122497.

Integrating autoencoder and heteroscedastic noise neural networks for industrial data analysis and soft-sensor design

Sam Kay^{a,†}, Harry Kay^{a,†}, Max Mowbray^{a,*}, Amanda Lane^b, Cesar Mendoza^b, Philip Martin^a, Dongda Zhang^{a,*}

^a*Department of Chemical Engineering and Analytical Science, University of Manchester, Oxford Road, Manchester, M1 3AL, UK.*

^b*Unilever Research Port Sunlight, Quarry Rd East, Bebington, C63 3JW, UK.*

†: *These authors contributed equally to this work.*

*: max.mowbray@manchester.ac.uk; dongda.zhang@manchester.ac.uk;

Abstract

Viscosity represents a key indicator of product quality but has traditionally been difficult to measure in-process in real-time. This is particularly true if the process involves complex mixing phenomena operated at dynamic conditions. To address this challenge, a promising solution to monitoring product viscosity is to design soft-sensors which correlate viscosity with easily measured process variables. In this study, we developed an innovative machine learning based soft-sensor construction framework by integrating different types of advanced artificial neural networks. The framework first employs a deep learning autoencoder to generate information-rich statistic latent variables by compressing high-dimensional industrial data, and then adopts a heteroscedastic noise neural network to simultaneously predict product viscosity and process variance based on the extracted latent features. To evaluate its accuracy and robustness, the data-driven soft-sensor was used to predict product viscosity for a number of industrial batches operated over different seasons. It is found that the soft-sensor has both high accuracy (prediction error <12%) and high robustness in most of the cases, indicating its great potential for industrial batch process monitoring and quality control.

Keywords: Machine learning, data analytics, dimensionality reduction, quality control, uncertainty estimation.

1. Introduction

Quality control is paramount to many industrial processes to avoid the viscosity of the final product falling outside the predetermined acceptable boundaries, otherwise, the entire batch must be discarded, effectively wasting the entire process time. Not only does this lead to excessive losses due to material wastage, but it also incurs costs involved with safe disposal of the defective products. Ideally, advanced fault detection technique should be available to monitor the progress of the batch over the process time however, there is currently no efficient approach to quickly measure the viscosity of complex liquid formulations during their industrial scale production.

Real time prediction of viscosity has historically been a challenge within the consumer goods industry. The difficulties stem from the current lack of understanding of rheology within the context of highly viscous fluids, making it impractical to derive any accurate

physical models for viscosity prediction. It is common for industrial processes to sample measurements during a process to directly measure the viscosity. However, this is time consuming and if poor batch quality is observed, there is little opportunity to adjust the process to prevent deviations of viscosity outside acceptable boundaries. A solution to this can be applied with the usage of data-driven models which find the underlying relations that lie within the data recorded from a series of sensors on a plant and the measured viscosity. These data-driven models are then used as a soft-sensor for new process prediction. In addition, to mitigate false confidence, these data-driven soft-sensors should be able to make accurate estimations and represent the aleatoric uncertainty present within the data, analogous to the uncertainty specified within the instrument used to record the viscosity values. A further challenge associated with data-driven models is the capability to generalise to different processes. To resolve this, an excess of data is required by the model from different processes. However, this data is usually not readily available and is time consuming to produce and sort.

2. Problem statement

This study focuses on a batch process for consumer goods product production. The aim is to develop a robust soft-sensor for final viscosity prediction using real-time process sensors' recordings. Three datasets are available and are referred to as Alpha, Beta, and Gamma dataset, respectively, where the first two were obtained from the same process line and the latter being obtained from a similar, but different process. These datasets contain 30, 16 and 11 batches, respectively. Each batch contains 28 sensors recording temperature, pressure, and flowrate in different locations in the process. The actual batch process generates around 7000 times series data points and real-time data is recorded once per second or once per two seconds. In our previous work (Hicks et al., 2021), partial least squares has been used to identify commonly important sensors and critical time regions within the datasets, leaving us with only information relevant to predicting viscosity. This process reduced all 3 datasets to a 3-rank tensor of n batches, 300 timesteps and 8 sensors (n being dependent on the dataset). Before being used for training or validation purposes, the 3-rank tensors are timewise unfolded to generate a 2-rank matrix for soft-sensor construction.

Dimensionality reduction is necessary in this case to mitigate problems associated with high dimensional data analysis (Min, 2005) and to improve accuracy of the data-driven soft-sensor. Dimensionality reduction techniques remove multi-collinearity, improves the capability of the model to generalise to new datasets and will remove redundant features whilst retaining important characteristics of the data. Here we offer the usage of an autoencoder for dimensionality reduction as opposed to traditional linear regression techniques such as PCA and PLS. Although the usage of autoencoders is well demonstrated in a number of recent studies, the implementation in chemical process soft-sensing and monitoring is sparsely explored. Once key process information is extracted through dimensionality reduction, we adopted a heteroscedastic noise neural network (HNN) to make accurate viscosity predictions and meaningful uncertainty estimations. Successful predictions are those which replicate the industrial measurement errors (~10%). In addition, to effectively identify the best structure for the autoencoder and HNN, Bayesian optimisation is performed to optimise the hyperparameters of these models through construction of a Gaussian process surrogate model.

3. Methodology

First, we will introduce the structure of the autoencoder. For a given input (the process data) x , we obtain a projection z , and a reconstruction x' . The autoencoder is defined by two neural networks (Fournier & Aloise, 2019):

An encoder – This is defined by a function $f(x) = z$, where x and z are the respective inputs and outputs of the network.

A decoder – This is defined by a function $f(z) = x'$, where z and x' are the respective inputs and outputs of the network.

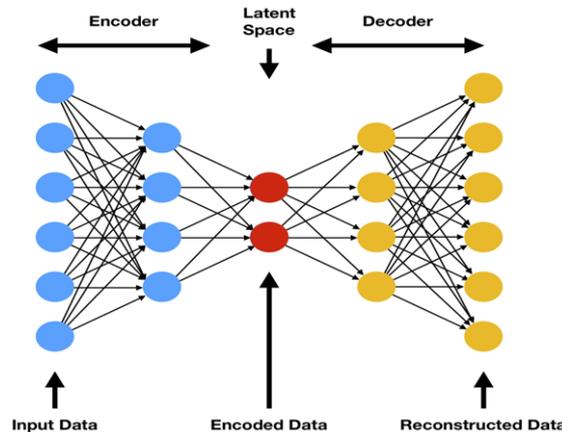


Figure 1: Diagram of a general autoencoder and its respective components.

The training objective is defined by a cost function in which the distance between the input and reconstruction error is to be minimised. Henceforth, the cost function assumed for the construction of the autoencoder will be the mean squared error; for an input of n datapoints, the error $E = \frac{1}{n} \sum (x - x')^2$ and the activation function applied to all hidden layers will be ELU (exponential linear unit) defined by $ELU(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases}$ (Clevert et al., n.d.). A condition must be enforced on the model to copy the input data as its output (i.e. reconstruct the dataset), in the interest of extracting useful properties and characteristics from the process data. The autoencoder found after hyperparameter optimisation and used throughout this paper is defined in Table 1.

Table 1: Representation of the parameters used for the construction of the autoencoder.

Hidden layer	Number of nodes	Learning rate	Epochs	Activation function
1	1404	0.001477	782	ELU (Exponential linear unit), $\alpha = 0.15$
2	94			
3	50			
4	16			
5	50			
6	94			
7	1404			

Let us now describe the structure of the heteroscedastic neural network (HNN) generating the viscosity predictions. The objective of the HNN is to accurately predict batch quality and associate each prediction with an uncertainty metric representing the confidence of the model in its predictions. Intuitively, it is essential for the model's uncertainty estimates to provide very high coverage probabilities for its predictions. Hereafter, the following metrics will be defined to establish a consistent method of comparison between soft sensor models: the mean average percentage error, $MAPE = \frac{|y - \mu(x)|}{\mu(x)}$; the percentage uncertainty, $PPU = \frac{3\sigma(x)}{y}$ and the coverage probability, $CP = P(PPU > MAPE)$ where $\sigma(x)$ is the standard deviation of the HNN's prediction, $\mu(x)$ is the output viscosity from the HNN and y is the measured viscosity. Using these predefined metrics, we can determine the best HNN structure based on its performance using cross validation. An important note is that the PPU was not considered to be a priority in determining the best performing model, as long as its estimated standard deviation is around 10%. It was however essential that the model's prediction errors were covered by the uncertainty estimations, so the following condition was created to ensure a viable soft-sensor, $CP > 0.8$. It is required to provide a meaningful uncertainty estimate for each prediction so, a gaussian negative log likelihood term has been implemented for the loss function of the HNN, where the loss, $L = \frac{1}{2} \ln(\sigma^2(x)) + \frac{(y - \mu(x))^2}{2\sigma^2(x)}$ (Hirschfeld et al., 2020). $\sigma(x)$ represents the uncertainty estimation of the HNN (one standard deviation), y is the measured values of viscosity and $\mu(x)$ is the predicted viscosity values.

Both the autoencoder and the HNN were optimised using Bayesian optimisation techniques assuming a GP surrogate function (Song et al., 2019), this allowed for efficient searching and mapping of the hyperparameter space, especially useful for when large numbers of hyperparameters must be optimised; such is true for the autoencoder. Once viable optima were found, manual refining was employed to further explore the regions surrounding each optimum to ensure the quality and stability of the solution.

4. Results and Discussion

Two-fold cross-validation on the alpha dataset was used to optimise the hyperparameters of the HNN soft-sensor using the latent space extracted from the autoencoder, averaging the MAPE and PPU over the 435 possible combinations. The final HNN structure, activation function, epochs, MAPE, PPU and CP are shown in Table 2 below.

Table 2: Parameters of the optimised HNN soft-sensor.

Number of hidden layers	2
Number of nodes [layer1, layer2]	[31, 3]
Learning rate	0.0125
Activation function	Sigmoid
Number of epochs	160
MAPE [training, validation]	[7.8, 10.0]
PPU [training, validation]	[28.9, 28.3]
CP [training, validation]	[1, 1]

The finalised model, as shown in Table 2 was used to predict the batch quality of both the beta and the gamma datasets.

The viscosities have been normalized to a range of 0-1 and the results are as shown below in Figures 2a and 2b.

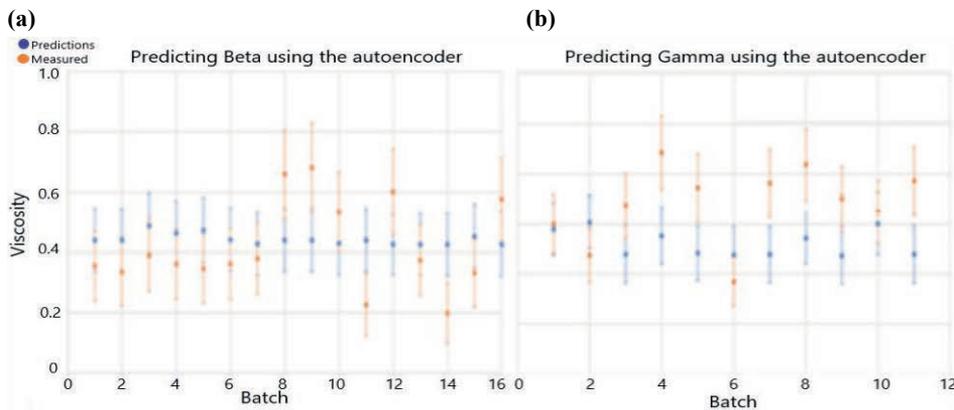


Figure 2: Plots of the soft sensor predictions against the measured values for batch data for datasets Beta (a) and Gamma (b). The error bars represent one standard deviation.

As shown in Figure 2 (a), the soft-sensor is capable of predicting Beta dataset’s batch quality to a reasonable degree of accuracy on data derived from the same process as the training set (Alpha), with an average MAPE of 11.3%. Similarly, the uncertainty estimations average to $\pm 26.0\%$ for three standard deviation (± 8.67 for one standard deviation). Notably, 87.5% of the datapoints for the measured and predicted viscosities have overlap between the error bars. This means that the soft-sensor’s predictions are lying within the acceptable range of values determined from experimental procedure; the expected standard deviation of measurements taken for the viscosity is $> \pm 10\%$ due to standard experimental errors (i.e. uncertainties of measurement equipment and human error in experiments). The significance of this is such that the soft-sensor seems to have successfully replicated the error within the process data in its predictions. The MAPE, PPU and CP for the HNN’s validation results of the beta dataset can be found in Table 3.

A similar conclusion can be drawn of the models capacity to predict the batch quality of the Gamma dataset as to that of the Beta dataset. The results indicate slightly worse performance with an average MAPE of 15.5% this, however is expected due to the data being derived from a different process. The model also seems to provide less overlap between the error bars with only 54.5% of them overlapping for one standard deviation. The MAPE, PPU and CP for the HNN’s validation results of the gamma dataset can be found in Table 3.

Table 3: Validation result of the HNN soft-sensor.

Dataset	MAPE	PPU	CP
Alpha	10.0	28.3	1
Beta	11.3	26.0	1
Gamma	15.5	26.8	1

Overall, it was anticipated that the HNN would provide better performance on the Alpha and Beta datasets than on the Gamma data due to them both being obtained from the same process line, meaning their characteristics should be similar. Eventhough the datasets were obtained at different times of the year, the performance on beta would indicate that this had little effect on the models capacity to accurately predict batch quality meaning that the inherent features of the process data was similar. On Figure 2(a), it can be seen that there is little variety in the estimation of viscosity made by the HNN. This could be attributed to the nature of the autoencoder as viscosity is not taken into account when extracting the feature space of dataset provided. This leaves the possibility that the important physical relations between the sensor data and the viscosity, that are necessary for making accurate predictions are partially lost when reducing the dimensionality. A second explanation could be that there is no identifiable difference between the features of each batch that would give rise to a reason for the model to predict largely different viscosities. This pattern also arises within the Gamma dataset, shown on Figure 2(b), however the extent of which is less severe.

5. Conclusion

In conclusion, autoencoders are a viable dimensionality reduction method and can be used in conjunction with machine learning regression models to help build a robust model to predict outcomes of industrial processes such as batch quality. The autoencoder is able to effectively produce small latent spaces which accurately represent large datasets thus removing the problems associated with high dimensionality. Through the use of data-driven models and non-linear dimensionality reduction techniques, it is possible to reduce the computational cost whilst identifying a high-quality solution to industrial problems. Through cross validation, it is observed that the HNN model has both high accuracy and high reliability when predicting final viscosity. The developed soft-sensor was also found to be able to predict different processes operated over a broad time span with a relatively large viscosity variation. This is the significant benefit over conventional linear regression based dimensionality reduction and modelling methods. Overall, this work demonstrates the innovative combination and potential impact of different machine learning techniques for high dimensional industrial data analysis and batch process monitoring.

References

- Clevert, D.-A., Unterthiner, T., & Hochreiter, S. (n.d.). *FAST AND ACCURATE DEEP NETWORK LEARNING BY EXPONENTIAL LINEAR UNITS (ELUS)*.
- Fournier, Q., & Aloise, D. (2019). Empirical comparison between autoencoders and traditional dimensionality reduction methods. *Proceedings - IEEE 2nd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2019*, 211–214. <https://doi.org/10.1109/AIKE.2019.00044>
- Hicks, A., Johnston, M., Mowbray, M., Barton, M., Lane, A., Mendoza, C., Martin, P., & Zhang, D. (2021). A two-step multivariate statistical learning approach for batch process soft sensing. *Digital Chemical Engineering, 1*, 100003. <https://doi.org/10.1016/j.dche.2021.100003>
- Hirschfeld, L., Swanson, K., Yang, K., Barzilay, R., & Coley, C. W. (2020). Uncertainty Quantification Using Neural Networks for Molecular Property Prediction. *J. Chem. Inf. Model*, 2020, 3770–3780. <https://doi.org/10.1021/acs.jcim.0c00502>
- Min, R. (2005). *A Non-linear Dimensionality Reduction Method for Improving Nearest Neighbour Classification*.
- Song, J., Chen, Y., & Yue, Y. (2019). *A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes*.

Economic risk analysis of emergent green hydrogen facilities – A Monte Carlo approach

^aBruno Gerard, ^bEduardo Carrera, ^aDenis Lun, ^cLu Zhao

^a*Envision Digital, 1 passerelle des reflets, 92400 Courbevoie, France*

^b*Capgemini Engineering, 2 Rue Paul Dautier, 78140 Vélizy-Villacoublay, France*

^c*Envision Digital, 1 Harbourfront Ave, #17-01 Keppel Bay Tower, Singapore*
denis.lun@envision-digital.com

Abstract

This paper presents an analysis of green hydrogen production from geothermal and solar energy sources through a Monte Carlo simulation approach, using a first version of a digital twin for the intelligent design of hydrogen facilities. The proposed methodology allows comparing hydrogen production based on three different configurations (case study 1, 2, and 3) by calculating the internal rate of return and the levelized cost of hydrogen. The trade-off between system availability and capital investment to reduce investment risks was also studied. The results show that hydrogen production from geothermal energy source presents business opportunities even in conservative scenarios. The valorization through batteries of surplus solar electricity implies significant investment risks. The cost of electricity and the investment costs of electrolysis were identified as the parameters with the highest impact on the LCOH. In addition, the best trade-off between system availability and return on capital has been studied, concluding that redundancy in the electrolyzers is sufficient for achieving it. Future research will need to improve the calculation of system availability and maintenance costs.

Keywords: Green hydrogen, Monte Carlo simulation, renewable energy, Digital Twins.

1. Introduction

Hydrogen is an emerging energy vector that has been identified as a key driver for achieving sustainable development goals. However, some challenges must be overcome. In this regard, the future of hydrogen economy depends on reducing costs (i.e., CAPEX and OPEX) and minimizing investment risks. Furthermore, hydrogen facilities are complex systems that involves the interaction between different energy sources and feedstock, technologies, hydrogen physical form, and storage modes (Carrera and Azzaro-Pantel, 2021), being required to address the uncertainty linked to the data available for its optimal design.

Digital Twin (DT) solutions could help to address these challenges. The DT concept is based on the development of multi-physical, multiscale and probabilistic models considering the integration of physical and virtual products through the exchange of data (Wang et al., 2020). Methodologies based on data-driven DT approaches have been used for optimal operation and simulation of green hydrogen technologies (Jaribion et al., 2020; Meraghni et al., 2021). However, this approach has not necessarily been applied to the design of entire facilities. To address these gaps, our study proposes a methodology based on stochastic simulations for the evaluation of green hydrogen facility designs. The main contribution of this work is thus to address data uncertainty and its propagation to KPIs, identify the parameters that have the highest impact on LCOH, and choose the

facility design that presents the best trade-off between system availability and CAPEX. The paper is divided into five sections following this introduction. The methods and tools are presented in Section 2 with a focus on the stochastic simulations. The case studies are described in Section 3. Finally, the main results are analyzed in Section 4, followed by the conclusions and perspectives presented in Section 5.

2. Methods and tools

A methodological framework for the analysis of hydrogen production facilities from a Monte Carlo approach is presented. It is summarized in the Fig. 1.a.

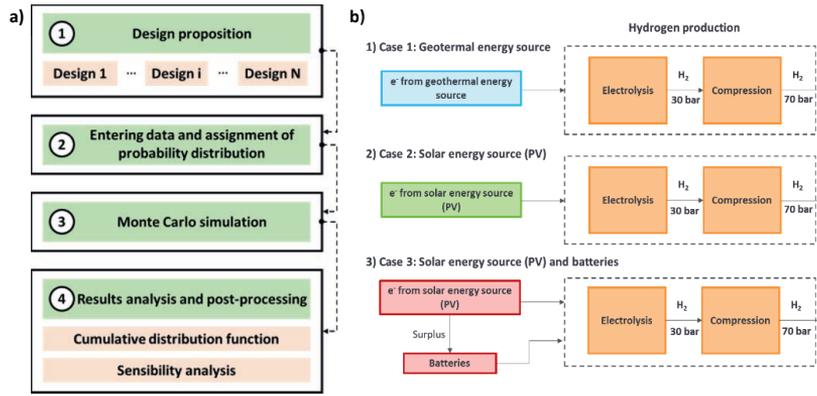


Fig. 1. a) Methodological framework; b) Case studies

2.1. Design proposition

Different hydrogen production configurations are proposed to analyze business opportunities, which may consider redundancy between equipment.

2.2. Entering data and assignment of probability distribution

The data considers the characteristic parameters of each component of the system. This involves capacities, CAPEX, OPEX, mean time to failure (MTTF), main time to repair (MTTR), among others. The MTTF and MTTR are used to calculate system availability, which is defined as the probability that the system will be operational at a given time "t" (Hou et al., 2015). The failure (λ_i) and repair rate (μ_i) are considered and calculated through Eq. (1.a) and (1.b). The availability of each equipment is then determined through Eq. (2) (Hastings, 2015). Finally, the system availability is calculated considering series (A_s) or parallel (A_p) configurations through Eq. (3.a) and (3.b), respectively (Hastings, 2015). Other input parameters considering the data that is independent of the equipment are also involved (i.e., indirect cost).

$$a) \quad \lambda_i = \frac{1}{MTTF_i} \quad ; \quad b) \quad \mu_i = \frac{1}{MTTR_i} \quad (1)$$

$$a_i(t) = \frac{\lambda_i}{\lambda_i + \mu_i} + \frac{\mu_i}{\lambda_i + \mu_i} \exp [(-\lambda_i + \mu_i)t] \quad (2)$$

$$a) \quad A_s(t) = \prod_{i=1}^n a_i(t) \quad ; \quad b) \quad A_p(t) = 1 - \prod_{i=1}^n (1 - a_i(t)) \quad (3)$$

2.3. Monte Carlo simulation

The Monte Carlo method (Johansen, 2010) can be summarized in the following steps. First, the cumulative distribution function (CDF, $F_X(x)$) (see Eq. (4)) of each input parameter is defined. It describes the probability that a random variable “X” with a given probability distribution will be found at a value less than or equal to x (Arora, 2017). Secondly, random values between 0 and 1 are chosen, allowing random values to be taken from the CDF. Then, the output variables of interest or KPIs are calculated. Finally, the process is repeated "n" times according to the precision required, to obtain the CDF for the outputs.

$$F_X(x) = P[X \leq x] = \int_{-\infty}^x f_X(u) du \tag{4}$$

2.4. Results analysis and post-processing

The CDF of the key performance indicator (KPI) of interest (i.e., internal rate of return, levelized cost of hydrogen, etc.) are analysed for each design, considering the “opportunity index”. This index indicates the probability of obtaining a value equal to or lower than the desired value for a given KPI.

3. Case studies

The methodology described above was applied to four case studies for green hydrogen production, which are shown in Fig. 1.b and described below:

- Case 1: Using geothermal energy (facility located in Iceland).
- Case 2: Using solar energy (facility located in southern Spain).
- Case 3: Using solar energy and storing the surplus electricity to be valorized in the electrolysis plant.
- Complementary studies from Case 2, keeping the same solar energy source but analyzing the business opportunities associated with equipment redundancy and availability of the system. Three configurations are studied, which are shown in Table 1.

The main assumptions of the case studies are:

- Energy facilities are entirely used to meet hydrogen demand.
- Triangular probability distributions using minimum, maximum, and average values are involved (See Fig. 2).
- Electrolysis plant capacity: 100 MW.
- Lifetime: 35 years.

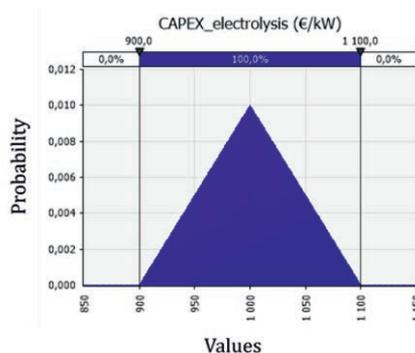


Fig. 2. Example of triangular probability distribution

Table 1. Green hydrogen facility design configurations for complementary studies of case 2

Design	Description
1 (Initial Case 2)	Project in the south of Spain
2 (Hypothetical Case 2.1)	Design 1 with 20% increase in electrolysis capacity.
3 (Hypothetical Case 2.2)	Design 2 with doubled compression capacity

4. Results and discussion

4.1. Case studies 1 to 3 - Evaluation of hydrogen production from geothermal and solar sources

The main results obtained in this study are discussed in this section. Fig. 3 shows the business opportunities associated with Case Studies 1, 2, and 3. The geothermal energy source presents reduced investment risk compared to the use of solar energy, since its IRR value remains positive even for a conservative scenario (opportunity index close to one). This is mainly due to the load factor of geothermal electricity production (about 92%) compared to the solar load factor (about 31%), which allows the IRR to rise by up to 48% in an optimistic scenario. The scenario with the highest investment risks would be case study 3, in which the probability of obtaining a positive IRR is 0.48. This implies that adding batteries to valorize the total available energy to produce hydrogen cannot be profitable according to our results. However, other case studies can be economically viable (i.e., isolated regions) (Marocco et al., 2021).

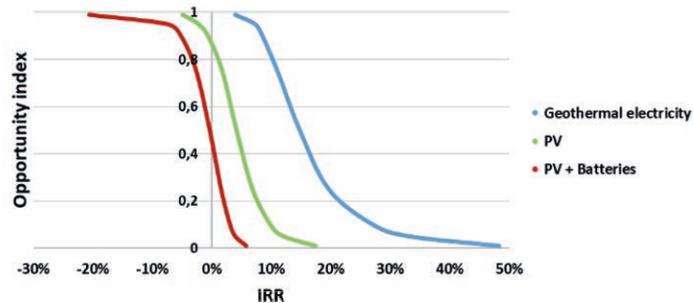


Fig. 3. Opportunity index vs IRR for case studies 1, 2 and 3

In addition, a sensitivity analysis on LCOH was performed for each major parameter, which is shown in Fig. 4. The LCOH varies between 2.11 and 3.29 €/kg H₂. The results indicates that electricity cost, electrolyzer capital cost, and system power composition are the major inputs influencing the LCOH. These should be prioritized from an optimal design and operational point of view. The investment cost of batteries has an important impact in Case 3, being the parameter that increases the LCOH compared to Case 2.

4.2. Complementary studies to case 2- Finding the best trade-off between CAPEX and availability

Several business opportunities associated with three production system designs and system availability (SA) are explored, which are indicated in Table 1. Fig. 5 shows the IRR for each design at the P50 point (opportunity index of 0.5) as a function of their CAPEX. “Case 2.1” offers an increase of a 0.26% in IRR despite having a higher CAPEX. This is achieved through increased system availability. In this regard, an improvement in the availability of the electrolyzers is sufficient to increase the probability of better economic performance. “Case 2.2” implies a decrease of 0.43% in IRR, degrading economic performance compared to Designs of Case 2 and 2.1.

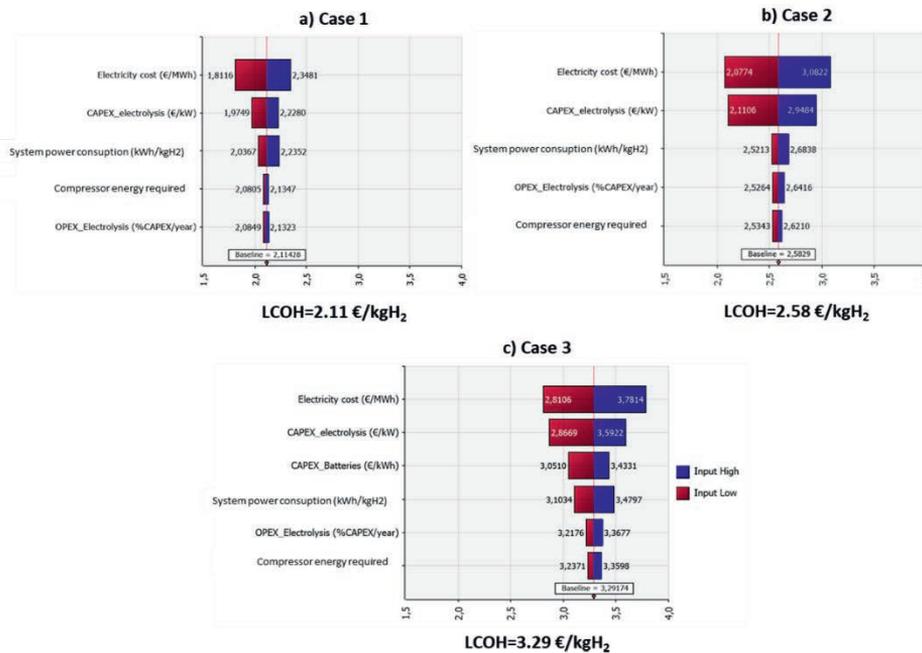


Fig. 4. Inputs ranked by effect on LCOH for a) Case 1, b) Case 2, and c) Case 3

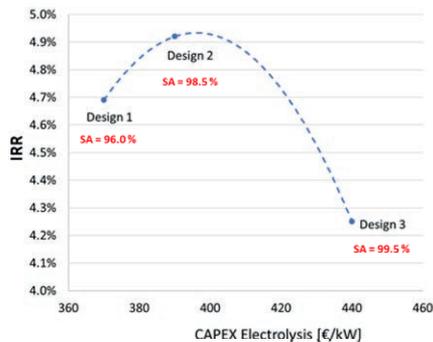


Fig. 5. IRR as a function of CAPEX at P50

5. Conclusion and perspectives

In this paper, a methodological framework for the economic risk analysis and design of green hydrogen facilities was proposed. It was demonstrated that an approach driven by Monte Carlo simulations generates substantial value-added insights to decision-makers. Firstly, the methodology addresses the uncertainty inherent in the parameters and its propagation to the KPIs. Secondly, the investment risk related to different facility designs is considered by an opportunity index and providing confidence intervals. Thirdly, sensitivity analyses can be carried out to identify the parameters with the highest impact on KPIs. Finally, several facility designs can be proposed and evaluated, identifying the best trade-off between the capital invested and the system availability.

The added value of the methodology was shown in several case studies. It was found that the geothermal energy source offers business opportunities even in conservative scenarios, compared to solar energy source. The valorization of the overproduction of electricity from batteries to produce hydrogen implies high investment risks. In addition, it was possible to choose between three design configurations including equipment redundancy to increase system availability. For this case study, we conclude that a redundancy in the electrolyzers is sufficient to increase the economic performance of the project, and then represents the best trade-off between CAPEX and system availability. Future research will require studying how to improve the calculation of system availability and O&M cost to precisely capture the stochastic behavior of the facility.

References

- Arora, J.S., 2017. Introduction to Optimum Design, Fourth ed. ed. Elsevier, Iowa. <https://doi.org/10.1016/C2013-0-15344-5>
- Carrera, E., Azzaro-Pantel, C., 2021. A methodological design framework for hydrogen and methane supply chain with special focus on Power-to-Gas systems: Application to Occitanie region, France. *Comput. Chem. Eng.* 153, 107386. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2021.107386>
- Grieves, M., 2014. Digital twin: manufacturing excellence through virtual factory replication. White Pap. 1, 1–7.
- Hastings, N.A.J., 2015. Physical Asset Management. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-14777-2>
- Hou, Y., Sallak, M., Schön, W., 2015. Availability analysis of systems using random set theory. *IFAC-PapersOnLine* 28, 1315–1320. <https://doi.org/10.1016/j.ifacol.2015.09.707>
- Jaribion, A., Khajavi, S.H., Öhman, M., Knapen, A., Holmström, J., 2020. A Digital Twin for Safety and Risk Management: A Prototype for a Hydrogen High-Pressure Vessel. pp. 369–375. https://doi.org/10.1007/978-3-030-64823-7_34
- Johansen, A.M., 2010. Monte Carlo Methods, in: International Encyclopedia of Education. Elsevier, pp. 296–303. <https://doi.org/10.1016/B978-0-08-044894-7.01543-8>
- Marocco, P., Ferrero, D., Martelli, E., Santarelli, M., Lanzini, A., 2021. An MILP approach for the optimal design of renewable battery-hydrogen energy systems for off-grid insular communities. *Energy Convers. Manag.* 245, 114564. <https://doi.org/10.1016/j.enconman.2021.114564>
- Meraghni, S., Terrissa, L.S., Yue, M., Ma, J., Jemei, S., Zerhouni, N., 2021. A data-driven digital-twin prognostics method for proton exchange membrane fuel cell remaining useful life prediction. *Int. J. Hydrogen Energy* 46, 2555–2564. <https://doi.org/10.1016/j.ijhydene.2020.10.108>
- Wang, Y., Wang, X., Liu, A., 2020. Digital Twin-driven Supply Chain Planning. *Procedia CIRP* 93, 198–203. <https://doi.org/10.1016/j.procir.2020.04.154>

Machine Learning-Based Surrogate Models and Transfer Learning for Derivative Free Optimization of HTPEM Fuel Cells

Luis A. Briceno-Mena^a, Christopher G. Arges^b, Jose A. Romagnoli^a

^a*Cain Department of Chemical Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States*

^b*Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, United States*
jose@lsu.edu

Abstract

Widespread adoption of high-temperature electrochemical systems such as polymer electrolyte membrane fuel cells (HT-PEMFCs) requires models and computational tools for accurate optimization and guiding new materials for enhancing fuel cell performance and durability. In this contribution, knowledge-based modelling and data-driven modelling are combined using Few-Shot Learning and implementing an Automated Machine Learning framework for the generation of Machine Learning-based surrogate models.

Keywords: Surrogate Models, Derivative Free Optimization, Transfer Learning, AutoML, data-driven modeling, high-temperature polymer electrolyte membrane fuel cells.

1. Introduction

The discovery of new materials like catalysts, polymeric membranes, and biomolecules, is driven by industrial needs such as improving reaction or separation selectivity, enhancing therapeutic effects on medical treatments, or reducing manufacturing costs. However, deployment of these advances in industrial applications is often hindered by the lack of models needed for design and optimization. Due to the novelty of the materials and devices, experimental data and first principles knowledge is scarce, making it hard to build models either via data-driven or knowledge based approaches. In this context, a way to efficiently combine domain knowledge with data could provide a pathway to streamline new materials discovery for industrial applications. Transfer Learning (TL) is an extension of Machine Learning (ML) in which knowledge learned for a particular task can be leveraged to ease the training for a new task. In terms of modeling for electrochemical systems, models developed for a given device or material can be leveraged to reduce the amount of data needed to accurately predict how new materials, operating parameters, and device configurations affect system performance. When the data for the initial training stage comes from a simulation, domain knowledge can be easily incorporated into the data-driven model, while at the same time reducing the number of experiments to be conducted – which can be timely and costly. This approach generates surrogate models that approximate the real behavior of the systems with adequate accuracy at a reasonable cost.

In this work, a TL framework for the development of Machine Learning-based surrogate models (MLBSM) to be used with a derivative free optimization algorithm is presented. Simulations from a compositional knowledge-based model validated using experimental data was used to generate training examples for a MLBSM. Then, TL learning was used to improve the performance of the MLBSM by using experimental data as training examples. Finally, an implementation of Particle Swarm Optimization (PSO) was used as demonstration of the applicability of this approach for derivative free optimization.

2. Transfer Learning

Transfer Learning (TL) (Pan & Yang, 2010) is a ML technique in which a data-driven model previously trained (general training) for a given task (source domain) is used as the base to build a model for a new task (target domain), with less data being required for the new training stage (task-specific training). Typically, the source and target domains are similar or closely related. Formally, given a source domain \mathcal{D}_s , and learning task \mathcal{T}_s , a target domain \mathcal{D}_T and a learning task \mathcal{T}_T , transfer learning aims to help the learning of the target predictive function $f_T(\cdot)$ for the target domain using the knowledge in \mathcal{D}_s and \mathcal{T}_s , where $\mathcal{D}_s \neq \mathcal{D}_T$ and $\mathcal{T}_s \neq \mathcal{T}_T$ (Yang, Zhang, Dai, & Pan, 2020). In this work, the source domain \mathcal{D}_s and learning task \mathcal{T}_s come from simulations generated using a low fidelity knowledge-based model. The target domain \mathcal{D}_T and learning task \mathcal{T}_T come from experimental data. A useful extension of TL is the so-called few-shot learning (FSL) in which the task-specific training stage uses a very small amount of data (i.e., on the order of 1×10^1) (Yang et al., 2020).

The overarching strategy of this implementation is described in **Figure 1**. A knowledge-based explicit equation model (EEM) for a HT-PEMFC was used as the source domain. An experimental dataset was used as target domain. A fully connected neural network optimized using a genetic algorithm-based Automated Machine Learning framework (AutoML) was implemented.

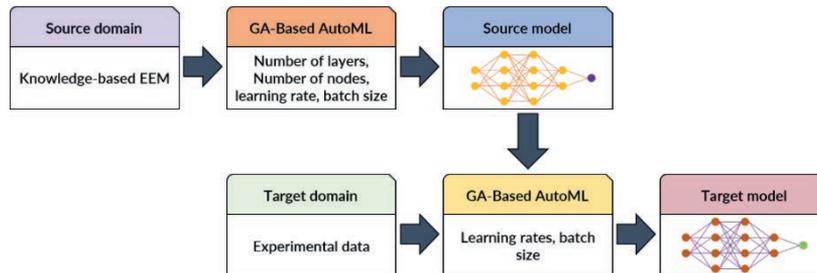


Figure 1. Transfer learning-based modeling strategy.

3. Data preparation

For the general training stage, simulated data was generated using the knowledge-based model reported elsewhere (Briceno-Mena, Venugopalan, Romagnoli, & Arges, 2021). To obtain a balanced dataset, the values for the input variables were structured following a full 3-level factorial experimental design using 11 variables (3^{11}). It is important to note that all HT-PEMFC models and data only consider pure oxygen as the oxidant. For each set of inputs, 5 points in the polarization curves were generated. The resulting datasets were arranged to generate a total of 887,735 input vectors of size 12 (the 11 input variables plus the current density) and the corresponding labels of size 1. For the target

domain, experimental data reported by Venugopalan et al. (Venugopalan et al., 2020) was used. Data were scaled using over min-max scaling in all cases.

4. Few Shot Learning and Automated Machine Learning Framework

Broadly, FSL strategies are based upon the notion of preserving information obtained through a first stage of training and reusing it in some fashion to develop a target model. FSL was implemented through regularization where the target model has the same architecture as the source model, and the learned parameters (weights) are adjusted during a new training stage, but some restriction is applied to prevent the model from overfitting the new data (Yang et al., 2020). In artificial neural networks (ANN), regularization takes the form of restricting how much the weights of the model can be updated. By applying setting different restrictions for each layer in the ANN, influence of the new data over the model can be controlled. The problem of obtaining the MLBSM can then be divided in two stages: (i) find a good source model and (ii) find the best way to perform the task-specific training. The first stage involves finding the proper architecture and training conditions (number of hidden layers (L), number of nodes (N), the learning rate (lr), and the batch size (n)) for the given source domain. The second stage requires finding a proper set of learning rates and batch size for a given target domain. Here, the selection of L , N , lr and n is defined as a Mixed Integer Nonlinear Programming problem and solved using a multi-objective evolutionary algorithm (Vishwakarma, Haghghatlari, & Hachmann, 2019), namely the Non-dominated Sorting Genetic Algorithm (NSGA-II) (Blank & Deb, 2020; Deb, Pratap, Agarwal, & Meyarivan, 2002). For tuning the hyperparameters of the neural network at source training, the multi-objective optimization problem (**Optimization problem 1**) is formulated as the minimization of $f_m(x)$ subject to $x \in R^4$ and $x_i \in \mathbb{Z}, i = 1, 2, 3$ where, x_1, x_2, x_3 are L, N and n respectively and x_4 is the learning rate. $f_1(x) = E_{training}^S$, $f_2(x) = E_{testing}^S$, and $f_3(x) = E_{validation}^S$ are the mean values for training, testing and validation in the 5-fold cross validation. The testing and validation errors are included in the multi-objective optimization problem to prevent overfitting. To obtain a model for a given source domain, a new neural network is built for each individual using the set of hyperparameters x (in addition to a fixed input layer $12:N$ and 2 fixed output layers $N: \lfloor \frac{N}{2} \rfloor : 1$) and then its performance is cross validated. For the FSL stage, a similar optimization problem (**Optimization problem 2**) can be formulated to obtain the appropriate learning rates ($lr_{input}, lr_{general}, lr_{task}$) and batch size (n) for a given target domain. In this case, x_1, x_2, x_3 are $lr_{input}, lr_{general}$ and lr_{task} respectively and x_4 is the batch size.

5. Derivative Free Optimization using MLBSMs

Once the surrogate model is obtained, the next step is use it to inform both the design of new materials for the fuel cell and its operating conditions. This task can be defined as an optimization problem (**Optimization problem 3**) in which the power density of the fuel cell estimated using surrogate model, $G(x)$, is the objective function and the physical realizability of the input variables ($x \in \mathbb{R}^{10}$) sets the upper and lower bounds for the problem variables. Since knowledge about the derivatives of the objective function is often unavailable, the problem is well suited for derivative free optimization (DFO) (Rios & Sahinidis, 2013). Furthermore, DFO benefits from the ability of ANN to perform well in high dimensional spaces (Bhosekar & Ierapetritou, 2018), opening the opportunity to exploit the flexibility of both MLBSMs to represent complex systems and DFO to

perform optimization using black-box models. In this work, Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995) is implemented as a demonstration of the potential of coupling MLBSMs obtained via transfer learning with DFO.

6. Results and Discussion

6.1. Source model training

From **Optimization Problem 1**, it was found that a neural network with 3 hidden layers, each with 36 nodes (12:36:36:36:18:1), a batch size $n = 331$, and a learning rate $lr = 9.94 \times 10^{-5}$ was the best architecture. For the 5-fold cross validation source model training using the optimized neural network the mean relative root mean squared error (rRMSE) values for training, testing and validation were 2.80%, 2.87% and 29.16% respectively. The simulated data was obtained from an explicit equations model first developed for MEA0 and reported elsewhere (Briceno-Mena et al., 2021). **Figure 2** shows the accuracies and their variability for the source model (**Figure 2A**) and the target model (**Figure 2B**). The high validation error for the source model stems from the inaccuracy in the original explicit equations model used to generate the simulations for training, namely the estimation of the activation polarization.

6.2. Target model training

When the source domain is a simulation and the target domain is real experimental data for the same system, TL can be used to obtain a better model. As shown in **Figure 2A**, although the optimized neural network reached very low errors for the training and testing data, the error for the validation data is still high. This is consistent with the fact that at pretraining the ANN have not been exposed to real experimental data and the source domain is known to have inferior performance at very low current densities. For the task-specific training, the optimized learning rates and batch size from **Optimization Problem 2** ($lr_{input} = 1.99 \times 10^{-8}$, $lr_{general} = 8.80 \times 10^{-6}$, $lr_{task} = 0.00097$, $n = 7$) (the number of nodes and layers remained unchanged) were used. Validation data corresponds to the 200 °C polarization curve. As it can be observed, the validation error decreases by a factor of 3 for the target model. These results demonstrate the applicability of TL to improve existing models using small datasets, thus reducing the burden of generating experimental data.

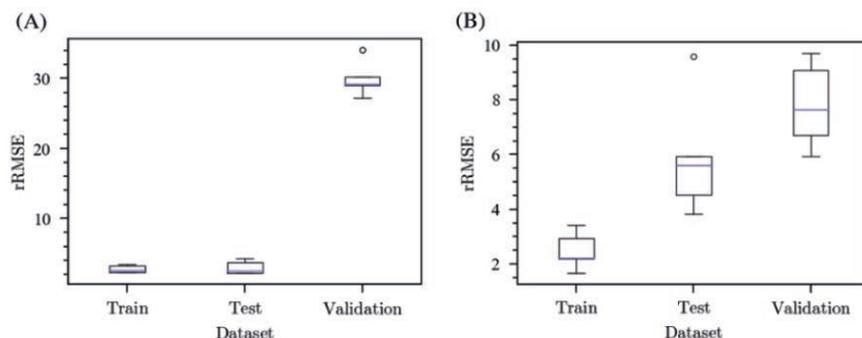


Figure 2. 5-fold cross validation results for model improvement. Boxes represent the standard deviation of the errors. (A) source model. (B) target model.

As the EEM models contains physical knowledge (e.g., Butler-Volmer kinetics for informing the activation overpotential and Ohm's Law for determining the ohmic overpotential from electrolyte conductivity) (Briceno-Mena et al., 2021), TL enables the introduction of this information into the MLBSM. An illustrative example of this is shown

in **Figure 3**. Here, the change in the polarization curve with changes in pressure is represented for both the source model (**Figure 3A**) and the target model (**Figure 3B**). Although the target domain does not contain information about the effect of pressure over the system (all examples correspond to the same pressure), the target model preserves some of this knowledge from the pretraining stage, which was introduced via the Nernst equation and physics-informed expressions for the exchange and limiting current density values. Therefore, new information on all the predictors in the source model is not required. This opens the opportunity for a faster development of models for new materials and device designs and optimal operating parameters, helping guide new materials by identifying the necessary properties for enhancing fuel cell performance (e.g., peak power density or power density at 0.7 V). Furthermore, the MLBSM can assist with future optimization activities such as pathways to reduce PGM loadings in the MEA while minimizing device performance losses such as power density.

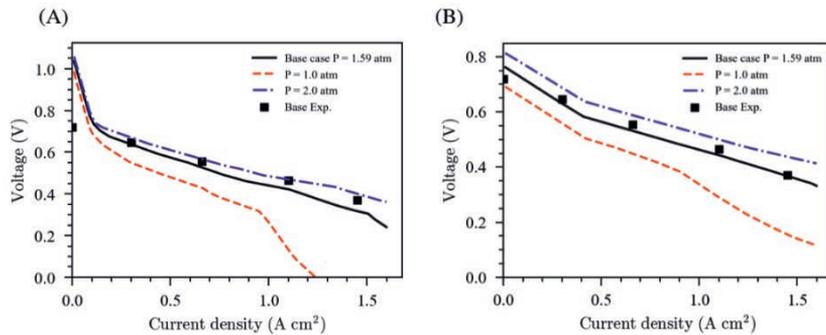


Figure 3. Effect of pressure as predicted by the source model (A) and the target model (B). The base case corresponds to 200 °C.

6.3. Derivative Free Optimization

Figure 4A shows the convergence of the optimization algorithm and **Figure 4B** shows the values for the problem variables corresponding to the optimal solution.

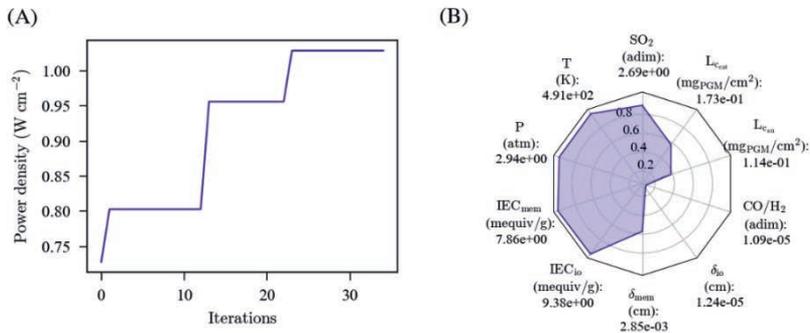


Figure 4. (A) Convergence of the optimization algorithm. (B) values of the problem variables for the optimal solution. The values are normalized by the upper bound in the optimization problem and grow in the direction of cost/difficulty of realization.

The optimization results suggest that the conductivity properties of both the membrane and ionomer binder (IEC_{mem} and IEC_{io}) are crucial for enhancing the performance of the fuel thus warranting further research on new materials to achieve higher conductivities. Furthermore, the ability to optimize operating conditions for a given set of materials could

allow researchers to test the potential of these materials *in silico* which can in turn reduce the cost of experiments and accelerate the timeline for bringing new materials to market.

7. Conclusions

A ML framework that exploits FSL in an AutoML framework for the generation for MLBSMs was demonstrated. Using FSL, the validation error was effectively reduced by a factor of 3, showcasing the capabilities of this approach to improve existing models using small datasets. The final surrogate model preserved physical knowledge included at pretraining, which is key advantage for the development of physics informed ML applications. Finally, the usefulness of the MLBSM was demonstrated by using it as the objective function in an optimization problem solved via a derivative free optimization algorithm. PSO converged after 25 iterations and rendered physically meaningful results that can inform future developments in materials for the components and operation conditions of the HT-PEM fuel cell.

8. Acknowledgements

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Advanced Manufacturing Office (AMO) Award Number DE-EE0009101. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Luis A. Briceno-Mena thanks the support received from Universidad de Costa Rica.

References

- Bhosekar, A., & Ierapetritou, M. (2018). Space mapping based derivative-free optimization framework for supply chain optimization. In M. R. Eden, M. G. Ierapetritou, & G. P. Towler (Eds.), *Computer Aided Chemical Engineering* (Vol. 44, pp. 985-990): Elsevier.
- Blank, J., & Deb, K. (2020). Pymoo: Multi-Objective Optimization in Python. *IEEE Access*, 8, 89497-89509. doi: 10.1109/ACCESS.2020.2990567
- Briceno-Mena, L. A., Venugopalan, G., Romagnoli, J. A., & Arges, C. G. (2021). Machine learning for guiding high-temperature PEM fuel cells with greater power density. *Patterns*, 100187. doi: 10.1016/j.patter.2020.100187
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197. doi:10.1109/4235.996017
- Kennedy, J., & Eberhart, R. (1995, 27 Nov.-1 Dec. 1995). *Particle swarm optimization*. Paper presented at the Proceedings of ICNN'95 - International Conference on Neural Networks.
- Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359. doi: 10.1109/TKDE.2009.191
- Rios, L. M., & Sahinidis, N. V. (2013). Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3), 1247-1293. doi: 10.1007/s10898-012-9951-y
- Venugopalan, G., Chang, K., Nijoka, J., Livingston, S., Geise, G. M., & Arges, C. G. (2020). Stable and Highly Conductive Polycation-Polybenzimidazole Membrane Blends for Intermediate Temperature Polymer Electrolyte Membrane Fuel Cells. *ACS Applied Energy Materials*, 3(1), 573-585. doi: 10.1021/acs.aem.9b01802
- Vishwakarma, G., Haghighatlari, M., & Hachmann, J. (2019). Towards autonomous machine learning in chemistry via evolutionary algorithms. *ChemRxiv*. doi:10.26434/chemrxiv.9782387.v1
- Yang, Q., Zhang, Y., Dai, W., & Pan, S. J. (2020). *Transfer Learning*. Cambridge: Cambridge University Press.

A Methodology for The Optimal Surrogate Modelling of Digital Twins Using Machine Learning

Andrea Galeazzi^a, Kristiano Prifti^a, Francesco Gallo^b and Flavio Manenti^{a,*}

^a*Politecnico di Milano, CMIC Department "Giulio Natta", Piazza Leonardo da Vinci 32, Milan, 20133, Italy*

^b*Itelyum Regeneration s.r.l., Pieve Fissiraga, Lodi, 26854, Italy*
flavio.manenti@polimi.it

Abstract

Process simulation and digital twins are of paramount importance to design new processes or optimize already existing plants and equipment. The main drawback with current process simulation software is the computational time required to obtain a solution convergence towards a new steady-state. Especially when the input or the output to a system are perturbed. This procedure may take up to minutes in large systems or with strong non-linear recycles. Surrogate modelling of digital twins offers the possibility to speed up the time to convergence required by process simulators by substituting the fundamental or rigorous models with machine learning methods and models. In this work, a surrogate modelling methodology is described for extracting a useful amount of data in a domain near the nominal steady-state of the plant for which a digital twin has been created. For each process variable, a plethora of machine learning models are trained and compared. The best-performing models are chosen to predict the behaviour of such process variables. The application of the surrogate modelling framework thus created has been successfully applied to a steady-state simulation, i.e. digital twin, of an acid gas dimethylamine washing process at the Itelyum exhausted oil refinery in Pieve Fissiraga (LO), Italy.

Keywords: surrogate model, machine learning, digital twin, amine washing

1. Introduction

Surrogate models are gaining more interest in the field of chemical process simulation (McBride and Sundmacher, 2019). Their strength lies in the ability to substitute rigorous fundamental models, thus accelerating computationally intensive tasks like superstructure optimization (Granacher et al., 2021) or dynamic non-linear process prediction (Shokry et al., 2020). One key step for accurate modelling using surrogates is the data sampling technique. Surrogate models have found applications in Bayesian optimization where data samples are generated mainly near the optimal conditions, thus making the surrogate able to predict optimal working conditions (Keßler et al., 2019). On the other hand, surrogate models have also been applied with wider data samples in order to explore working conditions outside of the canonical domain for a given process (Ganti et al., 2020). It is needless to say that each application requires a tailored solution and so does the data sampling procedure. Different data sampling techniques are described in Liu et al. (2017). This work focuses on the usage of surrogate models for the substitution of digital twins of real processes around nominal steady-state conditions. For this reason, the data sampling procedure adopted is both wide and narrow. Narrow, like in the case of the Bayesian optimization, for samples close to the nominal steady-state, and wide for the feasible domain around the steady-state. This is done, firstly, to enable the model to consistently predict the production near the steady-state, where

the plant is going to operate the majority of times. Secondly, data samples far from the nominal conditions should enable the model to be used for exploration purposes. An example is real-time steady-state optimization. In this case, it is advantageous to extend the computation to the cloud for faster results (Galeazzi et al., 2021). The methodology just described has been applied to a process of acid gas removal with amine washing using diethanolamine in the bio-oil regeneration refinery of Itelyum in Pieve Fissiraga (LO), Italy.

2. Data Generation

The data generation procedure is of key importance since a correct mapping of the operative domain of the process enables the surrogate model to make good predictions everywhere inside this domain. Especially when it is used for testing purposes rather than training. Several advanced methods exist for adaptively sampling the design space, an example can be found in Eason and Cremaschi (2014). For this work, the preferred method for data sampling is the Latin-Hypercube (McKay et al., 1979) design of experiment (DoE) since it spreads the samples uniformly within the constrained domain. The data thus generated is preprocessed to remove eventual outliers and, most importantly, non-converged simulations. These are all Aspen HYSYS simulations that did not arrive at a solution, with the desired operative conditions, in a reasonable amount of time at the imposed numerical tolerance. Generally, the data generation is slow due to the presence of non-linearities in the simulation and recycle loops. An optimization that has been implemented is to sort the samples, in ascending (or descending) order, generated through the DoE for the principal variables, e.g. the temperature of a stream entering a column. Thus, the starting point for the next sample of the DoE is much closer and the simulator should go to the solution more quickly.

3. Modelling framework

After the data is gathered and preprocessed the surrogate modelling framework starts the training of several machine learning (ML) algorithms. The type and amount of algorithms chosen are arbitrary and a specific case study may require a tailored solution. In this case, the proposed algorithms are, in inverse order of complexity, linear regression, polynomial regression, support vector regression (SVR), decision tree regression, random forest, AdaBoost, gradient boosting, and artificial neural networks. For this particular case study, every feature variable has been used to make a prediction about the target variables. Eventually, the framework should benefit from a previous feature selection to reduce the total number of variables given to each algorithm, thus speeding up the training time. Every ML algorithm is trained and compared against each other with a k -folds cross-validation method (with $k = 5$, in this case). The comparison of each model has been performed through the analysis of the mean absolute error (MAE), in Eq. 1, and the root mean squared error (RMSE), in Eq. 2.

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \quad (1)$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2} \quad (2)$$

where y is the target value, \hat{y} is the predicted value, n is the number of samples, and i is i -th data sample.

4. Case study

The surrogate modelling framework has been applied for the process of acid gas removal with amine washing in the exhausted oil refinery of Itelyum in Pieve Fissiraga (LO), Italy. The process flow

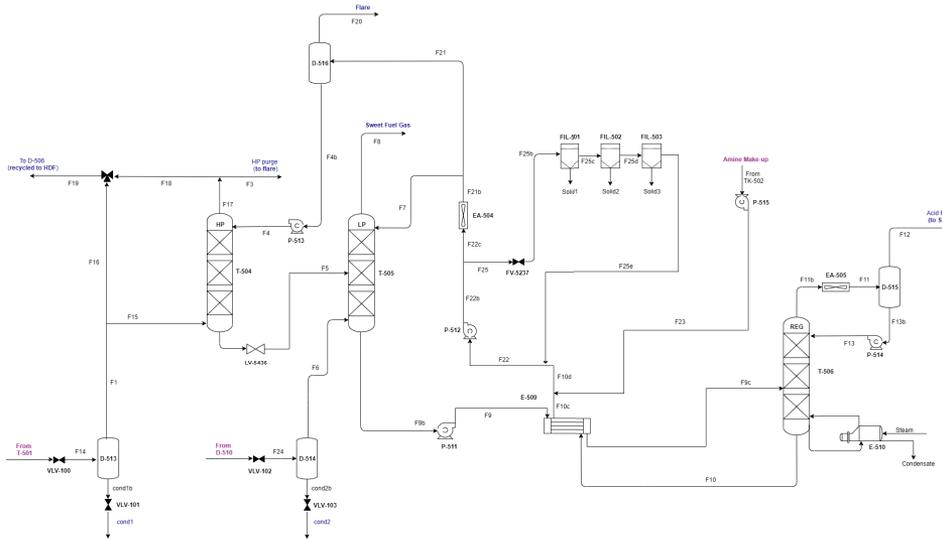


Figure 1: PFD of the amine washing section of Itelyum exhausted oil refinery in Pieve Fissiraga (LO), Italy

diagram (PFD) is shown in Figure 1. It is composed of two main sections, namely the absorbing section and the regeneration section. The absorbing section is composed of two absorbers, one being at high pressure (T-504, in Figure 1) and the other at low pressure (T-505, in Figure 1). The regeneration section is composed of a single column (T-506, in Figure 1) used to regenerate the spent diethanolamine (DEA) coming from the previous HP and LP absorbers. These unit operations are part of the Revivoil process, created by Itelyum and Axens, and the streams entering the amine washing subsection are coming from the hydrofinishing units. They are rich in light hydrocarbons, hydrogen, and hydrogen sulfide. Hydrogen is separated in the HP absorber and recycled back to the hydrofinishing while H_2S is sent to a sulfur recovery unit (SRU). The small amount of light hydrocarbons remaining after the separation are finally purged in the flare.

First, a steady-state simulation was created using Aspen HYSYS. Several simplifications have been implemented in order to reduce the time to convergence of the HYSYS simulator. Moreover, for computational reasons, the two main sections described above have been split into two different flowsheets as shown in Figures 2 and 3. The steady-state digital twin solution has been validated against plant data taken from the distributed control system (DCS) through Yokogawa’s Exaquantum platform. The simulation has shown a small deviation of 4.4 % on average from the process data. An example is shown in Table 1.

Table 1: Methane mass flow [kg/h] in amine washing process.

	F18	F9	F8	F5	F4	F7
Digital twin	297.47	0.00061	6.17	0.00029	0	0
Process data	297.44	0	7	0	0	0

The data are generated for each section independently. The data extracted from the simulation are the temperature, pressure, flow rate, and composition of each stream shown in Figures 2 and 3. A total of 4000 data points for each section have been generated with box constraints around the

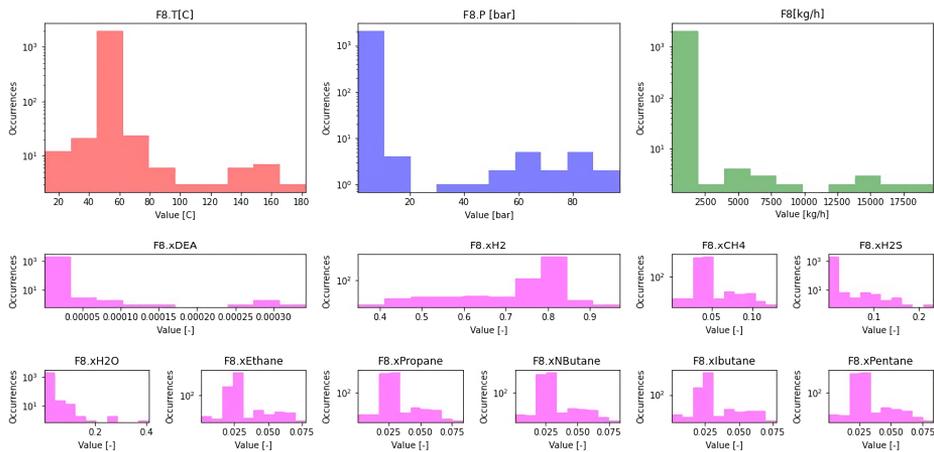


Figure 4: Histograms of the occurrences of data points in the range of process variables for stream F8.

nominal conditions and with a Latin-Hypercube DoE. Two different domains, i.e. box constraints, have been defined, the former being closer to the operational range with an arbitrary $\pm 10\%$ from the nominal conditions while the latter being wider with a range of $\pm 40\%$. Figure 4 shows an example of data extraction for stream F8. The split ratio for the data created in both domains is 50%. However, due to the relative width of the second domain ($\pm 40\%$), only a few data samples have found a convergence with the desired tolerance (64 out of 2000 for the case of the HP and LP absorbers). On the other hand, for the first box constraint of $\pm 10\%$ the case is inverted, and only a few operative conditions did not result in convergence (1947 converged out of 2000 for the case of the HP and LP absorbers). The convergence/non-convergence information is still insightful. In fact, by applying a binary classification algorithm, like logistic regression (for example), it is possible to understand the feasible domain for process convergence.

After this initial preprocessing, features and targets were defined. The features chosen for the modelling framework are all those inputs needed for fixing all the degrees of freedom of the simulation. Viceversa, the targets are all those process variables (temperature, pressure, composition, and flow rate) calculated by the same simulation. Once the variables are set, the training of each algorithm described in Section 3 starts, and the residual errors (MAE and RMSE) are taken for establishing the best-performing algorithm. The results of the best-performing algorithms chosen are shown in Table 2. It is worth noticing that not always the most complex model is the best performing one. For example, the linear regressor has always been chosen for predicting the pressure variables. This is explained by the fact that pressure drops are set as constants in the HYSYS simulation. Moreover, other more advanced models like artificial neural networks were always underperforming their lower complexity competitors, even though they took more computational time for training. Finally, the overall residual error for the first domain of exploration ($\pm 10\%$) is approximately 2% for the absorbers section and 5% for the regenerator. If we calculate the residual error also for the case of the wider domain ($\pm 40\%$) the percentage is higher by orders of magnitude. This is explained by the fact that the actual converged solutions in this domain are very few, and the models may neglect these measurements given the higher amount of data of the first domain. Solutions to this problem may range from the addition of weights that try to rebalance the domains disparity during the regression procedure, to the generation of more data inside the actual feasible domain, possibly with a binary classifier guiding the DoE inside the actual boundary of convergence.

Table 2: Number of occurrences for each best-performing algorithm for the regenerator and HP/LP absorbers sections

	Regenerator section		Absorbers section	
	Occurrences (#)	Occurrences (%)	Occurrences (#)	Occurrences (%)
Linear regression	2	7.7%	3	7.7%
Polynomial order 2	1	3.8%	-	-
SVR	12	46.2%	-	-
Decision tree	2	7.7%	-	-
Random forest	6	23.1%	28	71.8%
AdaBoost	3	11.5%	8	20.5%

5. Conclusions

In conclusion, the surrogate modelling approach has been successfully applied for a case study of a real industrial process. The surrogate modelling framework has been written in Python and, for communicating with the process simulator (Aspen HYSYS), an interface was developed in Excel via visual basic (VB) using the native libraries of HYSYS for VB. This approach is optimal for creating surrogate models of steady-state digital twins of industrial plants, given the data generation procedure that creates many data points near the nominal operating conditions of the process. With these surrogate models, it is now possible to apply optimization strategies or applications for real-time usage since the amount of time required to obtain a result from the surrogate model itself is lesser. This methodology should be further enhanced to make it able to treat dynamic digital twins. Furthermore, the final goal of this method is to substitute the need for a process simulator that generates the data and take this data directly from the DCS of the industrial plant, automating the process of surrogate models creation.

References

- J. Eason, S. Cremaschi, sep 2014. Adaptive sequential sampling for surrogate model generation with artificial neural networks. *Computers & Chemical Engineering* 68, 220–232.
- A. Galeazzi, R. Nasti, G. L. Bozzano, L. Verotta, S. Marzorati, F. Manenti, 2021. A cloud computing application for the supercritical carbon dioxide extraction using coffee grounds silverskin. In: 31st European Symposium on Computer Aided Process Engineering. Elsevier, pp. 1035–1040.
- H. Ganti, M. Kamin, P. Khare, sep 2020. Design space exploration of turbulent multiphase flows using machine learning-based surrogate model. *Energies* 13 (17), 4565.
- J. Granacher, I. D. Kantor, M. Lopez, F. Maréchal, 2021. Self-learning surrogate models in superstructure optimization. In: 31st European Symposium on Computer Aided Process Engineering. Elsevier, pp. 439–444.
- T. Keßler, C. Kunde, K. McBride, N. Mertens, D. Michaels, K. Sundmacher, A. Kienle, apr 2019. Global optimization of distillation columns using explicit and implicit surrogate models. *Chemical Engineering Science* 197, 235–245.
- H. Liu, Y.-S. Ong, J. Cai, jun 2017. A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design. *Structural and Multidisciplinary Optimization* 57 (1), 393–416.
- K. McBride, K. Sundmacher, jan 2019. Overview of surrogate modeling in chemical process engineering. *Chemie Ingenieur Technik* 91 (3), 228–239.
- M. D. McKay, R. J. Beckman, W. J. Conover, may 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239.
- A. Shokry, P. Baraldi, E. Zio, A. Espuña, jul 2020. Dynamic surrogate modeling for multistep-ahead prediction of multivariate nonlinear chemical processes. *Industrial & Engineering Chemistry Research* 59 (35), 15634–15655.

Bayesian Neural Network-Based Calibration for Urban Air Quality Sensors

Gustavo R. Taira,^a Adriano G. Leal,^b Alessandro S. Santos,^b Song W. Park^a

^a*Department of Chemical Engineering, Polytechnic School, University of Sao Paulo, Av. Luciano Gualberto 380, Sao Paulo 05508-010, Brazil*

^b*Institute for Technological Research of the Sao Paulo State, Av. Almeida Prado 532, Sao Paulo 05508-901, Brazil*
gustavo.taira@usp.br

Abstract

With the growing urbanization process in several cities around the world, air pollution mitigation has become one of the main environmental challenges of the present time. Recently, low-cost air pollution sensors have become a current trend in the air quality control area since they are an affordable alternative for deploying air quality monitoring systems with high spatial resolution. However, the main drawback of these sensors is that they tend to provide measurements with lower accuracy and reliability compared to traditional air quality monitoring stations. Therefore, periodical calibration of these sensors is essential to maintain the quality of their measurements. This work presents a novel air quality sensor calibration method based on a Bayesian neural network model. The proposed method is assessed using a real public available dataset. The test experiment results show that the method has a good accuracy performance, with a lower mean absolute error compared to other machine learning-based calibration methods applied to the same dataset. In addition, the method presents the advantage of directly providing estimations of the uncertainty of the calibrated measurements, which is an important metric used to assess the quality and reliability of data provided by air pollution sensors and that most other calibration methods usually cannot provide.

Keywords: air quality evaluation, sensor calibration, machine learning, Bayesian neural network.

1. Introduction

Air pollution mitigation has become one of the main environmental challenges of the present time. It is estimated that 96% of the urban population is exposed to air pollution levels that exceed the recommended limits (Lewis and Edwards, 2016) and, just in 2016, more than 4.2 million deaths in the world were caused by factors related to air pollution exposition (WHO, 2016). Therefore, the implementation of air quality monitoring systems that collect detailed data about air pollution concentration in urban areas and assist city managers to remediate problems related to air pollution has become crucial for large cities. Most cities, however, use air quality monitoring systems based on expensive monitoring stations that are sparsely distributed in their urban areas. As such, they can only monitor air pollution with low spatial resolution and they cannot collect detailed information about air pollutants in their whole urban area. In this context, low-cost air pollution sensors have recently become a trend in the air quality monitoring area, as they are a potential alternative for increasing the spatial resolution of the air quality monitoring systems. Since these devices can be acquired at a much lower cost compared to the

traditional air quality monitoring stations, they can be widely deployed in urban areas and support the traditional monitoring stations carrying out air pollution concentration measurements in regions that the stations cannot cover. However, the main drawback of these sensors is that they tend to provide lower accuracy and reliability measurements compared to traditional air quality monitoring stations. They suffer from cross-sensitivities between different air pollutants, they are affected by external factors (such as temperature and humidity), and their measurement performance degrades over time due to aging and poisoning effects on their components (Maag et al., 2018) (Concas et al., 2021).

To improve the accuracy of these sensors, a widely used strategy is to periodically calibrate them. The calibration of these sensors is traditionally performed in laboratories, by exposing them to controlled atmospheres with known concentrations of pollutants. However, as these devices are widely deployed in urban areas, this calibration approach tends to become unfeasible, since a massive number of sensors would have to be removed from their deployment site and then transported to laboratories. In this context, calibration methods based on machine learning models have recently emerged as excellent alternatives for air pollution sensor calibration as they allow in-field calibration. The general idea of these methods is to use measurements of a monitoring station (or an already calibrated sensor) located in the proximity of the low-cost sensor as references to train a machine learning model that can correct the errors of its measurements and output more accurate measurements (Concas et al., 2021).

Artificial neural networks have become commonly used machine learning models for air quality sensor calibration. They showed to be flexible multivariate calibration models, as they are capable of handling the cross-sensitivity between air pollutants, the influence of external conditions, and the nonlinear relationships between a sensor measurement and a reference measurement, which are challenging problems in the sensor calibration task (De Vito et al., 2018). These models, however, are known to suffer from overfitting problems, which can affect their generalization capability (Concas et al., 2021). Thus, this study proposes a novel low-cost air pollution sensor calibration method based on a Bayesian neural network model, which is a probabilistic neural network model that has the advantages of presenting natural mechanisms of regularization against overfitting and directly providing estimates of the uncertainty of its predictions (Jospin et al., 2020).

2. Methodology

2.1. Bayesian Neural Network

A Bayesian neural network (BNN) is a stochastic artificial neural network trained using Bayesian inference (Jospin et al., 2020). Stochastic neural networks are a type of artificial neural networks (ANNs) built by introducing stochastic components into their architectures that allow the estimation of uncertainty of their predictions. In the case of BNNs specifically, the network parameters θ are usually considered stochastic and this allows the networks to simulate multiple model parameter configurations according to their associated probability distribution $p(\theta)$. In this way, while training a BNN, instead of training one single model, a set of models is trained and their predictions are aggregated into probabilistic distributions. This characteristic avoids problems of overfitting for the neural network and gives access to uncertainty estimates of its predictions.

To design and train a BNN denoted by $y = F_{\theta}(x)$, the following Bayesian inference process is carried out. Given a training dataset with $X = [x_1, \dots, x_N]$ as its input variables and $Y = [y_1, \dots, y_N]$ as their respective target outputs, the first step is to choose a proper neural network architecture (i.e. feed-forward, recurrent, convolutional) for the problem

in hand. Then, the next step is to choose a prior distribution $p(\theta)$ over the model parameters, that describe the prior beliefs about the parameters before analyzing the dataset, and a likelihood distribution $p(Y|X, \theta)$, that describes a prior confidence in the predictive power of the model F_θ . In the case of a regression problem, such as the calibration sensor problem, a Gaussian distribution is normally used for the likelihood distribution (Gal, 2016):

$$p(y|x, \theta) = N(y; F_\theta(x), \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(F_\theta(x) - y)^2}{2\sigma^2}\right) \quad (1)$$

Then, applying the Bayes' theorem, the Bayesian posterior distribution can then be computed as:

$$p(\theta|X, Y) = \frac{p(Y|X, \theta)p(\theta)}{\int p(Y|X, \theta')p(\theta')d\theta'} \quad (2)$$

This posterior distribution describes which are the most likely values for the parameters θ to have generated the data, according to the combination of prior knowledge about θ and the information about θ provided by the training data. Finally, given this Bayesian posterior, it becomes possible to compute a marginal probability distribution of the network's output, given a certain new input x^* , which quantifies exactly the uncertainty of model predictions:

$$p(y^*|x^*, X, Y) = \int p(y^*|x^*, \theta')p(\theta'|X, Y)d\theta' \quad (3)$$

In practice, however, this predictive distribution is summarized by a few statistics computed using a Monte Carlo sampling approach. In this approach T samples of θ are sampled from the posterior distribution $p(\theta|X, Y)$ and then used to compute statistics of the predictive distribution. For instance, to summarize the prediction of a BNN used in regression problem, the usual procedure is to perform model averaging:

$$\tilde{y}^* = \frac{1}{T} \sum_{i=1}^T F_{\theta_i}(x^*) \quad (4)$$

While the uncertainty of the prediction is computed as follows:

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_{i=1}^T (F_{\theta_i}(x^*) - \tilde{y}^*)^2 \quad (5)$$

One problem while training a BNN is that the posterior distribution computation is usually an intractable problem, since computing the evidence $\int p(Y|X, \theta')p(\theta')d\theta'$ cannot be done analytically. To address this problem, Bayesian inference approximation methods are normally used to train BNNs in practice. A review of Bayesian inference approximation methods specific for BNNs is presented in Gal (2016). For this study, the variational inference method called Monte Carlo (MC) Dropout (Gal and Ghahramani, 2016) is adopted, since it allows to approximately perform Bayesian inference in BNNs with only simple adaptations of traditional neural networks training techniques.

2.2. Bayesian Neural Network Calibration Method

The BNN-based air quality sensor calibration method proposed for this study is schematized in Figure 1. As illustrated, the goal of this method is to train a BNN calibration model that takes as input a raw sensor measurement (x_{raw}) and external conditions data (x_{ext}) (such as temperature, humidity, and concentration of other pollutants sensors), and returns as output a calibrated concentration measurement of the target pollutant of the sensor (y_{cal}^*). To train the model, the Bayesian inference

approximation method MC-Dropout (Gal and Ghahramani, 2016) is used. Previously collected raw sensor measurements and external conditions data stored in a historical database are used as the input variables of the training dataset, and reference measurements provided by some reference meter close to the sensor (such as a traditional monitoring station or another previously calibrated sensor) are used as target outputs Y of the training dataset. Once trained, the BNN model is then deployed to calibrate new raw sensor measurements. Since the trained model is Bayesian, when receiving as input new inputs $x^* = (x_{raw}^*, x_{ext}^*)$, it returns a predictive probabilistic distribution of calibrated pollutant concentration values corresponding to the new raw sensor measurements $p(y_{cal}^* | x^*, \theta)$. To simplify the model output, only the mean of the predictive probabilistic distribution (\tilde{y}_{cal}^*) and its respective uncertainty value ($\tilde{\sigma}_{cal}$) are outputted by the model.

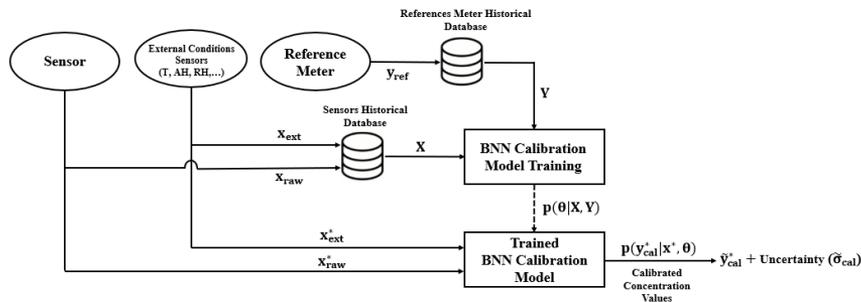


Figure 1 - BNN calibration method

3. Experiment

3.1. Data

To evaluate the proposed method's performance, an application experiment was carried out. For this experiment, the ENEA (National Agency for New Technologies, Energy and Sustainable Economic Development) dataset (De Vito et al., 2008) was utilized. This dataset was chosen for this experiment since it has already been used as a benchmark for evaluation of other machine learning-based sensor calibration methods (De Vito et al., 2018), thus enabling the performance comparison between the proposed method and methods already tested against this dataset. A complete description of this dataset can be found in De Vito et al. (2018).

3.2. Experimental Setting

The specific case of the CO sensor calibration was chosen for this experiment since the performance of other machine learning-based methods in the calibration of this specific sensor were described in De Vito et al. (2018) and it can be used for performance comparison. To allow a fair comparison between the proposed methods and the methods used in De Vito et al. (2018), a similar dataset split configuration was used: the first 504 instances were used to train the BNN model and the remaining instances were used to test the trained model. To assess the performance of the model, the mean absolute error (MAE) was used, since it was the performance metric used to evaluate the other machine learning models in De Vito et al. (2018).

3.3. Calibration Model

To calibrate the CO sensor, a Bayesian feed-forward neural network model was chosen. The model was structured to receive as input the responses from CO, NMHC, NOx, NO2, O3, temperature, relative humidity, and absolute humidity sensors and to return as output CO concentration values. To train the model, the CO ground truth concentrations values

provided by the reference monitoring station were used as target outputs. The model was implemented using the Python programming language and the Pytorch library (Paszke et al., 2019). The optimal model was determined using the Optuna framework (Akiba et al., 2019). The optimal Bayesian neural network model was structured with a single hidden layer containing 24 processing units that use the ReLU activation function and an output layer containing only one processing unit. This model was trained using a dropout rate equal to 0.163, a learning rate equal to 0.012, 400 training epochs, and an RMSprop optimizer (Tieleman et al., 2012).

4. Results and Discussion

To evaluate its performance, the optimal model was tested against the test dataset. The test experiment showed that the model has good accuracy in the calibration of the CO sensor with a good generalization capacity. As highlighted in Table 1, the model presented an MAE value of 0.47, which is lower compared to the MAE values obtained by the models evaluated in De Vito et al. (2018) for the calibration of the same sensor. Furthermore, as can be seen in Figure 2, uncertainty estimates for the calibrated measurement are directly provided by the calibration model. Uncertainty measures are an important metric in the context of validation of air quality sensors (Council of European Union, 2008), as they allow to evaluate the quality and reliability of the measurements provided by them. In the literature, only a few machine learning-based methods for the calibration of air pollution sensors address the obtainment of uncertainty estimates of calibrated measurements, and usually, in their approaches these estimates are not obtained directly by the calibration model, requiring the use of extra mechanisms to compute the uncertainty values (Esposito et al., 2016). Therefore, the proposed method proves to be novel in this sense.

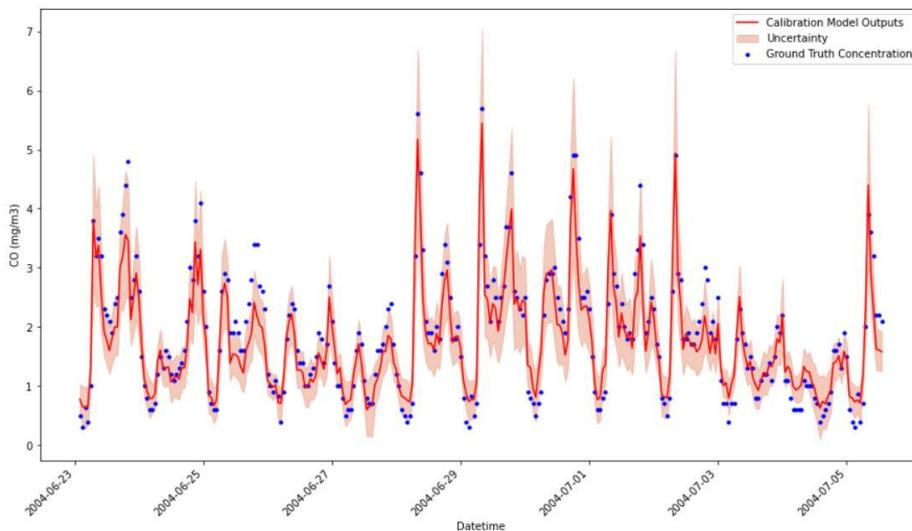


Figure 2 - Uncertainty estimates for the calibrated measurement

5. Conclusion

A Bayesian neural network calibration method for low-cost air quality sensors was proposed in this study. The experimental results show that the method has a good accuracy performance, with a lower mean absolute error compared to other machine

learning-based calibration methods applied to the same dataset. In addition, the method presents the advantage of directly providing estimations of the uncertainty of the calibrated measurements, which is an important metric used to assess the quality of data provided by air pollution sensors and that most other calibration methods usually cannot provide.

Table 1 - Performance comparison

Model	Mean Absolute Error (MAE)
Multiple Linear Regressors	1.09
Support Vector Machine	0.64
Gaussian Process Regression	0.55
Neural Network	0.61
Bayesian Neural Network	0.47

Acknowledgements: The authors thank The São Paulo Research Foundation (FAPESP) for the financial support (FAPESP Proc. 17/50343-2 and 19/08280-9).

References

- T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, 2019, Optuna: A next-generation hyperparameter optimization framework, In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data mining, 2623–2631
- F. Concas, J. Mineraud, E. Lagerspetz, S. Varjonen, X. Liu, K. Puolamäki, P. Nurmi, S. Tarkoma, 2021. Low-cost outdoor air quality monitoring and sensor calibration: A survey and critical analysis, *ACM Transactions on Sensor Networks (TOSN)* 17 (2), 1–44.
- Council of European Union, 2008. Directive 2008/50/ec of the european parliament and of the council of 21 May 2008 on ambient air quality and cleaner air for Europe, *Official Journal of the European Union*.
- S. De Vito, E. Esposito, M. Salvato, O. Popoola, F. Formisano, R. Jones, G. Di Francia, 2018. Calibrating chemical multi-sensory devices for real world applications: An in-depth comparison of quantitative machine learning approaches, *Sensors and Actuators B: Chemical* 255, 1191–1210.
- S. De Vito, E. Massera, M. Piga, L. Martinotto, G. Di Francia, 2008. On field calibration of an electronic nose for benzene est in an urban pollution monitoring scenario. *Sensors and Actuators B: Chemical* 129 (2), 750–757.
- E. Esposito, S. De Vito, M. Salvato, V. Bright, R. L. Jones, O. Popoola, 2016. Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems. *Sensors and Actuators B: Chemical* 231, 701–713.
- Y. Gal, 2016. Uncertainty in deep learning. Ph.D. thesis, University of Cambridge.
- Y. Gal, Z. Ghahramani, 2016. Dropout as a bayesian approximation: Representing model uncertainty in deep learning, In: *International Conference on Machine Learning*. PMLR, 1050–1059.
- L. V. Jospin, W. Buntine, F. Boussaid, H. Laga, M. Bennamoun, 2020. Hands-on bayesian neural networks—a tutorial for deep learning users, *arXiv preprint arXiv:2007.06823*
- A. Lewis, P. Edwards, 2016. Validate personal air-pollution sensors. *Nature* 535 (7610), 29–31.
- B. Maag, Z. Zhou, L. Thiele, 2018. A survey on sensor calibration in air pollution monitoring deployments, *IEEE Internet of Things Journal* 5 (6), 4857–4870.
- A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., 2019. Pytorch: An imperative style, high-performance deep learning library, *Advances in neural information processing systems* 32, 8026–8037.
- T. Tieleman, G. Hinton, 2012, Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude, *COURSERA: Neural networks for machine learning* 4 (2), 26–31.
- WHO, 2016, Ambient air pollution: A global assessment of exposure and burden of disease, World Health Organization.

Hyperconic Machine Learning to Predict Microbial Growth

J. Cortez-González^a, J.P. Serrano-Rubio^a, R. Murrieta-Dueñas^{a,*}, I. Segovia-Dominguez^{b,c}, V. López-Ramírez^a

^a *Tecnológico Nacional de México / ITS de Irapuato (ITESI), Carretera Irapuato-Silao Km. 12.5, C.P.36821 Irapuato, Guanajuato, México*

^b *University of Texas at Dallas, Mathematical Science department, TX 75080-3021, US*

^c *Jet Propulsion Laboratory, Caltech, Pasadena, CA 91109, US*

*corresponding author: rodolfo.md@irapuato.tecnm.mx

Abstract

The design and control of biological reactors depends upon the correct calibration and selection of kinetic parameters from the exponential phase in the microbial growth. Nowadays, kinetic microbial parameters reported in the literature of biochemical engineering are obtained based on specific growth conditions, thus limiting their possible application in a wide variety of dynamic models. In this paper, we propose an adaptive machine learning approach to predict microbial growth and provide information on the effect that variations of pH (5, 7, 9) and concentration of the culture media (10, 20, 40, 60, 80, 100%v/v) have on the growth rate. We study biological reactions and obtain a set of experimental exploratory data using the *Pseudomonas aeruginosa*. The versatile and robust metabolism of *P. aeruginosa* is responsible of its ability of growth in different environment conditions even at low nutrient and oxygen levels, in a sample range of temperatures (4°- 42° C) and polluted sites. As first contribution, we propose a technique to gather experimental data via measurements of optical density, microbial growth, from the Multiskan™ FC Microplate Photometer, Thermo Scientific. Our second contribution consist in integrating the Hyperconic Multilayer Perceptron (HCMLP) as computational and mathematical approach to predict microbial growth across all the set of conditions of the experimental design. HCMLP is a state-of-the-art method to define complex non-linear decision boundaries, in the parameters' space, using a mix of ideas from conformal geometry and neural networks, and focusing on quadratic hyper-surfaces through multiple hidden layers. As a consequence, we generate precise hyper-surface responses which predicts microbial growth even in values not evaluated during the experimental stage. Finally, our statistical testing and comparisons validate that the proposed experimental, mathematical and computational framework is robust and capable of predicting the dynamic growth of bacteria *P. aeruginosa* using two main operation conditions: pH and concentration culture media. In the future, we plan to apply our proposed methodology to other bacterial strains and advance HCMLP for forecasting dynamics of other multiple microbial measurements under a wide variety of conditions.

Keywords: Microbial Growth, Machine Learning, Fitting Experimental Data, Hyperconic Multilayer Perceptron

1. Introduction

The chemical, biochemical and pharmaceutical industries have always focused on the development of new products. However, the scaling of equipment is an important bottleneck in the transition from chemical, biochemical and pharmaceutical research to the development of processes for which a large amount of experimental data is necessary, which subsequently allows the generation of models to describe the behavior of the process. A crucial part of this procedure is the development of a kinetic model of the chemical or biological reaction system. The tools that engineers provide, in combination with the use of specialized software, allow us to reduce time and development costs through a reduction in laboratory experiments, rapid expansion and avoidance of the need for expensive tests pilot. In the particular case of microbial kinetics, existing models only govern processes for specific operating/growth conditions; that is, if the operating conditions change, the kinetic parameters must be obtained through new experiments, thus increasing both the material and monetary costs. Here, the development of more robust and precise kinetic models for predicting microbial growth emerge as a novel approach to build profitable and faster chemical process which, in turn, lower monetary costs for the end-user.

Modelling of microbial kinetic behavior through mathematical computing, particularly machine learning (ML), is a still vastly unexplored field. The biokinetic parameters for microbial growth of *P. aeruginosa*, assuming static growth conditions, are experimentally determined by Bakke et al. (1984); and Tuovinen (1984) and Beyenal et al. (2003) in biofilms, planktonic cultures and double substrate growth kinetics of *P. aeruginosa*, respectively. Tochampa et al. (2005) develops a model of xylitol production by the yeast *C. mogii* ATCC 18364 in the presence of glucose as the cosubstrate using a genetic algorithm, further improved by Sirisansaneeyakul et al. (2013) through mathematical optimization. Lee et al. (2015) provides growth kinetic models for microalgae cultivation using multiple factors but only under single operation conditions. Recently, Eze et al. (2018) designed a kinetic model to describe the growth of microalgae strain *Desmodesmus sp.* in wastewater cultures using numerical methods for differential equations. As shown in literature, although there are models of microbial kinetic behavior, most of them lack of adaptability using various dynamic factors and under multiple operation conditions.

Our contributions focus in both experimental design for data collection, and modelling through a novel machine learning technique, with particular emphasis on prediction of microbial growth. Section 2 describes the proposed framework, from biochemical experimentation to the machine learning application. Section 3 analyze and statistically validate the prediction under various conditions. Finally, Section 4 summarize our findings and states future work.

2. Methodology

The proposed framework include two distinctive phases: 1) biochemical experimentation, and 2) microbial growth data approximation. As first contribution, we study the microbial growth of the *P. aeruginosa* bacteria on selective media, and introduce a technique to analyze their cellular growth under various conditions of pH, culture media and agitation speed. We keep an oscillating temperature between 23° C and 32° C for kinetic inspection during all experiments. Thus, in our second contribution, the proposed kinetic model

predicts the microbial growth and its dynamics using pH, medium and agitation speed as inputs.

2.1. Biochemical experimentation

Experimental design takes into account two values in the culture medium, 80% and 100%, whilst pH is set to three different values, 5.0, 7.0 y 9.0. We perform four runs of each experiment in a Multiskan microplate spectrophotometer with a 5% measuring error, at most. To statistically verify that microbial growth is affected by pH values and the culture medium, we apply an analysis of variance ANOVA with 95% confidence interval. The *P. aeruginosa* strain (SA1-12) is isolated from Lerma river in Salamanca, Guanajuato, Mexico; and identified through 16S DNAr amplification and comparison analysis in public databases. To evaluate the growth of *P. aeruginosa* (SA1-12) at different pH and levels of nutrients and intensities of agitation during the first stages of development, we prepare overnight cultures of the bacteria in King's B Medium at 28° C and 100rpm. Bacterial pellets are recovered by centrifugation and resuspended in diluted culture media according the experimental condition (pH; 5.0, 7.0, 9.0; King's B medium; 10, 20, 40, 60, 80, 100% v/v). The fresh cultures media are inoculated with 0.01% of pre-inoculums and 200µL are deposited in 96-multiwell plate. The optical density is measured every 15min for 8h ($A = 620\text{nm}$, in a microplate reader Multiskan™ FC Microplate Photometer, Thermo Scientific) at two intensities of agitation, low and medium, at 28° C. We use a multiwell plate during our testing stage.

2.2. Microbial Growth Data Approximation

The second phase consist in learning data patterns using collected experimental data of microbial kinetics growth of *P. aeruginosa*. We develop a surface response model via modern approaches from machine learning based on geometric algebra. Well tested capabilities of modern artificial intelligence tools, e.g. novel versions of neural networks, provide us with an efficient way to predict microbial growth on unknown values of the exogenous variables.

The proposed Hyperconic Multilayer Perceptron (HC-MLP) includes neurons in the hidden layer named Hyperconic Neurons (HCN), which use transfer functions to define decision boundaries via hyperboloids, paraboloids, spheres, and ellipsoids. Figure 1a shows the architecture of the proposed HCMLP in this study. The architecture consists of three layers: 1) input layer, 2) hidden layer and 3) output layer. The input signals are propagated through the network from left to right, where the output scalar y comes from the final output layer. The calculation of the HCN includes the estimation of a conic. In this study, we use the HCMLP as function approximation and predictive model to estimate the absorbance y using as input the time x_1 and the percentage of the culture media concentration x_2 as is illustrated in Figure 1b. For practical purposes and for convenience of the reader, we present a brief description of the computed calculations for the HCN. By identifying points $x = (x_1, x_2) \in \mathbb{R}^2$ with points $\vec{x} = [x_1, x_2, 1]$, it is well known that given a symmetric 3x3 matrix A , the set of points $x = (x_1, x_2) \in \mathbb{R}^2$ such that $\vec{x}A\vec{x}^T = 0$ lie in a conic, i.e.

$$[x_1 \quad x_2 \quad x_3] \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = 0 \quad (1)$$

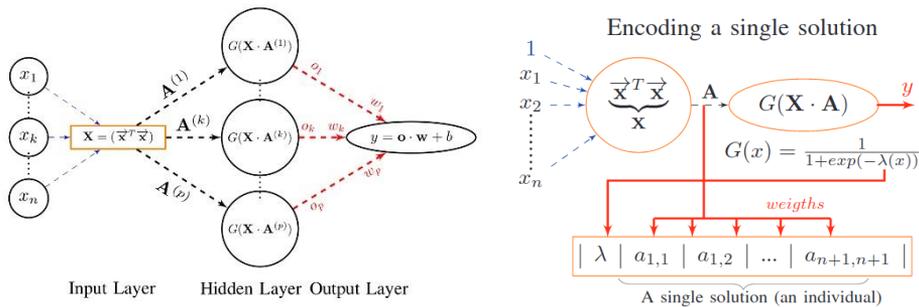
On the other hand, the relative position of a point (x_1, x_2) with respect to a conic can be obtained by means of the sign of the following number $(\vec{x}^T \vec{x})A$. Finding the best approximation of a solution to the equation $(\vec{x}^T \vec{x})A = 0$ for a set of points (x_1, x_2) is usually called the algebraic estimation of the conic passing through these points. Therefore, the output of the neuron in the hidden layer can be written as a function composition of a non-linear associator and the sigmoid function G (Serrano-Rubio et al., 2017). Thus, the output of a neuron o_i in the hidden layer is written as follows

$$o_i = \frac{1}{1 + \exp(-\lambda(\vec{x}^T \vec{x}) \cdot A^{(i)})} \tag{2}$$

Where $\lambda \in \mathbb{R}$ is the slope of the sigmoid function. Note that the machine learning problem, from an optimization perspective, consist in finding parameters $a_{i,j}$ and λ .

The Spherical Evolutionary Algorithm (SEA) is implemented for training the HCMLP (Serrano-Rubio et al., 2018). The population of SEA represents the potential solutions that determines an optimal approximation to the target outputs. These solutions are vectors which are encoded as individuals for the evolutionary algorithm. Therefore, the goal of the training is to find the vector of parameters that determines an optimal approximation to the target dataset. Figure 1b shows the encoding of a single solution.

The experimental setup is given as follows: The number of neurons in the hidden layer is varied to obtain the best topological-architecture for our Evolutionary Artificial Neural Network. The number of neurons is as follows: 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10. We use a dataset of 190 feature vectors. The population for the evolutionary algorithm was set to 30 individuals with uniform random initialization in the domain $[-1, +1]$. The stop criteria of the training occurs when the algorithm reaches an error of 0.0001 or 10000 fitness function evaluations. As fitness function we use the widely tested least square function during the training stage. Our model is validated using standard growth factors. We train the machine learning model using three pH values (5.0, 7.0 and 9.0) and multiple medium concentrations (100%, 80%, 60%, 40%, 20% and 10%). Our model predicts absorbance values for any configuration of pH in $[5.0, 9.0]$, concentration of medium in $[10\%, 100\%]$, and any timestamp, hours (h), between $0 h$ and $8 h$.



a) Hyperconic Multilayer Perceptron.

b) Technique to encode a single solution. Each solution consist of weights $a_{i,j}$ where $a_{i,j}$ represents the parameters to estimate for each conic neuron.

Figure 1. Hyperconic Machine Learning Framework.

3. Results

Common mathematical approaches, on microbial growth kinetics, assumes specific conditions for optimal bacterial growth. However, biological reactors cannot keep the same exact settings during prolonged periods of time due to inherent unpredictability on the reactor system. Thus, designing and deploying an approach capable of modelling the nonlinear behavior of microbial growth, lessen the amount of resources necessary for complex experiments in bioreactors. A key advantage is the reduction of spent time in laboratories because our proposed method can learn a characteristic surface, and predict outputs for conditions not present in the experimental dataset. In this study, we focus on the robust *P. aeruginosa* strain which flourish in pH values around 7.0. Our mathematical and computational model predicts the dynamic growth of bacteria *P. aeruginosa* when tuning two main operation conditions: pH and concentration media culture. Inputs/outputs of learning model F are as follows

$$\text{absorbance} = \beta = F(\text{pH}, [\text{media culture}], t),$$

Where β is the output-absorbance value, pH is the acidity or basicity of the solution, in our case of study pH=7.0, concentration media culture is the solution concentration and t represents a specific time. We perform statistical hypothesis testing to verify that microbial growth depends upon bioreactor's operation conditions. Here we use an ANOVA test with a significance level of $\alpha = 0.05$, thus confirming the impact on microbial growth when varying growth conditions. Figure 2 shows a visual comparison of experimental-data time series using different media culture during the 8h time interval, versus output surface response using our hyperconic multilayer perceptron.

Validation of the proposed model is performed in two different phases. In phase 1, we obtain the training error of 6.6939E-4, which refers to the difference between the value predicted by the HC-MLP and the experimental value, considering only the data set with which the network is trained. In phase 2, the testing error of 6.9530E-4 is obtained, i.e., the difference between the predicted value and the test-experimental value of some growth condition that is within the established interval for the input variables (pH, time and concentration media culture). Table 1 shows the comparison of the absorbance value predicted by the HC-MLP algorithm and the experimental value for the *P. aeruginosa* strain at pH=7.0. The calculated error presents an order of magnitude between 10E-3 and 10E-5. The order of magnitude of these errors can be compared with similar studies carried out in the area of food microbiology, regarding the prediction of microbial growth in foods exposed to environmental conditions. In this sense, the proposed mathematical models in literature receive pH and temperature as inputs (Pla et al., 2015). Such models are used to predict the growth of 3 strains: *Pseudomonas spp.* (Baranyi et al., 1999), *Listeria monocytogenes* (Oksüz and Buzrul, 2021) and *Bacillus cereus* INRA-AVTZ 415 (Pla et al., 2015) in different culture media obtaining in all cases errors in the adjustment of the order of 10E-3.

If the testing error of the HC-MLP algorithm is compared with the error obtained in the literature, it is evident the reduction of 1 order of magnitude in the fit of the data by the proposed algorithm. This means, that our tool improves on what is reported in the literature with respect to similar works for fitting experimental microbial growth data. Therefore, our most important contribution is the development of a methodology for the prediction of microbial kinetics based on robust and modern machine learning tools, and versatile enough to fit experimental data of any kind.

Table 1. Experimental value versus predicted value absorbance when training HC-MLP with $pH=7.0$

Exp	Microbial Growth condition		Absorbance (β)		Error (SSE)
	t (h)	[Culture Media] %	Experimental	Predicted	
1	0.25	80	0.0559	0.0650	8.281E-5
2	3.5	60	0.2728	0.2670	3.364E-5
3	5.25	40	0.3860	0.4210	1.220E-3
4	6.00	20	0.3879	0.3910	1.600E-5
5	7.25	10	0.2991	0.3380	1.140E-3

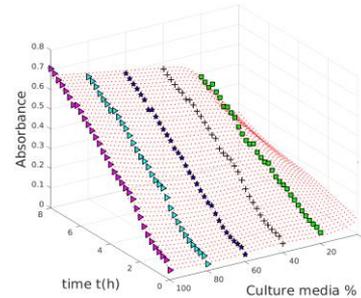


Figure 2. Surface response using our hyperconic multilayer perceptron

4. Conclusions

We have proposed the first model to predict the growth of the native bacterial strain *P. aeruginosa* based on Hyperconic Networks and conformal algebra. We have developed a mathematical and computational method to systematically select the structure of kinetic models and partially automatize the laboratory process. We have collected experimental data of our bacterial strain and statistically validated our approach to predict the microbial growth under operation conditions. Our studies indicates that an artificial intelligence-based predictor of microbial growth can be generated that takes into account changes in growth conditions. In the future we plan to apply our proposed approach to other bacterial strains and advance HCMLP to forecasting dynamics of multiple microbial measures. This methodology can be extended to adjust microbial growth data of any microorganism to obtain a characteristic growth surface as a response to changes in the conditions that most affect it.

References

- Bakke, R., Trulear, M. G., Robinson, J. A., & Characklis, W. G. (1984). Activity of *Pseudomonas aeruginosa* in biofilms: steady state. *Biotechnology and bioengineering*, 26(12), 1418-1424.
- Beyenal, H., Chen, S. N., & Lewandowski, Z. (2003). The double substrate growth kinetics of *Pseudomonas aeruginosa*. *Enzyme and Microbial Technology*, 32(1), 92-98.
- Eze, V. C., Velasquez-Orta, S. B., Hernández-García, A., Monje-Ramírez, I., & Orta-Ledesma, M. T. (2018). Kinetic modelling of microalgae cultivation for wastewater treatment and carbon dioxide sequestration. *Algal Research*, 32, 131-141.
- Serrano-Rubio, J. P., Hernández-Aguirre, A., & Herrera-Guzmán, R. (2017). Hyperconic multilayer perceptron. *Neural Processing Letters*, 45(1), 29-58.
- Lee, E., Jalalizadeh, M., & Zhang, Q. (2015). Growth kinetic models for microalgae cultivation: a review. *Algal research*, 12, 497-512.
- Robinson, J. B., & Tuovinen, O. (1984). Mechanisms of microbial resistance and detoxification of mercury and organomercury compounds: physiological, biochemical, and genetic analyses. *Microbiological reviews*, 48(2), 95-124.
- Sirisansaneeyakul, S., Wannawilai, S., & Chisti, Y. (2013). Repeated fed-batch production of xylitol by *Candida magnoliae* TISTR 5663. *Journal of Chemical Technology & Biotechnology*, 88(6), 1121-1129.
- Tochampa, W., Sirisansaneeyakul, S., Vanichsriratana, W., Srinophakun, P., Bakker, H. H., & Chisti, Y. (2005). A model of xylitol production by the yeast *Candida mogii*. *Bioprocess and biosystems engineering*, 28(3), 175-183.
- Serrano-Rubio, J. P., Hernandez-Aguirre, A., & Herrera-Guzman, R. (2018). An evolutionary algorithm using spherical inversions. *Soft Computing*, 22(6), 1993-2014.
- Pla, M. L., Oltra, S., Esteban, M. D., Andreu, S., & Palop, A. (2015). Comparison of primary models to predict microbial growth by the plate count and absorbance methods. *BioMed research international*, 2015.
- Baranyi, J., Pin, C., & Ross, T. (1999). Validating and comparing predictive models. *International journal of food microbiology*, 48(3), 159-166.
- Öksüz, H. B., & Buzrul, S. (2021). Monte Carlo analysis for microbial growth curves. *Journal of Microbiology, Biotechnology and Food Sciences*, 2021, 418-423.

Optimal Contract Selection for Contract Manufacturing Organisations in the Pharmaceutical Industry Under Uncertainty

Apostolos P. Elekidis,^a Michael C. Georgiadis^{a,*}

^a *Department of Chemical Engineering, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece*
mgeorg@auth.gr

Abstract

This work presents an integrated planning and scheduling framework for the optimal contract selection problem of Contract Manufacturing Organisations (CMO) under uncertainty. Considering a multistage, multiproduct, batch facility of a secondary pharmaceutical industry, an aggregated MILP planning model is firstly proposed including material balances and allocation constraints. Utilizing a rolling horizon approach, the production targets are then provided to a precedence-based MILP scheduling model to define batch-sizing and sequencing decisions in detail. To model demand uncertainty, a scenario-based approach is proposed, considering the Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) measures. Since large number of scenarios imposes significant challenge to computations, a scenario reduction framework is integrated to reduce the total solution time, when considering large-scale problem instances. The proposed methodology increases the profitability of CMOs by selecting the optimal contract combinations depending on their risk tolerance while considering the availability and optimal utilization of underlying production resources.

Keywords: Contract Manufacturing Organizations, Conditional Value-at-Risk, MILP.

1. Introduction

Over the past few years, large R&D pharmaceutical companies have increasingly outsourced non-core activities, such as manufacturing, to Contract Manufacturing Organisations (CMOs), which are companies without their own product portfolio. This policy enables multinational pharmaceutical industries to reduce their costs and emphasise on drug discovery and marketing, which are considered as key parts for their value chain. Typically, drug development is a time-consuming process, as it takes at least 10 years on average for a new medicine to be in the marketplace. Additionally, demand of newly developed pharmaceutical products is usually highly uncertain. Lower drug efficacy can affect the demand and total sales, while in the worst case, it can lead to the suspension or even the withdrawal of the drug. Under this dynamic and uncertain environment, a CMO must decide the best contract combination to accept, so as to maximize its profits (Marques et al.' 2020)

2. Problem statement

The problem under consideration is mainly concerned with the optimal contract selection of a secondary, multistage, pharmaceutical contract manufacturer (CMO). Secondary pharmaceutical manufacturing is focused on the further processing of active ingredients (APIs), by adding excipient inert materials to produce the final products, usually in SKU

form. Secondary manufacturers consist typically of three main production stages operating in batch mode: granulation, compression, and coating (Marques et al., 2020). In each stage, the production takes place in multiple parallel lines, while product-dependent changeovers occur between consecutive product batches, due to the necessary cleaning operations. To ensure the required purity and quality of final products, batch splitting or mixing is not typically allowed in these industries. Although there is no intermediate storage capacity between stages, if necessary, product batches can remain in a production unit after completing their process. The pharmaceutical plant operates 24 hours per day, for five days a week, in order to satisfy a weekly order-driven demand. Contract manufacturers (CMOs) don't produce their own product portfolio but serve other companies in the pharmaceutical industry on a contract basis to provide comprehensive services related to drug manufacturing. The length of contracts usually ranges from six months up to one year or more. A contract can include currently-developed products, characterized by highly volatile demand and high selling prices, or drugs with less uncertain demand and lower profit margins (Johnson. 2005). Currently-developed products are subject to unsystematic risk, while products that have been already placed into the market are affected only by systematic risk. Systematic risk is inherent to the whole market, reflecting the impact of economic, geo-political and financial factors, while unsystematic risk is unique to a specific drug or a group of products. Drugs that are subject only to systematic risk are characterized by similar demand fluctuations. Hence, an unexpected event can lead to an increase or a decrease in the demand for all mature drugs. On the other hand, new pharmaceutical products are characterized by strong demand fluctuations which are mainly related to their level of efficacy and unexpected side effects. In the worst case, strong side effects can lead even to the withdrawal of the drug. Considering these facts, decision-makers must determine the best contract combinations to be signed in order to increase expected profits, while considering the capacity of equipment, plant operational and design constraints, and the risk tolerance of the company.

3. Solution strategy

To model the problem under consideration a hierarchical approach is proposed based on the idea of rolling horizon framework (Johnson. 2005). An aggregated MILP model is proposed for determining planning level decisions such as weekly production and inventory targets. A general precedence MILP model, inspired by the work of Cerdá et al., (2020), is then proposed, for making batch sizing, timing, allocation and sequencing decisions. A feedback loop is included as well, to converge the produced amount of both decision levels.

Given a set of available and already agreed contracts, a CMO must define the best contract mixture to maximize its profits while avoiding high risk exposure. For each contract, demand uncertainty is modeled using several independent scenarios which represent different possible instances while they are associated with a weight representing the probability of the scenario realization. All contract combinations and all the individual scenarios are independent, thus the integrated planning and scheduling problem of each scenario can be solved separately (Johnson. 2005). The planning MILP model and the proposed solution algorithm are presented in the following sections in detail.

3.1. Planning MILP model

$$q_j^{\min} WV_{p,j,w} \leq Q_{p,j,w} \leq q_j^{\max} WV_{p,j,w} \quad \forall p \in P, j \in PJ_p, w \in W \quad (1)$$

$$\sum_{j \in JS_s} Q_{p,j,w} \leq \sum_{j \in JS_{s-1}} Q_{p,j,w} + I_{p,s,w-1} \quad \forall p \in P, s \in S, w \in W: s > 1 \quad (2)$$

$$I_{p,s,w-1} + \sum_{j \in (JS_s \cap PJ_p)} Q_{p,j,w} = I_{p,s,w} + d_{p,s,w} - B_{p,s,w} + B_{p,s,w-1} \quad (3)$$

$$\forall p \in P, s \in S, w \in W$$

$$N_{p,j,w} \geq \frac{Q_{p,j,w}}{q_j^{max}} \quad \forall p \in P, j \in PJ_p, w \in W \quad (4)$$

$$T_{p,j,w} = fx_{p,j} N_{p,j,w} + \frac{Q_{p,j,w}}{vt_{p,j}} \quad \forall p \in P, j \in PJ_p, w \in W \quad (5)$$

$$I_{p,s,w} \leq cap_{p,s} \quad \forall p \in P, s \in S, w \in W \quad (6)$$

$$\sum_p \sum_s I_{p,s,w} \leq wc \quad \forall w \in W \quad (7)$$

$$\sum_p T_{p,j,w} + \sum_p cl_{p,j} W V_{p,j,w} \leq h \quad \forall j \in J, w \in W \quad (8)$$

$$\sum_{j \in (JS_s \cap PJ_p)} T_{p,j,w} \leq avl_s \quad \forall s \in S, w \in W \quad (9)$$

$$\begin{aligned} \max \quad & \overbrace{\sum_c in_c}^{\text{initial contract payment}} + \overbrace{\sum_w \sum_p \sum_{j \in (JS_s \cap PJ_p)} W_{p,j,w} d_{p,s,w} pr_p}^{\text{total sales}} - \\ & \overbrace{\sum_w \sum_p B_{p,s,w} bc}^{\text{backlog cost}} - \overbrace{\sum_w \sum_p \sum_s I_{p,s,w} ic_p}^{\text{inventory cost}} - \overbrace{\sum_w \sum_p \sum_{j \in PJ_p} Q_{p,j,w} qc_p}^{\text{production cost}} - \\ & \overbrace{\sum_p fr_p rc_p}^{\text{raw material fixed cost}} - \overbrace{\sum_p rc_p \sum_{j \in PJ_p} \sum_w (Q_{p,j,w} - fr_p)}^{\text{raw material variable cost}} \end{aligned} \quad (10)$$

Constraints (1) impose upper and lower limits on the production of each product $Q_{p,j,w}$, equals to the maximum (q_j^{max}) and the minimum capacity (q_j^{min}) of the production unit j , respectively. According to the constraints (2), the amount of a product p at each stage s at the end of week w , should not exceed the amount that has been produced at the previous stages plus the amount being stored from the previous week, $I_{p, s,w-1}$. Material mass balances are expressed via constraints (3). The total produced and the stored amount of a product p , must be equal to the weekly demand ($d_{p,s,w}$) and the new stored amount, minus the unsatisfied demand of the current week (backlog, $B_{p,s,w}$), plus backlog from previous weeks. The number of batches of each product is expressed by integer variable $N_{p,j,w}$, which is determined by constraints (4). It is assumed that each production unit will process a certain number of batches at a full capacity and a single batch at flexible size at every stage, which is a valid assumption for realistic case studies. Furthermore, the processing time of each product at each processing unit j , $T_{p,j,w}$, is provided by constraint (5). The terms $fx_{p,j}$ and $vt_{p,j}$ express the coefficient for the fixed and the variable processing time, respectively. Constraints (6) and (7) guarantee that the stored amount doesn't exceed the total warehouse capacity of the plant wc , and the individual capacities of each product at each stage $cap_{p,s}$. According to constraints (8), the total processing time and

the average cleaning time $cl_{p,j}$ of each production unit must be lower than the available time horizon. Additionally, constraints (9) impose an extra upper bound on the total production time of each production stage. The upper bound is given by parameter avl_s , and its initial value is equal to the time horizon h . Since the planning model doesn't include timing and sequencing constraints, in some cases the production targets provided to the scheduling level could be infeasible. To converge the production amounts of the two MILP models, the upper bound is adjusted to the maximum production time defined by the scheduling level. Hence, the planning MILP model becomes more accurate. Finally, the objective function focuses on the maximization of the total profit of the plant. The main income is related to the initial payment of each signed contract and the total revenue of sales. On the other hand, costs include backlog, inventory, production and raw material costs. It is also assumed that if a contract is signed an initial amount of raw materials must be purchased regardless of the actual demand and the produced amount of products.

3.2. Solution algorithm

The proposed solution framework involves three phases. The first phase is concerned with the examination of the feasibility of each contract combination. In particular, the predominant scenario of each combination is solved by using the aggregated planning MILP model. If the generated solution leads to full demand satisfaction, the underlying contract combination is proven feasible. If a contract combination is shown to be infeasible, then any combination which is a superset of the former must also be infeasible. For example, if the combination of contracts C1 and C2 is infeasible then a combination with the contracts C1, C2 and C3 is also infeasible.

In the second phase, the predominant scenario of each combination that has been proven feasible in the first phase is solved by using the integrating planning and scheduling MILP framework. If the solution of any iteration of the rolling horizon framework leads to unsatisfied demand (generation of backlog) then the contract combination is proven infeasible and thus, any combination which is a superset of the former, deemed infeasible as well.

Since all contract combinations and the individual scenarios of each combination are independent, the proposed integrated planning and scheduling MILP framework can be solved separately for each combination and each contract. Thus, regarding the third phase of the proposed algorithm, the planning and scheduling problem of each combination is solved for each scenario. To cope with problems that involve large number of scenarios, a scenario reduction framework based on the work of Li and Floudas (2014), is utilized. Considering scheduling level decisions, the total profit is accurately reported for all scenarios and thus, risk metrics such as Value-at-Risk and Conditional Value-at-Risk can be calculated. As a result, decision-makers can determine the optimal contract combination based on their risk tolerance.

4. Application study

The efficiency and applicability of the proposed solution strategy is illustrated using a representative case study concerning a secondary pharmaceutical CMO. The problem is focused on a multistage batch facility, which consists of three stages (granulation, compression, and coating). Each stage includes multiple production units with varying capacity and production rates. The CMO must decide the best contract combination among 6 contracts, as it is presented in Table 1, while contract C1 is already agreed. Each contract has four demand scenarios: high, target and low demand, and failure. Contracts C1-C3 includes products with less demand volatility and smaller profit margins while

contracts C4-C6 are related to new products with higher selling prices. The time horizon is 1 year (52 weeks) and a weekly demand has to be satisfied for each product. The 6 contracts lead to 28 combinations with different demand distributions. Since contracts C1-C3 are subject only to systematic risk, their combinations consist of the same scenarios including high, target and low demand with the same realization probabilities. On the other hand, if a combination includes contracts with new drugs, the probability of combined scenarios is calculated as the product of the probabilities of the new drugs and the probability of the developed products.

Table 1 Contract data

contracts	Contract availability / Product type	Products	Demand multiplier for each scenario and probability of realization			
			High	Target	Low	Fail
C1	Agreed/developed	1,2	1.2 (10%)	1 (80%)	0.8 (10%)	0 (0%)
C2	non agreed/developed	3,4	1.2 (10%)	1 (80%)	0.8 (10%)	0 (0%)
C3	non agreed/developed	5,6,7	1.2 (10%)	1 (80%)	0.8 (10%)	0 (0%)
C4	non agreed/new	8,9,10	1.2 (15%)	1 (60%)	0.5 (20%)	0 (5%)
C5	non agreed/ new	11,12	1.4 (20%)	1 (50%)	0.4 (20%)	0 (10%)
C6	non agreed/ new	13,14	1.7 (25%)	1 (40%)	0.2 (15%)	0 (20%)

Table 2 Summary of results

Feasible Contact combination	Exp. Profit	VaR _{90%}	VaR _{95%}	CVaR _{90%}	CVaR _{95%}	Max. Profit
C1-C3	0.96	0.76	0.76	0.76	0.76	1.15
C1-C4	1.83	1.12	0.43	0.76	0.42	2.46
C1-C5	0.74	0.27	0.27	0.27	0.27	1.09
C1-C6	1.83	0.42	0.42	0.42	0.41	3.42
C1-C3-C5	1.50	1.15	1.15	1.10	1.06	1.99
C1-C3-C6	2.36	1.29	1.29	1.22	1.16	4.30
C1-C5-C6	2.66	0.76	0.76	0.76	0.76	4.23

*The values represent millions of relative monetary units (r.m.u.)

According to the first phase of the proposed solution algorithm, all combinations are solved using the aggregated planning model for the predominant scenario, by considering the target demand. Then, the feasible combinations are solved by applying the integrated planning and scheduling framework. The feasible combinations according to the second phase are presented in Table 2. Contract combinations C1-C3-C4 and C1-C4-C5 were proven feasible in the first but not in the second stage of the algorithm. Regarding the third phase of the solution framework, the planning and scheduling problem of each scenario is solved for all feasible combinations. To reduce the computational time, a scenario reduction algorithm is utilized so as to consider up to 10 scenarios for each combination. The expected profit, the Value-at-Risk, the Conditional Value-at-Risk, and the maximum profit for each combination are summarized in Table 2. Also, Figure 1 illustrates a bubble chart comparing expected profit, VaR and CVaR for all contract combinations. The diameter of the bubble represents the CVaR value of contract combinations. A favorable contract combination should be represented by a large bubble in the top right-hand corner of the diagram, implying high expected profit, VaR and CVaR values. A risk-neutral approach indicates that combination C1-C5-C6 is the optimal one since it leads to the highest expected profit. However, considering VaR_{90%} and CVaR_{90%}, the combination C1-C3-C6 seems more attractive, when a risk-averse policy is applied. In that case, the profit will exceed 1.29 million of relative monetary units (r.m.u) with a

90% confidence interval, while the mean of the lowest 10% profits will be equal to 1.16 million r.m.u.

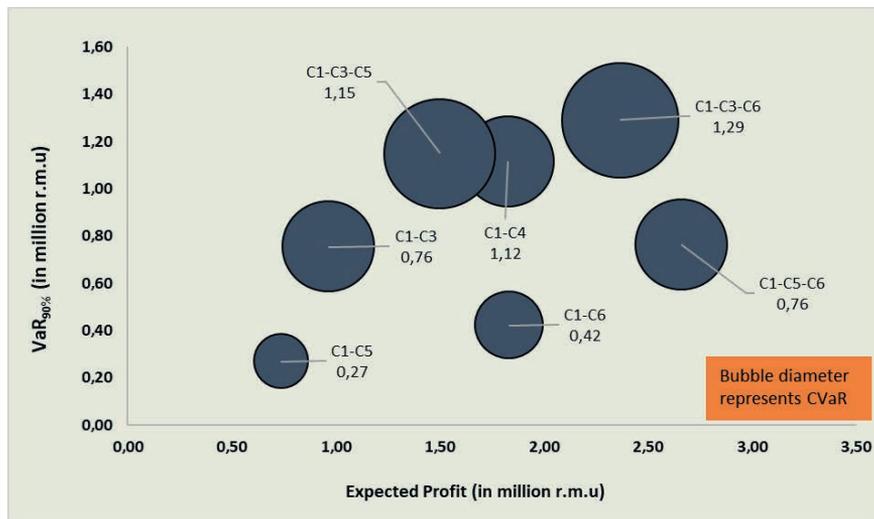


Figure 1 Expected profit, VaR_{90%} and CVaR_{90%} of each contract combination

5. Conclusions

In this work and optimization framework is proposed for the optimal contract selection problem of a CMO in the secondary pharmaceutical industry. Demand uncertainty is modeled by using discrete independent scenarios with known probabilities, while the consideration of scheduling level decisions into an hierarchical planning and scheduling approach leads to accurate profit estimations. In contrast with risk-neutral approaches, the proposed methodology allows decision makers to choose the optimal contract combinations depending on their risk tolerance.

Acknowledgements

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

References

- Cerdá, J., Cafaro, V.G., Cafaro, D.C., 2020. Synchronizing Operations in Multiproduct Batch Facilities Producing Semifinished and Final Products. *Industrial and Engineering Chemistry Research*. 59. 13113–13131.
- Johnson, R.J., 2005. Optimal Contract Appraisal for Fine Chemicals Contract Manufacturing Organisations. PhD thesis. Imperial College London.
- Li, Z., Floudas, C.A., 2014. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization. *Computers and Chemical Engineering*. 70. 50–66.
- Marques, C.M., Moniz, S., de Sousa, J.P., Barbosa-Povoa, A.P., Reklaitis, G., 2020. Decision-support challenges in the chemical-pharmaceutical industry: Findings and future research directions. *Computers and Chemical Engineering*. 134. 106672

A Framework for Economic Optimization of Carbon Capture and Sequestration from Italian Industrial Sources Under Seismic Risk Constraints

Federico d'Amore^a, Gianmarco Marcato^b, Paolo Mocellin^b, Fabrizio Bezzo^{b*}

^a*Politecnico di Milano, Department of Energy, via Lambruschini 4, IT-20156 Milano, Italy*

^b*CAPE-Lab - Computer-Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, IT-35131 Padova, Italy*
fabrizio.bezzo@unipd.it

Abstract

Carbon dioxide emission from the industrial sectors in Italy is second only to emissions from the energy sector. The implementation of carbon capture and storage chains, while considered an ideal solution to decarbonize the industry sector, is cause for additional concern due to the seismic profile of the Italian peninsula. A nationwide multi-objective carbon capture and storage chain is here optimized through a mixed integer linear programming framework, aiming at both cost minimization and reduction of the risk factor linked with the susceptibility of transport to local seismic activity. The resulting minimum specific CO₂ avoidance cost resulting from the economic optimization is 64.7 €/t for a 50% carbon reduction target, while the minimization of seismic risk determines an increase in infrastructural costs up to 79.4 €/t (+22.5%).

Keywords: Cement, steel, oil refining; Carbon capture and storage; Seismic risk; Multi-objective optimization; Mixed integer linear programming.

1. Introduction

Cement plants, steel mills and refineries are responsible for a significant share of carbon dioxide (CO₂) emissions (EEA, 2020), which are among the greenhouse gases responsible for the global threat of climate change. Implementing carbon capture and storage (CCS) technologies is key to reduce a significant portion of CO₂ emissions from these hard-to-abate sectors (Bui et al., 2018). A CCS system is an ensemble of technologies aimed at reducing anthropogenic CO₂ emissions. In particular, the capture of CO₂ consists in its separation from a process stream through different methods depending on the industry and the applied technology, then the transport step is necessary to move the captured CO₂ from emission sources towards areas suitable for sequestration, the latter carried out by injecting the CO₂ deep into underground geological basins. Given the inherent complex, multi-stage nature of CCS chains, the optimization of such systems proves fundamental to achieve a substantial level of decarbonization. Mixed integer linear programming (MILP) models are commonly regarded as powerful methodologies to optimize complex, multifaceted systems, such as energy systems, and they have been recently exploited for the case of CCS networks (Tapia et al., 2018). Several studies optimized CCS chains at different scales, such as region-to-country scale (Lee et al., 2017; Elahi et al., 2017)). Additionally, higher-level continent-wide analyses were proposed as well (Hasan et al.,

2015; d'Amore et al., 2021). One key issue in the deployment of CCS infrastructures is their social sustainability. Although being practiced for over 30 years, CCS is nowadays still rising public concern, especially when the possibility of leakage is taken into account in the vicinity of densely populated areas (d'Amore et al., 2018). In this context, seismic activity poses an additional requirement during the planning, installation, and operation of a CCS system, particularly with regards to pipelining (Psyrras and Sextos, 2018). In this work we will focus on optimizing a comprehensive CCS framework to abate the CO₂ emissions from the Italian industry sector. However, considering the widespread seismic activity in the Italian peninsula, our model will aim at simultaneously minimizing cost and risk (with respect to seismic phenomena) of a CCS system in Italy. Italian industrial stationary sources (i.e., cement, steel and refinery sectors) will be taken into account within a multi-objective MILP modeling framework.

2. Modelling framework

This section will describe the three echelons of the optimized Italian CCS system in terms of key model input parameters. The level of CO₂ emissions and location of industrial plants are taken from EEA (2020) and referred to the year 2017. The total number of emission points was reduced from that reported in the EEA database to decrease the computational burden while ensuring the representativeness of the sample (i.e., at least 80% of the emissions from each sector are included in the optimization). The reduced dataset includes a total of 27 CO₂ emitting nodes n , subdivided into 20 cement plants (described through subset c), 5 refineries (described through subset r) and 2 steel mills (described through subset s).

The CO₂ capture stage is modelled through a set $k = \{kc, ks_{1,2,3}, kr_{1,2,3}\}$ whose components refer to: (kc) oxy-fuel capture for cement plant; (ks_1) absorption from power plant stack at steel mill; (ks_2) absorption from (ks_1) and coke oven flue gas at steel mill; (ks_3) absorption from (ks_1+ks_2) and capture from sinter plant at steel mill; (kr_1) pre-combustion capture from steam methane reformer at refinery; (kr_2) capture at (kr_1) and post-combustion capture on power unit at refinery; and (kr_3) which comprises (kr_1+kr_2) and post-combustion capture from further emission points at refinery. A complete description of capture plants can be found in d'Amore et al. (2021), including the parameters for determining the CO₂ avoidance cost $CCA_{k,n}$ [€/t] of capture plant k installed in an industrial node n , comprising scale effects on capture plant size. In this study, the CO₂ can be transported through either onshore or offshore pipelines, which are discretized through set p into possible ranges of flowrate to account scale effects of total transport scale on unitary transport cost UTC_p [€/km/t] (Rubin et al., 2015). Offshore transport cost is increased by a $\gamma_{n,n'}$ [=1.71] (d'Amore et al., 2021). Differently, sequestration entails a fixed unitary cost USC [=7.2 €/t (Rubin et al., 2015)], which is increased by a θ_n [=2.5] factor to account for offshore basins (d'Amore et al., 2021). The positional and capacity data on storage are taken from Donda et al. (2011).

The seismicity-related parameters are obtained through a dataset containing the coordinates of polling points over the entire Italian peninsula, which are then averaged into seismic areas and implemented to calculate the risk specific to each pipeline in the transport stage of the CCS system (Gehl et al., 2014).

3. Mathematical formulation

The MILP model aims at minimizing both the total cost TC [€/year] of the CCS system and the total risk TR [ruptures/year] associated to the transport stage:

$$objective = \min\{TC; TR\} \quad (1)$$

The total cost of Eq.(1) is given by the contributions of capture (TCC [€/year]), transport (TTC [€/year]), and sequestration stages (TSC [€/year]):

$$TC = TCC + TTC + TSC \quad (2)$$

In particular, TCC of Eq.(2) is evaluated according to the captured CO_2 flowrate $IN_{k,n}$ [t/year] through capture plant k in node n :

$$TCC = \sum_{k,n} (IN_{k,n} \cdot CCA_{k,n}) \quad (3)$$

TTC of Eq.(2) depends on the transported flowrate $Q_{p,n,n'}$ from node n to n' (resulting from the mass balance among nodes), on the distance between nodes $LD_{n,n'}$ [km], and on unitary cost factors UTC_p [€/t/km]:

$$TTC = \sum_{p,n,n'} (Q_{p,n,n'} \cdot LD_{n,n'} \cdot UTC_p \cdot \gamma_{n,n'}) \quad (4)$$

Storage cost TSC of Eq.(2) is proportional to the yearly stored amount OUT_n [t/year] and to the unitary storage cost USC [€/t], the latter increased by θ_n if located offshore:

$$TSC = \sum_n (OUT_n \cdot USC \cdot \theta_n) \quad (5)$$

The overall stored amount is imposed to be larger than the set Italian carbon reduction target of 50% of yearly emissions from industry. TR of Eq.(1) is the sum of all the repair rates $RR_{n,n'}$ [ruptures/year] pertaining to the pipelines employed in the transport grid, the latter accounted through a binary variable $\lambda_{p,n,n'}$ associated to sizes p between n and n' :

$$TR = \sum_n (\lambda_{p,n,n'} \cdot RR_{n,n'}) \quad (6)$$

$$RR_{n,n'} = 0.002416 \cdot K_l \cdot LD_{n,n'} \cdot \max\{PGV_n; PGV_{n,n'}^{mid}; PGV_{n'}\} \quad \forall n, n' \quad (7)$$

where K_l is a parameter representing the effect of transport characteristics and construction material over repair rate, $LD_{n,n'}$ is the length of the arc, and PGV (calculated as its maximum among its values in n , at the center of the pipe, and in n') is the peak ground velocity, defined as the highest horizontal ground velocity during a seismic event expected in the next 50 years, with an exceedance probability of 10% (USGS, 2021). In particular, K_l [=0.6] was taken from ALA (2001) by assuming carbon steel pipelines transporting dry CO_2 , while PGV was evaluated from 16852 points (then discretized into 315 squares of 50 km of size) of spectral acceleration S_A [g] computed by the Istituto Nazionale di Geofisica e Vulcanologia – INGV (Stucchi et al., 2011), according to the formulation proposed by Allen and Wald (2007):

$$PGV = (386.4 \cdot S_A) / (2\pi \cdot 1.65) \quad (8)$$

4. Results

The MILP multi-objective problem was implemented in GAMS software and optimized through CPLEX solver (maximum optimality gap < 3.9 % and solution time < 2 hours). Results exhibit a clear trade-off between the two objectives (Table 1 and Figure 1), with the cheapest configuration (64.7 €/t) being characterized by the highest value of risk (13.9 ruptures/year) (Case A).

Table 1. Results in terms of specific total costs (TC [€/t]), specific costs breakdown (TCC , TTC , TSC [€/t]), total risk (TR [ruptures/year]), optimality gap ($Opt. gap$ [%]) and solution time ($Sol. time$ [s]).

	TC	TCC	TTC	TSC	TR	$Opt. gap$	$Sol. time$
Case	[€/t]	[€/t]	[€/t]	[€/t]	[rup./y.]	[%]	[s]
A	64.7	51.8	5.7	7.2	13.9	3.87	6631
B	75.6	64.2	4.2	7.2	1.7	0.01	300
C	79.4	67.8	4.4	7.2	1.3	0.01	54

The safest infrastructure (1.3 ruptures/year) is the most expensive (79.4 €/t) (Case C). Case B represents an interesting configuration, characterized by a reasonably balanced solution between costs (75.6 €/year) and risk (1.7 ruptures/year). The costs breakdown shows the significant contribution of the capture stage, independently from the optimized case study, with specific capture costs varying from a minimum of 80.1% (Case A) up to a maximum of 85.4% (Case C) of total costs, while transport and sequestration represent an expenditure in the range of 14.6% (Case C) – 19.9% (Case A) altogether.

More detailed analysis on the capture stage reveals for Case A (best economic network) the significant contribution of cement plants and steel mills in terms of captured CO₂, being these the cheapest options implemented in this model (Figure 1a). As for refineries, these are generally not employed as CO₂ sources, except for the site located in Sardinia, to which a capture plant kr_1 is installed to the steam methane reformer to exploit the beneficial effect on costs of this particularly large emitter. On the other hand, Case A makes extensive use of transportation arcs crossing areas with high values of spectral acceleration, with negative consequences on the performance of this network in terms of seismic risk results. Differently, the risk-oriented optimization takes advantage of the lack of seismicity in the Sardinian area by installing full-scale capture plants at its refinery, which determines the substantial increase in capture costs of Case C (Figure 1c). However, this increase in costs can be limited if considering Case B, which still takes advantage of the non-seismicity of the Sardinian area for safely transporting significant amounts of CO₂, but nonetheless suggests some minor modifications to the captured and transported flowrates to optimize the costs of these two stages (Figure 1b).

A sensitivity analysis on the minimum carbon reduction, which is set as lower bound for CO₂ capture (here varied between 20% and 80% of Italian emissions from industry, against the 50% target already discussed as base case) reveals a minimum specific cost (Case A) ranging between 59.0 €/t for a 20% capture target (-9.0% over base case) and 80.7 €/t for a 80% capture target (+24.5 % over base case). The cost increases alongside the capture target due to the necessity of exploiting progressively more expensive CO₂ sources and more complex transport infrastructures. Oppositely, the best networks in terms of risk minimization show a total risk varying between 0.1 ruptures/year for a 20 % capture target (-88.6% over base case) and a significant 9.9 ruptures/year for a 80% capture target (+675.8% over base case), the latter due to the vast network of pipelines that is needed to achieve such a challenging reduction target. Summarizing, the safest networks entail an increase in costs of +64.4% (20% capture target), +22.5% (50% capture target) and +0.7% (80% capture target) with respect to their respective best economic results.

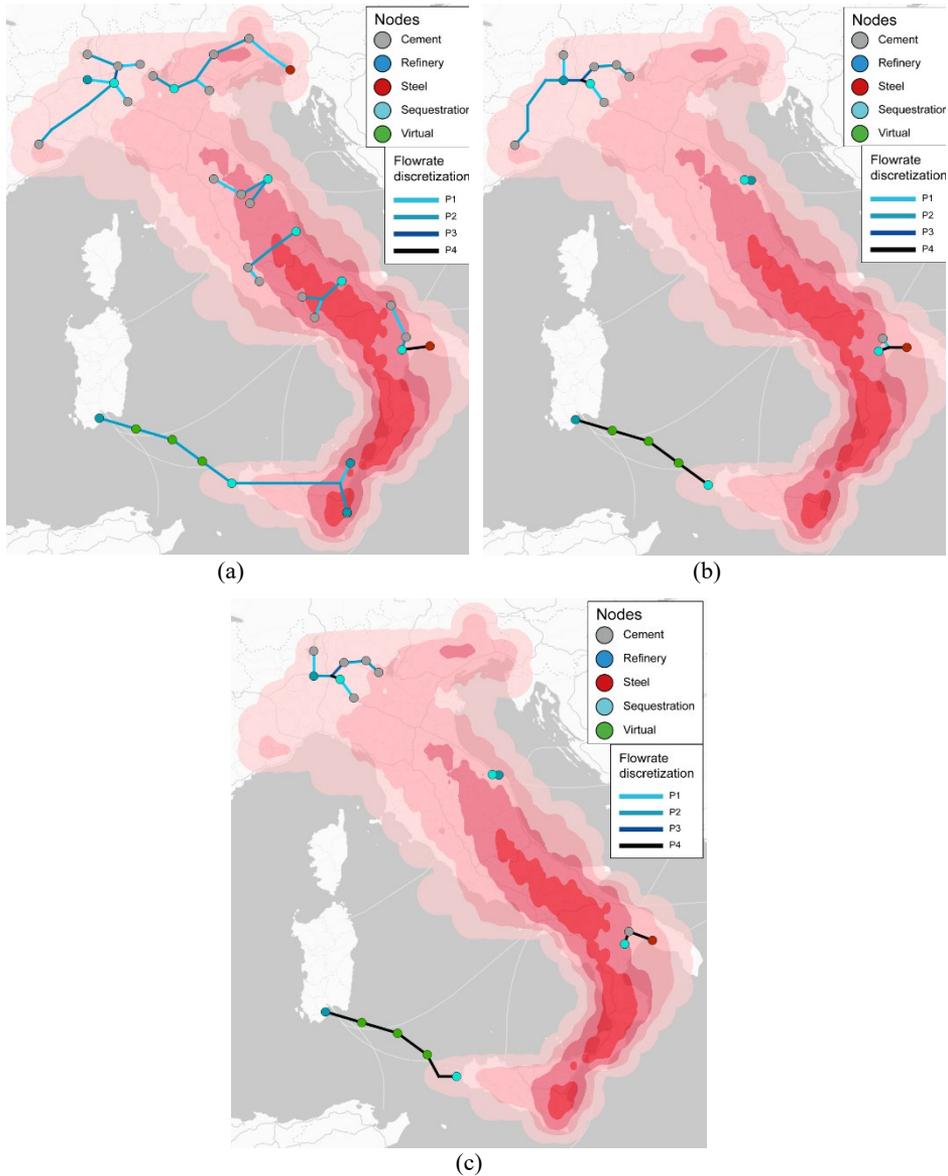


Figure 1. CCS chain configurations for (a) Case A (minimum cost), (b) Case B (trade-off result), and (c) (minimum risk).

5. Conclusions

A nationwide multi-objective carbon capture and storage chain was here optimized through a mixed integer linear programming framework for the context of decarbonizing the Italian industry. The model aimed at minimizing both the total cost (economic objective) and the total seismic risk (risk objective). Depending on the chosen carbon reduction target, the best networks in terms of economic result entailed a total cost of 59.0 €/t (20 % capture target), 64.7 €/t (50 % capture target), and 80.7 €/t (80% capture target). The minimization of risk determined an increase in costs in all analyzed cases due to the

necessity of diverting the transport infrastructure towards areas with low seismic activity and resulted equal to 97.0 €/t for 20% capture target (+64.4% with respect to best economic network), 79.4 €/t for 50% capture target (+22.5%) and 81.3 €/t for 80% capture target (+0.7%). Trade-off configurations were identified that could allow achieving an optimal balance between the two conflicting objectives hence simultaneously reducing costs and minimizing the exposure towards seismic risk.

References

- ALA, 2001, Seismic Fragility Formulations for Water Systems, Part I – Guideline, American Lifeline Association.
- T.I. Allen, D.J. Wald, 2007, Topographic slope as a proxy for seismic site conditions and amplification, U.S. Geological Survey Open-File Report 2007-1357, 69 p.
- M. Bui, C.S. Adjiman, A. Bardow, E.J. Anthony, A. Boston, et al., 2018, Carbon capture and storage (CCS): The way forward, *Energy & Environmental Science*, 11, 1062-1176.
- F. d'Amore, P. Mocellin, C. Vianello, G. Maschio, F. Bezzo, 2018, Economic optimisation of European supply chains for CO₂ capture, transport and sequestration, including societal risk analysis and risk mitigation measures, *Appl. Energy*, 223, 401-415.
- F. d'Amore, M.C. Romano, F. Bezzo, 2021, Optimal design of European supply chains for carbon capture and storage from industrial emission sources including pipe and ship transport, *Int. J. Greenh. Gas Control*, 109, 103372.
- F. Donda, V. Volpi, S. Persoglia, D. Parushev, 2011, CO₂ storage potential of deep saline aquifers: The case of Italy, *Int. J. Greenh. Gas Control*, 5, 327-335.
- EEA, 2020, European Pollutant Release and Transfer Register, <https://prtr.eea.europa.eu/#/pollutantreleases>.
- N. Elahi, N. Shah, A. Korre, S. Durucan, 2017, Multi-stage Stochastic Optimisation of a CO₂ Transport and Geological Storage in the UK, *Energy Procedia*, 114, 6514-6525.
- P. Gehl, N. Desramaut, A. Réveillère, H. Modaresi, 2014, Fragility Functions of Gas and Oil Networks. In: K. Pitilakis, H. Crowley, A. Kaynia (eds) SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk. *Geotechnical, Geological and Earthquake Engineering*, vol. 27, Springer, Dordrecht.
- M.M.F. Hasan, E.L. First, F. Boukouvala, C.A. Floudas, 2015, A multi-scale framework for CO₂ capture, utilization, and sequestration: CCUS and CCU, *Comput. Chem. Eng.*, 81, 2-21.
- S.Y. Lee, J.U. Lee, I.B. Lee, J. Han, 2017, Design under uncertainty of carbon capture and storage infrastructure considering cost, environmental impact, and preference on risk, *Appl. Energy*, 189, 725-738.
- N. Psyrras, A. Sextos, 2018, Safety of buried steel natural gas pipelines under earthquake-induced ground shaking: a review, *Soil Dyn. Earthq. Eng.*, 106, 254-277.
- E.S. Rubin, J.E. Davison, H.J. Herzog, 2015, The cost of CO₂ capture and storage, *Int. J. Greenh. Gas Control*, 40, 378-400.
- M. Stucchi, C. Meletti, V. Montaldo, H. Crowley, G.M. Calvi, E. Boschi, 2011, Seismic Hazard Assessment (2003-2009) for the Italian Building Code, *BSSA*, 101(4), 1885-1911.
- J.F.D. Tapia, J.Y. Lee, R.E.H. Ooi, D.C.Y. Foo, R.R. Tan, 2018, A review of optimization and decision-making models for the planning of CO₂ capture, utilization and storage (CCUS) systems, *Sustain. Prod. Consum.*, 13, 1-15.

Energy-water Scheduling Decisions for Agricultural Scenario Planning

Marcello Di Martino^{a,b}, Sarah Namany^c, Styliani Avraamidou^d, Tareq Al-Ansari^c, Patrick Linke^e and Efstratios N. Pistikopoulos^{a,b,*}

^a*Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, USA*

^b*Texas A&M Energy Institute, Texas A&M University, College Station, TX, USA*

^c*College of Science and Engineering, Hamad Bin Khalifa University, Qatar foundation, Doha, Qatar*

^d*Department of Chemical and Biological Engineering, University of Wisconsin-Madison, WI, USA*

^e*Department of Chemical Engineering, Texas A&M University at Qatar, Education City, Doha, Qatar*

stratos@tamu.edu

Abstract

Food, energy and water resources are heavily interconnected in many process systems and therefore need to be taken into consideration holistically for sustainable decision-making, via a systematic food-energy-water nexus (FEWN) framework. A FEWN study represents a promising decision-making support for regions characterized by scarce water resources, abundant renewable energy resources and significant population growth, as it tackles these different challenges simultaneously while taking into account the effect of each solution strategy on the resource systems at different scales. In this work, we investigate the impact of a FEWN strategy on small-scale farming within agricultural activities. Varying water demands are evaluated on the basis of renewable, and non-renewable based energy sources, along with varying water sources. To this end, a greenhouse connected to a reverse osmosis desalination plant is modeled as a mixed-integer non-linear optimization problem and optimized for an array of objectives to enable a cost-comparison on the basis of water scarcity. The decision-making tool is then extended towards energy-water scheduling decisions for the optimization of agricultural scenario planning.

Keywords: food-energy-water nexus, irrigation planning, energy-water scheduling, mixed-integer optimization.

1. Introduction

Due to a rising global population, rapid urbanization, expanding international trade, cultural and technological changes, diversifying diets, and economic growth the demand for food, energy and water are ever-increasing. It is estimated that by 2070 the global energy consumption will be at least doubled, whereas the global food production already has to be doubled by 2050. The water demand on the other hand is prognosticated to increase between 20% and 30% by 2050. To tackle these challenges sustainably, one has to consider food, energy and water as interconnected resources and model these holistically, which is summarized in the so called food-energy-water nexus (FEWN) (Nie et al., 2019). The necessity of FEWN solution generation is underlined by the following: Agriculture is the world's largest fresh water consumer and around 30% of the energy used globally is expended on food production and supply. Additionally, 90% of the global

power generation is water intensive. Therefore, it is essential to tackle challenges regarding these resources not as isolated problems but rather as interconnected systems, where solution strategies of one resource system influence the other two.

Especially for arid and semi-arid regions like Qatar where water is scarce, renewable energies are plentiful available, the country's food supply relies heavily on imports and the majority of vegetables can only be produced during a limited period of time due to harsh climate conditions, sustainable solution generation has to take into account food, energy and water as interconnected resource systems. Consequently, fully climatized greenhouses using available renewable energies and saline groundwater sources for agricultural applications can overcome these challenges by enabling yearround farming and therefore food supply. Subsequently, this work tackles the analysis and optimization of the energy consumption of a greenhouse connected to a reverse osmosis (RO) system for water purification. The goal is to enable qualitative and quantitative comparisons of varying greenhouse energy-water-food set-ups.

Energy-water systems for agricultural applications have already been investigated in the literature. Li et al. (2021) used a generic optimization approach of agricultural benefit and non-point pollution considering crop and livestock farming policies for distinct regions. Further, Lee et al. (2020) performed a country wide investigation of irrigation management regarding the impact of climate change on productivity, irrigation requirement and energy input of a representative crop. In these cases the focus of the work was set on region or country wide solutions for an agricultural sector based on energy-water systems. Work has also already been performed for interconnected greenhouse farming and RO systems. Rahimi et al. (2021) for example conducted region- and crop-specific techno-economic studies of nexus systems. In addition, detailed RO system modeling is often disregarded, especially when various farming scenarios are to be evaluated (Sadeghi et al. (2020)).

In this work, we introduce a generic framework to link water and energy supply systems for greenhouse farming. For the supply systems, detailed mathematical models are defined in which an array of possible objectives are considered for food-energy-water nexus trade-off analysis (Avraamidou et al., 2018), including cost and water utilization. Section 2 provides an overview of the framework. Section 3 describes the mathematical model for each subsystem, while Section 4 applies the framework to a Qatar based greenhouse case study.

2. FEWN Framework Overview

The framework integrates decisions regarding the energy supply in terms of renewable energy resources, the RO desalination system to supply purified water and the greenhouse farming as the agricultural application. The idea is to use the framework to minimize the overall energy consumption of the system while fulfilling various irrigation scenarios. Further, the model can also be applied to energy-water planning and scheduling decisions for multi-period greenhouse farming.

Thus, the framework consists of a water supply system, energy supply system and farming system, which altogether build an energy-water analysis tool for agricultural planning. These key components, as well as their interactions are illustrated in Figure 1.

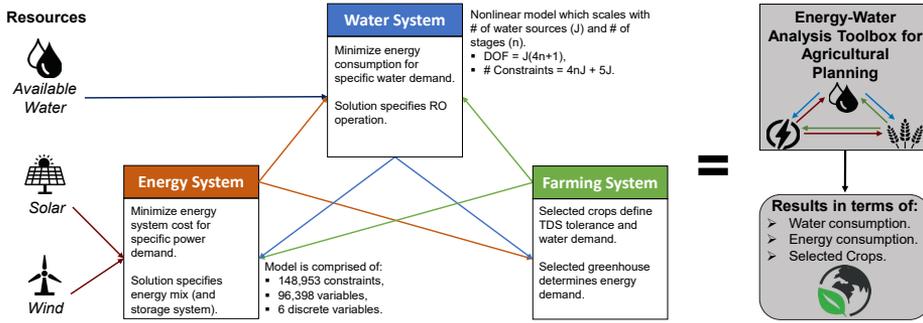


Figure 1: Visualization of the FEWN framework, the interconnected resource systems and their interactions.

3. Mathematical Modeling

The farming system determines the necessary water generation ($\frac{m^3}{day}$), the TDS constraint ($\frac{mg}{L}$), as well as the energy required to operate the greenhouse. For tomatoes for example the RO permeate concentration should not be higher than $1920 \frac{mg}{L}$ (or $3 \frac{mS}{cm}$). The daily irrigation requirement can be derived as specified in Eq. (1) or based on the average yearly water demand of the specified crop type in the respective farming environment. Regarding the required energy for the greenhouse operation the daily electricity consumption over a ten month period of a fully closed pilot greenhouse study in Qatar conducted by Hassad Food, Qatar Fertiliser Company and Yara International ASA is consulted (see section 4). Thus, the framework requires greenhouse specific information, i.e. the necessary energy for successful operation (kWh), and crop specific information, i.e. the TDS tolerance and the water demand.

$$Q_{sum}(\frac{L}{day}) \geq \frac{Yield(\frac{kg}{m^2}) \cdot Area(m^2) \cdot Water Used(\frac{L}{kg})}{Days of Operation} \quad (1)$$

The water supply system is based on the detailed mathematical RO desalination model as defined by Di Martino et al. (2021). The model optimizes for an array of objectives the necessary pressurization, the water recovery, the membrane surface area and the parallel flows for each stage and available water source, as well as the feed flow for each available water source. The employed model is nonlinear and scales with the number of water sources (J), as well as number of RO stages under investigation n , resulting in $J \cdot (4 \cdot n + 1)$ degrees of freedom, together with $4 \cdot n \cdot J + 5 \cdot J$ number of constraints. In this case the system design, i.e. number of stages and operational parameters, are based on the irrigation requirements of the crops under investigation. The key points of the selected water supply model are that a detailed mathematical model is employed which incorporates constraints based on the crops selection to minimize the system's energy consumption by adjusting the operational parameters of the RO process.

The overall energy demand is driven by the RO and greenhouse system. To fulfill the specified demands a combined mixed-integer linear model (MILP) is consulted to determine the optimal mix of solar panels and wind turbines based on supplied renewable energy data as described by Cook et al. (2022). The overall energy supply tool is summarized in Figure 2 and is comprised of 148,953 constraints, 96,398 variables, together with 6 discrete variables for a one year time horizon. The combined model determines a cost optimal energy and storage system to satisfy a

specified power demand over a predefined time horizon based on detailed mathematical modeling. To this end, solar direct normal irradiance (DNI) and wind speeds are preprocessed to obtain correlations of the total annualized cost vs. the desired power output for fixed angle and single axis tracking solar panels, as well as for wind turbines. On the basis of these surrogates together with the supplied climate data, the combined MILP model determines how the energy demand will be fulfilled. For the study conducted here, biomass as a possible energy resource has been disregarded.

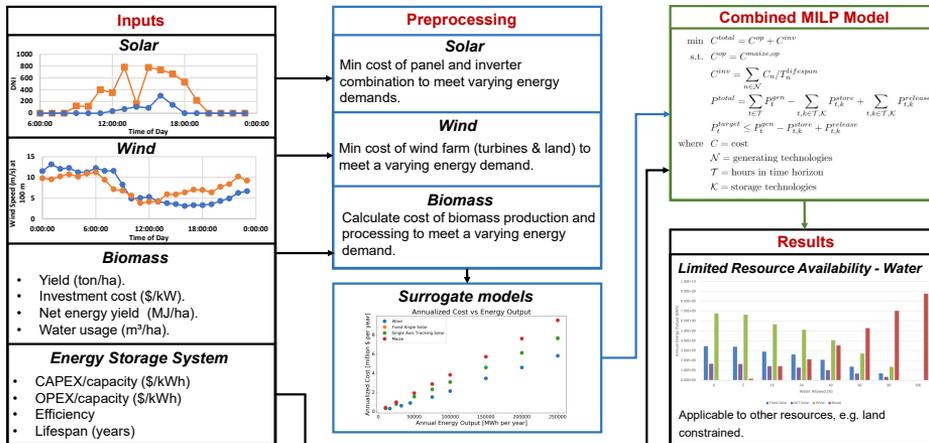


Figure 2: FEWN framework - The energy system.

Overall, this results in a mixed-integer programming model which can minimize the total system costs, the used energy and maximize the water usage for a specified energy-water-food mix and predefined time horizon, i.e. one year. Therefore, trade-offs and synergies between water and energy supply systems can be evaluated. Furthermore, by updating employed constraints, i.e. producing more water in the RO system, the overall optimization model can be modified according to the ϵ -constraint method enabling multi-objective optimization.

4. Case Study

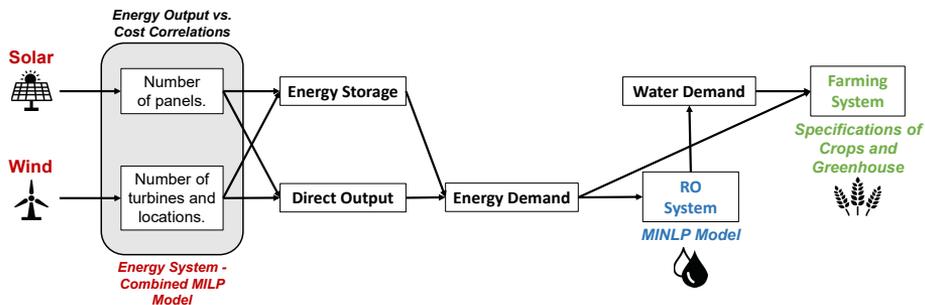


Figure 3: Simplified FEWN framework superstructure - Key features.

To illustrate the capabilities of the derived tool, we applied the simplified FEWN superstructure

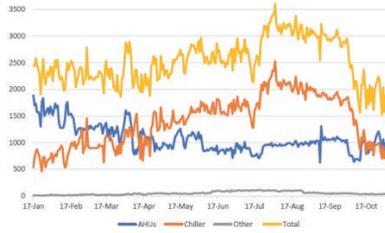


Figure 4: Daily consumption of electricity by components of the system, $\frac{kWh}{day}$, Qatar greenhouse trial 2020 (AHUs - Air handling units).

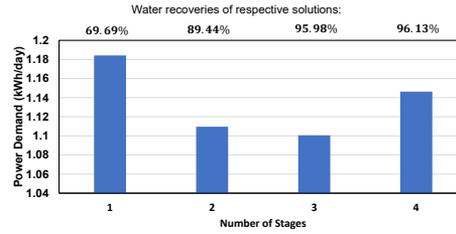


Figure 5: Comparison of results of RO operation optimization for minimizing energy consumption for different number of RO stages.

as shown in Figure 3 to a glass greenhouse using traditional farming practices to farm tomatoes ($C_p \leq 1900 \frac{mg}{L}$) on a $800 m^2$ farming area. Using the specified annualized yield of the previously mentioned pilot greenhouse in Qatar together with the specification of how much water per kg of produced tomatoes was used results in a constant daily water demand of $5 \frac{m^3}{day}$. The energy consumption of the greenhouse is assumed to be the same as the data provided in Figure 4.

Furthermore, it is assumed that only brackish water with a TDS concentration of $2560 \frac{mg}{L}$ is available. Additionally, the water recovery of the RO system is restricted to be at least 60%. The energy supply system has to provide energy for the RO process, as well as for the greenhouse farming system. The solar DNI and wind speed data have been obtained from the National Renewable Energy Laboratory (NREL). Moreover, the time horizon of this study has been fixed to one year. These specifications result in a three stage RO system in which the concentrate of a previous stage is used as the feed of a successive stage with an overall permeate output of $Q_{P,SUM} = 5 \frac{m^3}{day}$, an overall permeate concentration of $C_{P,SUM} = 1899 \frac{mg}{L}$ and a system water recovery of $WR = 95.98\%$ with a minimum overall energy requirement of the system of $1.1 \frac{kWh}{day}$ or $0.22 \frac{kWh}{m^3}$. The detailed results and operational parameters for each stage are summarized in Table 1. Further, in Figure 5 the results of minimizing the energy of the system for changing number of stages together with the respective water recoveries is shown. It can be deduced that indeed a three-stage system results in the lowest energy consumption closely followed by a two-stage process. One and four stages require the highest energy consumptions. In all cases the water recovery is comparably high, especially compared to the enforced lower bound water recovery restriction of 60%. After further analysis, i.e. minimizing the energy for step-wise increases of a newly introduced feed flow restriction while the energy requirement and water recovery of the system are monitored, it can be deduced that for this scenario the lowest possible feed flow is most advantageous for the energy consumption of the system, which results in high water recoveries. It is suggested that the energy for pressurization throughout the system is reduced by reducing the feed volume flow which offsets the possible energy recovery after the last stage.

Table 1: Results of RO operation optimization for minimizing energy consumption. m - parallel flows per stage.

Stage	$Q_f [m^3/d]$	$C_f [mg/L]$	$P [bar]$	$WR [%]$	$C_p [mg/L]$	$C_r [mg/L]$	m
Stage 1	5.21	2560	5	70.00	1392	5284	1
Stage 2	1.56	5284	5	62.99	2532	9968	1
Stage 3	0.58	9968	5	63.76	5221	18320	3

Based on the energy requirements of the RO system and the greenhouse, the energy supply system has to fulfill an overall power demand of $0.7 \cdot 10^6 \frac{kWh}{year}$ or alternatively an average demand of $80kW$. The results of minimizing the system's cost for these two cases are summarized in Table 2. In all cases only solar power as a generating technology is selected. Further, the first case does not require an energy storage system, whereas the second case requires a storage system with a capacity of $20300kWh$ and power rating of $530kW$. These evaluated cases underline the possible scenario analyses of the energy supply system, as well as the influence of renewable energy availability, together with the importance of the correct sizing of the energy generating technologies.

Table 2: Results of energy supply optimization for minimizing the total system cost to satisfy varying power demands. FA - Fixed angle solar panels, SAT - Single axis tracking solar panels.

Power demand	Power Output FA	Power Output SAT	Total Area	Cost
$0.7 \cdot 10^6 \frac{kWh}{year}$	$700000 \frac{kWh}{year}$	$0 \frac{kWh}{year}$	1.4 ha	$165500 \frac{\$}{year}$
$80kW$	$760250 \frac{kWh}{year}$	$99470 \frac{kWh}{year}$	1.72 ha	$284890 \frac{\$}{year}$

5. Conclusion

An energy-water decision tool for food production has been presented. The generic model has been applied to a greenhouse case study in Qatar. For this case the cost and energy optimal energy and water supply system has been determined for tomato farming. Future work will focus on extending the framework to (i) dynamic planning of irrigation schedules considering dynamic RO production schedules based on seasonality, and (ii) detailed greenhouse modeling to analyze and compare farming systems for varying agricultural scenarios.

6. Acknowledgements

This publication was made possible by the National Priorities Research Program (NPRP) grant No NPRP11S-0107–180216 from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work, and are solely the responsibility, of the authors. Further, the authors would like to thank the Texas A&M Energy Institute for its support.

References

- S. Avraamidou, A. Milhorn, O. Sarwar, E. N. Pistikopoulos, 2018. Towards a quantitative food-energy-water nexus metric to facilitate decision making in process systems: A case study on a dairy production plant. In: Computer Aided Chemical Engineering. Vol. 43. Elsevier, pp. 391–396.
- J. Cook, M. D. Martino, R. C. Allen, E. N. Pistikopoulos, S. Avraamidou, 2022. A decision-making framework for the optimal design of renewable energy systems under energy-water-land nexus considerations. Under Review.
- M. Di Martino, S. Avraamidou, J. Cook, E. N. Pistikopoulos, 2021. An optimization framework for the design of reverse osmosis desalination plants under food-energy-water nexus considerations. Desalination 503, 114937.
- S.-H. Lee, J.-Y. Choi, S.-O. Hur, M. Taniguchi, N. Masuhara, K. S. Kim, S. Hyun, E. Choi, J. hoon Sung, S.-H. Yoo, 2020. Food-centric interlinkages in agricultural food-energy-water nexus under climate change and irrigation management. Resources, Conservation and Recycling 163, 105099.
- M. Li, V. P. Singh, Q. Fu, D. Liu, T. Li, Y. Zhou, 2021. Optimization of agricultural water–food–energy nexus in a random environment: an integrated modelling approach. Stochastic Environmental Research and Risk Assessment 35 (1), 3–19.
- Y. Nie, S. Avraamidou, X. Xiao, E. N. Pistikopoulos, J. Li, Y. Zeng, F. Song, J. Yu, M. Zhu, 2019. A food-energy-water nexus approach for land use optimization. Science of The Total Environment 659, 7–19.
- B. Rahimi, M. Afzali, F. Farhadi, A. A. Alamolhoda, 2021. Reverse osmosis desalination for irrigation in a pistachio orchard. Desalination 516, 115236.
- S. H. Sadeghi, E. Sharifi Moghadam, M. Delavar, M. Zarghami, 2020. Application of water-energy-food nexus approach for designating optimal agricultural management pattern at a watershed scale. Agricultural Water Management 233, 106071.

Optimizing the allocation of resources for the security of the water-energy-food nexus

Brenda Cansino-Loeza,^a Aurora del Carmen Munguía-López,^a José María Ponce-Ortega^{a*}

^a*Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Francisco J. Mujica S/N, Ciudad Universitaria, Morelia, Michoacán 58060, México*

Abstract

Increasing demands and climate change are critical challenges to ensure water, energy, and food security. The water-energy-food nexus requires integrating tools to guide the allocation of resources and promote sustainability. This work presents a mathematical formulation for the optimal design and management of resources to enhance the water-energy-food nexus security. Resource security is measured through indicators related to the availability, access, and sustainability of the water, energy, and food sectors. Furthermore, the problem was analyzed under different allocation schemes (social welfare, Rawlsian, Nash, and Rawlsian-Nash) to maximize resource security and obtain the optimal design of the system. To show the applicability of the model, a Mexican state evaluated by regions was selected as a case study. Results show that through this approach, it is possible to increase 14%, 44%, and 15% the security of the water, energy, and food sectors, respectively, and 25% security of the water-energy-food nexus in the addressed case study. The proposed framework can be applied to any region with the corresponding data.

Keywords: Resource security; Allocation schemes; Sustainability; Optimization; Water-energy-food nexus.

1. Introduction

The water-energy-food (WEF) nexus introduced in the Bonn Conference (Hoff, 2011) has become a relevant topic to promote the study of synergies in the water, energy, and food sectors to achieve sustainable development. In the last years, growing resource demands and limited access and availability of resources have caused increasing concern about water, energy, and food security. Due to the multidimensional nature of the WEF nexus, decision-making methods have been used to address the challenges that the nexus faces and new approaches have been proposed to enhance the synergies between resources and face climate change events and globalization using quantitative analysis methods (Zhang et al., 2018; Radini et al., 2021). In this context, optimization models have been proposed for the sustainable design for resource management, the planning of the WEF nexus involving multiple objectives in conflict, and the performance of the WEF nexus when policies are incorporated (Ogbolumani and Nwulu, 2021; Chamas et al., 2021; Sušnik et al., 2021). However, despite the progress made towards the WEF nexus, there is still a relative lack of methods that allow quantifying, optimizing, and evaluating the nexus security in the planning and design of a WEF system.

It has been highlighted that a driving factor to achieve long-term sustainability is the security of the WEF nexus, therefore, the planning, distribution, and resource management policy should be improved. The allocation of resources is one of the main concerns in social planning. Typically, when optimization models are solved, resources allocation is guided by the maximization of the sum of the player's utilities or objectives in interest. However, it has been proved that solving problems based on this approach may lead to non-unique solutions since different allocations of resources can give the same total utility of the system (Sampat and Zavala, 2019). Therefore, some allocation or distribution schemes have been reported and have been associated with fairness measures. Through these distribution schemes, it is possible to solve the problem through different perspectives and find solutions that capture the scales of stakeholders or objectives in interest and that are unique. Nonetheless, fairness measures related to allocation schemes in integrated WEF systems have not been addressed. Therefore, the novelty of this work is the development of a model formulation for the integration of the WEF nexus involving indicators to measure the WEF security nexus, where the optimal distribution of resources is evaluated through different allocation schemes.

2. Problem statement

Water-energy-food nexus security has been threatened by climate change impacts and increasing resource demands. Allocation and distribution of resources are important factors in the decision-making for planning WEF nexus systems. In this context, several efforts have been made to explore synergies between resources and improve the management of these sectors. However, there are required tools to quantify resource security involving the optimal resource allocation to improve WEF nexus security and achieve sustainable development. In this sense, this work presents a mathematical model of a macroscopic water-energy-food nexus integration for a set of regions that involves water, energy, and food availability, accessibility, and sustainability indicators to quantify the WEF nexus security. The developed model is based on the superstructure (see Figure 1) that indicates the possible alternatives to integrate water, energy, and food resources.

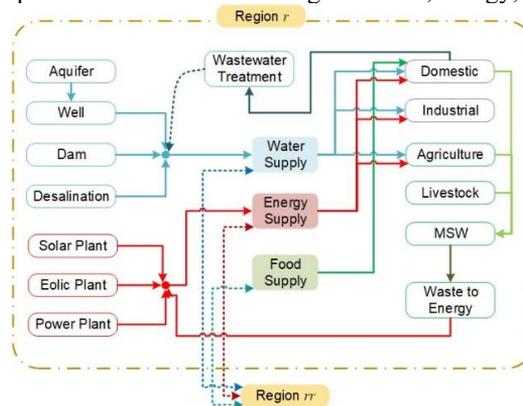


Figure 1. General Superstructure.

3. Water- energy-food nexus optimization framework

The mathematical model is composed of material balances and constraints to determine the existence of the technologies proposed in the superstructure in order to integrate the resources properly. Security indices related to water (I^{Water}), energy (I^{Energy}) and food (I^{Food}) are included, and they are based on availability ($I^{Availability}$), accessibility ($I^{Accessibility}$)

and sustainability ($I_{Sustainability}$) of resources. The average of these indicators equals the WEF nexus security index (I^{WEF}) which is maximized according to different allocation schemes.

3.1. Objective functions: Allocation schemes

A suitable way to generate solutions that can help decision-making and address a deemed fair resource distribution is through the implementation of allocation schemes (social welfare (SW), Rawlsian (R), Rawlsian Nash (RW-N), and Nash (N)).

The social welfare scheme is commonly used to allocate resources by maximizing the sum of the utilities of stakeholders (Equation 1). However, solutions including multiple allocations of resources could be found since this scheme may provide non-unique solutions, and thus, this method causes ambiguity. Furthermore, it cannot capture the scales of the stakeholders involved in the problem. For purposes of this work, a new function to solve the social welfare scheme is introduced, which consists in maximizing the sum of the water, energy, and food security indexes:

$$\phi^{sw} = I^{water} + I^{energy} + I^{food} \quad (1)$$

The Rawlsian scheme is an alternative to allocate resources; the principle of this scheme is to maximize the smallest utility of the stakeholders. One of the disadvantages of the Rawlsian scheme is that it does not capture scales properly; therefore, large stakeholders may be ignored. The formulation of the Rawlsian scheme for this problem consists in including a new variable (Θ) that is minimized:

$$\phi^{rw} = \Theta \quad (2)$$

$$-I^{water} \leq \Theta \quad (3)$$

$$-I^{energy} \leq \Theta \quad (4)$$

$$-I^{food} \leq \Theta \quad (5)$$

In the Nash scheme, it is possible to capture the scales of the stakeholders since a logarithm function is used. The Nash allocation generates a unique solution that is obtained by maximizing the sum of the logarithms of the water, energy, and food security indexes.

$$\phi^n = \ln(I^{water}) + \ln(I^{energy}) + \ln(I^{food}) \quad (6)$$

Finally, a combination of the Rawlsian and Nash schemes is included. The formulation of this scheme is similar to the Rawlsian scheme, but in this case, the constraints are modified to include the logarithm function, and another one is included to represent the Nash scheme:

$$\phi^{rw-n} = \Theta \quad (7)$$

$$-\ln(I^{water}) \leq \Theta \quad (8)$$

$$-\ln(I^{energy}) \leq \Theta \quad (9)$$

$$-\ln(I^{food}) \leq \Theta \quad (10)$$

$$-\left(\ln(I^{water}) + \ln(I^{energy}) + \ln(I^{food})\right) \leq \Theta \quad (11)$$

4. Case study

The Mexican state of Sonora was selected as a case study to demonstrate the applicability of the model. Sonora has been one of the most important economic entities in Mexico because it has a diversity of natural resources that facilitate the development of economic activities such as agriculture, livestock, fishing, mining, and services. Nevertheless, the

availability and distribution of resources in Sonora are unequal due to its geographical and hydrological conditions. The northern of the state is covered by the Sonoran Desert, which is the hottest desert in the country, while the south of the state has potential for agriculture activities. However, providing access to water for the different economic sectors represents a great challenge. Furthermore, considering that Sonora has potential for renewable energy, the design of an integrated WEF system is addressed to increase the security of the WEF.

In this work, the Sonora state was divided into four representative regions to evaluate the security of the WEF Nexus; the border region (R1), center region (R2), central-mountain range region (R3), and southern region (R4). **Table 1** presents the water, energy, and food indices associated with the availability, accessibility, and sustainability of resources to date. Currently, the central-mountain range region (R3) presents the lowest water availability, while the center region (R2) presents the highest water availability of the state; therefore, water from R2 is used to help satisfy water requirements in the near regions. On the other hand, it is shown that energy production is centered in R2, and R3 exhibits the lower energy production of the state.

Table 1. Current water-energy-food nexus security indices.

	Water			Energy			Food		
	$I^W_{Availability}$	$I^W_{Accessibility}$	$I^W_{Sustainability}$	$I^E_{Availability}$	$I^E_{Accessibility}$	$I^E_{Sustainability}$	$I^F_{Availability}$		
R1	0.465	0.981	0.465	0.400	0.988	0.400	0.751	I^{Water}	0.836
R2	2.494	0.991	1.000	2.171	0.978	2.171	0.751	I^{Energy}	0.814
R3	0.030	0.971	0.030	0.140	0.980	0.140	0.751	I^{Food}	0.751
R4	0.823	0.958	0.823	0.215	0.970	0.215	0.751	I^{WEF}	0.800

The Sonora's regions present unequal distribution of resources availability and production. It can be seen that R2 presents the highest water and energy production, which is associated with its industrialized nature. On the other hand, R3 corresponds to the least urbanized region of the state, therefore, technologies for energy production are not abundant. In addition, it is important to mention that this region does not have access to the sea, and the existence of desalination plants that increase the water production in the region is not an option. R2 has the highest availability, accessibility, and sustainability indicators of the regions of the state. Nevertheless, it is important to notice that the potential of Sonora for renewable energy can be exploited, and increasing the renewable energy capacity could improve security in the energy sector. In the same way, water reuse and irrigation techniques could enhance the security of the WEF substantially.

5. Results

The mathematical formulation for the analyzed fair allocation schemes was implemented in the software GAMS. The models of the schemes correspond to a Mixed Integer Linear Programming Model (MINLP) and were solved using the solver LINDOGlobal. **Table 2** shows the costs, freshwater consumption, and GHGE generated in the different allocation schemes. The social welfare scheme presents the lowest GHGE generation, but also it is the scheme with the highest costs and freshwater consumption. On the other hand, the Rawlsian scheme presents the lowest cost for the system and the lowest freshwater abstraction.

Table 2. Economic and environmental aspects of the allocation schemes.

	Social welfare	Rawlsian	Nash	Rawlsian-Nash
--	----------------	----------	------	---------------

WATER (hm³)	6,135.56	5,471.93	5,363.63	5,511.55
COST (MMUSD)	27,212.2	26,768.6	27,104.6	26,948.1
GHGE x 10³ (Ton CO₂)	20.864	24.103	20.95	24.523

The results of the security index for water, energy, and food for the social welfare, Rawlsian, Nash, and Rawlsian-Nash schemes are presented in **Table 3**.

Table 3. WEF Security Index for the schemes analyzed.

	Social welfare	Rawlsian	Nash	Rawlsian-Nash
I^{Water}	0.957	0.870	0.853	0.869
I^{Energy}	1.177	0.886	1.164	1.157
I^{Food}	0.869	0.870	0.825	0.869
I^{WEF}	1.001	0.875	0.947	0.965
Water				
$I^W_{Availability}$	1.086	0.839	0.774	0.831
$I^W_{Accessibility}$	1	1	1	1
$I^W_{Sustainability}$	0.784	0.772	0.784	0.775
Energy				
$I^E_{Availability}$	1.383	0.942	1.364	1.353
$I^E_{Accessibility}$	0.987	0.974	0.987	0.987
$I^E_{Sustainability}$	1.16	0.741	1.141	0.775
Food				
$I^F_{Availability}$	0.757	0.761	0.674	0.757
$I^F_{Accessibility}$	1	1	1	1
$I^F_{Sustainability}$	0.849	0.849	0.802	0.849

The results obtained in the social welfare scheme present the highest WEF Security Index (1.001) compared with the other schemes. The water, energy, and food security indexes were higher than those of the other schemes. In the case of the water sector, the availability index was approximately 29% higher than in the other schemes, this is mainly attributed to the water production in R2, which exceeds the demand and can be exported to other regions of the state to cover the water requirements. Similarly, in the case of the energy sector, the security index is also the highest in comparison with the other schemes, energy production in R2 is 3.33 times the energy demand in the region, and it is exported to other regions to cover their energy demands. In this scheme, even the availability of energy exceeds the demand of the region, the model determines that the energy is distributed to other regions without access to energy rather than to provide access to all the population in R2. On the other hand, in the food sector, a value of 1 was obtained for the food accessibility index, which indicates that all population has access to a balanced diet where food is supplied by the associated region and by the imported food.

The results of the system design for the Rawlsian scheme show a WEF security index of 0.875, which corresponds to the lowest security index resulting from the evaluated schemes. This is mainly attributed to security in the energy sector, which is around 30% lower than the energy security index of the other schemes. Specifically, energy production in region 2 (R2) is equivalent to 2 times the demanded energy in the region, while for the other schemes the energy production capacity was 3.33 times the energy consumed. The WEF security index in the Rawlsian-Nash scheme is equal to 0.965, and it is 3% lower than the WEF security index resulting from the social welfare scheme. However, it is important to mention that it is possible to obtain a unique solution by using the Rawlsian-Nash formulation, unlike the social welfare scheme. For the Nash scheme, the accessibility to water, energy, and food is completely covered for all the population as in the social welfare scheme. In the water sector, it is observed that R2 does not produce

more water than the demand in the region, therefore, there are no internal water imports between the regions.

6. Conclusions

This paper has presented a water-energy-food optimization framework to maximize the security of the nexus considering different allocation schemes, which supports the decision-making for better management of resources and performance of the sectors. As a case study, regions of the state of Sonora from Mexico were used to implement the proposed framework. By the application of the proposed approach, and comparing the results with the current water-energy-food nexus security indices, results show an increase in the security index of the water sector of 14%, 44% of the energy sector, and 15% of the food sector. This gives an increment of 25% in the WEF security index. Results for the design of the Rawlsian scheme present the lowest costs (\$26,768.6 MMUSD) and the lowest freshwater consumption (5471.93 hm³) among the different allocation schemes. Moreover, this scheme presents the lowest WEF security index of 0.875. On the other hand, the social welfare scheme presents a WEF security index of 1.001, in addition, its design presents the lower GHGE (20,864 ton CO₂) but the highest freshwater abstraction (6135.56 hm³) and costs (6,135.56 MMUSD). The main contribution of the proposed framework is the development of a model that provides the optimal allocation of water, energy, and food and quantifies and maximizes the security of the water, energy, and food sectors. The proposed water-energy-food nexus framework can be applied to any region at any scale with the corresponding data. However, its main limitation is the uncertainty related to the lack of time dimension.

References

- A.M. Sampat, V. Zavala, 2019. Fairness measures for decision-making and conflict resolution. *Optim. Eng.* 2019 204 20, 1249–1272.
- H. Hoff, 2011. Background paper for the Bonn 2011 Nexus Conference: The Water, Energy and Food Security Nexus. https://www.water-energy-food.org/uploads/media/understanding_the_nexus.pdf (accessed 3.4.20).
- J. Sušnik, S. Masia, D. Indriksone, I. Brēmere, L. Vamvakeridou-Lydroudia, 2021. System dynamics modelling to explore the impacts of policies on the water-energy-food-land-climate nexus in Latvia. *Sci. Total Environ.* 775, 145827.
- J. Zhang, P.E. Campana, T. Yao, Y. Zhang, A. Lundblad, F. Melton, J. Yan, 2018. The water-food-energy nexus optimization approach to combat agricultural drought: a case study in the United States. *Appl. Energy* 227, 449–464.
- O.A. Ogbolumani, N.I. Nwulu, 2021. Multi-objective optimisation of constrained food-energy-water-nexus systems for sustainable resource allocation. *Sustain. Energy Technol. Assess.* 44, 100967.
- S. Radini, E. Marinelli, Ç. Akyol, A.L. Eusebi, V. Vasilaki, A. Mancini, E. Frontoni, G.B. Bischetti, C. Gandolfi, E. Katsou, F. Fatone, 2021. Urban water-energy-food-climate nexus in integrated wastewater and reuse systems: Cyber-physical framework and innovations. *Appl. Energy* 298, 117268.
- Z. Chamas, M. Abou Najm, M. Al-Hindi, A. Yassine, R. Khattar, 2021. Sustainable resource optimization under water-energy-food-carbon nexus. *J. Clean. Prod.* 278, 123894.

Application of CAPE Tools into Prospective Life Cycle Assessment: A Case Study in Recycling Systems Design of Lithium-Ion Battery

Yasunori Kikuchi,^{a,b,c*} Aya Heiho,^b Yi Dou,^b Izuru Suwa,^c Chiharu Tokoro^{d,e}

^a *Institute for Future Initiatives, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan*

^b *Presidential Endowed Chair for “Platinum Society”, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

^c *Department of Chemical System Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

^d *Department of Resources and Environmental Engineering, Faculty of Science and Engineering, Waseda University, 3-4-1 Okubo Shinjuku-ku, Tokyo 169-8555, Japan*

^e *Department of Systems Innovation, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

Abstract

In this study, we are tackling systems design with assessments for emerging technologies. Because of the data limitation on the systems and processes adopting emerging technologies, computer-aided process engineering (CAPE) tools such as process design heuristics, process simulation, optimization, parametric analysis for characterizing sensitivity and alternative generation, and decision making with uncertainties have huge potential to compensate such data limitation and jump up to the deep technology assessments with quantified results. In this paper, we examine the applicability of CAPE tools for systems design and assessment adopting emerging technologies with a case study of recycling systems design of Lithium-ion battery (LiB).

Keywords: Spent lithium-ion batteries, Recycling, Positive electrode active materials.

1. Introduction

Recently, Lithium-ion batteries (LiBs) have been widely utilized not only for electronic devices such as mobile phones and laptops, but also for electric vehicles and stationary energy storage systems. Because this increase in LIB demand causes significant increases in demand for certain materials, i.e., lithium (Li), cobalt (Co), nickel (Ni), and manganese (Mn), it is necessary to secure these resources by recycling, as well as through mine developments. In recent years, there has been an increase in the number of case studies using prospective life cycle assessment (LCA), which take into account the future potential of the technology and aim to predict the environmental impacts on the technology under development (Arvidsson et al., 2018; Moni et al., 2020; Thonemann et al., 2020).

In this study, we are tackling systems design with assessments for emerging technologies. Because of the data limitation on the systems and processes adopting emerging technologies, computer-aided process engineering (CAPE) tools such as process design heuristics, process simulation, optimization, parametric analysis for characterizing sensitivity and alternative generation, and decision making with uncertainties have huge

potential to compensate such data limitation and jump up to the deep technology assessments with quantified results. In this paper, we examine the applicability of CAPE tools for systems design and assessment adopting emerging technologies with a case study of recycling systems design of Lithium-ion battery (LiB). The recycling of cathode particles and aluminum (Al) foil from positive electrode sheet (PE sheet) dismantled from spent LiBs was experimentally demonstrated by applying a high-voltage pulsed discharge (Tokoro et al., 2021). This separation of LIB components by pulsed discharge was examined by means of prospective LCA (Kikuchi et al., 2021). The indicators selected were life cycle greenhouse gas (LC-GHG) emissions and life cycle resource consumption potential (LC-RCP). We first completed supplementary experiments to collect redundant data under several scale-up circumstances, and then attempted to quantify the uncertainties from scaling up and progress made in battery technology. When the batch scale of pulsed discharge separation is sufficiently large, the recovery of cathode particles and Al foil from PE sheet by pulsed discharge can reduce both LC-GHG and LC-RCP, in contrast to conventional recycling with roasting processes.

2. Materials and methods

2.1. Application of CAPE tools into prospective LCA

Figure 1 shows the description of systems assessments applying CAPE tools for prospective LCA. In management activity and resource provider, data estimation and interpretations are assigned to CAPE tools considering the conditions in prospective LCA.

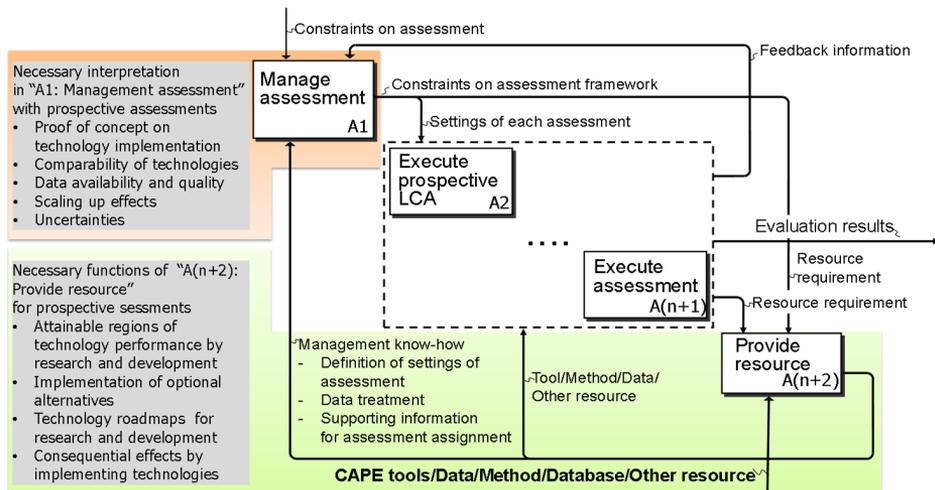


Figure 1 Description of multiple assessment activities with the necessary conditions for prospective assessments including LCA. The parameter n is the number of assessment methods. (Modified from previous studies (Kikuchi, 2014; Kikuchi et al., 2010; Kikuchi and Hirao, 2009))

Conventional LCA does not take into account changes in technology level, because it refers to information on the current technology level and specifically estimates the environmental impacts of each process related to the provision of products and services. On the other hand, efforts to tackle climate change have become more active in recent years, and new products and technologies are changing concepts and models more rapidly, making the transition to a low-emission society more urgent.

The significance of conducting a strategic LCA of emerging technologies for the 30-year time horizon up to the target year of 2050 arose regarding the issues on the climate change.

Emerging technologies, as defined by Rotolo et al. (2015), are; “innovative and rapidly growing technologies that have the potential to have a significant social and economic impact in the domains in which they are structured, with some degree of persistent coherence, actors, institutions, ways of interacting with them and related knowledge production processes. It is characterised by its potential to have significant social and economic impacts. However, its most prominent impact lies in the future and is therefore somewhat uncertain and ambiguous at the stage at which the technology emerges.” These technologies are characterized as “innovative”, “rapid growth”, “consistent”, “significant impact” and “uncertain”, which makes technology assessment difficult due to lack of existing data and knowledge.

Four main issues were identified as needing to be addressed in conducting prospective LCAs of emerging technologies (Thonemann et al., 2020; Moni et al., 2020). (1) comparability of technologies; (2) availability and quality of data; (3) scale-up challenges; and (4) uncertainty of assessment results. Process modeling and simulation are effective in estimating the missing process inventories in industrial scale production, because these technologies are under development in lab or pilot scale.

2.2. Case study: Recycling Systems Design of Lithium-Ion Battery

Figure 2 shows the life cycle model of conventional and the proposed alternative recycling systems. The dashed-line boxes are not included in the life cycle inventory analysis. (M-SO₄: metal sulfate, i.e., CoSO₄, NiSO₄, and MnSO₄; positive electrode active material (PEAM): Li(Ni_xCo_yMn_z)O₂, raw materials of positive electrode active material (rawPEAM): (Ni_xCo_yMn_z)(OH)₂, and R- (recycled-). The functional unit of this LCA was defined as the use of a positive electrode as a component of LIBs for vehicles. The LIB components assessed in this study are provided in the previous study (Kikuchi et al., 2021). The pulsed discharging process applies a novel technology to separate cathode particles, i.e., cathode black powder (Co and Ni) from Al foil (Tokoro et al., 2021). The technology readiness level (TRL) of this technology is the lab-scale demonstration, where the process inventory data required for LCA was not sufficiently obtained from the experimental demonstration considering the upscaling of throughputs of pulsed discharging. For such technology, modeling and simulation can be employed to fill the gap of foreground data (Tsoy et al., 2020). As for the prospective thinking for LIBs recycling applying the novel pulsed discharging separation process, three types of uncertainties should be addressed by the application of CAPE tools.

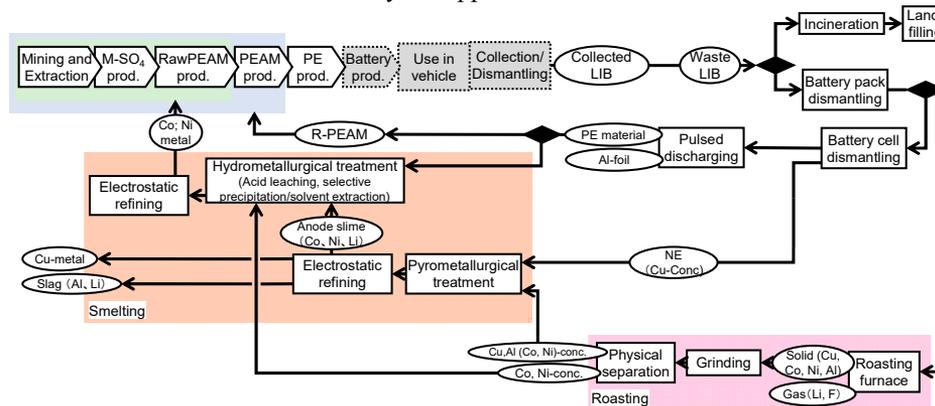


Figure 2 Life cycle model of conventional and alternative case settings to compare recycling processes for LIB and to consider the shift of elemental composition in positive electrode active materials. (Modified from the previous study (Kikuchi et al., 2021))

2.2.1. Settings on LCA

The operating ratios of precise cell dismantling and high-voltage pulsed discharge separation should be carefully examined. Because the pulsed discharge device contains high capacitance capacitors, the lifetime throughput is highly sensitive to the life cycle impacts (Kikuchi et al., 2021). The elemental composition in PEAM in LiB cathodes has changed with technological development. While PEAM of LiBs, which has been used for several years in the market, can be reused, the technology of the reused PEAM may become outdated during its period of use.

2.2.2. Application of CAPE tools

The process system combining precise cell dismantling and high-voltage pulsed discharge separation is semi-batch system. The residence times of contained unit operations were measured, and then the number of equipment items for cell dismantling and pulsed discharge was optimized. To improve the performance of the process system, the scale of pulsed discharge was re-designed and experimentally demonstrated to reduce the takt time of the treatment of unit LiB module. If the lifetime throughputs could be increased sufficiently, the environmental loads induced by the initial manufacturing of machines could become small due to the allocation of such initial environmental loads to the throughputs.

The increase of the lifetime throughputs determined by the processing scale of precise cell dismantling and pulsed discharge can reduce the environmental impact of equipment manufacturing per throughput. On the other hand, since the number of LiBs depends on the number of vehicles produced and disposed of in the market, it is not possible to determine whether the processing scale is necessarily sufficient. This could be considered by dynamic material flow analysis to forecast the future availability of spent LiBs from markets by assuming the shipments and wastes of products (Kikuchi et al., 2014; 2021).

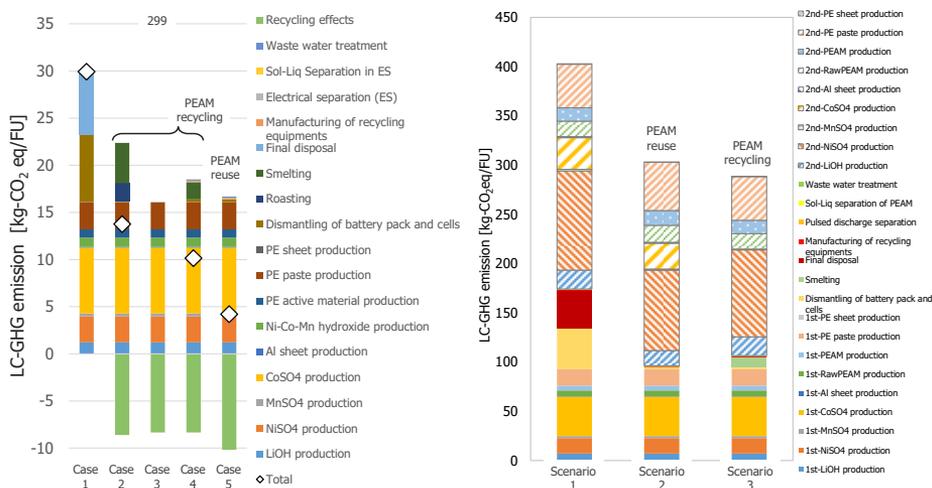
3. Results and discussions

Figure 3 shows the results of prospective LCA applying CAPE tools. The same tendencies were also observed for LC-RCP (Kikuchi et al., 2021), where (a) Results of LCA for treatment of unit amount of PE sheet, where Case 1 treats spent LiBs in incineration and land filling, Case 2 is roasting and smelting, Cases 3 and 5 are the recycling of PEAM applying pulsed discharging with low and high lifetime throughputs, respectively, and Case 4 is the reuse of PEAM applying pulsed discharging. (b) Results of LCA for scenarios with a shift in elemental composition in PEAM, where the NCM111, i.e., $\text{Li}(\text{Ni}_{1/3}\text{Co}_{1/3}\text{Mn}_{1/3})\text{O}_2$, cathode is first manufactured using 1 kg-Co in all scenarios and is then used in the LIB. In Scenario 1, all products are incinerated or landfilled (Case 1). In Scenario 2, 80% of the LIB is recovered and reused as PEAM for the NCM111 (Case 5). In Scenario 3, rawPEAM is recycled from the recovered LIB (Case 4) and PEAM of NCM811 is produced. Since the total amount of LIB capacity in each scenario is different, the capacity in Scenario 3, which has the largest storage capacity among these scenarios, is used as the reference flow of the functional unit, i.e., 4.83 kWh, and the shortage in Scenarios 1 and 2 is compensated for by newly manufactured NCM811. The storage capacities for NCM111 and NCM811 were assumed as 0.16 kWh/kg-Cathode and 0.25 kWh/kg-Cathode, respectively. As shown in Figure 3(a), lifetime throughput is an essential variable to understand the scale of the process and to design the process and to be dominant factor to the total LC-GHG. The lifetime throughput varied with the treated positive electrode sheet length at pulsed discharge, residence time of each item of equipment, and adjustment of the number of items of equipment for cell dismantling and pulsed discharge. The number of equipment items for cell dismantling and pulsed

discharge were optimized, considering the rate-determining step of their batch unit processes. This optimization is required for revealing the plausible effects of the newly implemented machines into society.

The development of the elemental composition in cathode particle resulted in the change of preferred recycling system. Fig. 3(a) shows that the reuse of PEAM had less environmental impacts than the recycling of rawPEAM. Comparing Figures 3(b), the LC-GHG of Scenario 3 were less than those of Scenario 2. This is because of the improvement of cathode capacity per Co use from NCM111 to NCM811. Even though the smaller circle recycling considered in Case 5 has less environmental impacts than the larger circle recycling in Case 4, as shown in Figure 3(a) for the same battery chemistry, the shift in elemental composition results in a different conclusion. As it is possible that the future trend of battery chemistry will be dominated by lithium nickel cobalt manganese oxide, or other battery chemistries, e.g., lithium iron phosphate, novel lithium-sulfur, or lithium-air (Xu et al., 2020), the need for lithium, nickel, cobalt, and manganese will remain for the next decades. Pulsed discharge separation can provide the option to recover such metal resources from spent LIBs. Based on the trends of battery chemistry, metal efficiency could be considered in the selection for reuse as PEAM or recycling of rawPEAM (Figure 3(b)) to meet future demand for batteries (Xu et al., 2020). Metal recycling with pulsed discharge separation can thus influence the direction of future technological developments of battery cathodes.

At present, there are more than 3,400 car dismantling plants in operation in Japan, and the scale of processing in Case 4 would be able to process about 1,000 spent LIBs per year; this would be sufficient to process about 3.4 million cars, assuming that the precise cell dismantling and pulsed discharge separation were installed in all dismantling plants. While the number of waste hybrid vehicle in Japan was estimated as from 0.11 million to less than 1 million, the uptake of low emission vehicles is promoted more and more recently as like European Union. Since the number of new vehicle registrations and sales in Japan in 2019 was about 5 million, the order of magnitude is about the same for processing.



(a) Treatment of unit amount of PE sheet

(b) Scenarios in composition in PEAM

Figure 3 Prospective assessments results on LC-GHG emission. (Modified from the previous study (Kikuchi et al., 2021))

4. Conclusion

CAPE tools can become methods applicable for acquiring data for prospective assessments. Prospective LCA should be applied into the technology assessment that employs modelling tools which focus on potential environmental impacts arising from various technologies even still at the R&D stage, i.e., low technology readiness level. With CAPE tools, the inventory data for prospective LCA can be connected with the design methods for optimizing the throughputs of unit operations, analyzing the upscaled process systems, and conducting the quantification of environmental loads with plausible process system design.

Through the case study on LiB recycling, it was demonstrated that the application of CAPE tools into prospective LCA enables the strategic technology assessments for systems design. Especially in the proof of concept on technology implementation can be verified and validated with the ranged values of uncertainties in emerging technology under development.

Acknowledgement

This work was supported by MEXT/JSPS KAKENHI Grant Number JP21H03660 and JST-Mirai Program Grant Number JPMJMI19C7. Activities of the Presidential Endowed Chair for “Platinum Society” at the University of Tokyo are supported by the KAITEKI Institute Incorporated, Mitsui Fudosan Corporation, Shin-Etsu Chemical Co., ORIX Corporation, Sekisui House, Ltd., the East Japan Railway Company, and Toyota Tsusho Corporation.

References

- R. Arvidsson, A.M. Tillman, B.A. Sandén, M. Janssen, A. Nordelöf, D. Kushnir, S. Molander, 2018, Environmental assessment of emerging technologies: recommendations for prospective LCA. *J Ind Ecol*, 22(6), 1286-1294.
- Y. Kikuchi, M. Hirao. 2009, Hierarchical Activity Model for Risk-Based Decision Making Integrating Life Cycle and Plant-Specific Risk Assessments, *J Ind Ecol*, 13(6) 945-964.
- Y. Kikuchi, K. Mayumi, M. Hirao, 2010, Integration of CAPE and LCA tools in environmentally-conscious process design: a case study on biomass-derived resin, *Comput. Aided Chem. Eng.*, 28, 1051-1056
- Y. Kikuchi, I. Suwa, A. Heiho, Y. Dou, S. Lim, T. Namihira, K. Mochizuki, T. Koita, C. Tokoro, 2021, Separation of cathode particles and aluminum current foil in lithium-ion battery by high-voltage pulsed discharge Part II: Prospective life cycle assessment based on experimental data, *Waste Manage.*, 132, 86-95.
- Y. Kikuchi. 2014. Activity and Data Models for Process Assessment Considering Sustainability, *Kagaku Kogaku Ronbunshu*, 40(3) 211-223.
- S.M. Moni, R. Mahmud, K. High, M. Carbajales-Dale, 2020, Life cycle assessment of emerging technologies: A review. *J Ind Ecol*, 24, 52–63.
- D. Rotolo, D. Hicks, B.R. Martin. 2015. What is an emerging technology? *Res. Policy*, 44(10), 1827-1843.
- N. Thonemann, A. Schulte, D. Maga, 2020, How to Conduct Prospective Life Cycle Assessment for Emerging Technologies? A Systematic Review and Methodological Guidance. *Sustainability*, 12(3), 1192.
- C. Tokoro, S. Lim, K. Teruya, M. Kondo, K. Mochizuki, T. Namihira, Y. Kikuchi, 2021, Separation of cathode particles and aluminum current foil in Lithium-ion battery by high-voltage pulsed discharge part I: Experimental investigation, *Waste Management*, 125, 58-66.
- N. Tsoy, B. Steubing, C. van der Giesen, J. Guinée. 2020. Upscaling methods used in ex ante life cycle assessment of emerging technologies: a review. *Int. J. Life Cycle Assess.* 25, 1680-1692.
- C. Xu, Q. Dai, L. Gaines, M. Hu, A. Tukker, S. Bernhard, 2020. Future material demand for automotive lithium-based batteries. *Commun. Mater.* 1, 99,

A Fair-Sustainable Approach for the Optimization of an Integrated Fuel Production System

Aurora del Carmen Munguía-López,^a Aurora de Fátima Sánchez-Bautista,^a
Mahmoud M. El-Halwagi,^b José María Ponce-Ortega^{a,*}

^a*Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Francisco J. Mujica S/N, Ciudad Universitaria, Morelia, Michoacán 58060, México*

^b*Artie McFerrin Department of Chemical Engineering, Texas A&M University, 386 Spence St, Jack E. Brown Engineering Building, College Station, Texas, United States, 77843*

jose.ponce@umich.mx

Abstract

To achieve sustainable development, certain objectives must be considered. A fair allocation of resources is key for long-term sustainability. Therefore, these objectives can include attaining environmental fairness and sustainability metrics. Besides considering the fair allocation of resources, equity in policies, implementation, and equality in outcomes need to be addressed. Recently, several fairness schemes have been applied as measures to allocate resources in multi-stakeholder systems. These schemes have been analyzed in different sectors, such as an agricultural system and an integrated residential complex. These studies have verified the deficiencies of certain approaches and highlighted the importance of comparing different allocation schemes. On the other hand, an economic approach to include sustainability in process integration projects has been reported recently. Here, a new sustainability metric is proposed. This metric is denominated as the Sustainability Weighted Return on Investment metric (SWROIM). The metric is based on the economic return on investment but also evaluates the contribution to sustainability. Extended forms of the metric have been proposed to include safety and resilience. In this work, we propose a fair-sustainable approach that involves applying fairness schemes along with the sustainability metric SWROIM for the optimal income allocation of an integrated system to produce fuels and simultaneously capture emissions. The analyzed schemes are the social welfare, Rawlsian welfare, and Nash approaches. The involved stakeholders in the integrated system include refineries, biorefineries, and eco-industries. To foster the reduction of emissions, we include economic compensations for the eco-industries. Furthermore, we analyze the avoided emissions (environmental function) and the generated jobs (social function) obtained through the different schemes and the sustainability metric. To demonstrate the applicability of the optimization approach, a case study for the future planning of the energy system in Mexico was presented. In the results, important differences are observed for the stakeholders' income allocation, as well as for the environmental and social functions under the schemes and the sustainability metric. These differences highlight the importance of exploring these different allocations since they can be key for decision-makers. The results also show that the Nash scheme can provide fair trade-offs among the stakeholders' income and the environmental and social functions.

Keywords: Optimization, Fairness schemes, Sustainability metric.

1. Introduction

Previous works have proposed approaches to include a fairness environment for stakeholders and an evaluation of sustainability metrics separately. Recently, fairness schemes for the allocation of resources in multi-stakeholder systems have been reported (Sampat and Zavala, 2019). These schemes have been applied to allocate resources in different systems (see for instance: Munguía-López et al., 2019). These studies have highlighted the importance of comparing different allocation schemes in diverse systems and verified some deficiencies of other approaches that may guide to solutions that are not fair. The desired aim would be to simultaneously foster equity, environmental protection, and economic growth (Rosa, 2009). However, these aspects represent a great challenge. In this regard, an economic approach to involve sustainability in process integration projects has been reported by El-Halwagi (2017). Here, sustainability is included by proposing a metric that is based on the economic return on investment but also evaluates the contribution to sustainability (SWROIM). Extended forms of the metric have been proposed to include safety and resilience (Guillen-Cuevas et al., 2018; Moreno-Sader et al., 2019). As a fair allocation of resources is key for long-term sustainability, we propose a fair-sustainable approach that includes applying the fairness schemes along with the sustainability metric SWROIM for resource allocation in complex engineering systems that include economic, environmental, and social objectives.

This work uses this fair-sustainable approach for the optimization of an integrated system to produce fuels and simultaneously capture emissions. The involved stakeholders in the integrated system include refineries, biorefineries, and eco-industries. In this approach, the eco-industries are taken as endeavors for forest plantations resulting in capturing emissions. To further foster the reduction of emissions, we include economic compensations for the eco-industries. The optimal allocation of income among these stakeholders is evaluated by using different fairness schemes (social welfare, Rawlsian welfare, and Nash approaches) and the sustainability metric SWROIM. Furthermore, the environmental and social functions (avoided emissions and generated jobs) obtained through the distinct schemes are analyzed.

2. System Description

In the proposed integrated system, there are three types of industries: refineries, biorefineries, and eco-industries. These industries are considered as the stakeholders that compete for income allocations (economic function). The integrated system includes the production of fuels and biofuels by refineries and biorefineries. Moreover, it includes the eco-industries or forest plantations that can capture part of the emissions generated by the fuels and biofuels production processes (an economic compensation for the eco-industries is involved). The avoided emissions by the eco-industries and the generated jobs by all the types of industries (environmental and social functions) are also evaluated through the analysis. Parameters such as potential locations for the industries, biomass types, availability restrictions, and climatic conditions are given. The objective of the work is to find the optimal allocations for the involved stakeholders through the fair-sustainable approach. Here, allocation schemes (social welfare, Rawlsian welfare, and Nash approaches) that have been previously used to allocate resources in a fair manner are evaluated. The sustainability metric SWROIM is also used for evaluating the allocations because it allows assessing the impacts on sustainability while considering the economic profitability. Each allocation scheme and the SWROIM metric result in different mathematical models since the objective function and some constraints vary. The schematic representation of the system is shown in Figure 1.

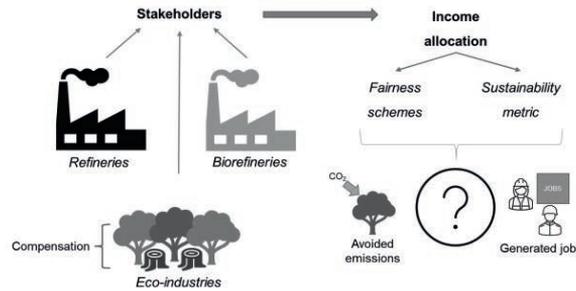


Figure 1. Schematic representation of the integrated system.

3. Model Formulation

The formulation of the proposed model includes the following steps. First, the mathematical model to describe the integrated fuel production system was developed. Here, the mass balances to represent the interactions among industries, the technical relationships to compute costs and incomes, and the disjunctions to find the optimal installation of industries were included. Also, the economic compensation for the eco-industries, the evaluation of the environmental and social impact of the system by estimating the avoided emissions and the generated jobs, and the introduction of the fairness schemes and the SWROIM metric as alternative objective functions were included. Then, the utopic solutions of the economic, environmental, and social variables were estimated. The environmental function refers to the emissions captured by the eco-industries from refineries and biorefineries ($Emcap_i^{Ref}$, $Emcap_b^{Bio}$).

$$AEM = \sum_e \left(\sum_i Emcap_{i,e}^{Ref} + \sum_b Emcap_{b,e}^{Bio} \right) \quad (1)$$

The social function in this work refers to the total generated jobs and it is defined as the sum of the number of jobs in refineries, biorefineries, and eco-industries:

$$TOTALJOBS = Jobs^{Refinery} + Jobs^{Bioref} + Jobs^{Ecoind} \quad (2)$$

The economic functions refer to the incomes of refineries, biorefineries, and eco-industries. The income of refineries and biorefineries includes the sold products and bioproducts. They are estimated considering the unitary sale cost of each product, the operating days, and the total flow rate of products and bioproducts.

$$Income^{Refinery} = \sum_i \sum_{p1} UC_{i,p1}^{p-ref} HY F_{i,p1}^{p-ref} \quad (3)$$

$$Income^{Biorefinery} = \sum_b \sum_{p2} UC_{b,p2}^{p-bio} HY F_{b,p2}^{p-bio} \quad (4)$$

The income of eco-industries depends on the captured emissions from refineries and biorefineries (see Equation (5)). Here, we can observe that the cost of the emissions generated by refineries ($CEmis_r^{Ref}$) and biorefineries ($CEmis_b^{Bio}$) is multiplied by the captured emissions by eco-industries for each type of refinery. Furthermore, the economic compensation is also included.

$$Income^{Ecoind} = \sum_e \left(\sum_i CEmis_i^{Ref} Emcap_{i,e}^{Ref} + \sum_b CEmis_b^{Bio} Emcap_{b,e}^{Bio} + Comp_e^{pl} \right) \quad (5)$$

Considering the economic functions of each stakeholder, the allocations through the fairness schemes are computed. First, the social welfare scheme is estimated by maximizing the sum of the utilities of the stakeholders as follows:

$$SW = \text{Income}^{\text{Refinery}} + \text{Income}^{\text{Bioref}} + \text{Income}^{\text{Ecoind}} \quad (6)$$

Similarly, the Nash allocation is given by maximizing the function N that corresponds to the sum of the logarithms of the utilities of the stakeholders.

$$N = \ln \text{Income}^{\text{Refinery}} + \ln \text{Income}^{\text{Bioref}} + \ln \text{Income}^{\text{Ecoind}} \quad (7)$$

Then, for the Rawlsian scheme, the objective function RW is minimized and the following constraints are included:

$$-\text{Income}^{\text{Refinery}} \leq RW \quad (8)$$

$$-\text{Income}^{\text{Bioref}} \leq RW \quad (9)$$

$$-\text{Income}^{\text{Ecoind}} \leq RW \quad (10)$$

This formulation corresponds to maximizing the smallest utility of the stakeholders. Regarding the sustainability metric SWROIM, the following objective function is maximized:

$$SWROIM = \frac{\left(\text{Income}^{\text{Refinery}} + \text{Income}^{\text{Bioref}} \right) + \text{Income}^{\text{Ecoind}} - TCI}{TCI} \left[\begin{array}{l} 1 + w^A \left(\frac{AEM}{AEM^{\text{target}}} \right) \\ + w^J \left(\frac{TOTALJOBS}{TOTALJOBS^{\text{target}}} \right) \end{array} \right] \quad (11)$$

Here, the economic, environmental, and social functions mentioned above are involved and TCI is the total capital investment.

4. Results and Discussion

To illustrate the applicability of the model, we addressed a case study of an integrated system that involves existing refineries in Mexico and potential locations for installing new refineries, biorefineries, and eco-industries (Sánchez-Bautista et al., 2017). Different biomass types for the biorefineries and availability restrictions for suppliers were included. Regarding the eco-industries, one forest plantation for each Mexican state was considered. First, Figure 2 presents the income of each stakeholder (refineries, biorefineries, and eco-industries) through the distinct schemes: SWROIM, social welfare (SW), Nash (N), Rawlsian welfare (RW), and their difference with the utopia point (UP). For the refineries, the differences between the schemes and the utopia point are minimal in percentage. However, since this income is high, a minimum difference can represent a great variation of income. The SWROIM scheme gives the closest income to the utopic solution (with 1.491E+07 USD MM/year of difference) while the Rawlsian approach attains the lowest income. This behavior occurs because the refineries' income corresponds to the greatest part of the total income. For the biorefineries, the differences in income are greater among the schemes and have the opposite behavior. The Rawlsian scheme provides a solution equal to the utopia point and the SWROIM analysis allocates no income to this stakeholder. Moreover, the social welfare scheme allocates almost half of the utopic solution, while the Nash approach allocates 14% more. The biorefineries' income is smaller than the income of the other stakeholders. Because of this, we observe that the Rawlsian approach favors this stakeholder. For the eco-industries, the SWROIM and Rawlsian schemes provide the worst solutions. However, the social welfare gives the same income as the utopia point. The Nash scheme provides a similar allocation since it attains only 12% less. The difference between these schemes is similar in percentage to their difference for the biorefineries' income, but opposite in direction. The results for the

integer decisions including the installation of new refineries, biorefineries, and eco-industries are represented in Figure 3. For the SWROIM scheme, the refineries are the only new industries that are installed (3 out of the 3 new possible refineries). For the rest of the schemes, the same number of new refineries is installed. However, in some approaches, this proportion is smaller than the proportion related to the installation of the other industries. For instance, for the social welfare and Nash schemes, the highest proportion of eco-industries is observed. Regarding the biorefineries, the possibility of installing 6 new industries of this type was considered. For the social welfare and the Rawlsian schemes, this maximum number of new biorefineries are installed.

Figure 4 shows the environmental and social functions under the distinct schemes and their difference with the utopia point. We observe that the behavior of the eco-industries' income (see Figure 2) is almost equal to the environmental and social functions. This occurs because the economic objectives of the eco-industries are not conflicting with these functions. On the contrary, they increase in the same direction since the eco-industries' income depends on the captured emissions that are the avoided emissions. The allocations for the number of generated jobs are also similar, only with a slight difference in the SWROIM and Rawlsian schemes. It should be noticed that through the SWROIM analysis no emissions are avoided. However, 30,715 jobs are generated due to the existing refineries as well as the installation of new refineries.

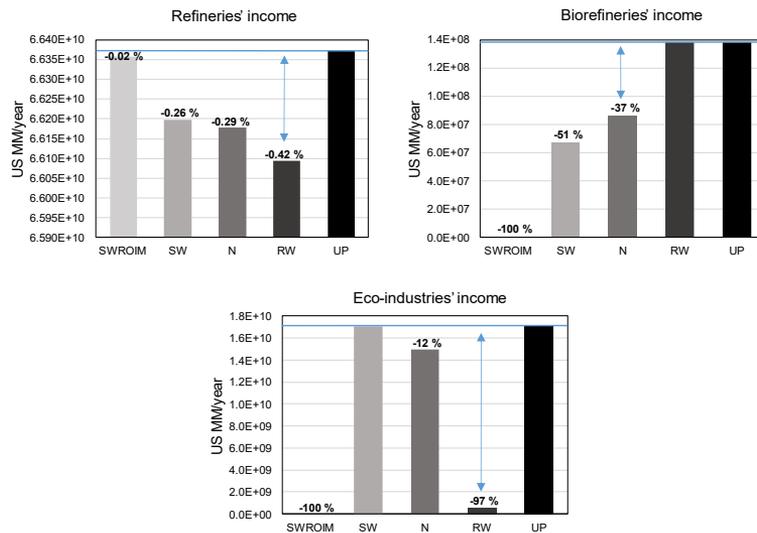


Figure 2. Refineries, biorefineries, and eco-industries income through the different schemes.

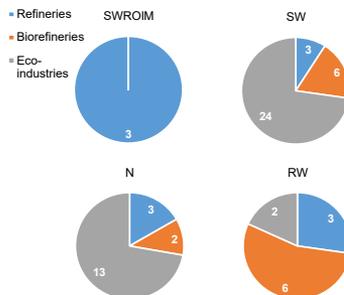


Figure 3. Proportion of the installation of new industries through the different schemes.

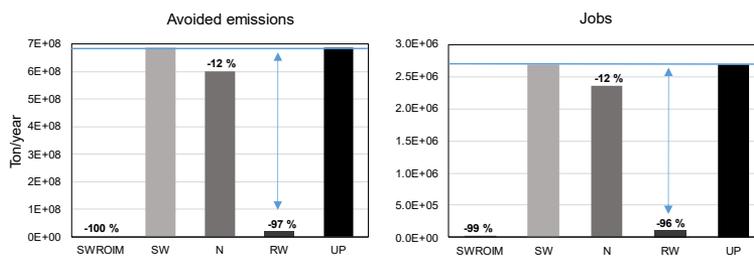


Figure 4. Avoided emissions and generated jobs through the different schemes.

5. Conclusions

This work presented an approach for the evaluation of fairness schemes (social welfare, Rawlsian welfare, and Nash approaches) and the sustainability metric SWROIM for the optimal allocation of income among different stakeholders (refineries, biorefineries, and eco-industries). The differences in incomes with each evaluated scheme highlight the importance of analyzing these possible allocations since they can be key for decision-makers. We found that certain schemes attain the utopic solution (or a value close to it) for one of the stakeholders; however, there is not a specific scheme that can provide the utopic solution for all stakeholders. Also, it was found that the SWROIM metric does not favor the environmental and social functions of the addressed system. On the contrary, this metric gives preference to the economic function and favors the stakeholder that contributes the most to the total profit. Furthermore, there is no installation of new biorefineries and eco-industries in the solution given by this scheme. Additionally, the Nash approach allows identifying trade-off solutions among the incomes of the stakeholders. Simultaneously, this allocation scheme provides trade-offs for the avoided emissions and the generated jobs. Moreover, through the Nash approach, an optimal design that allows the installation of biorefineries and eco-industries was found.

References

- A.C. Munguía-López, A.M. Sapat, E. Rubio-Castro, J.M. Ponce-Ortega, V.M. Zavala, 2019, Fairness-guided design of water distribution networks for agricultural lands. *Comput. Chem. Eng.*, 130, 106547.
- A.F. Sánchez-Bautista, J.E. Santibañez-Aguilar, F. You, J.M. Ponce-Ortega, 2017, Optimal design of energy systems involving pollution trading through forest plantations. *ACS Sustain. Chem. Eng.*, 5(3), 2585-2604.
- A.M. Sapat, V.M. Zavala, Fairness measures for decision-making and conflict resolution, 2019, *Optim. Eng.*, 20(4), 1249-1272.
- D.J. Rosa, 2009, Sustainability and infrastructure resource allocation, *J. Bus. Econ. Res.*, 7(9), 71-76.
- K. Guillen-Cuevas, A.P. Ortiz-Espinoza, E. Ozinan, A. Jiménez-Gutiérrez, N.K. Kazantzis, M.M. El-Halwagi, 2018, Incorporation of safety and sustainability in conceptual design via a return on investment metric, *ACS Sustain. Chem. Eng.*, 6(1), 1411-1416.
- K. Moreno-Sader, P. Jain, L.C.B. Tenorio, M.S. Mannan, M.M. El-Halwagi, 2019, Integrated approach of safety, sustainability, reliability, and resilience analysis via a return on investment metric, *ACS Sustain. Chem. Eng.*, 7(24), 19522-19536.
- M.M. El-Halwagi, 2017, A return on investment metric for incorporating sustainability in process integration and improvement projects, *Clean Technol. Environ. Policy*, 19(2), 611-617.

The advancement of zero-emission natural gas power plants and their role in future energy supply

Aibo Zhang¹, Haoshui Yu^{2*}, Emre Gençer³, Adem A.R. Nielsen⁴, Gürkan Sin⁴, Songjie Shi⁵

¹*Department of Advanced Design and System Engineering, City University of Hong Kong, Kowloon, Hong Kong*

²*Department of Chemistry and Bioscience, Aalborg University, Niels Bohrs Vej 8A, Esbjerg 6700, Denmark*

³*MIT Energy Initiative, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States*

⁴*Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Kgs., Lyngby, Denmark*

⁵*School of Electrical Engineering, Shenyang University of Technology, No. 111, Shenyang West Road, Shenyang, 110870, China*
hayu@bio.aau.dk

Abstract

Reducing the CO₂ emission from the power sector is one of the effective ways to mitigate global warming. Net-zero-emission power generation implies the control of not only pollutants such as NO_x, CO and organic air pollutants but also greenhouse gas CO₂. In traditional gas power plants, a huge amount of CO₂ is generated, which has to be captured to achieve the near-zero emission target. Carbon capture in the power industry is technically and economically challenging due to the energy penalty and high capital cost. CO₂ taxation among others is expected to accelerate the progress of zero-emission power plants. The concept of zero- and near-zero-emission power plants can be realized by combining advanced power cycles and carbon capture technologies. Scientists and engineers are striving hard to improve the efficiency of power cycles and decrease the cost of carbon capture. Novel systems combining carbon capture technologies and advanced power cycles are reviewed in this paper. Each country has its own roadmap and timeline to achieve the zero-emission target in power sector. Therefore, the role of the zero-emission natural gas power plants in the future energy supply depends heavily on the country. In this paper, the prospect of the zero-emission natural gas power plants in the USA, China and Denmark are analyzed.

Keywords: zero-emission, natural gas, power plants, future energy supply

1. Introduction

Global warming has become one of the most contentious scientific and technological challenges in the last two decades. Among the greenhouse gases, CO₂ is the largest contributor to global warming (NASA 2019). Reducing the CO₂ emission from power plants is one of the effective ways to limit global warming to less than 2°C. Previous attempts to capture CO₂ in power plants have decreased the efficiency of the power plant by about 10% (Fu and Gundersen 2012). Generally, carbon capture on a natural gas power plant is not economic due to the high operating cost and capital cost. However, with the increasing concern on climate change, carbon capture and storage

(CCS) is attracting more and more attention from both industry and academia. To achieve the target of the zero-emission power plant, advanced power cycles with CCS becomes the primary measure to take. There are many advanced power cycles and carbon capture technologies in power plants. In this paper, power cycles and carbon capture technologies are reviewed first. Based on this, the most promising natural gas power plants with zero- or near-zero-emission are reviewed and analyzed. The prospect of natural gas power plants in the future energy supply depends on the country. The role of natural gas power plants in the USA, China and Denmark in the near future is presented in the last section of this paper.

2. Power cycles

Commonly used power cycles in natural gas power plants are briefly reviewed. (I) Gas turbine cycle: The gas turbine cycle is the primary cycle for natural gas power plants. The high temperature and pressure flue gas expands in a gas turbine to generate electricity. The higher temperatures at the inlet of the turbine, the higher the thermal efficiency of the system (Unger and Herzog 1998). However, the materials cannot bear too high combustion temperatures in the gas turbine and thus the material is the main limitation of the gas turbine efficiency. (II) Steam Rankine cycle: Since the temperature of exhaust flue gas from the gas turbine is still very high, the steam can be generated in a heat recovery steam generator (HRSG). The generated steam can be utilized to drive a steam Rankine cycle. Therefore, the steam Rankine cycle is often integrated with the gas turbine cycle to form the well-known natural gas combined cycle (NGCC) plants. (III) Steam injection gas turbine cycle: The steam generated in the HRSG can be injected into the combustor to increase the power output of the simple gas turbine cycle (Nishida et al. 2005). The steam and air can mix at the outlet of the compressor or in the combustor and thus the thermal efficiency of a combined cycle can be improved. (IV) Allam cycle: This is a new, high-pressure, supercritical CO₂ cycle that generates electricity from fossil fuels with near-zero emissions. CO₂ is used as the working fluid rather than steam or air. All CO₂ generated by the system is produced as a high-pressure, pipeline-ready by-product. It is reported that Allam Cycle with carbon capture has a similar levelized cost of energy compared with NGCC power plants without carbon capture (Allam et al. 2014). (V) Organic Rankine cycle (ORC): ORC has been widely used for waste heat recovery (Yu et al. 2016). ORC can generate power at low-temperature or even cryogenic temperature levels. Therefore, ORC can improve the system efficiency with the recovery of low-temperature heat and LNG cold energy in a liquefied natural gas (LNG) fired power plant. (VI) Direct expansion cycle: LNG regasification process can be integrated with the power plant. In the natural gas direct expansion cycle, LNG is firstly pumped to very high pressure, then heated up to a gaseous or supercritical state, and finally expands in an expander to generate power (Franco and Casarosa 2015).

3. Carbon capture technologies

There are generally four categories of carbon capture technologies in power plants: (i) Pre-combustion carbon capture: Reforming of natural gas with oxygen produces a mixture of CO and H₂, and then the water-gas shift reaction converts CO into CO₂. Pre-combustion carbon capture refers to capturing CO₂ in a synthesis gas after the conversion of CO into CO₂. CO₂ can be separated by physical absorbent and hydrogen can be the fuel to generate electricity (Kanniche et al. 2010). (ii) Post-combustion

carbon capture: CO₂ is separated from the flue gas after the combustion. There are many different ways to separate CO₂ from flue gas, such as chemical absorption, adsorption, membrane separation and cryogenic CO₂ separation. Solvent-based absorption (MEA-CO₂ absorption) has become a benchmark for post-combustion carbon capture (Alie et al. 2005). (iii) Oxy-combustion carbon capture: Pure oxygen is employed in the combustion process. Therefore, the flue gas mainly consists of water and CO₂. Oxy-combustion provides an effective way to burn natural gas while allowing the capture of CO₂ through simple physical separation processes. The concentration of CO₂ in the flue gas can reach 80%, which is an advantage for the carbon capture process. However, an air separation unit (ASU) is required, thus both the operating cost and capital cost are higher compared with the conventional power plants. (iv) Chemical looping combustion (CLC): A metal oxide as the oxygen carrier is circulated between two interconnected fluidized bed reactors: an air reactor and a fuel reactor. CLC is similar to oxy-combustion where there is no direct contact between air and fuel. Oxygen is extracted from the air by oxygen carriers in the air reactor. The fuel is oxidized by the lattice oxygen of the metal oxide and produces CO₂ and vapor in the fuel reactor. The main advantage of chemical looping resides in the inherent separation of both CO₂ and H₂O from the flue gases. In addition, the NO_x formation is minimized since the combustion takes place in a nitrogen-free environment.

Among the above-mentioned carbon capture technologies, solvent-based chemical absorption post-combustion capture is the most popular choice because of its relatively low cost and the maturity of the technology. Pre-combustion, post-combustion and oxy-combustion carbon capture technologies are energy-intensive, resulting in a significant decrease in the overall efficiency and increase in the cost of produced electricity. CLC technology is energy efficient, but commercial scale-up of the CLC depends on the availability of the performance and stability of oxygen carriers.

4. Review of zero-emission natural gas power plant

To achieve zero- and near-zero-emission goal, both power cycles and carbon capture have to be considered simultaneously in a natural gas power plant. Bolland and Sather (1992) presented the analysis of natural gas-fired seawater cooled combined cycle power plants with CO₂ capture. 90% of CO₂ was captured by amine scrubbing. A fraction of the flue gas is recirculated back to the gas turbine compressor to reduce the volumetric flow and increase the concentration of CO₂ for the downstream CO₂ recovery plant. Shao et al. (1995) proposed a natural gas-fired power plant with virtually zero emission. The plant operates in a gas-steam turbine combined cycle with oxy-combustion mode. The liquid oxygen is used to liquefy CO₂ from the flue gas to save the compression work. Mathieu and Nihart (1999) proposed a novel MATIANT cycle, which is a gas cycle with CO₂ as the working fluid and O₂ as the oxidizer. In this system, CO₂ is compressed to very high pressure and then cooled down by cooling water and finally removed in the liquid state. The temperature of the combustion chamber is 1300 °C and the condensation temperature is 29°C at 70.5 bar. Deng et al. (2004) proposed a cogeneration power system with LNG regasification process. This system can not only generate electricity but also produce natural gas for other end users. The power system is an oxy-combustion power cycle and CO₂ in the flue gas is liquefied by LNG. Sanz et al. (2005) proposed an S-Graz Cycle to achieve the zero-emission goal of the power plant. They declared that S-Graz Cycle could be the most economical solution for CO₂ capture once the development of new turbomachinery

components is realized. Zhang and Lior (2006) proposed a quasi-combined cycle mode with a supercritical CO₂ Rankine-like cycle and a CO₂ Brayton cycle. The top Brayton cycle and the bottom supercritical CO₂ Rankine cycle are coupled. The LNG cold energy is reutilized to capture the CO₂ in the flue gas and cool down the inlet stream of compressors to save compression work. Zhang and Lior (2008) also proposed two novel systems for oxy-combustion natural gas-fired power plant integrated with steam reforming and CO₂ capture. Natural gas is converted into H₂ and CO to improve the fuel heating value and the turbine exhaust heat can be recovered by the reforming reaction. They declared that the net energy efficiency is in the range of 50–52%. Liu et al. (2009) proposed an oxy-combustion gas turbine cycle with LNG cold energy utilization, which is based on the flowsheet proposed by Deng et al. (2004). The primary difference is the integration of the LNG evaporation with the CO₂ condensation. Xiong et al. (2014) proposed a combined cycle with LNG cold energy recovery. The combined cycle consists of a gas turbine cycle and a steam Rankine cycle. LNG cold energy is utilized in ASU and CO₂ capture processes. The electrical exergy efficiency can reach 54.9% with 90.6% CO₂ recovery. Scaccabarozzi et al. (2016) performed thermodynamic analysis and optimization of Allam cycle. The maximum efficiency is 54.8% with 100% CO₂ capture. Chen et al. (2017) proposed a novel gas and steam mixture cycle. In this system, peak shaving, energy storage and CO₂ capture are considered along with power generation. Liquefied oxygen is produced during off-peak hours. The main advantage of the novel cycle is that the pressure of LNG and liquefied O₂ is increased to the combustion pressure by pumps instead of compressors, thus a huge amount of compression work is saved. Naqvi et al. (2004) proposed a chemical looping combustion natural gas power plant with CO₂ capture. The oxidation reactor outlet stream drives the gas turbine and the exhaust from the fuel reactor drives CO₂-turbine. The results show that an optimum efficiency of 49.7% can be achieved under given conditions with a CLC-combined cycle at zero emissions level.

5. The role of zero-emission natural gas power plants around the world

Each country is at a different point in the decarbonization journey. Therefore, natural gas plays different roles in decarbonization all over the world. The USA is the largest natural gas consumer in the world. Annual electricity generation from natural gas power plants in the USA increased by 31% in the Northeast region, by 20% in the Central region, and by 17% in the South region between 2015 and 2019. Natural gas power plants will continue to play an important role in the USA. Net power technology, which is a company focusing on zero-emission electricity, intends to build two natural-gas power plants in the U.S. that will have all its emissions captured and buried deep underground. Almost all new power plants built in 2021 will be carbon-free. However, for the existing natural gas power plants, revamping with carbon capture is expected to be done to achieve carbon neutrality in 2050.

China is the largest CO₂ emission country, accounting for about 30% of the world's total emission in 2013 (Friedlingstein et al. 2014). Currently, natural gas plays a relatively small role in China's power sector despite the growth in recent years. Coal remains the dominant fuel in China's power sector, but wind and solar generation have risen more rapidly than gas power. China has pledged to peak carbon dioxide emissions before 2030 and achieve carbon neutrality before 2060. China's energy consumption is dominated by coal. Natural gas will play an important and stable role in the Chinese

energy market with a share of 12% in 2030 and 11% in 2060. China has committed to increasing support for other developing countries in developing green and low-carbon energy, and not to build new coal-fired power projects abroad. Natural gas power plants are key to stable power systems in the energy transition. Natural gas is expected to be a key bridge fuel over the next two decades. China will fundamentally revamp its natural gas power plants to achieve the zero-emission target. Natural gas power plants are expected to be retrofitted with CCS by 2050 in China (Qin 2020).

As regards Denmark, the country has been playing a leading role in the carbon-neutral society. Copenhagen, the capital of Denmark, committed to becoming carbon-neutral in 2010-five years before the Paris Agreement and will become the world's first carbon-neutral capital by 2025 (Damsø et al. 2017). Denmark has set up two targets to achieve the carbon-neutral goal: year 2050 with 100% renewable energy from biomass, wind, solar and wave energy; year 2030 with 50% renewable energy, as the first important milestone on the way to carbon-neutral society (Lund and Mathiesen 2009). Currently, there are only two natural gas-fired power plants in Denmark. Towards the target of carbon neutrality, biogas will play an increasingly important role. Biogas is abundant in Denmark due to its advancement in biomass utilization (Korberg et al. 2020). Biogas can replace natural gas after upgrading. Thus with biogas as the fuel, the existing power plants do not need to be revamped with CCS. If CCS is implemented, a negative emission goal can be achieved. Natural gas power plants will be replaced by various sources of renewable energy in Denmark. In summary, natural gas power plants will be phased out in Denmark soon. The advancement of technologies in natural gas power plants seems to be of no interest to Denmark.

6. Conclusion

To combat climate change, decarbonization in the power sector is a necessity. Natural gas power plants play an important role in the power sector. The advancement of zero and near-zero-emission natural gas power plants is reviewed in this paper. Even though it is technologically viable to achieve the zero-emission target, there are more challenges in putting the concept into practice considering the cost of carbon capture. All the studies are focusing on reducing the cost for carbon capture while maintaining an acceptable thermal efficiency of the power plants. For the policymakers in different countries, zero-emission natural gas power plants play different roles on the journey to carbon neutrality depending on the country. The prospect of the zero-emission natural gas power plant in the USA, China, and Denmark are analyzed in this paper. Zero-emission natural gas power plants will still be of interest to the USA and China on the journey to carbon neutrality, while Denmark will phase out the natural gas-fired power plants in the near future.

References

- Alie, C., L. Backham, E. Croiset and P. L. Douglas (2005). Simulation of CO₂ capture using MEA scrubbing: a flowsheet decomposition method. *Energy conversion and management* 46(3): 475-487.
- Allam, R., J. Fetvedt, B. Forrest and D. Freed (2014). The oxy-fuel, supercritical CO₂ Allam Cycle: New cycle developments to produce even lower-cost electricity from fossil fuels without atmospheric emissions. ASME turbo expo 2014: turbine technical conference and exposition, American Society of Mechanical Engineers Digital Collection.
- Bolland, O. and S. Saether (1992). New concepts for natural gas fired power plants which simplify the recovery of carbon dioxide. *Energy Conversion & Management* 33: 467-475.

- Chen, Y., Z. Zhu, J. Wu, S. Yang and B. Zhang (2017). A novel LNG/O₂ combustion gas and steam mixture cycle with energy storage and CO₂ capture. *Energy* 120: 128-137.
- Damsø, T., T. Kjær and T. B. Christensen (2017). Implementation of local climate action plans: Copenhagen—Towards a carbon-neutral capital. *Journal of cleaner production* 167: 406-415.
- Deng, S., H. Jin, R. Cai and R. Lin (2004). "Novel cogeneration power system with liquefied natural gas (LNG) cryogenic exergy utilization." *Energy* 29(4): 497-512.
- Franco, A. and C. Casarosa (2015). Thermodynamic analysis of direct expansion configurations for electricity production by LNG cold energy recovery. *Applied Thermal Engineering* 78(Supplement C): 649-657.
- Friedlingstein, P., R. M. Andrew, J. Rogelj, G. P. Peters, J. G. Canadell, R. Knutti, G. Luderer, M. R. Raupach, M. Schaeffer and D. P. van Vuuren (2014). Persistent growth of CO₂ emissions and implications for reaching climate targets. *Nature geoscience* 7(10): 709-715.
- Fu, C. and T. Gundersen (2012). Carbon Capture and Storage in the Power Industry: Challenges and Opportunities. *Energy Procedia* 16, Part C: 1806-1812.
- Kanniche, M., R. Gros-Bonnivard, P. Jaud, J. Valle-Marcos, J.-M. Amann and C. Bouallou (2010). Pre-combustion, post-combustion and oxy-combustion in thermal power plant for CO₂ capture. *Applied Thermal Engineering* 30(1): 53-62.
- Korberg, A. D., I. R. Skov and B. V. Mathiesen (2020). The role of biogas and biogas-derived fuels in a 100% renewable energy system in Denmark. *Energy* 199: 117426.
- Liu, M., N. Lior, N. Zhang and W. Han (2009). Thermoeconomic analysis of a novel zero-CO₂-emission high-efficiency power cycle using LNG coldness. *Energy Conversion and Management* 50(11): 2768-2781.
- Lund, H. and B. V. Mathiesen (2009). Energy system analysis of 100% renewable energy systems—The case of Denmark in years 2030 and 2050. *Energy* 34(5): 524-531.
- Mathieu, P. and R. Nihart (1999). Zero-emission MATIANT cycle. *Journal of engineering for gas turbines and power* 121(1): 116-120.
- Naqvi, R., O. Bolland, O. y. Brandvoll and K. Helle (2004). Chemical looping combustion-analysis of natural gas fired power cycles with inherent CO₂ capture. *ASME Turbo Expo 2004: Power for Land, Sea, and Air*, American Society of Mechanical Engineers Digital Collection.
- NASA (2019). Global Climate Change-Carbon Dioxide. Available: <https://climate.nasa.gov/vital-signs/carbon-dioxide/>.
- Nishida, K., T. Takagi and S. Kinoshita (2005). Regenerative steam-injection gas-turbine systems. *Applied Energy* 81(3): 231-246.
- Qin, Y. (2020). Natural gas in China's power sector: Challenges and the road ahead.
- Sanz, W., H. Jericha, M. Moser and F. Heitmeir (2005). Thermodynamic and economic investigation of an improved Graz cycle power plant for CO₂ capture. *Journal of Engineering for Gas Turbines and Power* 127(4): 765-772.
- Scaccabarozzi, R., M. Gatti and E. Martelli (2016). Thermodynamic analysis and numerical optimization of the NET Power oxy-combustion cycle. *Applied Energy* 178: 505-526.
- Shao, Y., D. Golomb and G. Brown (1995). Natural gas fired combined cycle power plant with CO₂ capture. *Energy Conversion and Management* 36(12): 1115-1128.
- Unger, D. P. and H. J. Herzog (1998). Comparative Study on Energy R & D Performance: Gas Turbine Case Study, Energy Laboratory, Massachusetts Institute of Technology.
- Xiong, Y., P. Luo and B. Hua (2014). A novel CO₂-capturing natural gas combined cycle with LNG cold energy utilization. *Energy Procedia* 61: 899-903.
- Yu, H., X. Feng, Y. Wang, L. T. Biegler and J. Eason (2016). A systematic method to customize an efficient organic Rankine cycle (ORC) to recover waste heat in refineries. *Applied Energy* 179: 302-315.
- Zhang, N. and N. Lior (2006). A novel near-zero CO₂ emission thermal cycle with LNG cryogenic exergy utilization. *Energy* 31(10): 1666-1679.
- Zhang, N. and N. Lior (2008). Two novel oxy-fuel power cycles integrated with natural gas reforming and CO₂ capture. *Energy* 33(2): 340-351.

A Process Integration-Based Optimal Decarbonisation Policymaking Software Framework

Purusothmn Nair S Bhasker Nair,^a Michael Short,^{b*} Dominic C. Y. Foo,^a
Raymond R. Tan^c

^a*Department of Chemical and Environmental Engineering/Centre of Excellence for Green Technologies, University of Nottingham Malaysia, Broga Raod, 43500 Semenyih, Selangor, Malaysia*

^b*Department of Chemical and Process Engineering, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom*

^c*Department of Chemical Engineering, De La Salle University, 2401 Taft Avenue, 0922 Manila, Philippines*
m.short@surrey.ac.uk

Abstract

The criticality of climate change is such that any further delay in mitigation action would result in irreversible damage. This work develops a novel software framework for optimal decarbonisation in energy planning to determine the optimum deployment of renewable energy sources, alternative low carbon fuels, negative emission technologies (NETs) and CO₂ capture and storage (CCS). The mathematical programming-based tools in this work provide rigorous optimal solutions, subject to budget, demand, and emission constraints. The application of the software framework is demonstrated with a Malaysian energy decarbonisation case study. The results indicate a heavy reliance on NETs, alongside reductions in coal and natural gas use to achieve CO₂ neutrality by 2050.

Keywords: Multiperiod Energy Planning; Negative Emission Technologies; Process Integration; Policymaking; Decarbonisation Software

Introduction

Global leaders recently gathered at COP26 to agree on emission reduction targets. Limiting warming by 2100 to 1.5 °C is technically feasible but time is short to achieve this. An emissions cut of 45% must be achieved by year 2030 relative to 2010 (International Renewable Energy Agency, 2021) to meet the Paris Agreement targets. On a positive note, the recent plummeting costs of renewable energy sources offers hope for climate change mitigation. The challenge is to optimally deploy different decarbonisation measures to reduce emissions strategically and rapidly within economic and environmental constraints.

Sustainable energy planning via Carbon Emissions Pinch Analysis (CEPA) was pioneered by Tan and Foo (2007). Later, Tan et al. (2009) extended the CEPA approach by incorporating CCS as a decarbonisation option. The initial graphical approach was followed with the development of an automated targeting model by Lee et al. (2009), which was applied to CCS deployment by Ooi et al. (2013). The deployment of carbon dioxide removal (CDR) via negative emission technologies (NETs) was then addressed with graphical (Nair et al., 2020) and algebraic targeting tools (Nair et al., 2021). Equally, several energy planning tools e.g., MARKAL, TIMES etc. had been developed to analyse

energy and electricity systems (Ringkjøb et al., 2018). These models conduct short and long-term planning, considering the technical and economic constraints. This work develops a decision-making software framework based on mathematical programming for planning decarbonisation. A multiperiod model is used to allow progressive targets to be incorporated subject to technical and economic constraints. The novelty of this work is that it includes fuel substitution, NETs, and CCS as concurrent alternatives, along with process integration-inspired visualisation tools and detailed strategic scheduling and technology selection for a Malaysian case study.

1. Problem Statement

A superstructure mathematical programming formulation is developed to synthesise an optimal decarbonisation plan with the following specifications. For time-period $k \in K$, the CO₂ emission limit, L_k and energy demand, D_k are specified. Power plant $i \in I$ with lower, $F_{i,LB}$ and upper bound energy output, $F_{i,UB}$, and CO₂ intensity, CS_i constitute the energy planning system for period k . To meet CO₂ emissions and energy demands in period k , CCS $n \in N$ of different types can be employed at each power plant, as well as energy-producing NETs (EP-NETs) $p \in P$, energy-consuming NETs (EC-NETs) $q \in Q$ and compensatory renewable energy $c \in C$ that are considered for deployment. Additionally, alternative solid-based fuels $s \in S$ and gas-based fuels $g \in G$ are available to substitute the use of coal and natural gas respectively for power generation.

2. Mathematical Programming Formulation

Within period k , the energy outputs from all power plants $i \in I$ ($\sum_i FS_{i,k}$) must satisfy energy demands (D_k), as shown in Eq.(1).

$$\sum_i FS_{i,k} = D_k \quad \forall k \quad (1)$$

Carbon intensities of power plants with CCS technology n installed in period k ($CR_{i,n}$) are determined from Eq.(2) (Ooi et al., 2013).

$$CR_{i,n} = \frac{CS_i \times (1 - RR_n)}{1 - X_n} \quad \forall i \forall n \quad (2)$$

where RR_n and X_n are the CO₂ removal ratio and power loss associated with installing CCS.

Eq.(3) is used to calculate the net energy output from power plants with CCS retrofit (FNR). Note that there will be a reduced energy output from power plants with CCS retrofit due to the technologies' power consumption. Eq.(4) is a Big-M constraint associated with CCS selection.

$$R_{i,k,n} \times (1 - X_n) = FNR_{i,k,n} \quad \forall i \forall k \forall n \quad (3)$$

$$R_{i,k,n} \leq F_{i,UB} \times B_{i,k,n} \quad \forall i \forall k \forall n \quad (4)$$

where $B_{i,k,n}$ is a binary variable associated with selection to retrofit CCS technology n to power plant i in period k and $R_{i,k,n}$ is the extent of CCS retrofit.

The total extent of CCS retrofit of power plant i with all CCS technologies ($TR_{i,k}$) is obtained from Eq.(5). Additionally, the energy output from the power plant i cannot be less than the total extent of CCS retrofit with all CCS technologies, represented by Eq.(6).

$$\sum_n R_{i,k,n} = TR_{i,k} \quad \forall i \forall k \quad (5)$$

$$TR_{i,k} \leq FS_{i,k} \quad \forall i \forall k \quad (6)$$

In any period k , the total energy from a plant ($FS_{i,k}$) must be equal to the summation of the net energy output without CCS retrofit ($FNS_{i,k}$), the extent of CCS retrofit ($R_{i,k,n}$) and the alternative solid ($SD_{i,k,s}$) and gas-based fuels ($GS_{i,k,g}$); shown in Eq.(7).

$$FNS_{i,k} + \sum_n R_{i,k,n} + \sum_s SD_{i,k,s} + \sum_g GS_{i,k,g} = FS_{i,k} \quad \forall i \forall k \quad (7)$$

Eq.(8) ensures that in any period k , total energy output from all energy sources is equal to total energy demands, which includes the total power demands (D_k) combined with those required by EC-NETs ($EC_{k,q}$). Similarly, Eq.(9) enforces total CO₂ load from all energy sources is equal to total CO₂ emissions over period k (TE_k).

$$\begin{aligned} \sum_i \sum_n (FNS_{i,k} + FNR_{i,k,n}) + \sum_c FC_{c,k} + \sum_p EP_{k,p} + \sum_s SD_{i,k,s} \\ + \sum_g GS_{i,k,g} = \sum_q EC_{k,q} + D_k \quad \forall k \end{aligned} \quad (8)$$

$$\begin{aligned} \sum_i \sum_n (FNS_{i,k} CS_i + (FNR_{i,k,n} CR_{i,n})) + \sum_c FC_{c,k} CIC_{c,k} \\ + \sum_p EP_{k,p} CIEP_{k,p} + \sum_q EC_{k,q} CIEC_{k,q} \\ + \sum_s SD_{i,k,s} CISD_{i,k,s} + \sum_g GS_{i,k,g} CIGS_{i,k,g} = TE_k \quad \forall k \end{aligned} \quad (9)$$

where $CIC_{c,k}$, $CIEP_{k,p}$, $CIEC_{k,q}$, $CISD_{i,k,s}$ and $CIGS_{i,k,g}$ represent the carbon intensities of compensatory energy c , EP-NETs technology p , EC-NETs technology q , alternative solid-based fuel s and gas-based fuel g in period k respectively.

The total energy costs for period k (TC_k) are obtained using Eq.(10).

$$\begin{aligned} \sum_i \sum_n (FNS_{i,k} CT_{i,k} + (FNR_{i,k,n} CTR_{i,k,n}) + (CFX_{i,k,n} B_{i,k,n})) \\ + \sum_c FC_{c,k} CTC_{c,k} + \sum_p EP_{k,p} CTEP_{k,p} + \sum_q EC_{k,q} CTEC_{k,q} \\ + \sum_s SD_{i,k,s} CTSD_{i,k,s} + \sum_g GS_{i,k,g} CTGS_{i,k,g} = TC_k \quad \forall k \end{aligned} \quad (10)$$

where $CTR_{i,k,n}$ and $CT_{i,k}$ are costs of energy output by power plants with and without CCS technology, while $CTC_{c,k}$, $CTEP_{k,p}$, $CTEC_{k,q}$, $CTSD_{i,k,s}$ and $CTGS_{i,k,g}$ are the cost associated with compensatory energy c , EP-NETs technology p , EC-NETs technology q , alternative solid-based fuel s and gas-based fuel g . Meanwhile, $CFX_{i,k,n}$ represents the fixed cost of retrofitting power plant i with CCS n in period k .

Eq.(11) ensure that decisions taken to retrofit power plants with CCS in earlier periods is not reversed in later periods; the extent of CCS retrofit at later periods is at least equal to that in previous periods.

$$(R_i)_{k+1} \geq (R_i)_k \quad k = 1, 2, \dots, n - 1 \quad (11)$$

Total CO₂ emissions and total energy costs are subject to limits in Eq.(12) and Eq.(13), relating to emission limits L_k and budget allocation BD_k .

$$TE_k \leq L_k \quad \forall k \quad (12)$$

$$TC_k \leq BD_k \quad \forall k \quad (13)$$

Objective functions for the optimisation problem can either be Eq.(14) if the total costs are to be minimised subject to meeting certain emissions limits, or Eq.(15), where total CO₂ emissions are minimised subject to budgetary constraints.

$$\min TC_k \quad \forall k \quad (14)$$

$$\min TE_k \quad \forall k \quad (15)$$

This mixed-integer linear programming (MILP) model is implemented in Pyomo with a spreadsheet as an input interface and can be solved using the users' choice of solver. The code and supporting documents are available at <https://github.com/mchlshort/DECO2>.

3. Case Study

The application of the optimal decarbonisation software framework is demonstrated with a Malaysian energy planning case study. Malaysia is one of the fastest-growing economies within ASEAN (Association of South-East Asian Nations). This rapid economic growth makes Malaysia a carbon-intensive country, but a carbon-neutrality pledge by 2050 has been made with the 12th Malaysia Plan (Salim, 2021). This goal will require a revamp of the country's carbon-intensive grid. In year 2020, the power generation sector alone generated 109 Mt CO₂ (Energy Commission, 2020). Due to the Paris Agreement, renewable energy share is projected to hit 40% by year 2035 (The Straits Times, 2021). However, an increase in renewable energy share alone would not be sufficient. A recent study concluded that a carbon pricing policy must be implemented in Malaysia to achieve its target CO₂ emission reduction (Izlawanie, 2021). However, an ambiguous carbon tax framework and political instability would hinder its implementation. On the other hand, Malaysia, being one of the top palm oil producers in the world has significant potential for the use of crops as biomass (Loh, 2017). Given that coal plants are expected to be operational till at least year 2040, other alternatives i.e., fuel substitution and the use of NETs must be considered.

We consider energy planning for Malaysia across six 5-year periods starting in 2025, each with a specified demand, emission limit and budget allocation, based on forecasts. The CO₂ emission reduction is achievable with the potential deployment of three types of EP-NETs i.e., bioenergy with CCS and biochar and EC-NETs i.e., direct air capture and enhanced weathering, alongside two choices of CCS technologies and compensatory energy. Two types of biomass and biogas, each with a specified CO₂ intensity and cost are also available to replace coal and natural gas respectively. Table 1 presents the demand and CO₂ emission constraints for each period. The case study was solved for the minimum budget objective function (Eq.(14)) using CPLEX as the optimisation solver. Figure 1 presents the energy planning pinch diagram for period 6 (years 2045 – 2050).

Table 1: Energy Planning Data

Period k	Years	$D_k /$ TWh y^{-1}	$L_k /$ Mt y^{-1}
1	2022 - 2025	133	116
2	2025 - 2030	142	115
3	2030 - 2035	156	110
4	2035 - 2040	166	94
5	2040 - 2045	184	60
6	2045 - 2050	203	0

Figure 1 shows that the carbon-neutral target in 2050 can be achieved with the deployment of NETs, renewable energy and CCS.

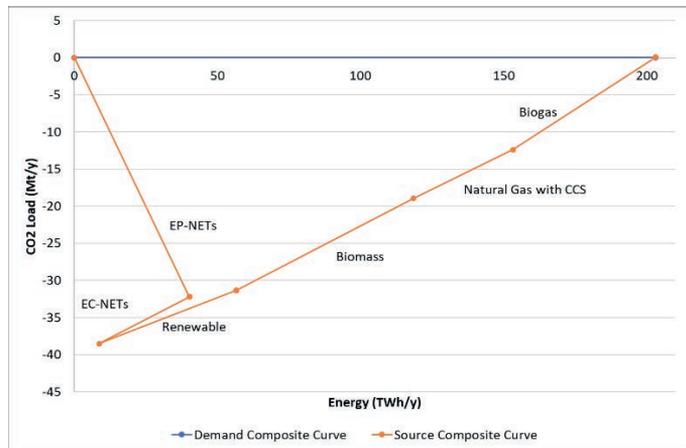


Figure 1: Energy planning pinch diagram for Period 6 (2045-2050)

The optimisation results demonstrate that all coal power plants would be decommissioned to be replaced by biomass by 2050. As of 2040, 13 of the available 19 natural gas plants were operational to generate 68.23 TWh electricity annually. In 2050, while one natural gas plant is decommissioned, four plants generating 43.5 TWh electricity are CCS retrofitted, with the remaining being transitioned to accommodate biogas as feedstock, generating 39.2 TWh electricity annually. The results of this software framework align with Malaysia’s aspiration for a gradual decommissions of coal power plants in light of its climate change target.

4. Conclusion

An optimal decarbonisation policy software framework was developed in this work based on mathematical programming models for energy planning. It can aid policymakers in crafting decarbonisation strategies. The MILP model provides solutions for optimal deployment of NETs, CCS, compensatory energy and fuel substitutions to meet emissions caps in each period. As demonstrated using a Malaysian decarbonisation case study, this software framework could contribute to the achievement of the net-zero emissions targets.

Results show that there is a need for energy planning models to consider fuel substitutions in existing plants, as well as techno-economic constraints down to the individual plant level. Future work on the open-source software will focus on incorporating uncertainty associated with parameters, as well as considering economy-wide CO₂ emissions from other industries/sectors e.g., petrochemical, cement etc. aside from incorporating the technical constraint i.e., availability, implementation time, cost etc. for each technology.

Acknowledgements

The team would like to offer their sincerest gratitude to the British Council for funding provided towards the COP26 Trilateral Research Initiative. Additionally, the team would like to offer thanks to Dr Yasunori Kikuchi and Dr Disni Gamaralalage from the University of Tokyo for their continuous support throughout this research.

References

- Energy Commission, 2020. Peninsular Malaysia Electricity Supply Industry Outlook 2019.
- International Renewable Energy Agency, 2021. World Energy Transitions Outlook 1.5C Pathway.
- Izlawanie, M., 2021. Challenges in implementing carbon pricing policy in Malaysia [WWW Document]. Bartlett, Univ. Coll. London. URL <https://www.ucl.ac.uk/bartlett/news/2021/nov/challenges-implementing-carbon-pricing-policy-malaysia> (accessed 11.19.21).
- Lee, S.C., Ng, D.K.S., Foo, D.C.Y., Tan, R.R., 2009. Extended pinch targeting techniques for carbon-constrained energy sector planning. *Appl. Energy* 86, 60–67. <https://doi.org/10.1016/j.apenergy.2008.04.002>
- Loh, S.K., 2017. The potential of the Malaysian oil palm biomass as a renewable energy source. *Energy Convers. Manag.* 141, 285–298. <https://doi.org/10.1016/j.enconman.2016.08.081>
- Nair, P.N.S.B., Tan, R.R., Foo, D.C.Y., 2021. A Generic Algebraic Targeting Approach for Integration of Renewable Energy Sources, CO₂ Capture and Storage and Negative Emission Technologies in Carbon-Constrained Energy Planning. *Energy* 235, 121280. <https://doi.org/10.1016/j.energy.2021.121280>
- Nair, P.N.S.B., Tan, R.R., Foo, D.C.Y., 2020. Extended Graphical Approach for the Deployment of Negative Emission Technologies. *Ind. Eng. Chem. Res.* 59, 18977–18990. <https://doi.org/10.1021/acs.iecr.0c03817>
- Ooi, R.E.H., Foo, D.C.Y., Tan, R.R., Ng, D.K.S., Smith, R., 2013. Carbon constrained energy planning (CCEP) for sustainable power generation sector with automated targeting model. *Ind. Eng. Chem. Res.* 52, 9889–9896. <https://doi.org/10.1021/ie4005018>
- Ringkjøb, H.K., Haugan, P.M., Solbrenke, I.M., 2018. A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renew. Sustain. Energy Rev.* 96, 440–459. <https://doi.org/10.1016/j.rser.2018.08.002>
- Salim, S., 2021. 12MP: Malaysia committed to becoming carbon-neutral nation by 2050, says PM [WWW Document]. Edge Mark. URL <https://www.theedgemarkets.com/article/12mp-malaysia-committed-becoming-carbonneutral-nation-2050-says-pm> (accessed 11.13.21).
- Tan, R.R., Foo, D.C.Y., 2007. Pinch analysis approach to carbon-constrained energy sector planning. *Energy* 32, 1422–1429. <https://doi.org/10.1016/j.energy.2006.09.018>
- Tan, R.R., Ng, D.K.S., Foo, D.C.Y., 2009. Pinch analysis approach to carbon-constrained planning for sustainable power generation. *J. Clean. Prod.* 17, 940–944. <https://doi.org/10.1016/j.jclepro.2009.02.007>
- The Straits Times, 2021. Malaysia targets renewable energy at 40 per cent of total capacity by 2035 [WWW Document]. URL <https://www.straitstimes.com/asia/se-asia/malaysia-targets-renewable-energy-at-40-per-cent-of-total-capacity-by-2035> (accessed 11.13.21).

Interplaying of industry 4.0 and circular economy in cyber-physical systems towards the mines of the future

Mohammed Yaqot^a, Brenno C. Menezes^{a*}, Robert E. Franzoi^a

^aDivision of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar

bmenezes@hbku.edu.qa

Abstract

Today's Industry 4.0 (I4) is re-designing industrial activities and re-shaping communities into the so-called Society 5.0 (S5), enabling the creation of a new industry-society-environment nexus. The application of novel technologies such as pervasive sensing, widespread internet of the things (IoT), artificial intelligence (AI), and robotics aims to sense, calculate, and actuate by employing data-driven architectures and automated decisions for reaching more efficient processes that reduce environmental impacts and transform manpower workloads. This work introduces a cyber-physical system (CPS) towards more innovative operations within the mining industry's mine-to-mill process, whereby the operations' decision-making and executions are often done by human intelligence. However, the increasing use of autonomous machines or mechatronics (MEC) such as trucks, drills, drones, and conveyor belts, combined with advanced modelling and solving algorithms (MSA), can achieve the necessary autonomous operations for handling such hazards and harsh environments, where no or reduced manpower is required. For that, information and computing technologies (ICT) facilitate innovative solutions to re-shape and re-design production systems. The interplaying of advanced technologies of the I4 mandate and circular economy (CE) ideology towards the mines of the future creates a smart connected mining industry that embeds vast amounts of data into predictive and fully integrated intelligent systems. This paper discusses how I4's technological developments in mining sector applications generate opportunities to re-design and re-execute stockpiling to conveyor-belt processes into enhanced production states from a CE perspective (environmental, economic, and social). CE theories within I4-S5 offer innovative sustainable industrial concepts driven by safer, more environmentally friendly towards precision and autonomous mining.

Keywords: Industry 4.0, Society 5.0, mining, circular economy, sustainable development.

1. Introduction

Human beings have engaged in mining since the stone age. Since then, we have used natural resources from the earth for utilitarian purposes to support global development. Mining is generally considered to be the activity of recovering minerals and other materials from the earth and is often associated with further processing of minerals to concentrate the amount of metal and to remove impurities. The combination of industrial organisations involved in traditional mining and oil and gas recovery is often described

as the extractives sector. There are many interpretations of mining and its associated activities, and some terms are used interchangeably. Mining can be classified according to various criteria, including the location of activities and the scale at which it occurs. Figure 1 summarises common mining pathways, including underground, quarrying, sand, artisanal small-scale, undersea, in addition to the promising lunar or beyond Earth mining.

As our world's population and economic growth increases, the demand for total global metal production continues to increase in scale with its associated impacts. Some metals experience cycles of demand as older uses are phased out, such as cryolite, while other historically less-exploited metals have come to prominence as a result of technological innovations, such as the increasing production of lithium due to its use in batteries (Martin et al., 2017). Since the world of mining is changing with new technologies reshaping mining practices, traditional value chains (input, exploration mine development, blasting load and haul, processing, transportation, further processing, and products) are disrupted and society's expectations of mining companies have never been higher. The digital transformation and the introduction of CPSs towards smarter operations within the mining industry have put huge ambition towards global prosperity and sustainable development. The interplaying of advanced technologies and CE ideology towards the mines of the future creates a smart connected mining industry, whereby vast amounts of data are embedded into predictive analytics capabilities for achieving fully integrated intelligent systems (Kelly and Menezes, 2019). Such Industry 4.0 (I4) autonomous machinery and tools are fundamental to provide safer and more efficient environments. For example, mechatronics (MEC) with sophisticated modelling and solving algorithms (MSA) are needed for autonomous handling of hazards and the severe environment when no or little human intervention is required. ICT enables novel solutions to reshape and re-design mining production processes towards a complete digitalization of the processes.

At a process level in the mining system (from the crushing and to the milling), improved stockpiling to conveyor-belt operations can be achieved by employing better processes for separating and transporting ore, preventing evaporation, and producing drier tailings through a hybrid dynamic control application. Thus, additional safety, stability, and predictability can be achieved to maximise the ratio of metal to ore output while minimising the environmental impacts and the operating and capital costs. In addition, this creates sustainable communities to identify socio-economic development opportunities. This paper addresses how I4 technological advances drive opportunities to re-design and re-execute operations at the process level (of stockpiling to conveyor-belt) into an improved state under CE ideologies within the scope of the environmental, economic, and social pillars. In this context, CE theories within I4-S5 introduce novel sustainable industry principles for innovative processes that are driven by safer and more efficient, accurate, precise, and autonomous mining processes.

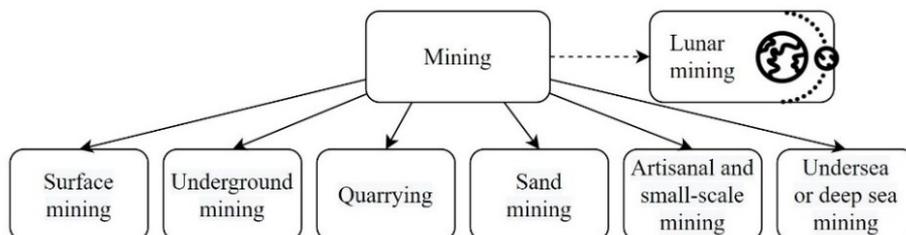


Figure 1: Mining pathways.

2. Mining challenges from a CE perspective

The life cycle assessment of mining and mineral processing as well as oil and gas industries use heavy equipment that consumes electrical, thermal and mechanical energy, which are key contributors to the overall industrial impacts on the environment (Farjana et al., 2019). The mining industry has been undergoing significant changes as companies adopt automation technologies to decrease costs, increase efficiency, and improve safety. The accident and ill-health record of the mining sector compares poorly to that of other economic sectors such as manufacturing, construction, and rail, leading to mining's reputation as the most hazardous industrial sector (Domingues et al., 2017). Mining poses significant health issues due to airborne contaminants such as silica and coal dust, as well as noise, heat, and vibration. Other severe health concerns include chemical hazards unrelated to the subterranean, such as air pollution, gases, skin diseases, ergonomic strains, and ionic, cosmic, and radioactivity radiations, among others. Although health problems can be mitigated by applying tight controls at the source in the workplace, developing such measures for mining operations provides significant hurdles given that dust and noise are created by mining operations themselves. The impacts of shifts in employment patterns and tedious working hours that increase site exposure are added to the complexity of assessing the likelihood of health risks. Many mineworkers have worked in mines for over two decades, with increasing risk of occupational diseases. Lately, the effectiveness of protective systems based on occupational exposure limits for eight-hour working days has been questioned (Komljenovic et al., 2017).

Besides social risks, the world faces the biggest environmental threat in generations, and low-carbon technologies are needed to resolve this issue. However, the development of such technologies increases the demand for the mining of raw materials. Electric vehicles, as an example, are made with around 70% of copper. The mining industry faces an interesting future with growing demand as well as severe challenges (Sánchez and Hartlieb, 2020). The current technological trend is to vastly increase the level of automation driven by enhanced I4 capabilities. Within this context, many challenges arise. First, local mining enterprises have become multinational giants and mine sites are usually very remote and with harsh climatic conditions. Recruiting and retaining skilled personnel to operate the mines in distant locations such as deserts or mountains is difficult. Second, decreased ore grades, which leads to exploring more ore bodies. As a result, the trade-off between drilling deeper with increased costs is to maintain profitable throughput. Third, despite continuous improvements, workforce safety incidents and accidents continue to occur, prompting policies to impose stricter laws, rules, and regulations. The safety of workers must be addressed, particularly in subterranean operations and on-ground areas where people working close to heavy machines are exposed to risks. Fourth, the mining sector has long been concerned about growing energy costs, efficiency, and measuring power consumption. Assuring tightly integrated plant processes and power units would be considered key for improving energy efficiency. Fifth, the lack of information unity, in which there are several independent pieces of equipment, machinery, systems, and subsystems in the mining site, and each has its own information and interfaces. Thus, decision-makers have separate pieces of information at hand with no holistic overview, known as "islands of automation". The lack of unity and integration imposes difficulties in performing proper decisions in a timely manner.

Complementary to the I4 adaptation in organisations, concerns in resource utilisation, conversation, and recycling, aligned with environmental issues worldwide, have transformed the old-fashion profitability objective to be maximised in the short-term

horizon to the sustainable economic growth targeting a long-term prospectus. With the support of the I4 technologies, the sustainable deployment, also referred to as CE in manufacturing systems and supply chains, implies efficient use of natural resources, reduced energy consumption, minimum generation of wastes, as well as high-performance production, logistics, and services. Therefore, identification and design opportunities for the resource-process-product within the I4-CE adoption are expected to increase (Song and Wang, 2018). Modern MEC-MSA-ICT fundaments of the I4 transform production and service operations models, reducing energy and optimising material flows and inventories. The consequence of such transformations is the reduction of waste and greenhouse gases (GHG) emissions. Such an evidence of GHG reduction by virtue of the capabilities of the augmented information age in transportation, logistics, smart cities, efficient buildings, and facilities, etc., demonstrate more sustainable environment while offering economic benefits (GelenbeErol and CaseauYves, 2015; Moyer and Hughes, 2012; Murugesan and Laplante, 2011).

3. Sense, calculate, and actuate cycle within CPSs for mine-to-mill process

From a laboratory-level to mature processes, advanced production in the industry is evolving to re-identify, re-design, re-execute, and re-evaluate new opportunities within research, development, and deployment (RD&D) stages into the I4 improved state. In the mining case, a real-time dynamic predictive control model based on I4 capabilities relies on the cycle of sensing, calculating (via optimisation), and actuating (SCA) within CPS in the stockpiling to conveyor-belt process. In this context, Menezes et al. (2019) address a conveyor-belt intermittently delivering crushed-ore to form multiple stockpiles considering targets of their inventories to avoid the so-called bridging, reducing the flow-out beneath the stockpiles, which feeds other conveyor-belts to the mills. In such a system, a dynamic online optimisation for shuttle-conveyor-belt tripper car is treated as a real-time hybrid model predictive control (HMPC) with shifting prediction time-horizon. The HMPC aims to enhance the performance of the stockpiles' level by autonomously adjusting the idle-time (actuation or manipulated variable) of the robotic apparatus of the tripper car that moves over each stockpile position (as seen in Figure 2). The motion from one stockpile to the next in a neighbour-to-neighbour fixed sequence forms multiple stockpiles instead of one per belt. The problem is solved as a mixed-integer linear programming (MILP) problem to minimise the deviation of the given targets to the actual or live inventory of the stockpiles (sensed by industrial radar technologies using ultrasound). In this coordinating of the conveyor-belt for multiple stockpiles, the robotic arm idle-time in the schematics of Figure 2 is the actuation or manipulated variable depositing the solids into each stockpile to automatically control its inventory. Then, the robotic apparatus goes back and forth as a smart sweeping movement since it counts on the smart manufacturing ground bases of the ICT, MSA, and MEC pillars to be implemented within a continuously controlled cycle of sensing, calculating (via optimisation), and actuating.

The bottlenecks of such a CPS in mining have been surpassed by the development of fast discrete optimisation in the calculating stage of the SCA cycle to continuously determine idle-time as a binary variable per time-window of the positions (Kelly and Menezes, 2019). By moving from one to multiple stockpiles per belt, CE implications of its three pillars exist. Economically, more efficient flow-to-stockpile increases the utilisation factor of the equipment, while reducing dust, raw material losses, and energy consumption. Nevertheless, the raise of such I4 technologies leads to social impacts by increasingly closing low-skilled positions and requiring high-qualified personnel.

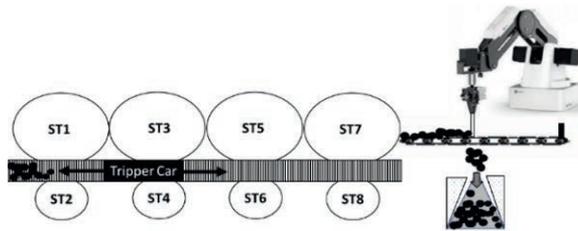


Figure 2: Robotic arm apparatus schematics feeding eight stockpiles in the multiple stockpiling per belt system.

Such control application considers advancements in manufacturing toward the I4 mandate involving ICT and MEC fundamentals evolving together with advances in MSA. In the latter, advances in network (flowsheet) optimisation, solving algorithms, and computer-aided resources allow faster and more efficient solutions (Franzoi et al., 2021). Moreover, the ICT expansion provides the demanded velocity of the communication of the online data measurement of complete process networks for the massive volume and variety of information from the plant, which permits the viability of the addressed control strategy. Such advances provide proper capabilities for the development and implementation of tools and sustainable strategies in mining operations with safer and more efficient process conditions towards the mines of the future.

By employing hybrid dynamic control applications in the mine-to-mill processes, stockpiling operations can be improved to better segregate and convey ore methods, reduce evaporation, and create drier tailings. As a result, increased human safety and predictability can be reached to maximise the metal to ore output ratio while reducing social and environmental consequences and operational and capital costs. Furthermore, this builds sustainable communities capable of identifying socioeconomic growth prospects.

4. Discussion and conclusion

After the hunter-gatherer, agrarian, and industrialisation stages of mankind, our current society in the wake of the S5 stage, referred to as an augmented society, moves towards the transformations of economic and environmental targets to the collection of individuals' and communities' well-being as a core of the new business models. Although the adoption of technologies is an economic added value, industries have faced difficulties in hiring skilled people since it requires specialised knowledge, abilities, and competencies in ICT, MEC, and MSA branches of the I4 age (Fitsilis et al., 2018). Mining is a labour-intensive industry with a harsh and dangerous working environment, and I4 facilitates replacing humans with autonomous machinery. The consequences affect employment profiles, but it ensures to maximise human safety in such remote and harsh places like deserts or mountains. Mine settings are particularly complex to handle since they deteriorate quickly and alter as the mining processes progress. Furthermore, mining emits hazardous pollutants into the atmosphere, and crushing rocks is intrinsically connected with dust and noises, whereby oxygen and lights must also be artificially provided in subterranean mines. Also, on- and under-ground miners operate heavy machinery and perform hard and tedious work in cramped contexts; thus, ergonomic risks are widespread in such industries. These hazards are connected to poor technical design and add higher safety risks (Komljenovic et al., 2017). The trade-offs debate between the pros and cons of increasing automation level depends on the society and its technological

evolution. If society's policies and procedures force industries to work towards the so-called CE ideology, an increased level of automation would create new job positions in remanufacturing and recycling activities with the high skills demanded in the I4 mandate.

The mining sector's accidents and ill-health records compare severely to those of other industry sectors, contributing to mining's reputation as being the most dangerous industrial sector. Intuitively, occupational injuries and illnesses records can negatively impact public safety, health, and the environment (SHE), where people's health and safety have monetary costs and environmental consequences. Furthermore, it connects initiatives to improve workplace health and safety to a wider societal agenda from a corporate social responsibility perspective, i.e., sustainable development. The experimental simulation of the current practice of stockpiling process shows a lack of resources utilisation. Increasing the level of automation by integrating technology solutions has the power of speeding stockpiling, maximising performance, improving safety, and reducing overall costs. The I4 elements can be implemented with a hybrid SCA dynamic control cycle to relocate feeding positions within seconds, increasing feeding performance, efficiently utilising resources, and performing multi-stockpile feeding. In the broader context of CE, healthy and safe working conditions are among the first expectations for sustainable development.

References

- M.S.Q. Domingues, A.L.F. Baptista, M.T. Diogo, 2017, Engineering complex systems applied to risk management in the mining industry, *International Journal of Mining Science and Technology*, 27, 611–616.
- S.H. Farjana, N. Huda, M.A. Parvez Mahmud, R. Saidur, 2019, A review on the impact of mining and mineral processing industries through life cycle assessment, *Journal of Cleaner Production*, 231, 1200–1217.
- P. Fitsilis, P. Tsoutsas, V. Gerogiannis, 2018. Industry 4.0: required personnel competences, *Industry 4.0*, 3, 130–133.
- R.E. Franzoi, B.C. Menezes, J.D. Kelly, J.A.W. Gut, 2021, A moving horizon rescheduling framework for continuous nonlinear processes with disturbances, *Chemical Engineering Research and Design*, 174, 276–293.
- E. Gelenbe, Y. Caseau, 2015, The impact of information technology on energy consumption and carbon emissions, *Ubiquity*, 1–15.
- J.D. Kelly, B.C. Menezes, 2019, Automating a shuttle-conveyor for multi-stockpile level control, *Computer Aided Chemical Engineering*, 46, 1153–1158.
- D. Komljenovic, G. Loiselle, M. Kumral, 2017, Organization: A new focus on mine safety improvement in a complex operational and business environment, *International Journal of Mining Science and Technology*, 27, 617–625.
- L. Martin, M. Höck, M. Bertau, 2017, Lithium market research – global supply, future demand and price development, *Energy Storage Material*, 6, 171–179.
- B.C. Menezes, J.D. Kelly, A.G. Leal, 2019, Identification and design of industry 4.0 opportunities in manufacturing: examples from mature industries to laboratory level systems, *IFAC-PapersOnLine*, 52, 2494–2500.
- J.D. Moyer, B.B. Hughes, 2012, ICTs: Do they contribute to increased carbon emissions?, *Technology Forecasting Society Change*, 79, 919–931.
- S. Murugesan, P.A. Laplante, 2011, IT for a greener planet, *IT Professional*, 13, 16–18.
- F. Sánchez, P. Hartlieb, 2020, Innovation in the mining industry: technological trends and a case study of the challenges of disruptive innovation, *Mining, Metallurgy and Exploration*, 37, 1385–1399.
- M. Song, S. Wang, 2018, Market competition, green technology progress and comparative advantages in China, *Management Decision*, 56, 188–203.

Environmental Impacts of Rice Husk-Derived Silica under Uncertainty: Is “Bio” better?

Ethan Errington^a, Miao Guo^{b*} and Jerry Y.Y. Heng^{a*}

^a*Department of Chemical Engineering, Imperial College London, London, SW7 2AZ, UK*

^b*Department of Engineering, King’s College London, London, WC2R 2LS, UK*
miao.guo@kcl.ac.uk (MG); jerry.heng@imperial.ac.uk(JH)

Abstract

Millions of tonnes of rice husk (RH) are produced annually as an agricultural waste. One area of interest for RH valorisation is to use rice husk ash (RHA, a by-product of RH combustion) as a replacement for mineral-derived synthetic amorphous silica (M-SAS). However, little information is available on the environmental benefit of this approach. This study details the first evaluation of the environmental benefits of producing RH-derived synthetic amorphous silica (RH-SAS). This is done by describing the life cycle of RH-SAS in terms of stages for which existing life cycle inventories can be linked and aggregated in a modular way. It is then shown how the physical meaning of linkages between modules are governed by both the characteristics of the RH feedstock and efficiencies of processes across the life cycle. To provide more robust findings, the sensitivity of predictions to model uncertainty are also considered. Finally, a case is provided for the benefit of RH-SAS production within the Asia-Pacific (APAC) SAS market.

Keywords: Silica, Rice Husk, Life Cycle Assessment, Uncertainty, Global Warming Potential

1. Introduction

An estimated 2.8 million-tonnes of synthetic amorphous silica (SAS) are produced per year by the “wet” industrial process Industry Experts (2019). This process commonly involves the reaction of a mineral-derived sodium silicate (SS) precursor following the stoichiometry in Equation 1 European Commission (2007).



However, many experimental investigations have now also been made into the potential of recovering SS and SAS from wastes produced by the construction and agriculture industries. Among these, a large subset consider the valorisation of rice husk ash (RHA) due to its high silica content Hossain et al. (2018). RHA is a product of the combustion of rice husk (RH), which is itself a by-product of rice grain agriculture. Yet, there is a gap in the literature with regards to understanding the environmental impact of RH-derived SAS (RH-SAS).

Consequently, this work develops models to compare the environmental impact of producing mineral-derived SAS (M-SAS) and RH-SAS via the wet method. This is done by the Life Cycle Assessment (LCA) of production methods for each feedstock case independently. To ensure robust comparisons, sensitivity of predicted impacts to LCA model uncertainty is also considered. This leads to the visualisation of a design space in which the demand for RH feedstock ($\text{kg}/\text{kg}_{\text{SAS}}$) and recoverable bioenergy ($\text{MJ}/\text{kg}_{\text{SAS}}$) of the RH-SAS life cycle can fall. Finally, a case study for the benefit of producing RH-SAS within the Asia-Pacific (APAC) region is provided. In all cases, comparisons are made in terms of the 100-year global warming potential (GWP).

2. Life Cycle Assessment

2.1. Scope, Functional Unit and Inventories

As illustrated in Figure 1, a cradle-to-gate scope has been used to develop inventories in this work. This has been done based on a functional unit of 1 kg of SAS.

For the M-SAS model, an inventory was developed based on the average inventory reported for the EU-15 region by the European Commission (2007). For the RH-SAS model, an inventory was developed by modifying the M-SAS model. This was done by representing the RH-SAS life-cycle in terms of four stages that are strongly linked to pre-existing industries. These were: Rice Grain Agriculture (RGA), RH combustion (RHC), SS production (SSP) and SAS production (SASP). For each stage, a relevant inventory was then selected from the EcoInvent Database (2019). The functional unit of each inventory was also converted to be consistent with 1 kg of SAS as illustrated in Figure 1 (see also Section 2.2). As RH is considered a waste of RGA, no environmental burdens of the RGA stage have been allocated to RH-SAS, following an economic allocation approach. Models and modifications used for each RH-SAS stage are shown in Table 1.

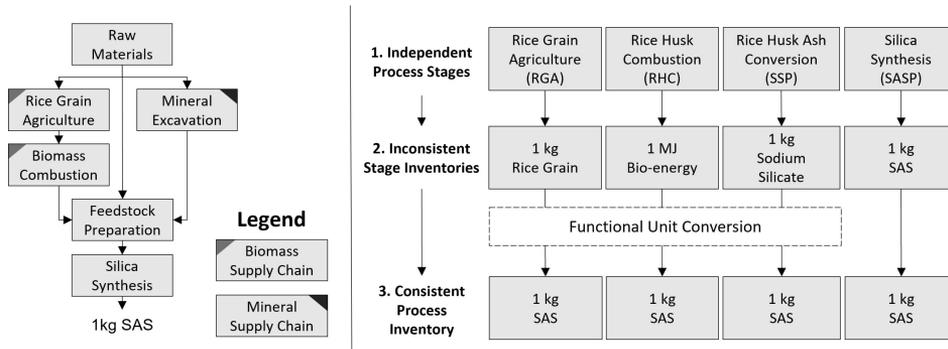


Figure 1: A summary of life cycle scopes considered for M-SAS and RH-SAS during impact assessment (left) and the method for specifying the entire RH-SAS inventory (right)

Table 1: Summary of inventory source associated with RH-SAS models

Stage	Inventory Source	Functional Unit	Modification
RGA	N/a (economic allocation = 0)	1 kg _{biomass}	N/a (economic allocation = 0)
RHC	Bauer (2007) ¹	1 MJ _{bioenergy}	references to biomass use removed
SP	Hirschier (2010) ²	1 kg _{silicate}	references to mineral use removed
SS	European Commission (2007) ³	1 kg _{silica}	references to mineral use removed

¹ "biomass combustion" process; ² "hydrothermal" process; ³ based on the M-SAS model in this work.

2.2. Life Cycle Impact Assessment

The total environmental impact of each independent life cycle model (M-SAS, RGA, RHC, SSP and SASP) was calculated using Equation 2. Impact assessment was carried out for GWP using the ReCiPe (H) method (v1.13) of Goedkoop et al. (2009). A cut-off value of 0.1% was applied.

$$EI_{kpi} = \sum_r \sum_s C_{r,kpi}^{in} F_{r,s}^{in} + \sum_c \sum_s C_{c,kpi}^{out} F_{c,s}^{out} \quad (2)$$

Where EI is the total impact with regards to some key performance indicator, kpi ; $C_{r,kpi}^{in}$ is the characterisation factor (impact/amount) for a resource, r ; $C_{c,kpi}^{out}$ is the characterisation factor (im-

fact/amount) for a compound, c ; $F_{r,s}^{in}$ is the amount of resource, r , arising from process step, s ; $F_{c,s}^{out}$ is the amount of compound, r , arising from process step, s .

The overall environmental impact of RH-SAS, EI_{kpi}^{RH-SAS} (impact/kg_{SAS}), was then calculated by making the individual impact predictions for each associated life cycle stage (RGA, RHC, SSP and SASP) consistent with the 1 kg of SAS functional unit - as shown in Equation 3.

$$EI_{kpi}^{RH-SAS} = EI_{kpi}^{RGA} \hat{M}^{RGA} + EI_{kpi}^{RHC} \hat{Q}^{RHC} + EI_{kpi}^{SSP} \hat{M}^{SSP} + EI_{kpi}^{SASP} \quad (3)$$

Where EI_{kpi}^{RGA} is the environmental impact of the RGA stage (impact/kg_{RH}); \hat{M}^{RGA} is the demand for RH-biomass from the RGA stage (kg_{RH}/kg_{SAS}); EI_{kpi}^{RHC} is the environmental impact of RHC stage (impact/MJ); \hat{Q}^{RHC} is the bioenergy recovered as electricity from the RHC stage (MJ/kg_{SAS}); EI_{kpi}^{SSP} is the environmental impact of the SSP stage (impact/kg_{feedstock}); \hat{M}^{SSP} is the amount of SS required for the SASP stage (kg_{SS}/kg_{SAS}); EI_{kpi}^{SASP} is the environmental impact of rice husk combustion stage (impact/kg_{SAS}) Note: RH mass (kg_{RH}) is considered on a dry basis.

Values of \hat{M}^{RGA} , \hat{Q}^{RHC} and \hat{M}^{SSP} were derived from process and biomass characteristics as:

$$\hat{M}^{RGA} = 1 / (\eta^{SSP} \eta^{SASP} x^{SiO_2} [1 - x^{H_2O}]) \quad (4)$$

$$\hat{Q}^{RHC} = (\eta^{RHC} / [\eta_p^{SSP} \eta_p^{SASP}]) \times \left(\left[\hat{Q}^{LHV} (1 - x^{H_2O}) - \lambda x_b^{H_2O} \right] / [x^{SiO_2} (1 - x^{H_2O})] \right) \quad (5)$$

$$\hat{M}^{SSP} = (1 + n) / \eta^{SSP} \quad (6)$$

Where η^i is the conversion efficiency of process stage, i (no units); \hat{Q}^{LHV} is the lower heating value of RH (MJ/kg_{RH}, dry basis); x^{H_2O} is the moisture content of RH (no units, mass basis); x^{SiO_2} is the silica content of RH (no units, mass basis); and n is the silicate number (no units) of SS (reported as $2 \leq n \leq 4$ by the European Commission (2007)). Values for each variable were taken as the midpoint of ranges reported in Section 2.5. Finally, the contribution of biogenic carbon emissions (arising from the RHC stage) to GWP was assumed to be zero due to the cyclical harvest and short crop rotation period typical of RGA.

2.3. Benefit of RH-SAS and Bioenergy Recovery

The benefit of RH-SAS has been quantified in two ways. Firstly, the benefit of RH-SAS when directly compared to M-SAS, B_{kpi}^{RH} (impact/kg_{SAS}), was calculated as shown in Equation 7. Secondly, the benefit of RH-SAS when using co-recovered bioenergy to offset grid-electricity use, B_{kpi}^{offset} (impact/kg_{SAS}), was calculated as shown in Equation 8.

$$B_{kpi}^{RH} = EI_{kpi}^{RH-SAS} - EI_{kpi}^{M-SAS} \quad (7)$$

$$B_{kpi}^{offset} = EI_{kpi}^{RH-SAS} - \hat{Q}^{RHC} EI_{kpi}^{grid} \quad (8)$$

Where EI_{kpi}^{M-SAS} is the environmental impact of M-SAS (impact/kg_{SAS}); EI_{kpi}^{grid} is the environmental impact of grid electricity (impact/MJ). Evaluation of EI_{kpi}^{grid} was done using the global medium voltage electricity model available within the EcoInvent Database (2019) (v3.6).

2.4. Uncertainty in Inventory Models

The sensitivity of impact evaluations to uncertainty in data quality was quantified by the pedigree matrix approach - in which, the quality of inventory data was rated against five criteria: reliability, completeness, and temporal, geographic and technological representativeness. Ratings were quantified as log-normal probability distributions based on the method of Frischknecht et al. (2004). Monte-Carlo simulations were used to estimate prediction uncertainty (sample size = 100,000).

2.5. Uncertainty in Feedstock and Process Characteristics

The sensitivity of \hat{M}^{RGA} and \hat{Q}^{RHC} to uncertainty in feedstock and process characteristics was derived from typical ranges reported in literature (see Table 2). It was assumed that the uncertainty for each variable was uniformly distributed. Values for each characteristic were randomly sampled 100,000 times (numpy package v 1.20.1 in python v3.7) to complement simulation of uncertainty from data quality (see Section 2.4). Following this, a further 900,000 random samples (i.e. 1-million total) were taken to visualise the feasible domain of \hat{M}^{RGA} and \hat{Q}^{RHC} .

Table 2: Uncertainty in Feedstock and Process Characteristics within the RH-SAS life cycle

Variable	Value	Ref(s).
RHC Stage Efficiency, η^{RHC} (%)	15-30	Bain (1993); Kawahara (2016)
SSP Stage Efficiency, η^{SSP} (%)	73-91	Pijarn et al. (2010); Kalapathy et al. (2000)
SASP Stage Efficiency, η^{SASP} (%)	95 ¹	European Commission (2007)
RH Heating Value, Q^{LHV} (MJ/kg _{GRH})	13 - 16	Natarajan et al. (1998)
RH Moisture Content, x^{SiO_2} (%)	10-14	Webb (1979); Van Hung et al. (2020)
RH Silica Content, x^{H_2O} (%)	15-20	Siddique and Cachim (2018)

¹ Data from European Commission (2007) implies certainty in the conversion efficiency of the SS stage

3. Results

3.1. Impact Evaluations and the Estimated Benefits of RH-SAS

Figure 2 illustrates the predicted GWP of M-SAS and RH-SAS as well as uncertainty in the benefits of RH-SAS (see Section 2.3). Single-value estimates (Figure 2 (left)) identify RH-SAS as having a lower GWP than M-SAS (2.05 vs. 1.77 kg CO₂-eq/kg_{SAS}). However, probability distributions in Figure 2 (right) suggest that the single-value estimate of B_{GWP}^{RH} is not significantly different from zero to any reasonable level of confidence. Consequently, the method could benefit from further minimising model uncertainty - either via data quality or processes and feedstock characteristics. This is also the case for B_{GWP}^{offset} , though this metric has a higher likelihood of being net GWP-negative. The lack of certainty in benefits being net-negative can also be related to the low contribution of feedstock to the GWP of both life cycles.

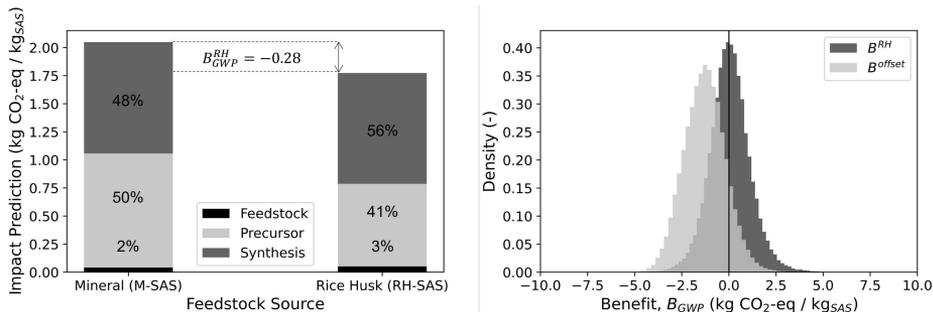


Figure 2: Global Warming Potential of SAS (left) and uncertainty in benefit of RH-SAS (right)

3.2. Uncertainty in Biomass Demand and Bioenergy Recoverable

Figure 3 illustrates the feasible design region for RH-SAS in terms of \hat{Q}^{RHC} and \hat{M}^{RGA} . Differences in the size of dotted regions illustrate the extent of importance in the absolute value of individual

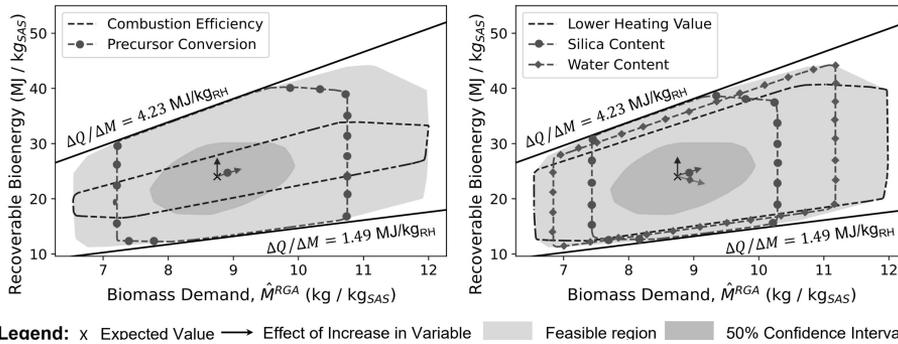


Figure 3: Feasible regions for Biomass Demand and Recoverable bioenergy (RH model)

Table 3: Summary of statistics relevant to potential production of RH-SAS within APAC.

Country	BD	CN	IN	ID	VN
RH Production (Mtonne/year) ¹	10.92	42.28	35.53	10.92	8.70
SAS Capacity (Mtonne/year)	1.30	5.04	4.23	1.30	1.04
Proportion of Market Demand ² (%)	81.63	316.14	265.66	81.66	64.97
Market Benefit, B_{GWP}^{RH} (Mtonne CO ₂ -eq/year)	0.36	0.44 ²	0.44 ²	0.36	0.29
Market, B_{GWP}^{offset} (Mtonne CO ₂ -eq/year)	2.12	2.58	2.58	2.12	1.69
Market GWP Reduction, B_{GWP}^{RH} (%/year)	11	13	13	11	9
Market GWP Reduction, B_{GWP}^{offset} (%/year)	65	79	79	65	52

¹ Based on rice paddy statistics for 2019 from the Food and Agriculture Organization of the United Nations (2020) and assuming 20 wt% of Rice Husk per mass of rice paddy Hossain et al. (2018), ² benefit limited to 100% of APAC market demand for consistency with real-world market

variables when held at their midpoint value. Therefore, smaller dotted regions reflect variables which the process is more sensitive to. Results show that \hat{Q}^{RHC} is most sensitive to changes in electrical efficiency of the RHC stage (i.e. η^{RHC}), while \hat{M}^{RGA} is most sensitive to the silica content of RH feedstock. Minimum and maximum gradients (4.23 and 1.49 MJ/kg_{RH}) observed also indicate the relatively low amount of energy recoverable per kg of RH as received, which is due to energy losses caused by the water content of the RH feedstock. Importantly, Figure 3 acts as a heuristic for understanding the effects of process design and feedstock selection on life cycle performance.

3.3. Implications for Asia-Pacific

A summary of key metrics for the use case of RH-SAS within the APAC region (based on top 5 producers of rice paddy by mass - Bangladesh, BD; China, CN; India, IN; Indonesia, ID; and Vietnam, VN) is provided in Table 3. Results indicate that each country produces RH in amounts which could serve SAS demand equivalent to more than 800% of the current APAC market. Moreover, it is estimated that the benefits from bioenergy co-recovery (within the RH-SAS life cycle) could lead to reductions of as much as 80% (i.e. 2.58 Mtonne/year) of the APAC SAS industry’s annual GWP footprint. However, as predictions rely on the values for B_{GWP}^{RH} and B_{GWP}^{offset} (Section 3.1), the uncertainty in values being significantly different from zero also applies here. Further refinement of each regional prediction will be achieved by using more specific regional models for energy use in future work.

4. Conclusion

This study details the first attempt to evaluate the GWP of RH-SAS. By comparing impacts with that of current industrial methods (M-SAS), our results suggest that RH-SAS may offer a pathway for reducing the GWP of SAS industry, including large reductions in APAC regions. However, high variability in the impact of both processes considered mean that certainty in the benefit of RH-SAS is low. This reflects the small contributions that RH and minerals make to the impact of RH-SAS and M-SAS respectively as well as the sensitivity of models to variability in process implementation in the wider life-cycle shared by each product. Thus, while areas for model refinement is discussed, it is believed that our findings are likely to provide a reliable answer to the question posed - i.e. "bio" is not generalisably "better". Further model refinement will be considered in future work. Finally, the methods described are generalisable for the study of alternate biomass feedstocks and impact metrics.

5. Acknowledgements

The authors acknowledge funding provided from Scottish Water and the UK EPSRC SSCP-DTP.

References

- R. L. Bain, 1993. Electricity from biomass in the United States: Status and future direction. *Bioresource Technology* 46 (1-2), 86–93.
- C. Bauer, 2007. electricity production, wood, future, Allocation at the Point of Substitution,ecoinvent database version 3.6. Tech. rep.
- EcoInvent Database, 2019. market group for electricity, medium voltage, GLO, Allocation at the Point of Substitution,ecoinvent database version 3.6.
- European Commission, 2007. Integrated Pollution Prevention and Control (IPCC): Reference Document on Best Available Techniques for the Manufacture of Large Volume Inorganic Chemicals-Solids and Others industry. Tech. rep.
- Food and Agriculture Organization of the United Nations, 2020. Food and Agriculture Data: Crops and Livestock Products - Rice, Paddy.
- R. Frischknecht, N. Jungbluth, H.-J. Althaus, G. Doka, R. Dones, T. Heck, S. Hellweg, R. Hischier, T. Nemecek, G. Rebitzer, M. Spielmann, 2004. The ecoinvent Database: Overview and Methodological Framework (7 pp). *The International Journal of Life Cycle Assessment* 2005 10:1 10 (1), 3–9.
- M. J. Goedkoop, M. A. J. Huijbregts, R. Heijungs, J. Struijs, 2009. Recipe 2008: Report 1 - A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level. Tech. rep., Netherlands Ministry of Housing for Spatial Planning and the Environment.
- R. Hirschier, 2010. sodium silicate production, hydrothermal liquor, product in 48% solution state — sodium silicate, without water, in 48% solution state — RER, APOS, U. Tech. rep.
- S. K. Hossain, L. Mathur, P. K. Roy, oct 2018. Rice husk/rice husk ash as an alternative source of silica in ceramics: A review. *Journal of Asian Ceramic Societies* 6 (4), 299–313.
- Industry Experts, 2019. Precipitated Silica - A Global Market Overview (CP082). Tech. rep.
- U. Kalapathy, A. Proctor, J. Shultz, 2000. A simple method for production of pure silica from rice hull ash. *Bioresource Technology* 73 (3), 257–262.
- Y. Kawahara, 2016. An overview on corrosion-resistant coating technologies in biomass/waste-to-energy plants in recent decades. *Coatings* 6 (3), 34.
- E. Natarajan, A. Nordin, A. N. Rao, 1998. Overview of combustion and gasification of rice husk in fluidized bed reactors. *Biomass and Bioenergy* 14 (5-6), 533–546.
- N. Pijarn, A. Jaroenworarluck, W. Sunsaneeyametha, R. Stevens, 2010. Synthesis and characterization of nanosized-silica gels formed under controlled conditions. *Powder Technology* 203 (3), 462–468.
- R. Siddique, P. Cachim, 2018. Waste and Supplementary Cementitious Materials in Concrete: Characterisation, Properties and Applications, 1st Edition. Elsevier, Woodhead Publishing.
- N. Van Hung, M. V. Migo, R. Quillooy, P. Chivenge, M. Gummert, 2020. Life Cycle Assessment Applied in Rice Production and Residue Management. In: *Sustainable Rice Straw Management*. Springer, pp. 161–174.
- B. Webb, 1979. Technical Aspects of Agricultural and Agroindustrial Residues Utilization. In: *Proceedings of UNEP/ESCAP/FAO Workshop on Agricultural and Agroindustrial Residue Utilization in Asia and Pacific Region*.

Economic Analysis of Novel Pathways for Recovery of Lithium Battery Waste

Khalid Amao, Bogdan Dorneanu, Harvey Arellano-Garcia

LS- Prozess- und Anlagentechnik, Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus D-03044, Germany

Abstract

Using the Umicore process, a current state-of-the-art recycling in the metal recovery industry for lithium battery waste, as a baseline, this contribution examines economic and environmentally friendly solutions for effective metal recovery from spent LIBs. At the same time, possible synergies between existing resource use from other manufacturing and waste treatment industries are considered as valuable input to metal recycling, while also reducing the amount of atmospheric carbon. This further presents a case for possible integration of various waste management approaches as a single business unit for economic incentive and profitability for possible investment.

Keywords: Lithium-ion battery, Waste recovery, e-Waste, Circular economy.

1. Introduction

Waste Management has become one of the most rapidly growing pollution problems globally as new technologies are increasingly exceeding millions of analogue techniques, which in turn result into their disposal in prescribed landfills with possible environmental consequences. Both landfilling and incineration are highly unsustainable remedies since they use large areas of land resulting to the release of significant environmental pollutants, including the greenhouse gases (GHGs) such as carbon dioxide (CO₂) and methane (CH₄) as the waste dissociates thereby causing health and environmental threats to developing countries (The World Bank, 2019). Waste electric and electronic equipment (WEEE/EEE) no longer suitable for their intended use are known as e-waste. TVs, telephones, radios, computers, printers, fax machines, DVDs, CDs, washing machines, refrigerators, dryers, vacuum cleaners, and other electronic devices are examples of e-waste. The composition of typical e-waste is depicted in Table 1. Approximately 50 million metric tons (Mt) of e-waste were produced globally in 2018 alone, compared to 44.7 Mt in 2016 and expected to rise in the future.

Due to a variety of factors such as technological challenges, implausible economic burdens, high environmental impact, complex chemical reactions, and so on, the optimum nature of the Lithium-Ion Batteries (LIBs) recovery process and technologies has yet to be determined (Azpagic, 1999). Creating a circular economy for e-waste could lower manufacturing costs, increase revenue streams, and provide tax advantages. Jobs could be created as a result of new and expanding markets. The Circular Economy (CE) is a regenerative strategy to waste reduction that aims to ensure the environmental sustainability of post-use items. For LIBs, a CE strategy is the central issue of more than 3,000 studies completed in the previous ten years, all of which have focused on the exploration of its primary steps (Yun et al., 2018). By enhancing a company's reputation and consumer confidence through environmentally responsible activities, a CE approach could improve market competitiveness. Reduced waste, energy use, and GHG emissions,

as well as the preservation of raw resources, could all be part of the greater environmental benefits. The goal is to develop a more sustainable e-waste recovery process that aligns with the global vision of reducing the amount of carbon dioxide emission and increase energy efficiency.

The Umicore process, a state-of-the-art commercial recycling process flow in the metal recovery industry for LIBs (Georgi-Maschler et al. 2012), is a global leading solution in environmentally friendly precious metal recycling and processing. It uses an integrated metals smelter and refining procedure to recover metals such as gold, silver, and the platinum group metals (PGMs: palladium, platinum, rhodium, iridium, ruthenium), special metals (selenium, tellurium, indium), secondary metals (antimony, tin, arsenic, bismuth), and base metals (antimony, tin, arsenic, bismuth). Sulphuric acid (from off gas-purification) and a depleted slag, which is used as a building material and in the concrete industry, are two other by-products of the factory.

In the following, environmentally friendly metal recovery from the spent LIB waste is explored using the Umicore process as a baseline. The introduction of solutions such as: anaerobic digestion of organic waste to produce methane, a biogas which can be sold for profit, mineralization of the recovered metals to form stable carbonates, through utilization of the carbon dioxide produced from the digester, as well as possible integration of captured carbon dioxide from energy intensive industries, are considered.

2. Methodology

This paper aims to provide understanding towards the development of an economically and environmentally feasible, efficient, and sustainable LIB recycling system for adoption in Europe. The stages of the Umicore process that fall within this scope are also represented schematically in Fig.1. The analysis, using data from the open literature (Li et al. 2018; Sommerville et al. 2021), was divided into 2 stages: firstly, an economic assessment of the state-of-the-art recovery process, broken down into its major stages, is performed, and secondly, an improved process is proposed, for which a similar study on the economics is carried out.

2.1. Economic analysis of current state-of-the-art LIB recovery process

For the purpose of this analysis, the Umicore process is broken down into its major sub-processes and the cost of each individual component is evaluated to establish the overall cost. Investment, operation and maintenance costs are not considered. All calculations are done on a yearly estimate, assuming a recycling capacity of 7,000 tons/year of LIB waste. The process involves the smelting of the LIBs in a smelter which recovers two main categories of high and low value metals. The high value metals are extracted as an alloy from the smelter, which would later be separated into individual pure components through leaching and solvent Extraction. The low value metals are extracted from the smelter as slag.

The following steps are considered in the analysis of the Umicore process:

- a) Breaking down the process into its main stages of Smelting, Leaching and Solvent Extraction.
- b) Calculate the cost of each stage using a functional unit of 1,000kg of LIB waste as input to establish an overall recovery cost.
- c) Calculate the carbon tax related to the smelting the processes.
- d) Calculate the mass of each high value metal and slag recovered.
- e) Calculate the market value of recovered metals and slag as revenue for the process.

- f) Calculate the profit from the difference between the cost of the process and revenue generated from the market value of the recovered metal and slag.
- g) Calculate the improved profit after a predicted 25% decrease in cost due to economies of scale at a recycling capacity of 7,000 tons of LIB per year.

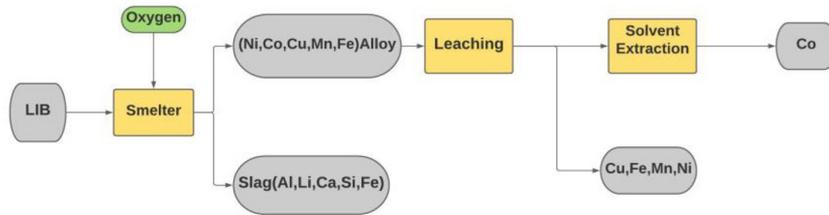


Figure 1. Flowsheet of the Umicore process

2.2 Economic analysis of improved process

In the second step, an improved and sustainable LIB recovery process is designed conceptually (Fig.2). New stages, such as mineralization, utilize CO₂ from an anaerobic bioreactor or absorption from the atmosphere for the conversion of the recovered metals and the slag to products which can be sold (e.g., Portland cement). In the case an anaerobic reactor is used, methane is also produced as biogas. Furthermore, the carbon tax acquired from the utilization of the CO₂ can also be sold as carbon credit. Thus, the recovery process creates new revenue streams, apart from the metals and LIB waste, by applying Circular Economy concepts, and creating new pathways for the reduction of atmospheric CO₂ emissions.

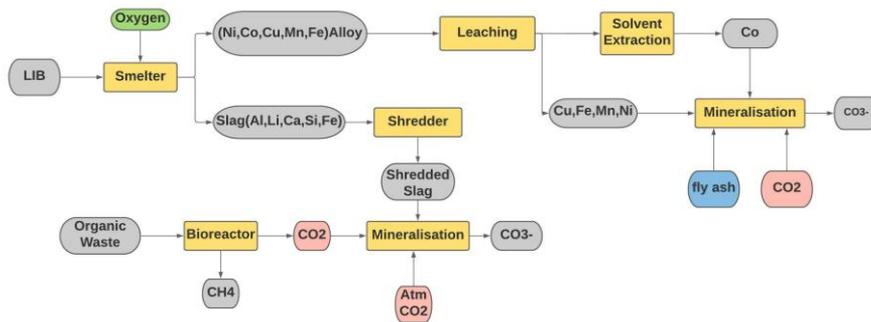


Figure 2. Flowsheet of the improved LIB recovery process

Once the process flowsheet is available, the following steps are followed to evaluate its efficiency and use the results for the comparison with the baseline:

- a) Calculate the cost of operating the anaerobic bioreactor and the shredder through their energy consumption.
- b) Calculate the cost of mineralization of the recovered metals and slag.
- c) Calculate the total cost of the overall improved process.
- d) Calculate the revenue generated from the sale of methane produced from the bioreactor as biogas.

- e) Calculate the revenue generated from the sale of slag Portland cement produced during the mineralization.
- f) Calculate the total carbon tax on the total CO₂ consumption for the whole processes which is to be sold as Carbon Credit.
- g) Calculate the total revenue of the overall improved process.
- h) Calculate the profit of the improved process after a 15% predicted reduction in cost due to economies of scale.

3. Results and Discussion

The recovery of the metals is divided into precious and non-precious metals, according to their market value. All precious metals are recovered through leaching and solvent extraction, while the non-valuable metals are extracted as slag during the smelting stage as. From the costing of the stages of the Umicore process (Table 1), the leaching contributed approx. 94% to the total recovery cost of \$19,858.16, while the solvent extraction and smelting contributed approx. 6.5 and 0.15%, respectively.

Table 1. State-of-the-art subprocess cost breakdown

Process	Cost (\$)	%
Smelting	1,266.28	6.38
Leaching	18,565.69	93.49
Solvent extraction	26,20	0.13
Total	19,858.16	100

The revenue generated from the recycling processes is comprised of the following streams: the selling of the recovered precious metals, the selling of the slag and the tipping fee (Table 2).

Table 2. State-of-the art subprocess revenue breakdown

Revenue breakdown	Valuation (\$)	%
Valuable metals	10,266.79	51.36
Slag resell value	223.02	1.16
Recycling fee per tonne	9,500	47.52
Total	19,989.81	100

As a commercial recycling process, it is predicted that a 25% reduction in recycling cost is expected due to economies of scale after the processing of 7,000 tons/year of LIB waste, which is the current operating capacity of the Umicore Process. This results in a profit of \$35,343,248 per year (Table 3).

The major cost of the improved process is the cost of mineralization, with a percentage of 99.81% for the cost of mineralizing the slag (0.04%) and the high value metals (99.77%), while the cost of the mechanical treatment by shredding contributes 0.15%.

Table 3. Cost-Revenue breakdown of current state-of-the-art process

	Value (\$)
25% Cost reduction	34,861,757
Improved cost	104,585,272
Revenue	139,928,701
Profit	35,343,428

Table 4. Cost breakdown of the improved process flow

Process	Cost (\$/year)	Percentage (%)
Bioreactor	61,205.76	0.04
Shredding	205,920.00	0.15
Slag mineralisation	52,992.92	0.04
Metals mineralisation	140,132,160.00	99.77
Total Cost	140,452,278.68	100
Reduction due to economy of scale (15%)	21,067,841.80	
Total	119,384,436.87	

Table 5. Revenue breakdown of the improved process flow

Revenue breakdown	Valuation (\$/year)	%
Carbon tax	1,414,011.40	1.50
Methane	19,225,728.00	20.38
Digester tipping	1,647,360.00	1.75
Mineralised metals	71,867,558.52	76.16
Slag Portland cement	203,218.02	0.22
Total	94,357,875.94	100

Table 6. Cost-Revenue breakdown of the improved process

Process economic parameters	Valuation (\$/year)
Total revenue	94,357,875.94
Cost	119,384,436.87
Loss	25,026,560.93

The total revenue from the improved metal recovery process focuses on applying a Circular Economy approach and the introduction of new revenue streams from the production of methane and slag Portland cement, in addition to the sale of stable

mineralized metals and the carbon credit achieved. As shown in Table 5, a total revenue of \$94,357,876 per year is achieved for the proposed process.

From Table 5, the revenue generated from methane and mineralized metal contributes 20.38% and 76.16%, respectively to the total revenue of the improved process. Additionally, the cost of mineralizing the metals contributes 99.77% to the total cost of the improved process (Table 4). This results in a loss of \$25,026,560 per year compared to a profit of \$35,343,428 from the Umicore process, as illustrated in Table 6. However, the improved process is able to use 46,270.01 metric tons of carbon dioxide per year.

4. Conclusion

This work analyzes the current state-of-the-art for the LIB recycling process from an economical perspective, by estimating its cost and revenue generated from the sale of the recovered metals. The results are used as a baseline to propose an improved conceptual pathway for the metal recovery process from both an environmental and economic perspective, with the introduction of new technologies that use carbon dioxide as input and improve the recovery efficiency.

This improved pathway for the recovery of metals from spent LIBs is efficient in recovering metals and recycling of organic waste which could have ended up in the landfill causing degradation of soil and water and increasing toxic and microbial levels in the environment. The mineralization of the recovered metals in the improved process uses up all the CO₂ from the digester, with captured carbon dioxide hypothetically making up for the deficit, with a total consumption of 46,270.01 metric tons per year. This serves as an environmental advantage in reduction of possible anthropogenic CO₂ released to the atmosphere compared to the current Umicore process.

Nonetheless, the new recovery pathway is more expensive compared to the commercially implemented process, with the mineralization of the high value metals taking up a huge percentage of the total cost. The revenue created by Circular Economy from the production of secondary materials (slag Portland cement) and reselling of the high value metals is not able to offset this high cost, which results in a loss of the overall proposed process. The high cost of mineralization can be reduced with improved investment in the integration of the chemical process into current carbon capture and storage technologies of current industrial processes.

References

- A. Azpagic, 1999, Life cycle assessment and its application to process, selection, design and optimisation, *Chem. Eng. J.* 73, 1-21
- T. Georgi-Maschler et al., 2012, Development of a recycling process for Li-ion batteries, *Journal of Power Resources* 207, 173-182
- L. Li et al., 2018, Economical recycling process for spent lithium-ion batteries and macro- and micro-scale mechanistic study, *Journal of Power Sources* 377, 70-79
- R. Sommerville et al., 2021, A qualitative assessment of lithium ion battery recycling processes, *Resources. Conserv. Recycl.* 165, 105219
- The World Bank, 2019. Solid Waste Management [Online] Available at: <https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management>
- L. Yun et al., 2018. Metallurgical and mechanical methods for recycling of lithium-ion battery pack for electric vehicles, *Resources. Conserv. Recycl.* 136, 198–208

Sensitivity Analysis and Risk Assessment for the In-Silico Design and Use of Optimized Cell Factories in a Xylitol Biorefinery

Nikolaus I. Vollmer,^a Krist V. Gernaey,^a Gürkan Sin^a

^a*Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 228A, 2800, Kongens Lyngby, Denmark*

Abstract

This work presents the results of investigating the design and use of optimized cell factories in a biorefinery. Based on a base-case process design created in an optimization-based framework, the cell factory models are analyzed through a sensitivity analysis to define and implement engineering targets. With the engineered cell factory models, flowsheet simulations are performed for a risk-based uncertainty analysis. The results show significantly improved economics of the biorefinery. A future outlook about integrating cell factory optimization into process design and optimization is given.

Keywords: Biorefinery, Mechanistic Modelling, Monte Carlo, Cell Factory Optimization

1. Introduction

Second-generation biorefineries are a key element in the transition towards more sustainable production processes and bio-based value chains. Despite harboring this immense potential, second-generation biorefineries are consensually hardly economically feasible (Ubando et al., 2020). Compared to first-generation biorefineries, the second generation commonly utilizes lignocellulosic biomass as feedstock, requiring an additional biomass pretreatment unit. This implies several challenges: firstly, the related capital and operational expenditures for the pretreatment unit directly affect the key performance indicators (KPIs) of the plant and, furthermore, both the sugar yield, as well as the yield of undesired inhibitor compounds created in the pretreatment influence the yield of the conversion processes and the downstream processing of the whole process (Vollmer et al., 2021b). Consequently, this also has a mediate effect on the KPIs, which often results in a poor overall economic performance.

On another note, there has been immense progress in cell factory optimization in the last decade. Firstly, rational design approaches in synthetic biology and genetic and metabolic engineering allow for optimizing yield, titer, and productivity of existing cell factories by manipulating metabolic pathways through, e.g., the overexpression or knockout of genes. Moreover, new cell factories can be created by inserting heterologous pathways by tools like, e.g., CRISPR-Cas9. This potential is currently unfolding and will further revolutionize biotechnology in the coming years. Secondly, irrational design approaches, e.g., adaptive laboratory evolution, can be employed to optimize cell factories by fast-forwarding the natural evolution process towards more robust or adapted cell factories for different operating conditions. Nonetheless, the translation of this paramount progress to bridge the gap into the process engineering domain for the design and optimization of

entire biotechnological processes is complex and highlights the need to integrate these approaches over different scales (Straathof et al., 2019)

In the scope of this work, a base-case process design for the production of xylitol and succinic acid is taken as the basis for investigating the potential of optimized cell factories in a second-generation biorefinery. For this, different tools from process systems engineering are used: Based on wild-type cell factories for the fermentation processes for both products, a sensitivity analysis is performed on both mechanistic models of the cell factories to identify targets for optimization in the cell factories to increase yield, titer, and/or productivities. Based on the results of the sensitivity analysis, the mechanistic models are redesigned accordingly to resemble the optimized cell factory. With these, a risk-based techno-economic analysis with a whole process model for the biorefinery is performed in order to investigate the economic potentials of the optimized-case process design. Ultimately, the conclusions of the analysis highlight potential future directions for the in-silico design of optimized cell factories and biotechnological processes to accelerate the transition towards bio-based value chains.

2. Methodology

2.1. Xylitol Biorefinery

2.1.1. Base-Case Process Design

The mentioned xylitol biorefinery was designed by Vollmer et al. (2021c) in an optimization-based framework – S3O – and assessed through a techno-economic analysis: The framework employs mechanistic unit operation models through surrogate models in the optimization procedure. The biorefinery involves a pretreatment unit for the lignocellulosic biomass, where the hemicellulosic fraction is separated and depolymerized. The resulting hydrolysate is upconcentrated in an evaporation unit. Subsequently, the sugars are used as substrate in a fermentation unit to produce xylitol. The downstream process involves an evaporation unit and two crystallization units. The solid residue of the pretreatment is subjected to enzymatic hydrolysis to separate the cellulosic fraction and depolymerize the sugars. These are used as substrate in a second fermentation process to produce succinic acid. The downstream process for succinic acid also involves an evaporation unit and two crystallization units. Lastly, the lignin fraction is combusted for steam and electricity generation to be integrated with the downstream processes (Vollmer 2021a, c). The flowsheet of the process is illustrated in Figure 1.

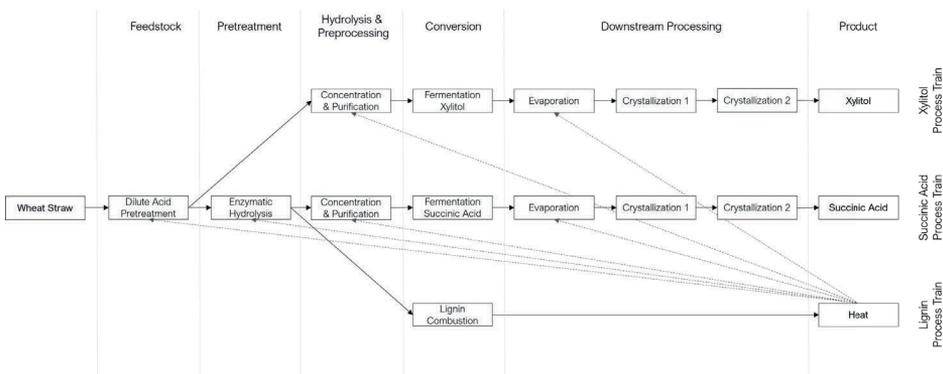


Figure 1: Flowsheet of the Base-Case Process Design of the Xylitol Biorefinery (adapted from Vollmer et al. 2021c)

2.1.2. Fermentation Models and Cell Factory Design

Both fermentation models are mechanistic models and based on a so-called black-box approach, referring to the reduction of intracellular reactions into three essential reactions, as depicted by Heijnen (2009): The first reaction describes the substrate uptake into the cell factory as shown in equation (1):

$$q_s = q_s \cdot \frac{C_s}{K_s + C_s} \cdot 1/I_s \cdot 1/I_p, \quad (1)$$

Here, q_s denotes the rate of substrate uptake, C_s the substrate concentration, K_s the substrate affinity constant of the cell factory and I_s and I_p terms for substrate and product inhibition. The second reaction is a Herbert-Pirt substrate distribution relation, summarizing the intracellular metabolism. Mathematically, this is described by equation (2):

$$q_s = a \cdot \mu + b \cdot q_p + m_s. \quad (2)$$

Here, a and b describe the yields of biomass and product over the substrate and μ the cell growth rate, q_p the product formation rate and m_s the cell maintenance rate.

Lastly, the third reaction is a Luedeking-Piret equation describing the product formation, given with equation (3):

$$q_p = \alpha \cdot \mu + \beta. \quad (3)$$

Here, α and β are fitted parameters, describing the product formation as a function of the cell growth (Heijnen, 2009). For the base-case process design, the model is fitted to wild-type strains of cell factories that naturally produce xylitol or respectively succinic acid. The reader is referred to the original work for a detailed description of the models and the fitted parameters (Vollmer et al., 2021c).

2.2. Monte Carlo Methods

2.2.1. Sensitivity Analysis

Sensitivity analyses attribute uncertainty in a model output to the different model inputs. In this work, we apply a variance-based sensitivity analysis based on Monte Carlo Simulations and implemented in the easyGSA package (Al et al., 2019). Both the first-order as well as the total sensitivity index are calculated. While the prior describes the sensitivity of the model output to uniquely this parameter, the latter describes the sensitivity of the output to the model parameter when considering all interactions with other input parameters (Vollmer et al., 2021b).

2.2.2. Uncertainty Analysis

The complementary to sensitivity analyses are uncertainty analyses trying to quantify the model output with varying model input. The methodology applied in this work is equally based on Monte Carlo Simulations. After defining input uncertainties and ranges, a sampling method is chosen, and Monte Carlo simulations are performed. The results are subsequently analyzed.

3. Results

3.1. Sensitivity Analysis of Cell Factory Model

As a first step, sensitivity analyses for both fermentation models are performed to determine the model parameters with the highest sensitivities. Each parameter can uniformly vary by $\pm 50\%$, equivalent to downregulating or overexpressing genes for specific pathways in the microorganism. For the xylitol model, $N = 8192$ and for the succinic acid model $N = 4096$ Monte Carlo samples are performed, using Sobol sampling. The results for the total sensitivity indices are shown in Figure 2.

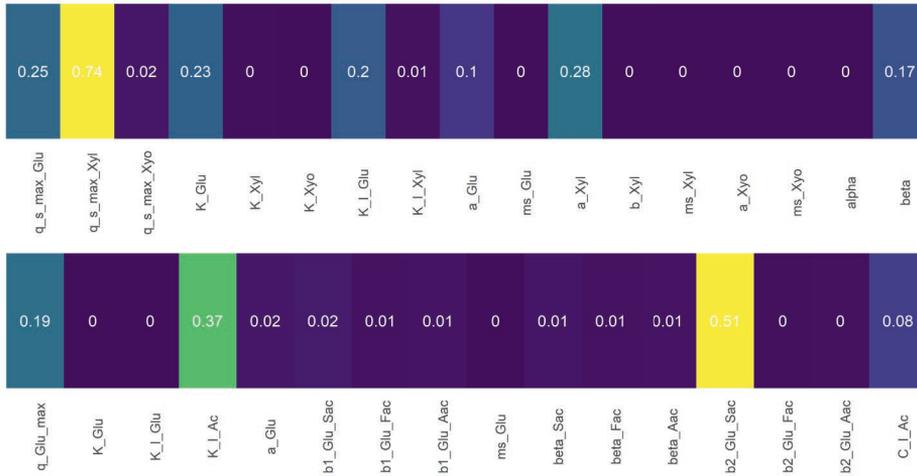


Figure 2: Heatmap of the total sensitivity indices for the fermentation model parameters.

For the xylitol cell factory, the most sensitive parameter is the substrate uptake rate of xylose. An increased uptake rate directly influences the productivity of the cell factory as the throughput through the cell factory increases as a function of the uptake rate. Secondly, also the operational yield of biomass over xylose a_{xyl} and the constant parameter of the product formation equation β influence the output, increasing the biomass growth and thus indirectly increasing the yield of xylitol over xylose. Lastly, the parameters related to the glucose uptake and growth on glucose influence the output, as the uptake of xylose is catabolically repressed by the glucose uptake, which influences productivity. For the succinic acid cell factory, the most influential parameters are the operational yield of succinic acid over glucose, increasing the yield of succinic acid directly, and the inhibition constant for the acids, which induces product inhibition in the cell factory. Secondly, the glucose uptake rate influences the output as increased substrate uptake also increases productivity.

3.2. Risk Assessment of Economic Feasibility

Based on the results presented in section 3.1, the following cell factory design is conceptualized and translated into the model parameters for retrofitting the fermentation models to engineered cell factories: The cell factory optimization for both models aims at increasing the yield of the respective product over the respective substrate. The product yields over substrate for both wild-type cell factories lie around 40 – 55%. The engineering strategy for the xylitol cell factory targets the growth metabolism on xylose, e.g., by overexpressing genes of relevant enzymes and the excretion of xylose out of the cell factory, e.g., by overexpressing the genes for the transporter. Mathematically, this is done by increasing the value of a_{xyl} and β by 75%. This translates to a new product yield of around 80% for the microorganism. The engineering strategy for the succinic acid cell factory targets the acid tolerance I_p to decrease the product inhibition, e.g., by performing adaptive laboratory evolution in media with low pH. Furthermore, the operational yield of succinic acid over glucose b_{Suc} is increased by e.g. knocking out respective genes for enzymes through a metabolic engineering approach. Mathematically, this is done by decreasing the value of I_p and increasing the value of b_{Suc} by 20%. This translates to a new product yield of around 80% for the microorganism. Both engineering strategies

have been proven to be realistic in different laboratory studies by the mentioned or similar engineering strategies (Hernández-Pérez et al., 2019 and Mancini et al., 2020).

Employing both retrofitted models in the S3O framework, an updated-case process design can be simulated to perform a new techno-economic analysis. Based on the results from Vollmer et al. (2021c), the four most promising candidate process topologies are investigated. Equally to the original study, uncertainties in the capital and operational expenditures (CAPEX and OPEX), including all equipment and utilities considered, and the product prices are assumed. The CAPEX varies in the interval $[-50, +100]\%$ and the OPEX varies in the interval $[-20, +50]\%$, assuming a triangular distribution for both. The product price of the xylitol varies between $4.29 \text{ \$/kg}$ and $5.42 \text{ \$/kg}$, and the succinic acid price between $3.18 \text{ \$/kg}$ and $3.24 \text{ \$/kg}$, assuming a uniform distribution for both. When performing an uncertainty analysis with these, $N = 1000$ Monte Carlo simulations and analyzing the net present value of the xylitol biorefinery as, the failure rates of investment result as $R_f = 48.9\%$, 86.9% , 92.4% and 55.5% (Vollmer et al., 2021c). Repeating the same uncertainty analysis with the engineered cell factory models and otherwise unaltered parameters, the metrics change significantly, as illustrated in Figure 3.

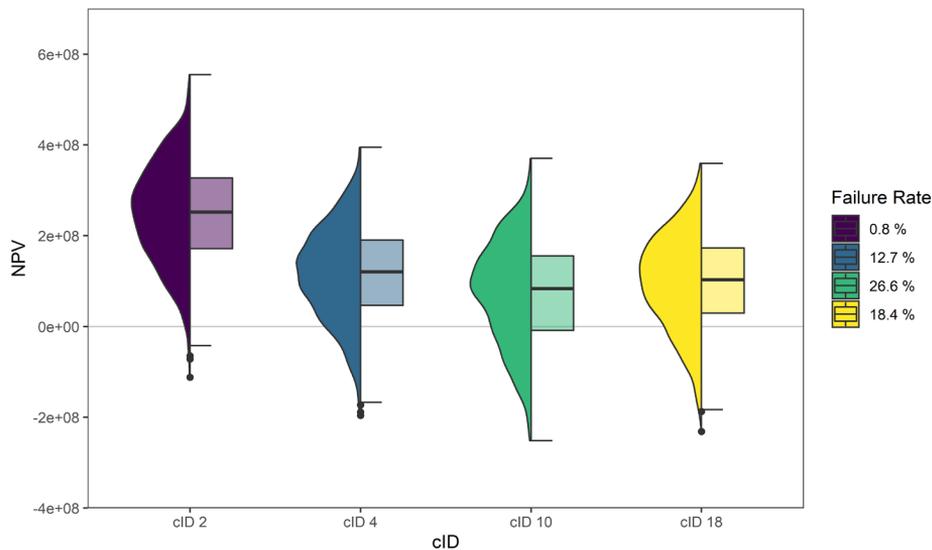


Figure 3: Results from the uncertainty analysis of the net present value of the xylitol biorefinery with engineered cell factories

Prominently visible is the increased economic feasibility for all four candidate process topologies. In particular, process configuration cID 2 shows an outstandingly low failure rate. This directly correlates with the increased yields of the engineered cell factories, consequently increasing the product sales without altering CAPEX or OPEX.

4. Conclusions

In this work, we showcased the potential in-silico design of cell factories based on mechanistic models for the use in a second-generation biorefinery. Based on a sensitivity analysis, the most sensitive model parameters of the cell factories were defined, and based on approaches in cell factory optimization, engineering targets were defined to increase

the product yield of the cell factories. The retrofitted cell factory models were used in flowsheet simulations to perform a risk-based uncertainty analysis on the net present value of the biorefinery to analyze the failure rate of the investment. Compared to the base-case process design, the optimized-case process design shows minimal failure rates, indicating a profitable investment.

Directions for future work could be the actual optimization of the model parameters by mathematical methods to guide potential cell factory optimization efforts. Ultimately, the used mechanistic models of the cell factories should be replaced by full genome-scale metabolic models, containing detailed information about the metabolism and potentially even information about transcriptional, metabolic, and other constraints. In Systems Biology, such models are constantly developed and improved and used for dynamic flux balance analysis to simulate fermentation processes (Sánchez and Nielsen, 2015). Integrating these models in full process models in a framework like the S3O framework would allow the simultaneous optimization-based design of cell factory, process, and value chain in a multi-scale approach. Ultimately, this approach allows bridging the gap between cell factory optimization and process design to accelerate the transition towards more bio-based processes and value chains for more sustainable industrial production.

Acknowledgments

The research project is part of the Fermentation-Based Biomanufacturing Initiative funded by the Novo Nordisk Foundation (Grant no. NNF17SA0031362).

References

- R. Al, C.R. Behera, A. Zubov, K.V. Gernaey and G. Sin. 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – An application to the BSM2 model. *Computers & Chemical Engineering*. Vol. 127. 233-246.
- J. Heijnen, Black-box models for Growth and Product formation, in: C. Smolke (Ed.). 2009: *The metabolic pathway engineering handbook – fundamentals*. CRC Press.
- A.F. Hernández-Pérez, P.V. de Arruda, L. Sene, S.S. da Silva, A.K. Chandel and M.G.A. Felipe. 2019. Xylitol bioproduction: state-of-the-art, industrial paradigm shift, and opportunities for integrated biorefineries. *Critical Reviews in Biotechnology*. Vol. 39 (7). 924-943.
- E. Mancini, S.S. Mansouri, K.V. Gernaey, J. Luo and M. Pinelo. 2020. From second generation feedstocks to innovative fermentation and downstream techniques for succinic acid production. *Critical Reviews in Environmental Science and Technology*. Vol. 50 (18). 1829-1873.
- B.J. Sánchez and J. Nielsen. 2015. Genome scale models of yeast: towards standardized evaluation and consistent omic integration. *Integrative Biology*. Vol. 7 (8). 846-858.
- A.J.J. Straathof, S.A. Wahl, K.R. Benjamin, R. Takors, N. Wierckx and H. Noorman. 2019. Grand Research Challenges for Sustainable Industrial Biotechnology. Vol. 37 (10). 1042-1050.
- A.T. Ubando, C.B. Felix and W.-H. Chen. 2020. Biorefineries in circular bioeconomy: A comprehensive review. *Bioresource Technology*. Vol. 299. 122585.
- N.I. Vollmer, R. Al, K.V. Gernaey and G. Sin. 2021a. Synergistic optimization framework for the process synthesis and design of biorefineries. *Frontiers of Chemical Science and Engineering*. In press.
- N.I. Vollmer, J.L.S.P. Driessen, C.K. Yamakawa, K.V. Gernaey, S.I. Mussatto and G. Sin. 2021b. Model development for the optimization of operational conditions of the pretreatment of wheat straw. *Chemical Engineering Journal*. In press. 133106.
- N. I. Vollmer, K.V. Gernaey and G. Sin. 2021c. Conceptual Process Design of an Integrated Xylitol Biorefinery with Value-Added Co-Products. *Frontiers in Chemical Engineering*. Under review.

Optimal design and planning of supply chains for viral vectors and RNA vaccines

Dauda Ibrahim,^a Zoltán Kis,^a Kyungjae Tak,^a Maria Papathanasiou,^a Cleo Kontoravdi,^a Benoît Chachuat,^a and Nilay Shah^a

^aThe Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, UK

Abstract

This work develops a multi-product MILP vaccine supply chain model that supports planning, distribution, and administration of viral vectors and RNA-based vaccines. The capability of the proposed vaccine supply chain model is illustrated using a real-world case study on vaccination against SARS-CoV-2 in the UK that concerns both viral vectors (e.g., AZD1222 developed by Oxford-AstraZeneca) and RNA-based vaccine (e.g., BNT162b2 developed by Pfizer-BioNTech). A comparison is made between the resources required and logistics costs when viral vectors and RNA vaccines are used during the SARS-CoV-2 vaccination campaign. Analysis of results shows that the logistics cost of RNA vaccines is 85% greater than that of viral vectors, and that transportation cost dominates logistics cost of RNA vaccines compared to viral vectors.

Keywords: SARS-CoV-2 vaccines, vaccination campaign, mathematical programming, economic analysis, vaccine availability.

1. Introduction

The COVID-19 pandemic has accelerated the research and development of new platform technologies for the production of vaccines against infectious diseases, including the novel corona virus, also known as Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2). Platform technologies such as viral vectors (Voysey et al., 2021) and RNA (Walsh et al., 2020) have been used to develop vaccine candidates to combat COVID-19. However, these new vaccines pose both logistics and distribution challenges. For example, unlike conventional vaccines, certain RNA-based vaccines require ultra-low temperature throughout the distribution network to avoid loss of potency.

Vaccine supply chain is a complex network that facilitates the transport of vaccines from manufacturing plants to administration points, such as GP surgeries, hospitals, pharmacies, clinics, and mass vaccination centers. The storage and transport conditions required by vaccines determine the type of cooling technology (e.g., fridge, freezer, or ultra-low freezer) to be installed at storage locations. During transportation, ultra-low temperature is maintained using thermal shipper loaded with dry ice (liquified CO₂).

The structure of a typical vaccine supply chain comprises manufacturing plants or import locations, fill-finish plants, warehouses/central stores, regional stores, and administration points. A manufacturing plant, aka primary manufacturing, consists of several unit operations used in the production of drug substance, which is the main ingredient in vaccines, while a fill-finish plant, also known as secondary manufacturing, inserts

vaccines (drug substances and excipients) into sterile glass vials or bags that are further packaged into cartons.

The design and planning of vaccine supply chains addresses the following: optimal selection of storage locations, production planning at manufacturing and fill-finish plants, inventory management/storage capacity planning, distribution planning, selection of routes and transport types, etc. In the recent past, Lee and co-workers (2014; 2015, 2016) have applied a simulation-based analytical tool, known as HERMES–Highly Extensible Resource for Modeling Supply Chains, to assess and re-design vaccine supply chains in low- and middle-income countries. However, HERMES does not support optimisation of supply chains, leading to solutions that could be suboptimal. Cavalho et al. (2019) developed a mixed-integer linear programming (MILP) model for optimal design and planning of a sustainable vaccine supply chain. The key performance indicators used to assess candidate supply chains are related to economic, environmental, and social performance. Kis et al. (2019) developed a supply chain model for the distribution of vaccine candidates such as RNA vaccines, outer membrane vesicle vaccines with genetically customizable membrane antigens virus-like particle vaccines with genetically configurable epitopes, and humanized yeast-produced vaccines. The model optimizes both supply-chain configuration and delivery type. Recently, Georgiadis & Georgiadis (2021) developed a MILP model for the distribution of COVID-19 vaccine and proposed an efficient method to solve the complex model. Nevertheless, the model does not account for quality control checks, fill-finish plants, production planning, selection of transport mode, and more importantly, management of vaccine thermal shippers.

This work develops a multiple-product MILP supply-chain model that supports the planning and distribution of viral vectors and RNA-based COVID-19 vaccines, from manufacturing plants to vaccination centers where vaccines are administered to targeted individuals. Unlike previous work, the proposed optimisation-based supply chain model comprises of five echelons, including manufacturing plants, fill-finish plants, imports, warehouses, regional stores, and administration points. A recycle loop for vaccine thermal shippers from warehouses to administration points and back to warehouses is implemented to ensure efficient management of vaccine thermal shippers. Including these entities within the supply-chain model allows efficient distribution of viral vectors and RNA-based vaccines, in addition to setting production targets at manufacturing and fill-finish plants. The performance of candidate supply chains is assessed using relevant key performance indicators such as vaccine availability, logistics cost, logistics cost per fully immunized patients, etc.

2. Methodology

2.1. Vaccination scheduling

Vaccination campaigns against infectious diseases are mostly preceded by scheduling of targeted individuals as well as healthcare personnel needed to administer vaccines. For the COVID-19 vaccination campaign, targeted individuals are selected according to risk of mortality and hospitalization when exposed to the highly contagious disease. Next, the targeted individuals are stratified, giving rise to cohorts to be schedule for the vaccine administration. The cohorts are prioritized according to risk-factors and vulnerability to COVID-19. Based on the COVID-19 vaccine regimen, the total number of doses required is estimated, and hence the vaccine demand profiles. By dividing the daily/weekly number of vaccinations by workload, it is possible to estimate the number of healthcare

workers that should be schedule on daily/weekly basis in order to ensure a successful vaccination campaign. This calculation may be carried out separately or embedded within the supply-chain model described in Section 2.2.

2.2. Design and planning of vaccine supply chains

The proposed vaccine supply-chain model supports the distribution and administration of vaccine candidates developed using various platform technologies, including viral vectors and RNA-based vaccines. The proposed supply chain comprises five echelons: manufacturing plants/imports, fill-finish plants, central stores, regional stores, and administration points (see Figure 1). Vaccines flow from manufacturing plants to vaccination centers via transportation modes, which can be a refrigerated van, a refrigerated truck, or an airplane. For vaccine candidates requiring ultra-low cooling temperatures, vaccine “thermal shippers” are used to ensure that the recommended temperature is not compromised during transportation. A recycle loop from warehouses to clinics and back to warehouses is implemented to enable efficient management of thermal shippers. The model inputs and outputs are summarised in Table 1.

Table 1. Vaccine supply chain inputs and outputs

Inputs	Outputs
i. vaccine demand profile	i. optimal supply chain configuration
ii. proposed supply chain structure (optional)	ii. transport mode per route
iii. minimum and maximum inventories (manufacturing and fill-finish, warehouses, regional stores, and administration points)	iii. backlog in each time period
iv. minimum and maximum capacities of manufacturing plants, fill-finish plants, and import rate	iv. vaccine availability and vaccine wastage at administration points
v. minimum and maximum capacities of transportation modes	v. vaccine supplied to administration points per time period
vi. operating cost and capital cost factors (manufacturing and fill-finish, warehouses, regional stores, and administration points)	vi. vaccine import rate and production rates in manufacturing and fill-finish plants
vii. travel distances and times	vii. capacity of quality control facilities
	viii. capital cost, operating cost, and total annualised cost of supply chain facilities
	ix. total transportation cost and transport cost per route
	x. inventories of vaccines in manufacturing, fill-finish, warehouses, regional stores, and administration points
	xi. inventories of vaccine thermal shippers (full and empty) in warehouses, regional stores, and administration points

Note that thermal shippers are needed only for vaccines stored and transported at ultra-low temperature. Also, constraints are included in the supply chain model to ensure that vaccines stored at clinics do not stay longer than their shelf-life. The objective function used to assess the performance of candidate supply chains include logistics cost, logistics

cost per fully immunized patient, and total cost of supply chain components. The model is built in GAMS, where the solver CPLEX is used to optimise the MILP model.

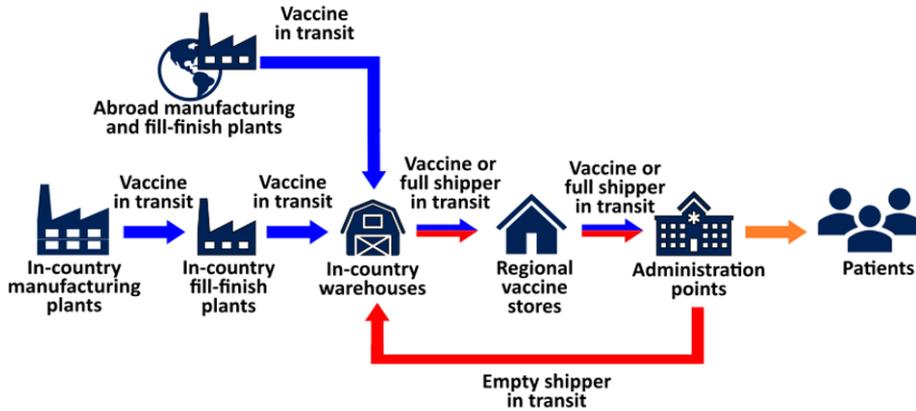


Figure 1. Schematic of the proposed vaccine supply chain comprising internal and external manufacturing and fill-finish plants, in-country warehouses, regional vaccine stores and administration.

3. Application

This section illustrates the capabilities of the proposed multi-product vaccine supply-chain model against COVID-19 in the UK, considering both viral vectors (AZD1222) and RNA vaccines (BNT162b2). The targeted individuals, as recommended by the UK Joint Committee on Vaccination and Immunization (JCVI), includes: care home residents, residential care workers, age 80 plus, health care workers, social care workers, age group 75-79, age group 70-74, clinically extremely vulnerable (under 70), age group 65-69, at risk (under 65), age groups 60-64, 55-59, and 50-54. For both vaccines, individuals are required to take two doses to be fully immunized. Vaccine doses are administered at least three weeks apart, leading to a vaccination timeframe of 38 weeks. The optimal supply chain distribution network is determined by the supply-chain model together with production targets at manufacturing and fill-finish plants, as well as vaccine inventory at storage locations, distribution planning, selected transportation mode, number of shippers and quantity of dry ice needed, etc. The total logistics cost and cost of supply chain components for viral vectors and RNA vaccines are shown in Figure 2 and Table 2, respectively.

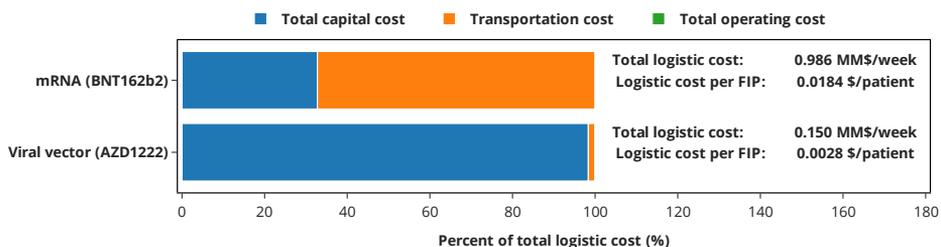


Figure 2. Total logistics cost required to deliver vaccines from manufacturing plants to administration points in England, Scotland, Wales, and Northern Ireland. The costs of precuring AZD1222 and BNT162b2 are not included, since logistics cost is the sum of total annualized capital cost, transportation cost, and total operating cost.

Figure 2 compares the logistics cost and logistics cost per fully immunized patient (FIP) for viral vectors and RNA-based vaccines. Note that there is no inventory of vaccines at warehouses and regional stores throughout the vaccination campaign, leading to zero operating cost, as shown in Figure 2. In this case study, vaccine inventory accumulates at administration points only. Note that this work considers a fixed supply chain structure in order to mimic the existing vaccine distribution network in the UK. Therefore at the design level, the supply chain model selects the transport type between echelons and their corresponding delivery frequency, which are discussed below.

From the results, the logistics cost of RNA vaccines is 6-fold larger than that of viral vectors. Similar trend is observed for logistics cost per FIP. Logistics cost is the sum of total annualized capital cost, transportation cost, and total operating cost. The high logistics cost observed for RNA vaccines is due to the large delivery frequency required to deliver sufficient vaccines to administration points during the vaccination campaign. RNA vaccine candidate, BNT162b2, has a shelf life of five days when stored at 2-8°C or kept in thermal shippers (without re-icing) and two hours at room temperature. Hence vaccine inventory can be kept to satisfy demand for five days only in order to avoid loss of potency and vaccine wastage. On the contrary, viral vector candidate, AZD1222, have a shelf life of six months at 2-8°C, thus warehouses, regional stores, and administration points can hold vaccine inventory for a longer period, consequently reducing the delivery frequency, but increasing the total annualized capital cost (see Figure 2). The increase in total annualized capital cost is due to the larger cold chain storage facility required to hold vaccine inventory. Another factor that leads to high transport cost for RNA vaccines is that BNT162b2 is produced in Pfizer-BioNTech manufacturing plant in Puur, Belgium, and have to be transported using an expensive transport mode (airplane) to warehouses in London, Edinburgh, Cardiff, and Belfast. On the other hand, viral vector candidate, AZD1222 is produced in the UK transported using refrigerated truck.

In addition to logistics cost and logistics cost per (FIP), the supply-chain model estimates the cost of other vaccine supply- chain components (see Table 2), including thermal shipper, dry ice, vaccinator wages, vaccine procurement, and quality control checks.

Table 2. Total cost of vaccine supply chain components such as vaccine thermal shippers, dry ice, vaccinator wages, vaccine procurement, and quality control checks.

Item (in million \$)	AZD1222	BNT162b2
Cost of vaccine shipper	-	7.72
Cost of dry ice	-	20.30
Cost of vaccinating individuals	1880.0	1880.0
Cost of vaccinating individuals at care home	24.10	24.10
Cost of vaccine procured	450.0	2000.0
Cost of quality control checks	59.90	59.90
Total cost	2420.0	3990.0

In Table 2, the total supply chain components cost for AZD1222 and BNT162b2 are \$2.42 and \$3.99 billion respectively. The selling price of BNT12b2, \$18 per dose, is higher than that of AZD1222, \$3 per dose, leading to 78% increase in the procurement cost of vaccine needed to vaccinate the entire UK target population of approximately 53 million individuals with two doses. AZD1222 does not require ultra-low cooling during

transportation, hence no thermal shipper and dry ice are required. Apart from the recycle loop for vaccine thermal shippers, the supply chain structure for ADZ1222 and BNT162b2 is similar. Also worth mentioning is that the storage technology for the two supply chains differs, i.e., refrigerators (2-8°C) for AZD1222 and deep freezers (-80°C) for BNT162b2. Deep freezers lead to high operating cost as a result of energy consumption needed for ultra-low temperature cooling.

4. Conclusions

A multi-product MILP vaccine supply chain model has been developed considering the essential features of a typical vaccine supply chain. The model can be used to design, plan, and optimize the distribution and administration of vaccine candidates developed using the most advance platform technologies, i.e., viral vectors and RNA. A case study compared the logistics costs when either viral vectors or RNA-based vaccines are used during a vaccination campaign against COVID-19 in the UK. The results show that the logistics cost of RNA-based vaccines (BNT162b2) is far greater than that of viral vector (AZD1222). Transportation cost dominates logistics cost of RNA-based vaccines as a result of high delivery frequency needed to supply sufficient vaccines to administration points. The long shelf life of viral vectors allows this vaccine type to be stored at administration points for a period of up to 6 months, consequently increasing capital cost related to storage facility.

References

- S. T. Brown, B. Schreiber, B. E. Cakouros, A. R. Wateska, H. M. Dicko, D. L. Connor, P. Jaillard, M. Mvundura, B. A. Norman, C. Levin, J. Rajgopal, M. Avella, C. Lebrun, E. Claypool, P. Paul & B. Y. Lee, 2014, The benefits of redesigning Benin's vaccine supply chain. *Vaccine*, 32(32), 4097–4103.
- G. P. Georgiadis & M. C. Georgiadis, 2021, Optimal planning of the COVID-19 vaccine supply chain. *Vaccine*, 39(37), 5302–5312.
- Z. Kis, M. Papatthaniou, R. Calvo-Serrano, C. Kontoravdi, & N. Shah, 2019, A model-based quantification of the impact of new manufacturing technologies on developing country vaccine supply chain performance: A Kenyan case study. *Journal of Advanced Manufacturing and Processing*, 1(3).
- B. Y. Lee, L. A. Haidari, W. Prosser, D. L. Connor, R. Bechtel, A. Dipuve, H. Kassim, B. Khanlawia & S. T. Brown, 2016, Re-designing the Mozambique vaccine supply chain to improve access to vaccines. *Vaccine*, 34(41), 4998-5004.
- B. Y. Lee, B. Schreiber, A. R. Wateska, D. L. Connor, H. M. Dicko, P. Jaillard, M. Mvundura, C. Levin, M. Avella, L. A. Haidari & S. T. Brown, 2015, The Benin experience: How computational modeling can assist major vaccine policy changes in low and middle income countries. *Vaccine*, 33(25), 2858-2861.
- M. I. Carvalho, D. Ribeiro & A. P. Barbosa-Povoa, 2019, Design and Planning of Sustainable Vaccine Supply Chain. *Springer Nature*.
- M. Voysey, S. A. C. Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, B. Angus, V. L. Baillie, S. L. Barnabas, Q. E. Bhorat, S. Bibi, C. Briner, P. Cicconi, A. M. Collins, R. Colin-Jones, C. L. Cutland, T. C. Darton, K. Dheda, C. J. A. Duncan, P. Zuidewind, 2021, Safety and efficacy of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet*, 397(10269), 99–111.
- E. E. Walsh, R. W. Frenck, A. R. Falsey, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, K. Neuzil, M. J. Mulligan, R. Bailey, K. A. Swanson, P. Li, K. Koury, W. Kalina, D. Cooper, C. Fontes-Garfias, P.Y. Shi, Ö. Türeçci, K. R. Tompkins, W. C. Gruber, 2020, Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *New England Journal of Medicine*, 383(25), 2439–2450.

Different approaches to epidemic modeling – The Covid-19 case study

Davide Manca^a

*^aPSE-Lab, Process Systems Engineering Laboratory, Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
davide.manca@polimi.it*

Abstract

The paper focuses on three different modeling approaches to describe the dynamic evolution of an epidemic. These models have been extensively tested in the field and applied to Italy, one of the countries that most suffered from Covid-19.

Keywords: Covid-19, pandemic, modeling, early-warning, vaccination policies.

1. Introduction

Covid-19 pandemic with its prolonged and iterated waves has caused emergencies at the hospital and political levels calling for responsible and prompt decisions. The “every day counts” saying is meaningful when one has to make choices and resolutions that have an impact on human lives. At the very beginning of Covid-19, there was no knowledge of the outbreak dynamics and rather simple questions about the expected allocation of resources for hospital wards and intensive care units (ICU) were difficult to answer. If one was able to answer those questions, the replies shocked both medical doctors and general managers of hospitals because the exponential growth of the pandemic phenomenon had never been experienced before in most continents of the world. The doubling time of infected, hospitalized, intubated, and dead individuals were unprecedented and often as short as two-three days.

Similar questions about the time evolution and intensity of the outbreaks were raised by local and central government representatives who were interested in making timely decisions to contain the spreading and relieve the pressure on all the organs responsible for the health and safety of citizens. Finally, the world of information focused most of its attention on the pandemic and on any news capable of predicting, quantifying, and assessing fragments of the whole framework. Mass media urged insistently on ever updated and critically analyzed pandemic data.

Most of the questions focused on what-if scenarios and resource allocation, pressure, and duration. The queries were first formulated for short-term horizons but quickly expanded towards medium- and long-term perspectives. The short-term horizon involved decisions for the allocation of hospital resources such as beds, wards, resuscitation personnel, and scarce material. The medium-term horizon called for decisions about non-pharmaceutical protective actions such as masks, swabs, social distancing, and lockdowns. These decisions had to find a balance between health/safety/treatment and economy/security/freedom of the population. Finally, both hospitals and governments called for a quantitative prediction of the pandemic on a long-term horizon to psychologically start seeing the light at the end of the tunnel (i.e. hope, perseverance, and resilience), to dynamically (re)allocate hospital means to elective medicine rather than

emergency treatments, and to loosen the strict lockdown rules in favor of a progressive return to normality. In the meanwhile, there were new requests for a quantitative prediction of pandemic-related themes when the first wave came almost to an end but suddenly a second outbreak caught most of the countries unprepared. Afterward, there have been further pandemic waves and some nations have experienced the fourth and fifth waves with the onset of new virus variants. What it was deemed to be understood became suddenly new and unknown. The lessons learned were no more sufficient and a new call for quantifying the ever-changing and multifaceted pandemic became evident and compelling. In the meanwhile, the first vaccines were released. Initially, for several months, the doses were few and the fragile population was exposed to more infectious and lethal virus variants (e.g., alpha, beta, delta, omicron to cite the most circulated). The vaccination problem was hyper-constrained as several categories claimed the right to be vaccinated first and every government adopted different sanitary policies. These governments chose different criteria to vaccinate the population and to inject the vaccine with different time intervals between doses. There was a new call for understanding the optimal vaccination policy and finding a suitable compromise between population coverage in number and category to reduce both the overall pandemic risk and mortality, as well as achieve more robust sanitary and public force systems. The paper shows how the PSE community with its tools, algorithms, models, and methods can cover, analyze, and quantify the different topics and issues that were listed above.

2. Short-term pandemic models

At the very beginning of the Covid-19 outbreak, the hospital heads of resuscitation wards were overwhelmed by SARS-CoV-2. The most stringent resource in hospitals was ICU beds. Resuscitation heads had often to double and even increase three-, four-, five-folds the number of beds in a few days alongside the resources connected (e.g., air-breathing lines, respiratory devices, monitors, space allocation, dirty and clean areas/paths). Human resources were the determining step, which called for specialized ICU doctors and nurses. Indeed, these specialized operators can be neither increased nor hired in the characteristic time of an epidemic. It calls for years of planning and formation. Regarding the dynamics of Covid-19 ICU beds, they start from zero at the very beginning of an epidemic (or from a positive known value in case of subsequent waves) and they increase steadily up to a maximum value. In the following, the ICU wards begin emptying and they reach either a final null value at the end of the epidemic or a minimum nonzero plateau before a new outbreak triggers the inversion of deflation with a new inflation trend. Similar qualitative behavior occurs also for daily positives, hospitalized, healed, and dead. Conversely, the total cumulated values of those quantities are curves that are monotonically increasing and reach either temporary or permanent plateaus between two waves or at the end of the epidemic respectively. Manca *et al.* (2020) proposed three different curves that describe the inflation trend of a pandemic wave: the exponential, the logistic, and the Gompertz functions (Eq.s 1-3). The mathematical description of the exponential model is:

$$y = a10^{bt} \quad (1)$$

The mathematical description of the logistic model is:

$$y = \frac{a}{1 + b \exp(-ct)} = \frac{a}{1 + b e^{-ct}} \quad (2)$$

The mathematical description of the Gompertz model is:

$$y = a \exp(-b \exp(-ct)) = a e^{-b e^{-ct}} \quad (3)$$

The adaptive parameters of Eq.s 1-3 can be evaluated through a nonlinear regression of real values that are periodically measured at both regional and national levels. For the sake of clarity, the logistic curve performs comparably and sometimes worse than the Gompertz model, which is the preferred one. Equally, the exponential model is very compact and effective at the very beginning of each outbreak but it soon overpredicts the epidemic dynamics. Hence, Gompertz is the most robust model and its analytical features allow also evaluating the inflection point, $t_{infl} = \ln(b)/c$, and the halfway point, $t_{halfway} = -\ln(\ln(2)/b)/c$. In addition, a is the asymptote of the dynamic inflation phenomenon. Once the plateau is reached, the curve starts decreasing and the reverse Gompertz function (Eq. 4) describes the deflation region:

$$y = a \left(1 - e^{-b e^{-c(t-t_0)}} \right) \tag{4}$$

Equally, the around-the-maximum period is well described by Eq. 5 which is an exponentially modified Gaussian (EMG) curve:

$$y = 10^{at^2+bt+c} \tag{5}$$

Figure 1 shows the typical dynamic behavior of both direct/reverse logistic and Gompertz functions, and the EMG curve for the ICU patients of Lombardy (the most populated region of Italy with over 10 million citizens) during the first pandemic wave.

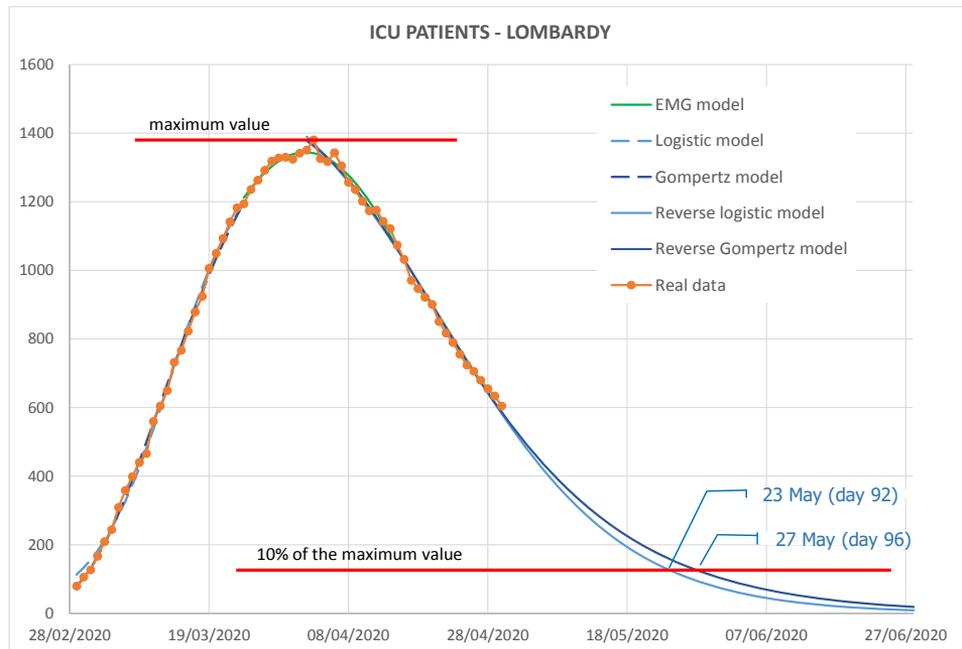


Figure 1: ICU patients in Lombardy and parametric models during the first Covid-19 wave.

3. Medium-term pandemic models

Another family of epidemic models can describe the contagion dynamics on a medium-term horizon. These models are based on interconnected compartments that summarize the condition of individuals in a given population. The simplest compartmental model for

an outbreak is the SIR model that is based on three compartments (susceptibles, infected, and removed) that describe the evolution of a population plagued by the epidemic. The population features N individuals, whose number is assumed constant (this hypothesis holds if N is sufficiently large and one assumes that the number of newborns equals that of fatalities and the prediction horizon is not too prolonged). Initially, $N-1$ individuals are susceptible (*i.e.* they belong to the S compartment of those vulnerable to the virus) and one person is infected and belongs to the I compartment of those who can transmit the contagion with a reproduction number, β . Once an infected individual either recovers or dies, they move to the R compartment, with a rate proportional to γ . Eq. 6 summarizes the SIR model based on an ordinary differential system with suitable initial conditions:

$$\begin{cases} \frac{dS}{dt} = -\beta \frac{I}{N} S & \frac{dI}{dt} = \beta \frac{I}{N} S - \gamma I & \frac{dR}{dt} = \gamma I \end{cases} \quad (6)$$

The adaptive parameters of the model (β, γ) can be identified by a nonlinear regression that minimizes the distance between the real data and the model predictions. The continuing availability of periodically measured real data allows improving the quality of the model and better predicting the dynamics of the pandemic. The complexity of the SIR model can be increased (Giordano *et al.*, 2020) by introducing further compartments (*e.g.*, those infected but not yet contagious, those who died) or contributions (those who have recovered but may lose their immunity and move back to the susceptible compartment). The adaptive parameters may change in time according to virus mutations, improved capacity of treating the infection, and different behavior of the population induced by nonpharmaceutical measures such as lockdowns, curfews, and social distancing. For the sake of simplicity, Eq. 7 shows some modifications to the adaptive parameters that account for lockdown measures and modification of the healing rate. These modifications call for the evaluation of additional adaptive parameters that pave the way to further simulations of what-if scenarios in case of government decisions such as non-pharmaceutical interventions.

$$\beta(t) = \begin{cases} \beta_0, t < t_{\text{lockdown}} \\ \beta_0 \exp\left(-\frac{t - t_{\text{lockdown}}}{\tau_\beta}\right) \end{cases} \quad \gamma(t) = \gamma_0 + \frac{\gamma_1}{1 + \exp(-t + \tau_\gamma)} \quad (7)$$

Compartmental models can be effective in predicting the expected dynamics of an epidemic. The so-called parametric runs allow quantifying the effect of different measures on the outcomes in terms of infected, recovered, and fatalities. Understanding the role played by adaptive parameters enhances the quality of decision-making aimed at prompt and effective resolutions for the health safety of citizens. Equally, compartmental models do not show a high degree of precision, rather they indicate the trend of the epidemic. However, their predictive capability usually increases towards the end of an outbreak once a good amount of real data allows refining the adaptive parameters.

A further differential equation can account for the vaccination of the population, which for most of the industrialized countries started in the middle of the second/third Covid-19 pandemic wave: $dV/dt = V_d M$. The first differential equation of system (6) must be modified accordingly to account for the individuals that once vaccinated are assumed not to be susceptible anymore and therefore neither recover nor die but move directly to the V compartment: $dS/dt = -\beta S \cdot I/N - V_d M$. V_d is the vaccination rate and M is the

fraction of individuals who were not infected and are vaccinated. V_d is a dynamic input value that measures the capacity of the system (either regional or national) to daily inject the vaccine and decrease (after one or two doses) the number of susceptible individuals (Figure 2). Again, the model can be improved by removing some oversimplifications and accounting for reinfections and partial protection from infection as a function of the number of doses and vaccine maker. Parametric runs can quantify the effect of increasing the rate of vaccination and population coverage by computing the number of saved lives.

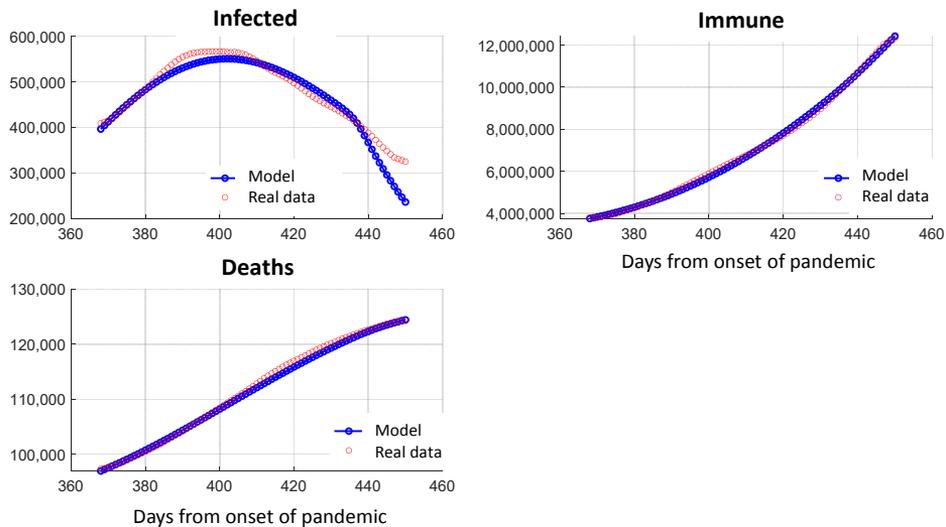


Figure 2: SIRDV model based on susceptible, infected, recovered, dead, and vaccinated compartments. The immune diagram collects both recovered and vaccinated individuals. Italy, third Covid-19 wave.

4. Long term predictions

Equations (2-5) can be extrapolated to long-term horizons to determine when the asymptotic condition for a maximum or minimum plateau is reached. Equally, it is possible to determine the time taken to reach and overtake a given threshold for both the inflation and deflation phases. These last bits of information are valuable when hospital managers have to either prepare for emergencies (during the inflation period) or progressively allocate resources to return to elective medicine (during the deflation period). Likewise, Equation (6) allows evaluating the asymptotic condition for rather long integration times. However, the user should be aware of the risks connected to the extrapolation of long-term models that depend heavily on adaptive parameters.

5. Alternative vaccination policies

In the case of a vaccination campaign against an epidemic, initially, the available doses are limited and a suitable vaccination policy should be chosen and implemented. The most effective indicator to account for is the number of fatalities. If the decision-makers try to minimize the fatalities toll then the pressure on the limiting resources of hospitals such as ICU beds and Covid-19 wards gets relieved and both medical doctors and nurses can cure the patients more efficiently, which results in a synergistic reduction of mortality. In the case of Italy, the first available doses (January 2021) were reasonably administered

to high-priority categories (such as hospital operators, health doctors, and fragile subjects). Nonetheless, it is well-known that Covid-19 mortality follows a monotonically increasing curve with age. In Italy, almost 99% of the officially documented 133,000+ victims of Covid-19 (as of November 2021) were people aged 50 and older. Unfortunately, the following doses did not strictly respect a reverse-order of age policy aimed at minimizing the fatalities. There were some diversions to lower priority categories that, thanks to their health status and age, were less exposed to either serious or fatal consequences (e.g., school and university personnel, armed forces, lawyers, scientific informants, social workers, customs and airport personnel, and funeral home staff).

An *in silico* vaccination simulator can quantify the difference between real fatalities and what might happen if an alternative vaccination procedure were adopted. That simulator relies on the immunization degree after injection (Creech *et al.*, 2021) and the death distribution curve, which both allow assessing the expected number of saved lives. After having covered the high-priority categories, if one implements a rigid approach to vaccination based on the reverse order of age ranges (i.e. the elderly first, the young last) every dose can be injected to individuals in the following decreasing sequence of 90+, 80-89, 70-79, 60-69, 50-59, 40-49, 30-39, 20-29, and 12-19 years old. The rate of injected doses is available daily and the optimal reallocation of doses allows evaluating the number of lives that would be saved. In addition, the simulator can solve what-if scenarios such as what would happen if (i) the daily delivered doses increased by some percent, or (ii) the vaccination campaign to the elderly were anticipated of a few days, or (iii) some doses were diverted from lower priority categories to the elderly, or (iv) the interval between first and second dose were increased or decreased. Table 1 reports some results of different vaccination policies on the expected number of saved lives.

Table 1: Saved lives in Italy by alternative vaccination policies. Rows report the expected saved lives under the hypothesis of administering the same amount (100%) or 10% and 20% more daily doses than those injected in reality. The columns report the calculated saved lives under the three hypotheses of 0, 10, and 20 days anticipation with respect to reality (which started on 18-Feb-2021). The numbers in brackets quantify the increase over the basic case study of 18-Feb at 100% administrations. The assessment was carried out at the end of May 2021.

Saved lives – ITALY	Start on 18-Feb-2021	Start 10 days before	Start 20 days before
100% of the real administrations	3969	4465 (+12.50%)	5317 (+33.96%)
110% of the real administrations	4905 (+23.58%)	5414 (+36.41%)	6311 (+59.01%)
120% of the real administrations	5712 (+43.92%)	6245 (+57.34%)	7083 (+78.46%)

References

- Creech, C. B., Walker, S. C., & Samuels, R. J. (2021). SARS-CoV-2 Vaccines. *JAMA - Journal of the American Medical Association*, 325, 1318-1320.
- Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., & Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 26, 855-860.
- Manca, D., Caldiroli, D., & Storti, E. (2020). A simplified math approach to predict ICU beds and mortality rate for hospital emergency planning under Covid-19 pandemic. *Computers and Chemical Engineering*, 140.

An integrated and sustainable Ethanol-Hydrogen-Ammonia-Urea plant

Artur S. Bispo^a, Leonardo O. S. Santana^a, Fernando L. P. Pessoa^a, Ana L. B. de Souza^a, Ewerton E. S. Calixto^a

^a*Centro Universitário SENAI-CIMATEC, Avenida Orlando Gomes, 1845 Piatã, Salvador 41613-036, Brazil.*
Email: Artur.bispo@outlook.com.br

Abstract

Considering the growing need for alternative and cleaner fuels to replace the ones based on fossil energy sources, and to reduce the environmental impacts associated, this work aims to present a novel process that replaces natural gas (NG) as a source of hydrogen (H₂) in the production of ammonia (NH₃). The generated CO₂ is fed to the urea synthesis unit, which reduces the greenhouse gas (GEEs) emissions through an integrated ethanol-ammonia-hydrogen-urea plant. To achieve these goals, an integrated plant was designed where the H₂ from NG was replaced by the H₂ from the ethanol reform ("moss hydrogen") and electrolysis process (green H₂) where it receives electricity from the co-generation ethanol system. The capturing process of the CO₂ produced during the fermentation and bagasse burning is carried out in the sugarcane photosynthesis. Besides, part of the CO₂ emitted during the ammonia section (in the ethanol reform stage) is used as feed in the urea plant, supplying all the demand for the urea production from the outputs of the process itself and minimal carbon emission to the atmosphere ($\approx 44\%$). Finally, after an evaluation performed in the Aspen Plus process simulator and based on metrics used in green chemistry associated with sustainability, it was concluded that the proposed plant can be considered sustainable with great potential, especially in regions rich in biomass.

Keywords: Ammonia, Ethanol, Urea, Green hydrogen, CO₂ emission.

1. Introduction

The path to a low carbon economy involves the development of new processes and the use of new raw materials to produce all the necessary products to maintain a modern standard of living, meeting the sustainability tripod. Thus, in addition to fuels and other chemical products that used to come from fossil sources, new routes to produce fertilizers, such as NH₃ and urea, must be developed. (Konig, et al. 2015)

A renewable source of chemicals and energy is biomass, where Brazil is one of the world leaders in the use it, especially sugarcane, used in the production of ethanol, sugar and electricity through the burning of bagasse and vinasse. Cogeneration of energy in sugarcane mills plays a fundamental role in the sustainability of the production process, which remains independent of external fuels. On the other hand, electricity is often generated in excess, and sold to the local energy distributor. (Mehmeti, et al. 2018)

Currently, according to Mehmeti, et al. (2018), the ammonia and urea feedstocks are natural gas, CO₂ and air. Developing new technological routes, with sustainable sources of CO₂ and hydrogen, is necessary for a cleaner and more sustainable process. This work aims to present a new route to produce green ammonia and green urea from clean energy sources widely available in Brazil, (biomass energy and sugarcane ethanol). (Erdemir & Dincer, 2020)

The hydrogen sources come from the electrolysis of water, using the excess of electrical energy generated in the plant, and ethanol, which undergoes a steam reforming process. The source of CO₂ is ethanol (EtOH), also through steam reforming, implementing the capture and storage process within the plant. Nitrogen continues to come directly from air. This new concept plant was modeled and simulated in Aspen Plus v.12. (Erdemir & Dincer, 2020)

2. Methodology

In this work, the methodology consisted, firstly, in carrying out the plant simulation of the proposed process, which integrates the stages of production of ethanol, green H₂, ammonia and urea and whose main objectives are to change the source of H₂ to a more sustainable one and the reduction of CO₂ emission. Then, an analysis was carried out to assess and quantify the proposed plant's capability to reduce CO₂ emissions.

The simulation only includes the ammonia, green H₂ and urea production sections (figure 1). As for the ethanol section, it was considered a plant with the capacity to process 1150 ton/h of sugar cane, producing 1766 kmol/h of ethanol and 287.5 ton/h of bagasse. In the following topics, the step-by-step performed to simulate the green H₂, ammonia and urea production sections will be elucidated. (Wirti, 2016)

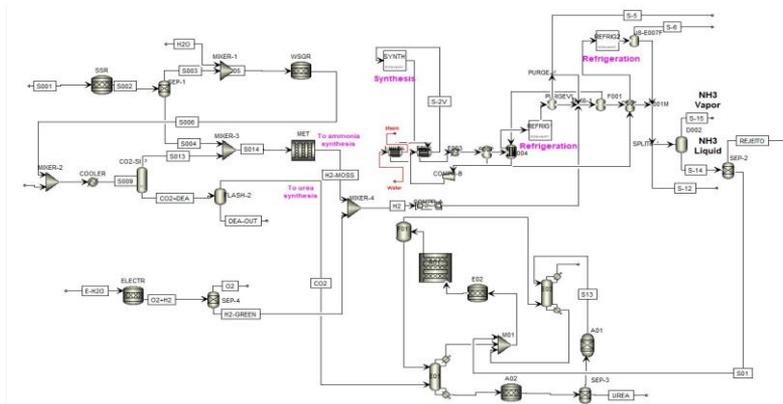


Figure 1. Integrated process flowsheet

2.1. Green H₂ Unit - Simulation

Green H₂ is produced through the alkaline electrolysis of water which, in this case, was simulated using an RSTOIC (Stoichiometric Reactor Module), named "ELECTR" and an SEP (Component Separation Module), named SEP-4.

In the reactor, the reaction shown in equation 1 was added and the yield attributed to the fractional conversion of H₂O into H₂ and O₂ was 99.9%, according to Faria (2018). In addition, the reaction temperature and pressure conditions were 80°C and 1 bar and the electricity consumption required by the electrolyzer (ELECTR) per Nm³ of H₂ produced is around 4.3 kWh.



2.2. Ammonia Unit - Simulation

According to Garcês (2021), ammonia is produced from the Haber-Bosch reaction (equation 2), which consists of reacting H_2 and N_2 to form NH_3 . However, on an industrial scale, some additional steps are necessary for this synthesis.



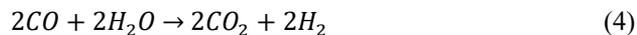
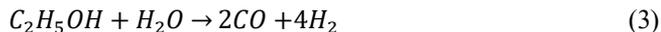
In this plant, different from the technologies commonly found in the literature ammonia production (BASF (1910), Kellogg (1960), KBR (Kellogg modified by Braun), Haldor Topsoe), the hydrogen source is the green H_2 process – produced from the electrolysis of water – and “moss” H_2 – produced from the reform of ethanol. (Pattabathula et al., 2016)

Therefore, to explain the ammonia section simulation process, the explanation will be divided into two parts: ethanol reforming and ammonia synthesis.

2.2.1. Ethanol Reforming

The ethanol reforming process consists of generating H_2 from ethanol to produce ammonia and, in addition, capturing and recovering the produced CO_2 , so that it can be used to in urea production plant.

The process basically takes place in two stages (SSR and WSGR), where the reactions shown in equations 3 and 4 occur, respectively, the first at high temperatures ($650^\circ C$) and the second at relatively lower temperatures ($200^\circ C$). Furthermore, according to Teixeira (2016), the fractional conversions for both reactions are, respectively, 93.7% and 99.5%.



The stream rich in CO_2 , H_2 and H_2O is sent to the CO_2 capture and recovery, as carbon dioxide is extremely poisonous to the ammonia production catalyst and needs to be separated from H_2 . (Sunny et al., 2016)

The stripping strategy adopted in the work was the capture of CO_2 by alkaline amine, more specifically DEA (Diethanolamine). The process was simulated in a “Hierarchy” block, composed of two RSTOICS and a FLASH, as seen in figure 2.

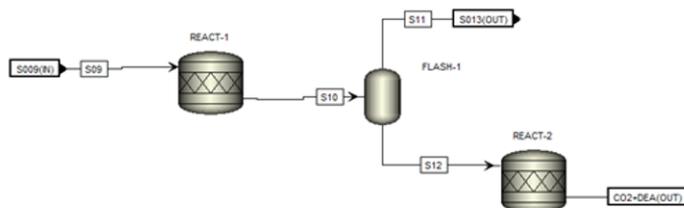
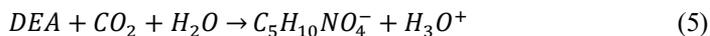


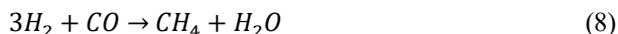
Figure 2. CO_2 Removal

As mentioned by Oh et al. (2010) and Mendieta (2011), in the first reactor (REACT-1), at low temperature and pressure conditions, the reaction between water, carbon dioxide and diethanolamine occurs (eq. 5), followed by a flash drum that separates H_2 from the other product components and, finally, in the REACT-2, the inverse reaction of the one

described above occurs (eq. 6), due to the high established temperature and pressure conditions.



Next, the separation of CO₂ from the DEA takes place, so that it can be directed to the urea plant, in addition, due to the residues of carbon oxides in the H₂ stream (S014), it goes to a methanation step, where the concentrations of CO₂ and CO are drastically reduced to the order of 0.2% and 0.02% (mol/mol), respectively. The reactions of the step can be seen in equations 7 and 8. (Sunny et al., 2016)



2.2.2. Ammonia Synthesis

As previously mentioned, the main reaction that occurs in the synthesis step is the Haber-Bosch (eq. 2), reacting the H₂ - obtained in the ethanol reforming and in the production section of Green H₂ - and N₂ removed from the atmosphere. This step uses Haldor Topsoe's S-250 ammonia conversion strategy, which consists of a two-bed radial flow converter with indirect heat exchange between the two beds. In this way, efficient use of the converter volume, low pressure drops, and high conversion occurs - due to indirect cooling. In these reactors, the NH₃ formation reaction was defined through the subroutine "NH3SYN", found in the software files. (Sunny et al., 2016)

Finally, the ammonia fed into the refrigeration unit is liquefied and goes direct to its intended purpose, whether it is for storage or, in this case, for feeding the urea plant.

2.3 Urea Section - Simulation

Finally, the urea synthesis step starts, simulated from the example flowsheet present in the Aspen Plus files. However, some changes were made to this flowsheet, such as: considering the formation of reactor biuret (unwanted by-product); add pure component - using the NIST database - and binary interaction parameters - defined as the same as urea - for biuret, in addition to DHFORM and DGFORM data; add vapor pressure data - PLDTDEP - for urea and biuret; finally, define the urea and biuret formation reactions as "POWERLAW", placing their kinetic data according to Chinda (2019).

3. Results

Having carried out all the methodology previously presented, obtaining the simulation of all the steps of the proposed process, an evaluation of the achieved results was carried out.

As mentioned, the ethanol production plant used as a base a processing capacity of 1150 ton/h of sugarcane, producing around 1766 kmol/h of ethanol and 287.5 ton/h of bagasse. All the bagasse was used to generate energy, first to supply the plant itself and then to produce Green H₂ by electrolysis. It is noteworthy that the produced CO₂ released into the atmosphere at this stage - from the fermentation and burning of bagasse - does not

enter the final balance, since the sugarcane manufacture itself will absorb it. (Oh et al., 2010)

After burning the bagasse and supplying the ethanol plant, 172.5 MW of energy were directed to the electrolyzer, being able to produce, with this amount of energy, 1790.8 kmol/h of green H₂ - entirely directed to the synthesis of ammonia. (Faria, 2018)

The ethanol produced is directed to the reform, so that the moss H₂ is produced. The table 1 shows the main composition data for this section, prior to ammonia synthesis.

Table 1. Streams composition – Ammonia section

Comp.	Flow – kmol/h								
	SRR		WSGR		CO ₂ - SEP			MET	
	In	Out	In	Out	In	Top	Bottom	In	Out
Ethanol	1766,0	111,3	-	-	-	-	-	-	-
H ₂	-	6619,0	-	3292,9	3292,9	3276,0	16,9	9895,0	9194,7
CO	-	3309,5	3309,5	16,5	16,5	16,2	0,3	16,2	1,8
CO ₂	-	-	-	3292,9	3292,9	182,5	2999,2	182,5	18,3

Then, all hydrogen produced - approximately 11000 kmol/h, adding the streams of "green H₂" and "moss H₂" - is sent to the ammonia synthesis, where it reacts with the N₂ removed from the atmosphere and, through the Haber-Bosch reaction, it forms NH₃, which is then sent to the synthesis of urea. At the end of the process, the ammonia section releases 2624.92 kmol/h of NH₃ and 2999.25 kmol/h of CO₂. (Garcês, 2021)

Due to the integration of the ammonia and urea sections, the amount of carbon dioxide that would be released into the atmosphere is reduced, since part of the CO₂ produced is transferred so that urea is produced. Considering that the NH₃ feed from the urea section is 2624.92 kmol/h, the simulated plant is capable of processing about 1317 kmol/h of greenhouse gas, reducing its emission by approximately 44%. The table below shows the overall balance of the proposed integrated plant.

Table 2. Overall Balance

Green H2 Section		Ammonia Section		Urea Section	
H ₂ O	In	EtOH	In	NH ₃	In
	1790,8 kmol/h		1766,0 kmol/h		2624,9 kmol/h
Energy	In	CO ₂	Out	CO ₂	In
	172,5 MW		2999,2 kmol/h		1316,9 kmol/h
H ₂	Out	NH ₃	Out	UREA	Out
	1790,8 kmol/h		2624,9 kmol/h		1312,0 kmol/h

Furthermore, it is worth noting that, in the ethanol steam reforming step, approximately 8300 kmol/h of water were processed, used both in the SSR and WSGR steps and in the capture of CO₂ by DEA. Besides, simultaneously with the production of urea fertilizer, there is also the production of $8.5 \cdot 10^{-3}$ kmol/h of biuret which, although low, must be considered (given the harmful character of the by-product for plants).

4. Conclusion

Finally, it can be concluded that, from the data found in the literature (operational conditions, reaction yield, mass, and energy balances of existing plants), it was possible to carry out the simulations of the proposed plant that integrated the production sections

of green H₂, green ammonia - produced from hydrogen from the reform of ethanol - and urea. As a result, it is observed that the results were satisfactory, showing the possibility of a plant with an alternative source of hydrogen (moss H₂) for the ammonia production, in addition to being able to reduce 44% of the CO₂ emission of this step and use the waste from the ethanol plant (bagasse) as an energy source to produce green H₂.

Acknowledgments

To SENAI CIMATEC for the financial support and the simulation software used. Also, to CNPq and the FAPERJ Institute for financial support.

References

- A. Carolina Rodrigues Teixeira et. al, 2016. UMA REVISÃO SOBRE A REFORMA DE ETANOL A VAPOR COM FOCO NA PRODUÇÃO DE HIDROGÊNIO. Programa de Pós-Graduação em engenharia mecânica, Pontificia Universidade Católica de Minas Gerais. Belo Horizonte.
- A. Mehmeti, et al. 2018. Life cycle assessment and water footprint of hydrogen production methods: from convencional to emerging technologies. *Environments*, 5, 24.
- A. Sunny, P. A. Solomon, K. Aparna, 2016. Syngas production from regasified liquefied natural gas and its simulation using ASPEN HYSYS. *Journal of Natural Gas Science and Engineering*, v. 30, p. 176 – 181.
- D. Erdemir & I. Dincer, 2020. A perspective on the use of ammonia as a clean fuel: challenges and solutions. *International journal of energy research*, 4, 45.
- D. Goulart Faria, 2018. Captura, armazenamento e utilização de dióxido de carbono na indústria de cimento. Dissertação de mestrado. Departamento de engenharia química. Universidade federal de Minas Gerais, Minas Gerais.
- D. Konig, et al. 2015. Simulation and evaluation of a process concept for the generation of synthetic fuel from CO₂ and H₂. *Proceedings of the ICE- energy*, 91, 833-841.
- J. Oh, D. Jeong, I. Yum, Y. Lee, 2010. A numerical analysis for CO₂ recovery from aqueous absorbent solution by hollow fiber membrane contactor. *Desalination and Water Treatment*, v. 17, p 218 – 226.
- J. Garcia, 2010. Emissão de gases de efeito estufa no ciclo de vida do etanol: estimativa nas fases de agricultura e industrialização em Minas Gerais. *Eng Sant Ambient*, v. 15, n. 3, p. 217 – 222.
- L. Hereda Garcês, 2021. SIMULAÇÃO, OTIMIZAÇÃO E INTENSIFICAÇÃO DO PROCESSO DE PRODUÇÃO DE AMÔNIA. VI SAPCT. Salvador, Brazil.
- L. Mendieta, 2011. Estudo da absorção de H₂S e CO₂ do gás combustível. Trabalho de conclusão de curso. Universidade Federal do Rio Grande do Sul. Departamento de engenharia química. Rio Grande do Sul.
- R. Chinda, 2015. Simulação da seção de síntese de uma unidade de produção de ureia – processo stamicarbon. Dissertação de mestrado. Programa de pós-graduação em tecnologia de processos químicos e bioquímicos. Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- R. Chinda, 2019. Process intensification applied to urea production process. Tese de doutorado. Programa de pós-graduação em engenharia de processos químicos e bioquímicos. Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- S. G. Wirti, 2016. Balanço de massa e de energia e estimativa da composição dos resíduos do setor de geração de energia a partir do bagaço de cana-de-açúcar em uma usina de açúcar e álcool, Karlsruhe Institut für Technologie, UFRGS.
- V. Pattabathula, J. Richardson. Introduction to Ammonia Production. CEP – Back to Basics. American Institute of Chemical Engineers (AIChE), p. 69 – 75.

A methodology based on social life cycle assessment for social hotspots identification

Andreia Santos,^a Madalena Veloso,^a Ana Carvalho,^a Ana Barbosa-Póvoa^a

^a*Centre for Management Studies of Instituto Superior Técnico, Av. Rovisco Pais, 1049-101 Lisboa, Portugal*

Abstract

Although there are studies where the economic and environmental impacts of different cork supply chains were assessed, no study exists where the social impacts of these supply chains have been analyzed. The main goal of this study is to narrow this research gap by assessing the social performance of the Portuguese natural cork stoppers' supply chain, using the social life cycle assessment methodology. Furthermore, this study provides suggestions on how this social performance can be improved based on the processes and regions where the most relevant potential social impacts are likely to arise (i.e., social hotspots). Five social issues were recognized as relevant and one hotspot, Cork harvest in Portugal, was identified for these issues. Based on these results, a set of industry-specific recommendations is provided for Portuguese cork suppliers to improve their social performance. If implemented, these recommendations will contribute to three Sustainable Development Goals, namely Goal 4 – Quality Education, Goal 8 – Decent Work and Economic Growth, and Goal 16 – Peace, Justice, and Strong Institutions.

Keywords: Sustainable supply chain management; Social life cycle assessment; Social hotspots; Sustainable Development Goals; Natural cork stoppers supply chain.

1. Introduction

The economic and environmental dimensions of sustainable development have been comprehensively covered in the field of sustainable supply chain management (SSCM), while the social dimension is addressed in a fairly simplified manner (Seuring, 2013). Similarly, a recent review found that the social dimension of sustainability is the most overlooked in assessment studies applied to forest wood supply chains (Santos, et al., 2019). While some studies have attempted to narrow this research gap, including the study conducted by Santos et al. (2020) where the social impacts of a Portuguese pulp and paper supply chain were quantified, there is still work to be done. For example, no study has been conducted where the social performance of a cork product supply chain has been assessed. Hence, the main goal of this study is to assess the social performance of the Portuguese natural cork stoppers' supply chain since the cork industry is one of Portugal's biggest industries and natural cork stoppers (NCS) are the most manufactured product in this industry (ICNF, 2018). Social life cycle assessment (LCA) was chosen in this study to assess the social performance of the selected supply chain because it has been identified as a promising approach to assess the social performance of forest wood supply chains (Santos, et al., 2019). Furthermore, this study also provides suggestions on how the social performance of the Portuguese NCS supply chain can be improved by identifying the social hotspots associated with this supply chain. Social hotspots are processes in a region where the most relevant potential social impacts are likely to arise (UNEP, 2020). Thus, identifying these hotspots allows determining where changes need to be made to reduce

the most critical social issues associated with a supply chain. The suggestions provided will assist natural cork stoppers' manufacturers and other supply chain partners dealing with the relevant social issues highlighted in improving their social performance and, as a result, the social performance of their supply chain. By implementing these suggestions, the different NCS supply chain partners will address critical social issues and help achieve specific Sustainable Development Goals (SDGs) (United Nations, 2021).

2. Social Life Cycle Assessment

The social LCA methodology is used to analyze the potential social impacts associated with the life cycle of a system (i.e., the specific object under analysis). This methodology is comprised of four steps (ISO, 2006): Goal and Scope Definition, Life Cycle Inventory (LCI), Life Cycle Impact Assessment (LCIA), and Results Interpretation. These steps are explained below together with an explanation of how they were applied to the supply chain under analysis.

2.1. Goal and Scope Definition

The first stage of this step is to establish the goal of the social LCA. In this study, the goal was to assess the social performance and identify the social hotspots associated with the supply chain of natural cork stoppers manufactured in Portugal.

The second stage of this step is to establish the scope of the social LCA which requires establishing the functional unit and system boundary. The functional unit is a representative element of the system(s) under analysis (EC-JRC-IES, 2010). One ton of natural cork stoppers was the functional unit selected in this study. The system boundary defines which part of the supply chain and respective processes will be considered (EC-JRC-IES, 2010). Given the goal previously defined (i.e. social hotspots identification), it is very important to select a boundary that considers the entire supply chain since choosing a restrictive system boundary could prevent the social hotspots from being identified if these were located outside the boundary. Hence, the system boundary chosen was cradle-to-grave, which includes four distinct life cycle stages: Raw materials' extraction, Products' manufacture, Products' distribution, and Products' end-of-life.

2.2. Life Cycle Inventory

The purpose of the second step is to collect a list of social flows (e.g., number of worker hours in bad working conditions) which will be converted into potential social impacts in the following step. To determine these flows, the system(s) under analysis must be modeled using foreground (i.e., specific) and background (i.e., generic) data related to the different life cycle stages included within the system boundary previously defined. Foreground data refers to data that has been collected for the specific system under analysis (UNEP, 2020) and it usually includes information on (Norris, et al., 2019):

- Materials (i.e., raw materials, utilities, and transports) used in the different life cycle stages included within the system boundary;
- Economic sectors (GTAP, 2019) to which the materials belong to;
- Region (e.g., country) from which the materials are sourced from;
- Cost of the materials.

Background data refers to data that has not been collected for the specific system under analysis (UNEP, 2020) and it can be found, for example, in databases such as the Social Hotspots Database (SHDB) which includes information on (Norris, et al., 2019):

- Supply chain composition – Using the Global Trade Analysis Project (GTAP) global economic equilibrium model version 9, which contains information on the trade flows between 57 economic sectors (GTAP, 2019) from 140 regions

for the reference year 2011, the SHDB database can provide information on supply chain composition. Based on the economic sector, origin, and cost of each material identified in the foreground data collection, the SHDB determines the trade flows (in US dollars) from other sectors/regions involved;

- Labor intensity – Using GTAP data on total wage payments (US dollars) per US dollar of output by the sector/region average wage (US dollars/hour), the SHDB database can determine the worker hours involved per dollar of output for the reference year 2011 for each of the 57 economic sectors in each of the 140 regions previously mentioned;
- Social risks – Using data obtained from different sources such as country statistics, academic papers, and intergovernmental databases (e.g., the International Labor Organization database), the SHDB has information on over 160 social impact indicators for 244 regions and 57 sectors.

In this study, the foreground data was collected considering the four life cycle stages included within the system boundary (see subsection 2.1) and the background data was collected using the Social Hotspots Database. The collected foreground and background data were used to model the Portuguese natural cork stoppers supply chain using the SimaPro 8.4.0 software. The developed model is presented in Figure 1.

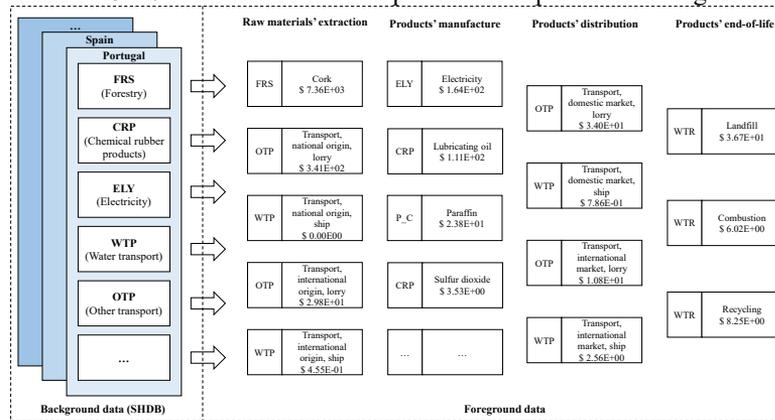


Figure 1 - Portuguese natural cork stoppers supply chain model.

The output of the model presented in Figure 1 is a list of social flows associated with the Portuguese natural cork stoppers supply chain.

2.3. Life Cycle Impact Assessment

The third step of a social LCA consists of converting the social flows collected in the previous step into potential social impacts. The Reference Scale Approach is the most developed impact assessment approach (UNEP, 2020). This approach assesses the potential social impacts based on reference scales typically comprised of 1 to 5 levels (i.e., performance reference points). Therefore, a reference scale with clearly defined levels needs to be developed for each social inventory indicator collected in the previous step (UNEP, 2020). One impact assessment method that uses the Reference Scale Approach is the Social Hotspots Index (SHI) available in the SHDB. This was the method chosen in this study to convert the inventory collected in the previous step into potential social impacts. The SHI method considers 25 impact subcategories (e.g., Child Labor, Occupational Toxics and Hazards, and Gender Equity) (Norris, et al., 2019). These impact subcategories are assessed using the 160 social impact indicators mentioned in subsection 2.2, with some being assessed using only one indicator, while others are assessed using

several indicators. Each indicator has a clearly defined reference scale of four risk levels (very high, high, medium, and low). A characterization factor is attributed to these levels (10, 5, 1, and 0.1, respectively), representing the relative probability of an adverse situation (Norris, et al., 2019). By combining these characterization factors with the social flows obtained in the LCI step, it is possible to quantify the potential social impacts in medium risk hours equivalent (MRH eq.). These impacts can be added together to determine the overall potential social impact (i.e., single score) of the system(s) under study after normalized and weighted using normalization and weighting factors, respectively.

2.4. Results Interpretation

The fourth and last step of a social LCA consists of examining and interpreting the results of the LCA study to accomplish the goal defined in Step 1. Since the goal of this study is not only to assess the social performance of the Portuguese natural cork stoppers supply chain but also to provide suggestions on how this performance can be improved, the interpretation of results will focus on social hotspots identification. According to the definition provided in the Introduction, social hotspots are processes in a region (e.g., cotton production in India) where the most relevant potential social impacts are likely to arise (UNEP, 2020). Following this definition, the first step to identifying social hotspots is to determine the most critical potential social impacts. In this study, the most relevant social impacts were identified using the “80/20” rule which asserts that 80% of outcomes (i.e., total social impact) result from 20% of all causes (i.e., impact subcategories). Hence, the 25 impact subcategories were sorted in decreasing order of normalized and weighted results and the top 20% (i.e., the five impact subcategories with the highest normalized and weighted results) were chosen. The next step to identifying social hotspots is to determine the most relevant processes and regions. In this study, for each relevant impact subcategory identified, the most relevant processes and regions were determined by sorting the processes and regions in decreasing order of characterized results and selecting the top result.

The results of applying the social LCA to the Portuguese natural cork stoppers supply chain are presented in the next section. These results will be used to provide informed suggestions on how the social performance of this supply chain can be improved.

3. Results Analysis and Discussion

From the normalized and weighted results obtained for the Portuguese natural cork stoppers supply chain, it is possible to conclude that the most relevant impact subcategories (i.e., highest normalized and weighted results) associated with the supply chain under analysis are Injuries and Fatalities, Occupational Toxics and Hazards, Corruption, Migrant Labor, and Children Out of School. The same social hotspot, Cork harvest in Portugal, was recognized for the five relevant impact subcategories identified since this is the process and region with the highest characterized results in these subcategories. Given this information, Portuguese cork suppliers are key to improving the overall social performance of the Portuguese NCS supply chain. Due to the background data available in the SHDB, it is possible to determine which economic sector and region is most responsible for the bad social performance of Cork harvest in Portugal (Figure 2). The results presented in Figure 2 can assist in providing suggestions on how the social performance of the Portuguese NCS supply chain can be improved. In the case of the two most relevant social issues, Injuries and Fatalities and Occupational Toxics and Hazards, the social hotspot identified was Cork harvest in Portugal. The sector/region most responsible for this result is the Forestry sector in Portugal (Figure 2) which

indicates that it is the activities/behaviors of the Portuguese cork suppliers the main responsible for the bad social performance of Cork harvest in Portugal in the two aforementioned subcategories. Consequently, to improve the social performance of the Portuguese NCS supply chain, Portuguese cork suppliers should improve their working conditions by implementing different strategies such as (1) identifying where common accidents occur and which health hazards may be present in the workplace; and (2) defining measures to minimize occupational accidents and prevent employers from being exposed and harmed by occupational hazards (ILO, 2021). If implemented, these measures will contribute towards achieving *SDG 8 – Decent Work and Economic Growth*.

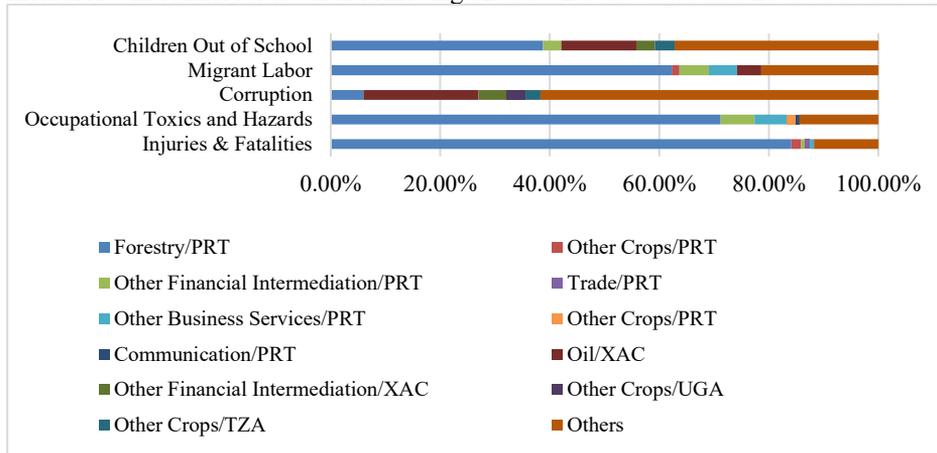


Figure 2 – Contribution of different sectors/regions to the characterized result of “Cork, Portugal” in the five relevant impact subcategories.

In the case of Corruption, the social hotspot identified was Cork harvest in Portugal and the sector/region most responsible for this result is the Oil sector in the South Central Africa (i.e., XAC) region (Figure 2). This result indicates that it is not the activities/behaviors of the Portuguese cork suppliers the main responsible for the bad social performance of Cork harvest in Portugal in the Corruption subcategory but activities/behaviors of oil providers from South Central Africa with whom the Portuguese supply chain partners deal with. To improve the social performance of the Portuguese NCS supply chain, Portuguese cork suppliers should contact the oil providers from South Central Africa and make sure they implement different measures such as (1) publishing financial accounts for each country of operations, including what is paid to each government in taxes and other contributions, and (2) disclosing and regularly monitoring anti-corruption systems throughout all operations and those of subsidiaries (Transparency International, 2017). If implemented, these measures will contribute towards achieving *SDG 16 – Peace, Justice, and Strong Institutions*. Similar to the conclusion reached for the two most relevant impact categories, the social hotspot identified for Migrant Labor was Cork harvest in Portugal and the sector/region most responsible for this result is the Forestry sector in Portugal (Figure 2). Portuguese cork suppliers should implement measures to support the migrant workers employed by them such as (1) developing policies and procedures that promote equality of opportunity in the workplace; and (2) providing equality and diversity training for all workers and having a suitable induction process that includes introducing the new workers to their team and touring the workplace (NiBusinessInfo, 2021). If implemented, these measures will contribute towards achieving *SDG 8 – Decent Work and Economic Growth*. In the case of Children Out of

School, the social hotspot identified was Cork harvest in Portugal and the sector/region most responsible for this result is the Forestry sector in Portugal (Figure 2). Portuguese cork suppliers should implement business actions such as (1) creating programs (e.g., internships) that give students early access to the corporate environment, which could reduce the number of students that drop out of school (SDG Compass, 2015). If implemented, these actions will contribute towards achieving *SDG 4 – Quality Education*.

4. Conclusion

In this study, the social performance of the Portuguese natural cork stoppers' supply chain was assessed and the social hotspots of this supply chain were identified using the social life cycle assessment methodology. Based on the hotspots recognized, a set of suggestions were provided to assist natural cork stoppers' manufacturers and other supply chain partners in improving their social performance and, as a result, the social performance of the Portuguese NCS supply chain. If implemented, these suggestions will contribute to three Sustainable Development Goals: *Goal 4 – Quality Education*, *Goal 8 – Decent Work and Economic Growth*, and *Goal 16 – Peace, Justice, and Strong Institutions*.

References

- EC-JRC-IES, 2010. International Reference Life Cycle Data System (ILCD) Handbook - General Guide for Life Cycle Assessment, Luxembourg: Publications Office of the European Union.
- GTAP, 2019. 57 Detailed Sectoral List. [Online] Available at: <https://www.gtap.agecon.purdue.edu/databases/contribute/detailedsector57.asp> [Accessed 07 January 2021].
- ICNF, 2018. Síntese económica 2018., Lisbon, Portugal: Institute for Nature Conservation and Forests.
- ILO, 2021. Occupational Safety and Health - A Guide for Labour Inspectors and other stakeholders. [Online] Available at: <https://www.ilo.org/global/topics/labour-administration-inspection/resources-library/publications/guide-for-labour-inspectors/lang--en/index.htm> [Accessed 15 September 2021].
- ISO, 2006. ISO 14040:2006 Environmental management - Life cycle assessment - Principles and framework., Geneva, Switzerland: International Organization for Standardization.
- NiBusinessInfo, 2021. How to support migrant workers in your business. [Online] Available at: <https://www.nibusinessinfo.co.uk/content/how-support-migrant-workers-your-business> [Accessed 21 September 2021].
- Norris, C. B., Bennema, M. & Norris, G., 2019. The Social Hotspots Database - Supporting documentation, Maine: New Earth B.
- Santos, A. et al., 2019. Assessment and optimization of sustainable forest wood supply chains – A systematic literature review. *Forest Policy and Economics*, Volume 105, pp. 112-135.
- Santos, A., Norris, C. B., Barbosa-Póvoa, A. & Carvalho, A., 2020. Social Life Cycle Assessment of Pulp and Paper Production – A Portuguese Case Study. *Computer Aided Chemical Engineering*, Volume 48, pp. 15-20.
- SDG Compass, 2015. SDG 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. [Online] Available at: <https://sdgcompass.org/sdgs/sdg-4/> [Accessed 21 September 2021].
- Seuring, S., 2013. A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, 54(4), pp. 1513-1520.
- Transparency International, 2017. Six ways business can help deliver the sustainable development goals. [Online] Available at: <https://www.transparency.org/en/news/six-ways-businesses-can-help-deliver-the-sustainable-development-goals> [Accessed 16 September 2021].
- UNEP, 2020. Guidelines for Social Life Cycle Assessment of Products and Organizations.. Nairobi, Kenya: United Nations Environment Programme (UNEP).
- United Nations, 2021. The 17 Goals. [Online] Available at: <https://sdgs.un.org/goals> [Accessed 23rd July 2021].

A game-theoretical approach for the analysis of waste treatment and circular economy networks

Antonis Kokossis and Evaripidis Melampianakis*

School of Chemical Engineering, National Technical University of Athens, Zografou Campus, 9, Iroon Polytechniou Str. GR-15780, Athens, Greece

**akokossis@mail.ntua.gr*

Abstract

The paper presents a game-theoretical approach for the analysis of waste management and circular economy networks. Players include waste producers, competitive technologies, and the governing authorities. The variables include interactions of the players, choices for valorization technologies and different cost models. In search of Nash equilibria, an interesting conclusion has been that the players often converge at errant and deviant equilibria unless waste valorization is involved. In other words, circularity brings stability. In the case where valorization is involved, bilevel optimization has been applied to evaluate the dependence of profit share on market elasticity, options to subcontract waste upgrade to third parties, the cost of processing technologies, and the choices of products. Results indicate a strong dependence of the profit share on market conditions and player interactions. Solutions include cases where leaders choose to subcontract followers, cases that indicate the impact of controls to divert profit share, and the impact of cost and process efficiency in the network development.

Keywords: circular economy, game theory, bilevel optimization, waste management

1. Introduction

Circular economy is emerging as a sustainable alternative to clean and waste management technologies in that waste is upgraded as a resource with a purpose to further use it as feedstock either to the same or to other industries. Major drivers include consumer needs, resource shortages, and technological breakthroughs. Other than materials and energy though, circular economy involves different stakeholders who have similar or different roles and entertain a different level of interactions with each other; the state, as government or municipality, is a major stakeholder and a player assigned with a role to incentivize or manipulate interactions. To understand the development of interactions, the use of game theory can be both insightful and powerful. Traditional applications in game theory typically follow Stackelberg formulations that feature players in open markets, with identical or similar roles. They compete each other on how they can access markets, by means of cost or market functions that dictate their profits. Approaches involve mathematical formulations produced as bilevel optimization problems optimized to determine the profit share of each player. Quite often, mathematical models take the form of multi-objective optimization approaches in which one optimization function balances trade-offs with another competitive function.

Game theory can be applied to address interactions and roles in emerging and circular economy networks. Torres and Stephanopoulos [1] have been among the first to illustrate the use of game theoretical methods to analyze interactions in chemical engineering. Palafox-Alcantar et al. [2] have specifically pointed the significance of game theory in circular economy networks. These are cases where markets are shaping up, and are not

ideal. Indeed, many researchers ignore the fact that profits are primarily made at the purchase of raw materials (waste) rather than the production of marketable products. As the ownership of value chains involves many actors, partner payoffs depend not only on the choices of a single actor but on the choices of others. Research by Y Jin et al [3] indicates that compensation policy may promote cooperation between players. D Salmon et al point [4] out instead that unqualified government subsidy to recyclers may not be sufficient to shift choices; instead, a more targeted approach should benefit recyclers that engage honestly with downstream material or resale markets. Drivers other than competitive advantage include a better management of threats, social and environmental benefits, and policies to establish a fair share of profits. Parlar et al [5] use mathematical programming to model the networks for control and on-line decision purposes. Instead, this paper addresses strategic decisions in the problem. Nash equilibria rather than Pareto optimal constitute more attractive and sensible objectives. Both in waste management and circular economy networks, there are two profit lines to consider: one at the receiving end and, potentially, a second one at the production as based on the valorization path that is selected. While in the context of a classical Stackelberg approach the competition relates to the product, in circular economy networks the competition is mainly for the feedstock; products may diversify with the players' choice to valorize feedstock. Moreover, while the profit share is important to study, it is equally important to study the equilibrium of the overall system in the context of Nash as the network are subjected to several threats and opportunities from the players involved.

The paper presents a game-theoretical approach for the analysis of waste management and circular economy networks. Players include waste producers, competitive technologies, and the governing authorities. The variables include interactions of the players, choices for valorization technologies and different cost models. In search of Nash equilibria, an interesting conclusion has been that the players often converge at errant and deviant equilibria unless waste valorization is involved. In other words, circularity brings stability. In the case where valorization is involved, bilevel optimization has been applied to evaluate the dependence of profit share on market elasticity, options to subcontract waste upgrade to third parties, the cost of processing technologies, and the choices of products. Results indicate a strong dependence of the profit share on market conditions and player interactions. Solutions include cases where leaders choose to subcontract followers, cases that indicate the impact of controls to divert profit share, and the impact of cost and process efficiency in the network development.

2. Games in extensive forms with complete information

Let us use game theory to model a typical problem of a waste producer (A) and a contractor (B). The model is formulated in extensive form with complete information. Let us use backward induction under the assumption that both players are rational and that each player is trying to maximize her/his own payoff. For each unit of waste, Player A

(A.1) is required to dispose waste to the contractor at a gate fee α (€/t) and receive a receipt

(A.2) may discharge his waste without receipt at a gate fee α' (€/t)

(A.3) may risk and dump his waste without cost

Besides, Player B is required to process waste and this would cost β' (€/t)

(B.1) will be taxed on his profits by $\gamma\%$

(B.2) may risk dump the waste he/she received without processing

The backward induction tree for the game is shown in Fig. 1. Table 1 is the payoff matrix of the game; x_1 and x_2 represent possible government subsidies for A and B.

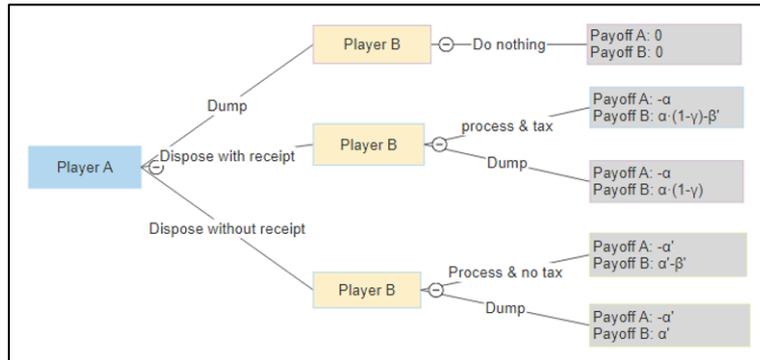


Figure 1. Backward induction tree for the game

Based on Table 1, without any government subsidies, the lawful and desirable choice is (A.1, B.1). However, payoffs for Player A are best for A.3; A.2 is the second-best choice. Similarly, payoffs for Player B are better for B.2; B.1 is the worst choice. In other words, based on game theory payoffs, both players are inclined for errant and unlawful choices. Provided payoffs are available, (a) player A would choose A.1 over A.2 & A.3 once $x_1 > a - \alpha'$ & $x_1 > a$, and (b) Player B would choose B.1 over B2 once $x_2 > \beta'$. Such conditions imply that, lawful choices are ensured once the government is prepared to take up all the costs of waste management.

		Player B	
		B.1	B.2
Player A	A.1	$-\alpha + x_1, \alpha \cdot (1 - \gamma) - \beta' + x_2$	$-\alpha + x_1, \alpha \cdot (1 - \gamma)$
	A.2	$-\alpha', \alpha' - \beta'$	$-\alpha', \alpha'$
	A.3	0, 0	0, 0

Table 1. Payoff matrix of the game

3. Continuous games with multiple players

The previous section considered games with players featuring different roles (e.g. waste producers, contractors) and a finite number of choices (e.g. select legal or lawbreaking action). This section considers players featuring similar roles and continuous choices as this often happens in open markets. More precisely, we consider games with two players as have been extensively studied already in Stackelberg's models of duopoly. Our incentive would be to assess whether game theoretical models for waste management and circular economy are straightforward extensions of Stackelberg problems or they feature differentiating aspects.

Let A and B be respectively leaders and followers in a context of Stackelberg duopolies. A and B compete in collecting waste, Q ; they may further valorize waste (circular economy) to increase revenues at a cost, C , that depends on technology and downstream markets. A and B profit by selecting waste (probably more profitably) and/or in processing waste. *A and B subsequently compete in two markets:* in the first market (waste) they compete against each other; in other markets they may compete, but they may also cooperate instead (e.g., subcontracting waste discharge on behalf of the other). Assume A and B share Q in capacities q_1 and q_2 . Instead of processing waste, A may opt to discharge all or part of q_1 , as q_{12} , over to B (at a lower price so he can profit). Once B

is prepared to process the waste, the total capacity would be then q_2+q_{12} . Another difference over the conventional Stackelberg, is thus that *A and B have a clear incentive to eventually cooperate*.

Fig. 2 illustrates the interactions between A and B. Let p be the collection price of waste; p_1, p_2 sale prices of end products from waste; p' the resale price of waste; c_1, c_2 the costs to process waste to products.

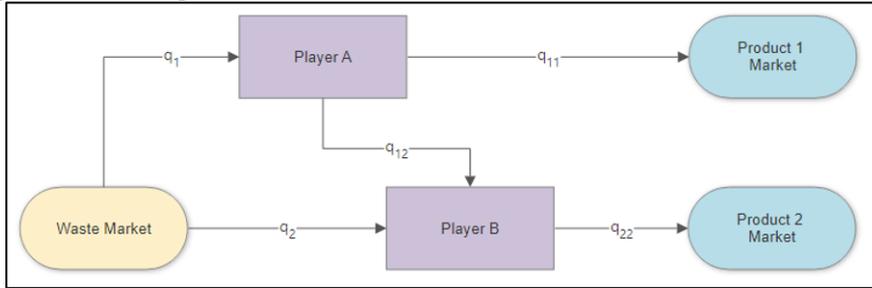


Figure 2. Player interactions in continuous game

Let us further assume linear markets for waste collection and treatment (M_W) as well as those valorizing products to markets (M_A and M_B respectively for A and B). The prices in M_W, M_A and M_B are set by

$$p = a - b \cdot Q \quad (1); \quad p_1 = \alpha_1 - b_1 \cdot q_{11} \quad (2); \quad p_2 = \alpha_2 - b_2 \cdot q_{22} \quad (3)$$

Where α, β are market parameters. Mass balances in Fig 1 translate to

$$Q = q_1 + q_2 \quad (4); \quad q_{11} = q_1 - q_{12} \quad (5); \quad q_{22} = q_2 + q_{12} \quad (6)$$

Economic balances subsequently translate to

$$P_1 = p \cdot q_1 + p_1 \cdot q_{11} - C_1 \quad (7); \quad P_2 = p \cdot q_2 + p_2 \cdot q_{22} + p' \cdot q_{12} - C_2 \quad (8)$$

$$C_1 = c_1 \cdot q_{11} + p' \cdot q_{12} \quad (9); \quad C_2 = c_2 \cdot q_{22} \quad (10)$$

The solution to the game-theoretical model that is set up by A and B results in a bilevel mathematical formulation with one leader and one follower that takes the form

$$\begin{aligned} & \max_{q_{11}, q_{12}} P_1(q_{11}, q_{12}) \\ & \text{s.t. Eqs. (1), (2), (4), (5), (7), (9)} \\ & q_2, q_{12} \text{ solves } \max_{q_{11}, q_{12}} P_2(q_2, q_{12}) \\ & \text{s.t. Eqs. (1), (3)-(6), (8), (10)} \end{aligned}$$

The bilevel optimization problems has been formulated and solved using the Extended Mathematical Programming (EMP) of GAMS and solvers JAMS and BARON.

4. Results

The bilevel optimization model of the previous section is applied to evaluate typical cases in waste management (primal) markets with gate fees closely regulated by the state and serviced by two contractors without upper bounds on the amount of waste they collect and process (typical to larger markets). The purpose would be to evaluate: (a) the split of profit and the market share, and (b) the emerging interactions and exchanges that differentiate the game from conventional cases. Sets of parameters are selected accordingly for low elasticity in primal markets (small $b, b_1 = b_2 = 0.005$), large gate fees (large a as compared to α_1 and α_2), and competitive processing costs $c_1 = c_2 = 10$. The impact of parameters is outlined in Fig 3-5.

(i) Fig. 3 illustrates the impact of waste management economics, namely the impact of gate fee on market shares. As expected, A holds the largest share in all cases, either picking up all waste ($a < 120$), or share with B (once $a > 120$). Without exception, A

makes use of B as a subcontractor profiting by diverting his waste to B. Subcontracting peaks as B enters the primal market; thereafter it decreases to lower volumes.

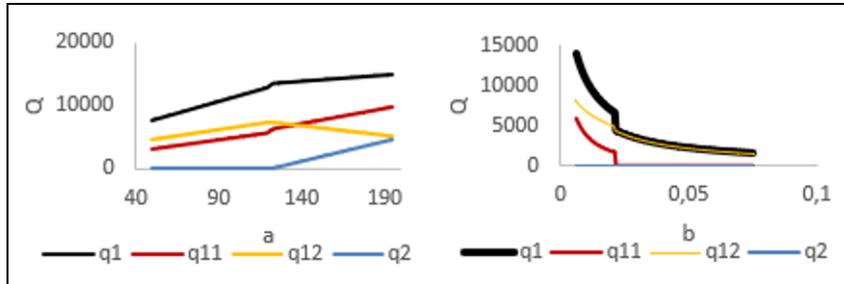


Figure 3. The impact of gate fees on market shares.

- (ii) Fig. 4 illustrates the impact of circular economy choices as made by players, namely the product types selected for valorization. Quite naturally, A increases his/her share as higher value products are selected. A continues to subcontract to B with a rate that increases with the product value (but at lower rate than the increase of his share in A. As the product value of B increases, the gap in the market share initially decreases ($a1 < 25$) but then increases again as A picks up waste for resale purposes (essentially A gradually transforms into a trader, processing no waste and profiteering by his position as a leader).

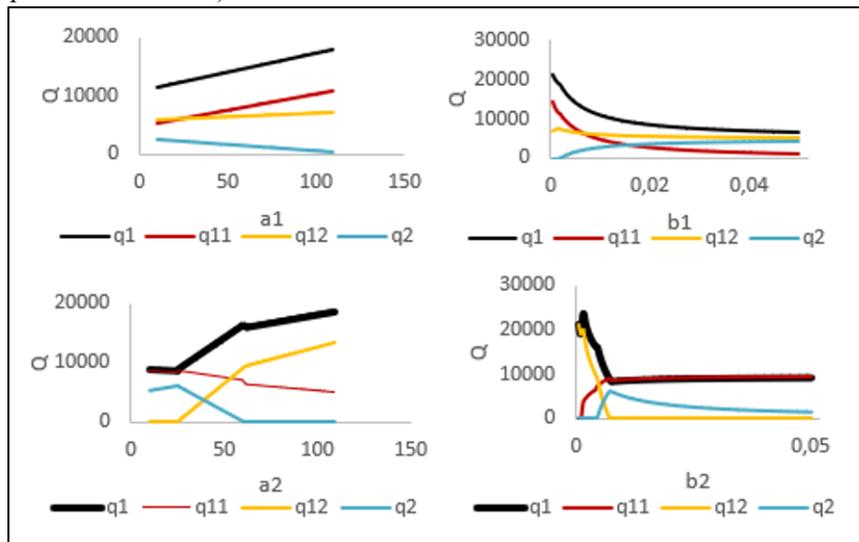


Figure 4. Impact of circular economy choices made by the players, namely product types selected for valorization the market.

Fig. 5 illustrates the impact of process technology costs. As expected, with an increase in the processing costs of A, the share of B in the primal market increases. Reversely, as the processing costs of B increase, both the market share of A and B decrease since the scope of A to use B as a subcontractor also decreases.

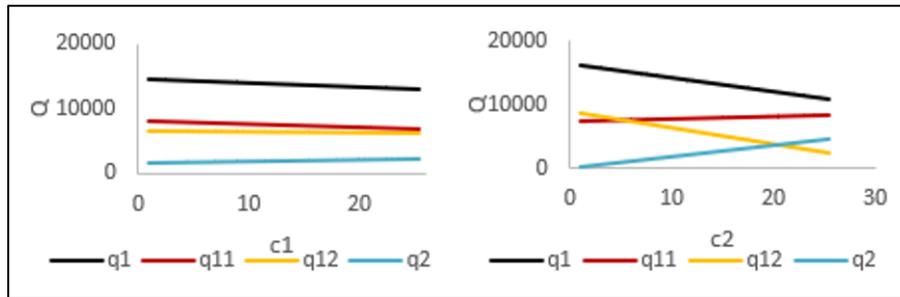


Figure 5. Impact of process technology costs on market share and interactions.

5. Conclusions

The paper applied a game theory in an extensive form using a backward induction model (discrete decisions, complete information) as well as a bilevel optimization problem exploring similarities and differences with a conventional Stackelberg approach. The discrete decision model has revealed that, unless there is an incentive to valorize waste, players are encouraged to bypass the law. Waste management can only be possible through full subsidies by the government. Contrary to the claims by Jin et al. [3] and more aligned to the conclusions of Salmon et al. [4], subsidies are not providing incentives unless with a scope to install circular economy models. The bilevel optimization formulation has similarly revealed important differences that characterize circular economy and waste management problems from conventional market problems: (a) economics are reversed as feedstocks make profits instead of products; (b) competitors are naturally and fully entitled to cooperate (as one is encouraged to subcontract for another), and (c) the type of products produced and the technologies used could have a significant impact.

6. Acknowledgments

The authors acknowledge financial support by the H2020-EU.1.3.3. Grant RENESING II (H2020-778332)

References

1. A. Torres and G. Stephanopoulos, Design of Multi-Actor Distributed Processing Systems: A Game-Theoretical Approach, June 2016, *AICHE Journal* 62(9)
2. P.G. Palafox-Alcantar, D.V.L. Hunt, C.D.F. Rogers, The complementary use of game theory for the circular economy: A review of waste management decision-making methods in civil engineering, *Waste Management*, Volume 102, 1 February 2020, Pages 598-612
3. Y. Jin, Z. Tang, Q. Zhou, H. Zeng, S. Mo, A government value compensation model of waste recycling in an industrial park: A game theory approach, *Journal of Cleaner Production*, Volume 275, 2020
4. D. Salmon, C.W. Babbitt, G.A. Babbitt, C. E. Wilmerad, A framework for modeling fraud in E-waste management, *Resources, Conservation and Recycling*, Volume 171, August 2021, 105613
5. M. Parlar, M. Sharafali, M. Goh, Optimal control and cooperative game theory based analysis of a by-product synergy system, *Journal of Cleaner Production*, Volume 233, 1 October 2019

Agile Process Systems Engineering (PSE) education – 1. What should be taught to achieve desired outcomes mastery by graduates?

Emilia M. Kondili,^a Ian T. Cameron,^b Grégoire Léonard,^c Daniel R. Lewin,^d
Seyed Soheil Mansouri,^e Fernando G. Martins,^f Luis Ricardez-Sandoval,^g
Hirokazu Sugiyama,^h Edwin Zondervanⁱ

^a*Mechanical Engineering, University of West Attica, Greece*

^b*Chemical Engineering, University of Queensland, Australia*

^c*Chemical Engineering, University of Liège, Belgium*

^d*Chemical Engineering, Technion. I. I. T., Israel*

^e*Chemical and Biochemical Engineering, Technical University of Denmark, Denmark*

^f*LEPABE, Chemical Engineering Department, University of Porto, Portugal*

^g*Chemical Engineering, University of Waterloo, Canada*

^h*Chemical System Engineering, The University of Tokyo, Japan*

ⁱ*Chemical Engineering, University of Twente, Netherlands*

Email for correspondence: ekondili@uniwa.gr

Abstract

Our two-part paper investigates the current status and future trends and suggests a framework for teaching of Process Systems Engineering (PSE) topics that addresses *what* should be taught and *how* these topics should be taught effectively in a classroom setting. This first part concerns the “what” – i.e., which specific key PSE topics should constitute the core requirement of chemical engineering education – either a BSc, but in many cases, an MSc, and which application areas should be included. We surveyed existing courses on novel aspects of PSE applications, as well as polling PSE stakeholders to ascertain their opinion of what is taught and the degree to which graduates skills match their expectations. Existing gaps and interesting prospects have been revealed by the surveys leading to suggestions for the future.

Keywords: PSE education, curriculum, active learning

1. Rationale and Background of the Work

Process Systems Engineering (PSE) is the branch of the chemical engineering discipline that exploits computational methods and tools for the analysis, design, control, optimization, and effective operation of processing systems, and the design of products, across different scales and dimensions. The field of Process Systems Engineering (PSE) in the context of Chemical Engineering (CE) has been active for more than 50 years. Prof. Roger Sargent, founder and pioneer of the PSE field, defined it in the mid-60’s as the development of systematic techniques for process modelling, design and control. Subsequently, a large number of academics and researchers have made significant developments and contributions to advance and expand PSE principles in many directions. Recently the interest in PSE education has increased with the development of comprehensive works, as described in Section 2.

Several inspiring works have been recently published that highlight prospects and critical points that need attention for the future of PSE education. Grossmann and Harjunkski (2019) describe in detail the current status, discuss the future academic and industrial perspectives of PSE, and summarize the results of their survey on the standing of PSE in academia and industry. The authors also outlined critical issues such as the disconnect between academia and industry with regards to the appreciation of PSE and the role of stakeholders to disseminate the crucial role of PSE in the academic content and the profession of graduates. In a recent work, Pistikopoulos et al (2021) analyzed the needs of PSE for the next generation in terms of basic principles, research, practical implementation as well as education. Their work adopts a hierarchical model representing the education needs, starting from the core and proceeding to the outer layers. The authors consider the Circular Economy as the framework for future PSE expansion and developments. Cameron et al (2019) have expressed - possibly for the first time so clearly - the relevance of PSE to the so-called Grand Challenges that require holistic approaches. Their work provides the insight of PSE as an integrative discipline in chemical and process engineering. The authors suggest an integrated framework for the design of PSE curriculum, mainly aiming at the development of technical knowledge as well as the mindset to approach problems in the PSE way.

This increased interest in educational needs for PSE highlights the need to match educational activities to a rapidly changing engineering world as well as the recognition of the impact PSE may have in all the great challenges in the years to come. In that respect, the present work should not be considered as just one more contribution. It supports the ideas in play and takes the discussion one step further. The methodology that has been followed in the present paper consist of: (a) a comprehensive survey of the actual topics and application areas covered in the courses taught in academia, and (b) an online survey that received responses from PSE stakeholders around the world (developers, researchers, and management) that map the education perceptions of PSE.

2. Current Status and Contents in PSE Education

This section surveys the PSE related modules being taught in some selected undergraduate and postgraduate courses in Europe, Asia, USA, and Canada. Clearly, there are Chemical and Biochemical Engineering courses with a strong presence of PSE related content without the presence of PSE modules as such. As pointed out by Cameron et al (2019), the presence, breadth and depth to which PSE is included in the curriculum of a Chemical and Biochemical Engineering program depends strongly on the number of faculty members with a background or research focus on PSE. Other relevant parameters may be the presence of a strong process industry, the attitude and focus of academics and industrials to cooperate (e.g. forms of industry – academia cooperation, e.g. in Imperial College), the recognition of PSE expansion in other strong areas such as the energy field (e.g. Texas A&M, USA), and the creation of, and local activity in, computer science and technology capacity. The modules present in almost all undergraduate Chemical and Biochemical Engineering undergraduate curricula that have PSE content are the ubiquitous courses covering integrated process design, process dynamics and control, and often courses in process modelling, simulation as well as in process optimization. While it has been widely recognized that process integration is key for the successful operation of chemical systems (Baldea and Harjunkski, 2014), this aspect is still missing in most of the chemical engineering curriculums, i.e., offer courses that integrate fundamental PSE tasks such as process design, control, scheduling and planning strategies.

In a few universities, PSE focused postgraduate studies are also carried out in dedicated MSc Courses, also including research dissertations in the PSE field. These PSE dedicated MSc courses have started appearing in universities programs during the last decade, clearly revealing the strong interest in the field and its expansion in other areas. However, it should be emphasized that the relevant PSE education content in the key societal issues such as energy, environment, water, pharmaceuticals, water-energy-food-environment nexus is often lagging behind the relevant focus of PSE research in these fields. A number of courses have also been found, mainly in MSc programs, such as big data methods and modelling in CE, supply chain planning and scheduling, process and energy integration, energy systems optimization and process intensification, knowledge based systems and AI, process safety and operations integrity, advanced environmental engineering; transition to a low carbon economy, modelling of biological systems, advanced bioprocess engineering, multivariate statistics in CE. These courses are found mainly in postgraduate courses since the capacity of undergraduate courses is usually limited to more basic subjects.

3. Survey Content and Methodology

As a complement to the review of actual teaching practice, we carried out an online survey to obtain the opinions from PSE stakeholders (developers, researchers, users) to map the education perceptions of PSE around the world. The survey asked the respondents to establish their position regarding the content of PSE education and its response to current and future needs. This questionnaire was termed the “WHAT” survey, delivered using Google Forms, and distributed via email links to the global PSE community. The second part of our contribution analyses the methods and tools (the HOW issues) to educate engineers more efficiently to respond to present and future needs (Lewin et al, 2022 The “WHAT” survey consisted of 16 questions and was organized into three main categories:

- A. Information about the responder (Q1-Q3 on the nature and size of business, geographic location, Q4 on how responder views PSE skills as critical)
- B. Questions relating to the responder’s position on aspects of the PSE and CE in general curriculum outcomes.
- C. Questions relating to how the respondents’ organizations use PSE methodology.

To collect as many responses as possible, we reached out to the following communities: the EFCE CAPE Working Party members, EURECHA members, the Energy Section of the EFCE, the AIChE CAST Division, the CACHE Corporation, the master list used to promote PSE 2018, the Systems and Control Division from Canada and the Japanese PSE community. The total number of respondents was 142: 92 from academia and research and 50 from a wide range of industries (i.e., process industry, software development, and consulting companies). The respondent’s geographical distribution is: 40% Europe, 21% USA, 27% Asia, 10% Central and South America, 2% Australia/New Zealand. About 50% of the responses were from very large organizations (more than 5,000 employees, with a further 25% from large organizations (500-5,000).

4. Survey Results

The responses to the 16 questions establishing the positions of respondents in Parts A, B and C of the survey were distributed on a 5-point Likert scale, where 1 indicated strong disagreement, 3 indicated a neutral position and 5 indicates a strong agreement. Tables 1-3 summarize the average and standard deviation of the received responses to the 16 position questions.

Table 1. Statistics and response distributions of the received responses to the position questions of the “what” survey, Part A: Your position on the importance of PSE.

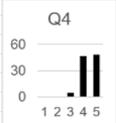
Question	Question statement	Ave	STD	Histogram
Q4	To what extent do you consider PSE skills as useful in developing a career as a chemical engineer?	4.42	0.62	
<p>Not surprisingly, given the PSE-related organizations who were polled, 94% of the respondents considered PSE skills as either important or vitally important to developing a career in chemical engineering. Some respondents made comments related to the reliability of the results due to the very wide nature of questions. One comment is quoted due to its significance: <i>“The students have high critical thinking, but it is not only due to PSE... We are very familiar here with ‘advanced’ tools and methods from PSE, but this area lost some way in some sense. Even with Optimization, Design and Cyber Physical Fusion, the real sense is very important. All computer aided design tools must be used with some discretion and engineering judgment on the part of the designer. This judgment mainly comes with experience. The art and practice of design cannot be learned from books. The intuition and judgment necessary to apply theory to practice will come only from practical experience.”</i> (G. Towler)</p>				

Table 2. Statistics and response distributions of the received responses to the position questions of the “what” survey, Part B: Your position on aspects of PSE curriculum outcomes.

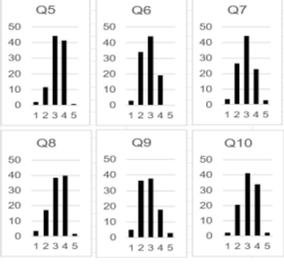
Question	Question statement	Ave	STD	Histogram
How close to your expectations are the skills of freshly-graduated chemical engineers in the following areas:				
Q5	Fundamentals of chemical engineering	3.27	0.76	
Q6	Practical engineering design capabilities	2.79	0.78	
Q7	Critical thinking	2.95	0.87	
Q8	Mastery of process analysis and synthesis software packages	3.18	0.86	
Q9	Computer programming skills	2.77	0.90	
Q10	Professional and personal skills (group work, presentation, writing)	3.13	0.84	
The last question in this part of the survey refers to the impact of teaching innovations on skills				
Q11	Do you see a positive effect of innovative teaching methods on the capabilities of freshly-graduated chemical engineers?	3.23	0.97	

Table 3. Statistics and response distributions of the received responses to the position questions of the “what” survey, **Part C: How your organization uses PSE methodology.**

Question	Question statement	Ave	STD	Histogram
To what extent is your industry/research involved in/using computation tools in:				
Q12	Enhancing sustainability or addressing climate change	3.39	1.25	
Q13	Production planning, scheduling or supply chain management	3.20	1.22	
Q14	Safety and risk management	2.87	1.17	
Perceived value of PSE				
Q15	How do you value of PSE in driving business?	3.52	1.03	
Q16	Extent of in-house capability to conduct PSE projects	3.72	1.09	
Q17	Degree of in-house development or implementation of Industry 4.0 projects	3.19	1.20	
Q19	Would you consider bringing in experts from outside to address PSE-related questions?	2.96	1.07	

5. Discussion of the results

The responses to the specific questions are presented in Tables 1-3. One of the interesting issues of this part of the survey is that it has been answered by respondents from both academia and industry, and thus, are weighted averages from both communities. However, separate analysis for the two groups indicates that there is no significant difference between them. The responses from industry in general avoid extreme opinions (i.e., “not at all” and “very much”). Almost all the respondents (94%) believe that PSE is a key factor that contributes to the knowledge and capabilities of freshly graduated chemical engineers. The respondents certainly are PSE experts or with PSE knowledge and this affects the survey’s outcome. But the general opinion is that young graduates appear to be rather moderate in chemical engineering fundamentals, practical engineering knowledge, personal and professional skills, and particularly low in design, critical thinking, and programming skills. On the other hand, it is very clear from the respondents that they have been using PSE methods and tools to face climate change and sustainability issues but to a lesser extent, supply chain management and planning/scheduling problems. In general, PSE is considered very important for their activities, and they are much interested in hiring specialised graduates to accomplish the PSE related projects.

The above results indicate serious issues not only for PSE related education but more generally for the ability and knowledge of graduates to understand the complexity of today’s industry and professional environment. Indeed, paraphrasing Pistikopoulos et al (2021), there is no chemical engineer that does not utilize PSE every day, everywhere. Therefore, PSE education needs to be totally integrated into the entire CE undergraduate curriculum to gain depth in different areas of application of chemical engineering and provide cases and expansion to all the relevant fields at a later stage. Courses in engineering computational tools, numerical methods, statistics and engineering

economics, essential in the PSE field, should be integrated to the engineering core of each curriculum.

To perform successfully, the chemical engineer must be able to design, analyse, operate and control processes to produce useful and desirable products from less valuable raw materials in an efficient, economic, and socially responsible way. The integrated approach of PSE and its focus on modelling and systems thinking make a very important framework and it is not only a matter of separate modules but a way of thinking and a mindset that should be introduced straight from the beginning in the Chemical Engineering curriculum.

6. Conclusions

Our two-part contribution suggests a framework for academic activity in preparing the next generation of engineers and researchers to be better aligned with the needs of academic research, industry, and society. This first part presents surveys of current practice and the perceptions of practitioners in both academia and industry, concerning the appropriate content of PSE education that would provide prospects and professional advancement to graduates as well as significant added value to industry for the most sustainable solution of their crucial problems. Outcomes of the research and suggestions for the future include the encouragement of the PSE community to expand to fields related to novel challenges in classical problems such as sustainable supply chains or the societal issues of circular economy, water, food, energy engineering, where the PSE approach has a lot to offer, and PSE-specialised engineers will find new professional prospects. Following the principal conclusion of Part 2 of this contribution, which refers to active learning enabled by novel teaching methods, perhaps the time has come for educational modules to be developed to facilitate the introduction of these new areas into the PSE curriculum worldwide, for the benefit of both students and the PSE community.

More work should follow on the investigation of the real PSE needs from the side of the demand and the users. It is believed that this will provide a much more complete insight where we should go. Furthermore, the accommodation of new PSE applications – expansion in the great social challenges will further enhance its role and contribution amongst industry and society in general.

References

- Baldea M., Harjunkoski I., 2014, Integrated production scheduling and process control: A systematic review, *Computers and Chemical Engineering*, 71, 377-390.
- Cameron, I. T., Engell, S., Georgakis, C., Aspiron, N., Bonvin, D., Gao, F., Gerogiordis, D. I., Grossmann, I., Macchietto, S., Preisig, H. and Young, B., 2019, Education in process systems engineering: why it matters more than ever and how it can be restructured, *Computers and Chemical Engineering*, 126, 1781-1792.
- Grossmann, I.E, Harjunkoski H., 2019, Process Systems Engineering: Academic and Industrial Perspectives, *Computers and Chemical Engineering*, 126, 474-484.
- Lewin, D. R., Cameron, I. T., Kondili, E., Léonard, G., Soheil Mansouri, S., Martins, F. G., Ricardez-Sandoval, L., Sugiyama, H., Zondervan, E., 2022, Agile Process Systems Engineering (PSE) education – 1. What should be taught to achieve desired outcomes mastery by graduates?, Montastruc, L., Negny, S. (Eds.), *Book of Abstracts of the 32nd European Symposium on Computer Aided Process Engineering (ESCAPE32)*, Toulouse, France.
- Pistikopoulos, E.N, Barbosa-Povoa A., Lee, J.H., Misener, R., Mitsos A., Reklaitis G.V., Venkatasunbramanian V., You F., Gani R., 2021, Process systems engineering – The generation next?, *Computers and Chemical Engineering*, 147, 107252.
- Sargent, R W H (2004) 25 years of progress in Process systems Engineering, *Comp Chem Eng* 28, 437-439.

You have nothing to lose: Worry-free flipping for PSE

Daniel R. Lewin

Chemical Engineering, Technion, Israel Institute of Technology, Haifa, Israel
Email for correspondence: dlewin@technion.ac.il

Abstract

Even though the COVID-19 pandemic forced all PSE instructors to move their teaching completely online, this has largely not changed most teaching practice, which has retained face-to-face lecturing as the main activity during contact between lecturers and students. This paper presents the case for shifting teaching to active learning by students, most effectively accomplished using the flipped classroom. With the advantages of adopting the flipping approach clear, one is left with two issues to resolve, both of which are the focus of the contributions of this paper: (a) how to effectively produce the necessary supporting materials (recorded lectures, online quizzes, and class activities), and (b) how to continuously foster engagement by the students. These contributions are supported by examples from the authors' extensive experience of successful applications of active learning approaches relying on the flipped classroom in process design and process control courses.

Keywords: PSE education, active learning, flipped class

1. Introduction

All topics taught under the PSE umbrella benefit from more time-on-task made available for students to learn by class discussion, experimentation, and cooperative solution of open-ended problems – basically by “getting their hands dirty” – as a key component of their own learning process. Courses taught using the conventional teacher-centered, lecture-based approach have less time available for these crucial activities, and thus may achieve lower levels of learning outcomes.

The flipped format moves the lecture material online, to be completed by students as homework. Thus, the main justification to move to flipped format is the desire to increase the proportion of the student-staff contact time in which students are actively learning, rather than just listening to lectures (Crouch and Mazur, 2001; Felder and Brent, 2015). This makes better use of the shared time between teacher and students, which has a huge impact on students' engagement, as does aiming to maximize the degree to which students are participating actively with the teacher and with each other, rather than passively listening to lectures. There are many existing studies that provide quantitative evidence that active learning improves course outcomes (e.g., Freeman et al, 2014, Velegol et al, 2015; Lewin and Barzilai, 2021). In an extended study involving secondary and post-secondary education, van Allen et al. (2019) found that the flipped classroom has a small effect of learning outcomes, but uncertain effect of student satisfaction, noting that the results depend strongly on how flipping is implemented.

Since the COVID-19 pandemic forced all teaching to move completely online, one would have thought that this would have motivated the transition to active methods in teaching. In fact, teaching pedagogy has largely not been affected by the potential of

technology, with much online teaching still teacher-centered, relying on synchronous lectures delivered over Zoom. Lewin et al. (2022) ran a survey that received 82 responses, mostly from veteran lecturers on PSE topics, which found that the main obstacles to change are the following: time taken away from research activities (65%), lack of available institutional funding (46%), and student dissatisfaction with new forms of teaching (32%). Teachers are clearly discouraged both by the significant investment of time and effort required to prepare quality online materials (prerecorded lectures and online exercises), and by the initial resistance of some students to active learning. Not surprisingly, there is reduced outcomes performances from the non-participants/non-engagers; Quantifiable lower outcomes are attained by students who engage less with the online materials and with class activity (Lewin, 2022).

The paper next discusses the link between active learning and learning outcomes. Next, in Section 3, the components of the proposed flipped class approach are presented, as well as details of what is required from the point of view of the lecturers and the benefits for students, with a focus on guidelines for practice that works. Next, working experience is shared regarding evidence for the contribution of the proposed teaching methodology to learning outcomes (exam results), followed by conclusions.

2. The Link Between Active Learning and Learning Outcomes

PSE topics are challenging to teach and to master since they all address the three top tiers in Bloom's Taxonomy (Bloom, 1956): analysis, synthesis, and evaluation. Ideally, a combination of examinations and group project assignments are the vehicles for teaching and assessing students' knowledge and competencies in all PSE topics. The utilization of project outcomes for assessing individual assessment requires careful checking to ensure all team members are truly contributing. For example, most process design courses also include a competitive design project component, calling for a demonstration of team-effort in addition to individual mastery. While both team and individual capabilities are important, examinations provide a reliable measure of the crucial mastery of individual students (Turton et al., 2013). Bloom (1968) postulated that the degree to which students achieve mastery depends on four conditions:

1. *Clear definition of what constitutes mastery.* It is the responsibility of the course instructor to clearly state the learning objectives in a manner that defines precisely what students need to achieve to demonstrate mastery.
2. *Systematic, well-organized instruction focused on student needs.* Our approach is based on pre-prepared, clear presentations of the course materials in which online lectures are composed of short video segments interspersed with practice activities, to enable students to actively control their initial acquisition of the basic materials. Then in the class meetings and active tutorials, students practice on more complex and advanced example exercises, first in cooperation with the course staff, and then on their own and with their peers. This sequence of actions leads to weekly cycles of systematic learning.
3. *Assistance for students when and where they experience difficulties.* The active tutorials are the ideal vehicle to aid students when they need it most: the first time when they attempt to solve example problems for themselves. This increases the likelihood that mastery will be achieved in less time.
4. *Provision of sufficient time for students to achieve mastery.* This implies the need to increase the time allotted to active tutorials at the expense of time expended in teacher-centered lectures and demonstrations. This is the reason for the shift

to switch lectures to online homework activities, which is the basis of the flipped classroom.

Bloom (1968) reports the modes of learning that improve outcomes, with the most significant obtained by personal tutoring, which increases the degree of mastery as exhibited by exam grades up to two standard deviations higher than for students taught by a conventional lecture-based approach. Amongst other factors indicated by Bloom (1984) as having significant positive effects on achieving learning mastery, are positive reinforcement and praise from the instructors, student classroom participation and time on task. Bloom reports that all these factors improve results by approximately one standard deviation higher than achieved by conventional lecture-based instruction.

The main justification to move to flipped format was the desire to increase the proportion of the student-staff contact time in which students are actively learning rather than just listening to lectures. This format makes better use of shared time between teacher and students to significantly impact students' engagement, as does aiming to maximize the degree to which students are participating actively with the teacher and with each other, rather than passively listening to lectures.

3. Flipped-class Enabled Active Learning

This section describes the three phases of the flipped approach, the demands on teachers who choose to adopt it in their courses, and the benefits to their students.

3.1. The recommended flipped approach

Three PSE topics, process design, process control and plant design, have been taught annually at the Technion using a three-phase flipped approach, the first of which since 2015, and all three online since the COVID-19 pandemic struck. In this approach a weekly cycle consists of three steps:

- (a) Asynchronous assignments in which pre-recorded video lessons are completed in advance of the week's activity by students as homework. Moodle *lessons* (<https://moodle.org>) are used as a framework for these, with each lesson being composed of a series of questions in which short video segments of lecture material are embedded.
- (b) Synchronous class meetings, using Zoom or in F2F/hybrid sessions, in which students interact with the lecturer and with each other. Typically, these involve review of concepts from the online lesson, discussions generated by quiz questions, and open-ended problem solving.
- (c) Synchronous active tutorials, using Zoom, in which students solve example problems for themselves. These usually begin with a brief review by the teaching assistant followed by problem solving by students working separately or in groups, utilizing breakout rooms. Our experience is that active tutorials run in Zoom breakout rooms are more effective than tutorials in regular F2F settings.

3.2. Requirements from the lecturer

The flipped format implemented imposes significant effort on the part of the lecturer:

- (a) Online materials, namely, the prerecorded lectures involving 5-15 minute video segments and associated quiz questions need to be prepared, most effectively using a video editor such as Camtasia® (<https://www.techsmith.com>). Each course typically requires of the order of 100 of each, which constitutes a huge investment. However, this effort is only invested once: the author prepared the materials for the process design course in 2015, for the process control course in 2017, and for the plant design course in 2020. No additional preparatory materials have been required for either course since then.

- (b) The lecture time freed by moving lectures to online homework tasks for students to do on their own needs to be occupied by suitable activities. The main difficulty for many teachers, especially those who are used to just lecturing, is the required change in their mind-set, which shifts the contact time between lecturer and student from being teacher-centered to being student-centered. This means that class materials should be designed to support open-ended problem solving performed by the lecturer but stressing class participation, the use of pop-quizzes to generate class discussion on their solutions, and any other activity that will enhance students' understanding. The development of these activities will take time to get right, and will likely evolve over time, as teachers become more accustomed to this mode of instruction.
- (c) The active tutorials could be as simple as just having students solve what used to be homework exercises working in teams in recitation time, or more involved and specially designed exercises. The main objective is to ensure that the students are doing most of the work for themselves.
- (d) It is important to continuously monitor the activity of each student, and follow-up on those students who are less active. This task is facilitated by a myriad of tools that are available in learning support systems.¹

3.3. Benefits to the students

Low-performing students typically do not significantly engage during the semester, leaving most of their effort for cramming just before final exams. This behavior is unlikely to achieve mastery of the taught materials. In each week of a course taught in flipped format, students need to prepare for class meeting and active tutorials by covering the pre-prepared materials ahead of time. They then benefit from participating effectively in the class meetings, by responding to the pop-quizzes, contributing to class discussion and brainstorming during the open-ended problem solving. Finally, they participate in active tutorials where they solve exercises for themselves, mentored by the course staff. This combination of activities increases the performance of the entire class, as will be described next.

4. Class Experience with the Teaching Method

As a teacher of process design for more than 20 years, I have taught using the full spectrum of possible approaches. This began with a traditional teacher-centered instruction, in which the course materials were transmitted via lectures and demonstration-recitations to my students, termed Phase I. The first transition was to active tutorials, where at least students were actively engaged in problem-solving for themselves, termed Phase II. The last major change, Phase III, was the move to the flipped classroom paradigm in 2015, which freed even more time for students to get to grips with the course material for themselves. I also teach process control and plant design, both of which are now also taught in the flipped format as previously described.

As described in Lewin and Barzilai (2021), the outcomes achieved by the students of the process design course have incrementally improved over the last 15 years, as illustrated in Figure 1, which shows a bubble plot showing disks whose diameters are in proportion to the fraction of the high-performing students in each year's class, p , centered on coordinates, whose ordinate and abscissa are the average grades of the high- and low-

¹ For more implementation details about our flipped classrooms, see the YouTube video: https://www.youtube.com/watch?v=O3hoSIYaGo4&list=PLW3u28VuDAHIOHECm8Vq30_Et_2chWFyO (13.5 min)

performing subsets of the class in each year, μ_1 and μ_2 respectively. These parameters are determined by fitting a bimodal distribution to actual exam grade distributions using the approach of Lewin (2021). Note that as $\mu_1 \geq \mu_2$ by definition, all disks have to be centered under the dashed line $\mu_1 = \mu_2$ indicated in Figure 1. The best performing classes would be those represented either by disks of any diameter in the top right of the plot (high average exam grades of both high- and low-performers, irrespective of their proportions), or lower on the right with large diameters (high average exam grade of high-performers, whose proportion dominates the population). Conversely, classes represented by disks on the left would be characterized by low average exam grades of the high-performers, and even lower average grades of the low-performers. As can be seen in Figure 1, in Phase I, the class disks, shown in black, are on the left. The transition to Phase II indicates a shift to the right, maintained after the transition to the flipped class in Phase III.

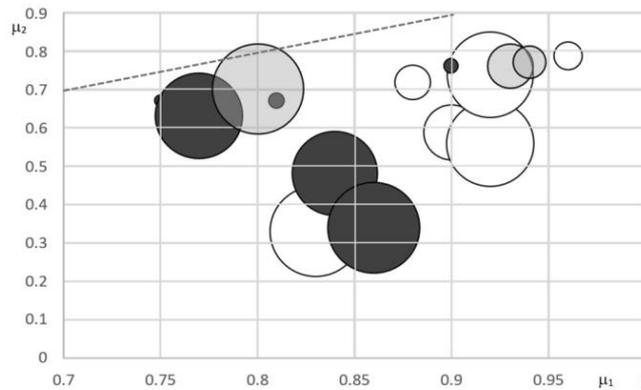


Figure 1. Bubble chart summarizing binomial distribution diagnosis (adapted from Lewin and Barzilai, 2021). The statistics for each year are centered on the $\mu_1 - \mu_2$ plane, with the disk diameter proportional to p . The disks are color-coded according to period: black – Phase I: 2005-2010 (before active tutorials), grey – Phase II: 2011-2013 (before flipping), white – Phase III: 2015-2020 (after flipping). The dashed line indicates $\mu_1 = \mu_2$.

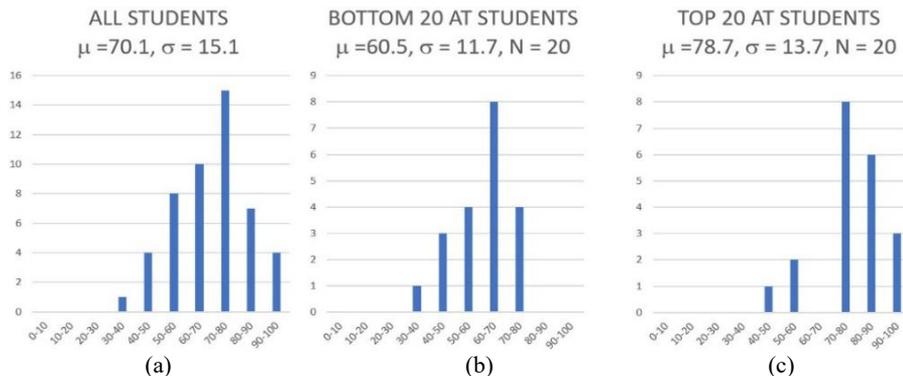


Figure 2. (a) 2021 Process control course outcome distribution for the entire class and separate distributions for (b) the 20 students that attended active tutorials the least and (c) the 20 that attended the most, out of a total of 50 students examined. In the histograms, abscissae indicate numbers of students, plotted against exam grade bins.

What is the cause of the bimodal distribution present in all exam grade histograms? As discussed in Lewin (2021), these could be inherent heterogeneous capabilities of the students in the class, or the consequence of the degree of engagement in the course materials. Certainly, there is evidence for the link between engagement and final exam grades, as shown in the example data in Figure 2 in which the exam grade distribution in the final exam of the process control course in 2021 is shown in comparison with separate distributions – one for the 20 students that attended active tutorials the least, and other distribution for the 20 that attended the most. Note that whereas the average exam grade for the entire course was about 70.1%, the average grade of the 20 students that attended the most tutorials was 78.7%, while that for the 20 least attending students was 60.5% -- more than one standard deviation lower than that achieved by the most attending students.

5. Conclusions

This paper advocates a change in teaching practice of PSE – from teacher-centered to student-centered instruction. It is worthwhile to consider moving much of the teaching materials from the lecture room to an online setting and require students to cover these materials on their own in preparation for class and tutorial activities. The effort is worthwhile in the long run, as better-prepared students will learn more effectively with the instructors and TAs, especially if they are expected to take an active part in the problem-solving sessions in class. The paper has provided evidence of the outcome improvements that can be expected.

References

- Bloom, B.S. (1956) Taxonomy of educational objectives: The classification of educational goals (Handbook One). David McKay. New York, NY.
- Bloom B.S. (1968). Learning for mastery. *UCLA - CSEIP - Evaluation Comment*, 1(2):1-12.
- Bloom B.S. (1984). The 2-sigma problem: The search for methods of instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6): 4-16.
- Crouch C.H. and Mazur E. (2001). Peer instruction: Ten years of experience and results, *American Journal of Physics*, 69 (9): 970-977
- Felder, R.M., and Brent, R. (2015). "To flip or not to flip," *Chem. Eng. Educ.*, 49(3), 191-192.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. S., Okorafor, N., Jordt, H. and Wenderoth, M. P. (2014). "Active learning increases student performance in science, engineering, and mathematics," *PNAS*, 111(23), 8410-8415.
- Lewin, D. R. (2021). "What Can We Learn from Exam Grade Distributions?" *International Journal for the Scholarship of Teaching and Learning*, 15(2), Article 7.
- Lewin, D. R. (2022). "Teaching process and plant design in the year of COVID-19," *Chem. Eng. Educ.*, 56(1), 57-67.
- Lewin, D. R. and Barzilai, B. (2021). "Teaching process design to chemical engineering undergraduates – an evolution," *Chem. Eng. Educ.*, 55(3), 157-172.
- Lewin, D. R., Cameron, I. T., Kondili, E., Léonard, G., Soheil Mansouri, S., Martins, F. G., Ricardez-Sandoval, L., Sugiyama, H., Zondervan, E. (2022). "Agile process systems engineering (PSE) education – 2. How to teach to achieve desired outcomes mastery by graduates?," Montastruc, L., Negny, S. (Eds.), *Book of Abstracts of the 32nd European Symposium on Computer Aided Process Engineering (ESCAPE32)*, Toulouse, France.
- Turton, R., R. C. Bailie, W. B. Whiting, J. A. Shaeiwitz and Bhattacharyya, D. (2013). *Analysis, Synthesis, and Design of Chemical Processes*, Pearson Education, Upper Saddle River, NJ.
- van Alten, D.C.D., Phielix, C., Janssen, J. and Kester, L. (2019). "Effects of flipping the classroom on learning outcomes and satisfaction: a meta-analysis," *Educational Research Review*, 28 (2019) 100282. <https://doi.org/10.1016/j.edurev.2019.05.003>.
- Velegol, S. B., Zappe, S. E. and Mahoney, E. (2015). "The evolution of a flipped classroom: evidence-based recommendations," *Advances in Engineering Education*, 4(3), 1-35.

Agile Process Systems Engineering (PSE) education – 2. How to teach to achieve desired outcomes mastery by graduates

Daniel R. Lewin,^a Ian T. Cameron,^b Emilia M. Kondili,^c Grégoire Léonard,^d
Seyed Soheil Mansouri,^e Fernando G. Martins,^f Luis Ricardez-Sandoval,^g
Hirokazu Sugiyama,^h Edwin Zondervanⁱ

^a*Chemical Engineering, Technion. I. I. T., Israel*

^b*Chemical Engineering, University of Queensland, Australia*

^c*Mechanical Engineering, University of West Attica, Greece*

^d*Chemical Engineering, University of Liège, Belgium*

^e*Chemical and Biochemical Engineering, Technical University of Denmark, Denmark*

^f*LEPABE, Chemical Engineering Department, University of Porto, Portugal*

^g*Chemical Engineering, University of Waterloo, Canada*

^h*Chemical System Engineering, The University of Tokyo, Japan*

ⁱ*Chemical Engineering, University of Twente, Netherlands*

Email for correspondence: dlewin@technion.ac.il

Abstract

This study is the second part of a contribution addressing the appropriate content of process systems engineering (PSE) curriculum (the “what”) and how best it should be taught, so that our graduates are equipped to effectively apply their knowledge (the “how”). The two-part contribution presents the collective views and perceptions of the responders from all over the world to surveys as well as the collective view of the authors. This paper presents the “how” – how best to instruct our students in all matters PSE, given the availability of teaching technologies, and the time available to effectively train our students.

Keywords: PSE education, curriculum, active learning.

1. Introduction

This two-part paper proposes a “game plan” for effective teaching of process systems engineering (PSE) topics that addresses *what* should be taught and *how* these topics should be taught effectively. We base our recommendations on: (a) A survey of the teaching methods used by professors teaching PSE concepts and tools in universities around the World (the “how”), and (b) Surveys of the actual topics and application areas covered in the courses taught in academia (the “what”). This second part concerns the “how” – the most effective methods that should be used to achieve learning objectives that enable graduates of the first degree in chemical engineering – either a Bachelor’s, but in many cases, a Master’s degree. In particular:

- Students should be taught fundamental concepts in detail, ideally self-paced. This is achieved more efficiently using prerecorded materials.

- Students need to be exposed to computer programming packages ranging from those solving chemical process flowsheets to those specialized in data analysis, optimization, and spreadsheets.
- Students need to understand why some software or computer packages may not provide accurate answers in some instances. The fact that they run and converge to a solution does not necessarily mean that the solution is correct.
- Students should become familiar with optimization tools earlier in their academic career, so that they can use them to solve practical problems in their senior years.
- Students should apply multivariate statistical and artificial intelligence tools for solving real problems.
- Students should be required to develop critical thinking skills, i.e., to question their solutions/methods and ask themselves if other (attractive) approaches could be used to tackle a particular problem.
- Students should develop professional and personal skills such as teamwork, communication, project management.

This study consists of a survey of teaching practices aimed to assess the degree to which active methods are used in practice, to understand their benefits, limitations, and potential reasons as to why they are not implemented, and to identify circumstances or the conditions under which these methods may be more effective. The two-part paper provides a working plan for academic activity in preparing the next generation of engineers and researchers to be better aligned with the needs of both academic research, industry, and society, without requiring additional time-on-task beyond that allocated currently for the coverage of PSE topics.

2. Teaching PSE using Active Learning

For students to attain mastery in the critical understanding and application of the PSE core materials, time needs to be allocated for them to experiment, get things wrong and understand why; thereby, repeating this process as many times as needed. Such student-centered, active approaches to learning require time, which in a conventional teacher-centered approach is often allocated to lecturing. Several methods have been advocated that free class time for students to engage in active learning such as project-based learning, blended teaching, and flipped class approaches. The flipped class paradigm, detailed by Lewin (2022), moves the transmission of basic information to online preparatory tasks, which students complete in advance of class activities. This freed class time enables the four key agile values to be incorporated into the class environment, i.e.,

1. Student-centered flipping inherently focuses on the learner rather than following traditional teaching processes, which are teacher-centered.
2. Student-staff contact time is mostly used to work problems cooperatively and for project work, rather than for the transmission of information.
3. The contact time is largely reserved for collaborative work between staff and students and transforms the staff member to take on the role of mentor and motivator.
4. Staff can respond to the feedback and needs of students.

3. Method

3.1. Design

We used a survey to map the teaching perceptions of PSE academic teachers around the world. The survey asked the responders to establish their position regarding the application of active teaching methods, and then describe the extent to which active

methods are used in one of the courses taught by the responder. Moreover, the responders were asked to discuss how they see their teaching will evolve, and to define barriers to future evolution of their teaching approaches. This questionnaire was termed the “HOW” survey, delivered using Google Forms, and distributed to the global PSE community.

3.2. Material

The survey consisted of 26 questions organized into five main categories:

- A. Information about the responder (i.e., geographical location, teaching experience)
- B. Questions regarding the responder’s position on aspects of the PSE curriculum
- C. Questions regarding the responder’s position on how teaching should be carried out and, to the degree to which active learning should be applied.
- D. Questions regarding the responder’s teaching practice, and to what extent active learning is employed.
- E. Questions regarding the responder’s future adaptation of teaching methods.

3.3. Participants

To cover as many individuals as possible, we reached out to the following communities: the EURECHA members, the Energy Section of the EFCE, the AIChE CAST mailing list, the CACHE mailing list, the master list used to promote PSE 2018, the Canadian Systems and Control Division mailing list, and the Japanese PSE community.

3.4. Procedure

On 15th October 2021, a request for feedback with a link to the survey was emailed to all potential responders on the mailing lists described above, with a reminder sent on 22nd October. We received 82 responses from academic lecturers from all over the world, with the following geographical distribution: 47.6% Europe, 20.7% North America, 15.9% Asia and 13.4% Central/South America. Most of the responders (83%) had more than 10 years of experience teaching PSE courses, with an additional 10% having between 6 and 10 years of experience.

4. Analysis

The responses to the 18 questions establishing the positions of responders in Parts B-E (see Section 3.2) were distributed on a 5-point Likert scale (1 indicates strong disagreement, 3 indicates a neutral position and 5 indicates strong agreement). Tables 1-4 summarize the averages and standard deviations and presents histograms of the received responses to the position questions.

Table 1. Statistics and response distributions of the received responses to the position questions of the “how” survey, **Part B: Your position on aspects of the PSE curriculum.**

Question	Question statement	Ave	STD	Histograms
Q3	Enrich Courses should be enriched with external sources or guest lecturers	4.17	0.83	
Q4	Should offer courses using open access software	3.94	0.93	
Q5	Should offer courses on advanced statistics	4.13	0.84	
Q6	Should offer courses on AI and ML	4.02	0.75	
<p>Summary: There is general agreement about the need to enrich PSE courses using external sources (Q3), and to maintain curriculum up-to-date with regards to the usage of open-access software (Q4), and the introduction of courses in advanced statistics, artificial intelligence, and machine learning (Q5 and Q6).</p>				

Table 2. Statistics and response distributions of the received responses to the position questions of the “how” survey, **Part C: Your position on how should teaching be carried out in practice.**

Question	Question statement	Ave	STD	Histograms
How course materials should be transmitted to students, and how best to utilize contact time				
Q7	Classes should be organized so teachers mostly lecture, and students listen	3.16	1.18	
Q8	Lecture materials should be reviewed by students on their own as homework in preparation for class activity	3.52	1.00	
Q9	Most of the contact time between teachers and students should be used for student activity	3.56	1.03	
The responses indicate ambivalence regarding face-to-face lecturing as the main transmission vehicle (Q7). There is slightly more support for moving materials online for students to cover on their own (Q8) and for class time to be used more for student activity (Q9). However, it is fair to indicate that this support was not overwhelming.				
How should most of recitation time be best utilized?				
Q10	Instructors demonstrating solutions and students listening	2.68	1.01	
Q11	Students solving problem sets with staff providing hints/motivating	4.04	0.80	
If responses are consistent, the two questions should have response distributions that are mirror images. The responses to Q10 are bimodally distributed, with more disagreement than agreement – more responders do not approve of student passivity, but clearly many are comfortable with it. Moreover, Q11 indicates that there is strong support of student activity in recitations. However, the distributions are clearly not inversed, as they should be for consistency.				
How should PSE course outcomes be assessed?				
Q12	Include a significant portion of project-based learning (teamwork)	4.35	0.62	
Q13	Students should do individually graded homework exercises	3.79	0.96	
Q14	Using one or more exams involving open-ended problem solving	3.51	0.89	
The need for project-based assessment (Q12) received strongly positive response. There were mixed feelings about the other two issues – while there is a slightly positive position regarding the need to check individual students’ formative abilities by grading homework (Q13), the position does not have overwhelming support. The support for summative assessment (exams, Q14) is moderate. The main issue is which kind of formative/summative assessment methods are the most appropriate given students’ time limitations. What proportion of assessment should be team or individual is crucial. That is likely to be more significant than the methods.				

Table 3. Statistics and response distributions of the received responses to the position questions of the “how” survey, **Part D: How do you teach?**

Question	Question statement	Ave	STD	Histograms
Q16	Graduate (G) or undergraduate (UG) course: 28% UG, 53% Combined, 19% G			
Q17	Subject taught: 48.5% Process design, 33.3% Process control, 15.2% Optimization, 3% Numerical methods			
Q18	Class size: : 8.6% <10, 58% 10-50, 25.9% 51-100, 7.4% >100			
Slightly more than half of the responders teach mixed classes of graduate and undergraduate students, with 28% teaching only undergraduates and 19% teaching only graduates. Most responders are teaching either process design (49%) or process control (33%). Most class sizes are either medium-sized with 10-50 students (58%) or large with 51-100 students (26%).				
The following questions were directed at the type and mode of learning environment that exists in the responders’ institution.				
Q19	Student-centered (>50% of contact time is student activity).	3.16	1.10	
Q20	Teacher-centered (>50% of contact time, students are listening to me).	3.02	1.06	
Q21	Students are required to work independently and not in groups.	2.74	1.07	
Q22	Student activities are included into the lectures and recitations.	3.70	1.02	
The responses to Q19 and Q20 regarding how responders run their lectures are very similar and are both bimodally distributed. The responses are almost split 50/50 between those who teach in the traditional teacher-centered method (teacher talks – students listen) and those who apply student-centered, active learning in their classes. More detailed analysis indicates that much of the support for student-centered activity was by teachers of process design, and independent of the class size. Q21 discloses the responders’ views on the need for students to work independently rather than in groups. The majority of the responders disagree with this statement, indicating there is some application of group effort in many of the responders’ courses. However, many of them (27%) still consider it important for students to spend time working problems on their own. Q22 discloses the responders’ choice to include student activities into lectures and recitations. The responses to this question are somewhat in conflict with the responses to Q19 and Q20. On the one hand, about 50% of the responders adopt teacher-centered classes, yet here there appears to be more than 50% support for student-centered class activities. Perhaps the question was poorly posed as it refers both to lectures and recitations.				

Table 4. Statistics and response distributions of the received responses to the position questions of the “how” survey, **Part E: Your position on adaptation of teaching methods.**

Question	Question statement	Ave	STD	Histograms
Q23	I regularly investigate the engineering education literature	3.63	0.94	
Q24	It is important that I increase my use of real-world situations	4.12	0.87	
The surprising response to Q23 indicates relative support for keeping in touch with state-of-the-art engineering education literature, while Q24 indicates strong support for increasing use of real-world situations in the classroom.				
Q25	Most important barriers for me to innovate in my teaching role	65%	Time taken away from research activities	
		46%	Lack of available institutional funding	
		32%	Student dissatisfaction with new methods	

5. Responder Opinions Summarized and Conclusions

Lecturers are ambivalent concerning whether teacher-centered instruction is appropriate (Q7) but tend to support the idea of students reviewing materials to prepare for class (Q8) as well as the use of contact time to enable student activity (Q9). Regarding recitations, the responders largely disagreed that staff should run them in demonstration mode (Q10), and strongly supported the idea of active tutorials, where students do the problem-solving themselves (Q11). Regarding outcomes assessment, there was high agreement that project-based learning (PBL) should be employed (Q12), agreement that individually graded homework should be assigned (Q13), and that open-ended exams should be used for formative assessment (Q14). Note that requiring both preparation before classes, as well as individual homework and project work may overload students, and this needs to be considered. When it comes to their own teaching practice in lectures, the responders were more ambivalent, with responses almost evenly split between those taking teacher- and student-centered approaches (Q19 and Q20). Regarding student activities, there are more instances of group work than individual work (Q21), and a majority of the responders include student activities in lectures and recitations (Q22). Furthermore, the responders indicated that a majority regularly update their teaching materials (Q23), and a large majority indicated the importance of using real-world situations in their teaching, which links to the importance of PBL indicated by Q12. Finally in addition, the main obstacles to teaching innovation (Q24) are identified as the lack of time (65%), followed by the lack of institutional support (46%), but it also appears that not all students are welcoming innovative teaching methods (32%). In addition to the above, the survey also included two questions (Q15 and Q26) giving responders the opportunity to make general remarks. Because of space limitations, these are not being addressed in this paper, but will be summarized in the full version, as will summaries of interviews with leading members of the worldwide PSE community.

The survey disclosed that there is a gap between the technological capabilities that can be harnessed to the teaching of PSE and practice for many of the responders, most of whom see this as a burden since research time is sacrificed to perform this activity. Perhaps additional incentives are required to promote the move to more active teaching. In conclusion, the objectives of our studies are to provide information and suggestions to improve learning outcomes (i.e., the “what”) but also their efficient transmission to students (i.e., the “how”). The process systems engineering community needs to openly share best practices and resources, otherwise we will be back talking on this subject in 5- or 10-years' time (Cameron and Lewin, 2009, Cameron et al., 2019, Kiss and Grievink, 2020).

References

- Cameron, I. T. and Lewin, D. R., 2009, Curricular and pedagogical challenges for enhanced graduate attributes in CAPE, *Computers and Chemical Engineering*, 33, 1781-1792.
- Cameron, I. T., Engell, S., Georgakis, C., Aspiron, N., Bonvin, D., Gao, F., Gerogiordis, D. I., Grossmann, I., Macchietto, S., Preisig, H. and Young, B., 2019, Education in process systems engineering: why it matters more than ever and how it can be restructured, *Computers and Chemical Engineering*, 126, 1781-1792.
- Kiss, A. A. and Grievink, J., 2020, Process systems engineering developments in Europe from an industrial and academic perspective, *Computers and Chemical Engineering*, 138, 106823.
- Lewin, D. R., 2022, You have nothing to lose: Worry-free flipping for PSE, Montastruc, L., Negny, S. (Eds.), *Book of Abstracts of the 32nd European Symposium on Computer Aided Process Engineering (ESCAPE32)*, Toulouse, France.

Decision-making framework for improved educational resilience under pandemic events

Robert E. Franzoi,^a Noof AlQashouti,^a Brenno C. Menezes,^{a,*}

^aDivision of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar Foundation, Qatar

bmenezes@hbku.edu.qa

Abstract

The recent pandemic events have significantly affected people and institutions worldwide. Multiple issues and difficulties arise, with an increasing number of challenges. In this work, we address the impact of pandemic events on educational resilience, and we provide guidelines for addressing such concerns by using a structured framework assisted by data-driven and decision-making capabilities. The educational resilience framework is comprised of five steps: data collection, data analysis, gaps formulation, solution development, and implementation planning. First, a data-driven strategy collects data from the internet, literature, surveys, and previous knowledge. Second, analyses are carried out to draw patterns and insights that can serve as indicatives of potential improvements. Third, the most critical gaps are analysed and classified according to a cost-effectiveness criterion. Fourth, guidance is provided to handle these gaps, whereby proper solutions are developed considering the availability of resources (time, effort, money) and outcomes (benefits, accomplishments, profit). Finally, a deployment plan is built using the structured solution. From the proposed guidelines, educational resilience improvements can be achieved for people, academia, industry, and society, in a wide variety of problems and applications and with multiple significant benefits. The results and conclusions derived from this work illustrate how a decision-making framework can be effectively and interestingly employed towards easier and more efficient educational strategies, methodologies, and policies.

Keywords: Educational resilience, Education in PSE, Learning capabilities, Decision-making framework, Modelling and optimisation, Pandemic events, COVID 19

1. Introduction

Resilience can be understood as the potential of adaptation, or the ability to maintain or regain healthy status when experiencing adversity (Herman et al., 2011). Within a community, resilience measures the ability of responding to and recovering from adverse situations. This allows adaptation and growth after disaster strikes (Magis, 2010). In this work, we address the topic of educational resilience, defined as the likelihood of success in the educational life accomplishments, despite environmental adversities brought by traits, conditions, experiences, and unforeseen events (Waxman et al., 2003). This concept focuses on dynamic variables such as motivation, smart-work (rather than just hard-work), and learning environment (McMillan and Reed, 1994).

The literature on the topic of educational resilience has been extensively addressed (Downey, 2008; Torsney and Symonds, 2019), but only a few works have focused on

scenarios of crisis (Annarumma et al., 2020; Zamfir, 2020; Naidu, 2021) such as the COVID-19 pandemic that has affected the society worldwide. This is a novel topic that has emerged due to the current pandemic scenario and which presents several gaps that require attention. Tackling these gaps represent multiple potential opportunities towards more efficient educational resilience, with several benefits towards real-life educational applications.

This work aims to introduce a decision-making framework to systematically address important and impactful educational gaps derived (or augmented) from times of crisis and to provide guidance towards implementing computational-aided decision-making (e.g., modelling, simulation, optimisation) to achieve enhanced educational resilience. Such decision-making tools have increasingly assisted a wide variety of applications in multidisciplinary areas and provide proper capabilities for improved solutions (Franzoi et al., 2021). The foundation of the proposed framework is twofold and relies on data-driven and solution-driven concepts. Reliable and representative data are required to allow proper analyses, which can be done based on previous knowledge on the topic, literature review, further research, surveys with students, researchers, faculty, workers, etc. Moreover, the development of an efficient educational framework is closely related to solution-driven objectives, which relies on data gathering, data-driven analyses, identification of educational gaps, development of solutions, and implementation of deployment plans required for an enhanced educational system.

This paper is structured as follows. Section 2 discusses educational resilience, including the main drivers, resources, and beneficiaries. Section 3 presents the proposed educational resilience framework. Section 4 provides an illustrative example. Section 5 highlights the main findings and future guidelines derived from this work.

2. Educational resilience

Educational resilience concerns a wide variety of gaps, issues, and improvements required towards a more efficient educational system. Its importance is significantly augmented in times of crisis, with increased difficulties and new challenges. Under the presence of unforeseen pandemic events, which have recently caused significant impacts worldwide, the general concept of resilience becomes particularly important. In this section, we present the: a) main drivers that affect educational resilience of individuals, organisations, and communities; b) resources required for improved resilience; and c) beneficiaries and benefits from a more resilient educational system.

2.1. Main drivers

There are three important drivers that significantly impact educational resilience: investments, learning capabilities, and collective intelligence.

1. Investments are typically funded by the government or partnerships between universities and companies. They concern project and research grants, scholarships, infrastructure, resources, tools, software, etc.
2. Learning capabilities can be described as the mechanisms and management structures that can be implemented to promote better learning environments (Goh et al., 2012). Examples of learning skills include organisational, communication, collaboration, critical-thinking, and creative skills.

3. Collective intelligence arises from a group of individuals acting intelligently towards collective goals, such as in collaborations between postgraduate students (Menezes et al., 2021).

The concepts and some key foundations of those drivers are illustrated in Figure 1.

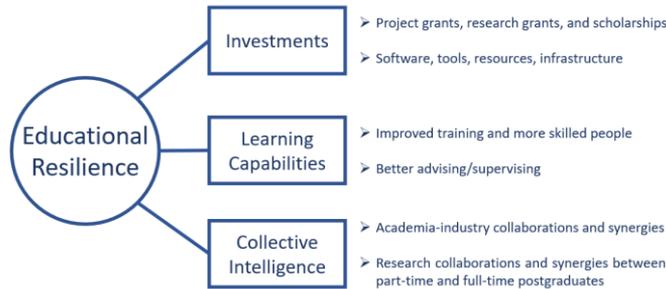


Figure 1: Educational resilience drivers.

2.2. Required resources

Three resources are fundamental to properly improve the educational system. The first are the data to be collected mostly through research and surveys, which provide insights on the most critical educational issues derived from the pandemic, either qualitative (i.e., according to peoples’ opinions) or quantitative (i.e., numerical data and statistical analyses). The second is the effort needed to analyse data, define gaps, propose efficient solutions, and draw implementation plans. The third concerns the capabilities required for implementing the deployment plan (e.g., investments, time, effort).

2.3. Beneficiaries

Improved educational resilience provides benefits for recipients (students, researchers, faculty, staff), academia (schools, research centers, laboratories, universities), industry (companies), and society (people, public system, government), as shown in Table 1.

Table 1: Drivers and outcomes, beneficiaries, and benefits.

Drivers and Outcomes	Direct Beneficiaries	Benefits
Investments: grants, scholarships	People and academia	Resources for studying and developing research, which can be applied for industrial and public sector applications.
Investments: software, resources, infrastructure	People	Resources for developing high-quality research, which also benefits universities.
Learning capabilities: training and skills	People	Better efficiency and accomplishments (also extend to universities).
Learning capabilities: better supervising	People	Improved performance, enhanced learning efficiency, and higher accomplishments.
Collective intelligence: academia-industry synergies/collaborations	Academia and industry	Development of industrial-driven research, scholarships, better environment.
Collective intelligence: research collaborations	People	Exchange knowledge (full-time researchers) and experience (part-time researchers).

3. Educational resilience framework

The educational resilience framework proposed herein is expected to identify the potential gaps that lead to (or augment) educational crisis, and to guide the implementation of efficient solutions seeking enhanced educational resilience. This provides assistance in understanding the importance of the education system, especially under pandemic events that bring many difficulties and impose restrictions worldwide (e.g., remote access, distance learning, educational limitations). Figure 2 illustrates the proposed educational resilience framework.

The framework is comprised of five steps: data collection, data analysis, gaps formulation, solution development, and implementation planning. This data-driven approach employs distinct sources of data: a) surveys to be carried out with students, researchers, faculty, and workers; b) previous knowledge from the research team to provide practical examples where educational gaps exist; c) literature review aiming to collect useful information and data related to the topic; and d) search for recent news and information from reliable sources that indicate real scenarios and examples of impactful educational issues. Multiple procedures and analyses may be carried out on the collected data, including aggregating data to provide quantitative information, drawing insights, performing statistical analyses, and verifying whether the data are accurate and sufficient for the purpose of the research. Then, educational gaps are identified and classified according to a cost-effectiveness criterion, whereby listing for each gap their respective pros and cons, risks and probabilities of success, and expected outcomes. Decision-making capabilities, which may include any simulation- or optimisation-based tools, are employed for properly tackling the gaps. This is fundamental to manage the resources available to be used, and to consider potential issues, risks, and damages that may jeopardise the feasibility and success rate in the implementation phase. Finally, a deployment plan is carefully developed to guide the solution implementation. An efficient implementation is mandatory and can be performed through a structured and organised plan, whereby avoiding delays and minimising risks.

The objective of the proposed framework is to assist in identifying the most impactful educational gaps and leveraging the capabilities required for enhanced educational resilience. Such measures target people, academia, industry, and government, and concern the implementation of strategies such as public sector investments, assistance to people for easier and better adaptation against pandemic events, incentives for companies and universities to build partnerships, etc. From the guidelines provided by such a framework, decision-making based on mathematical modelling, simulation, and optimisation approaches can provide proper capabilities for achieving improved educational resilience, whereby solutions can be implemented in the real-life environment according to requirements, restrictions, and limitations (e.g., availability of people and resources, investments). To illustrate the foundation of such educational resilience framework, we provide an illustrative example that follows the sequence of steps shown in Figure 2.

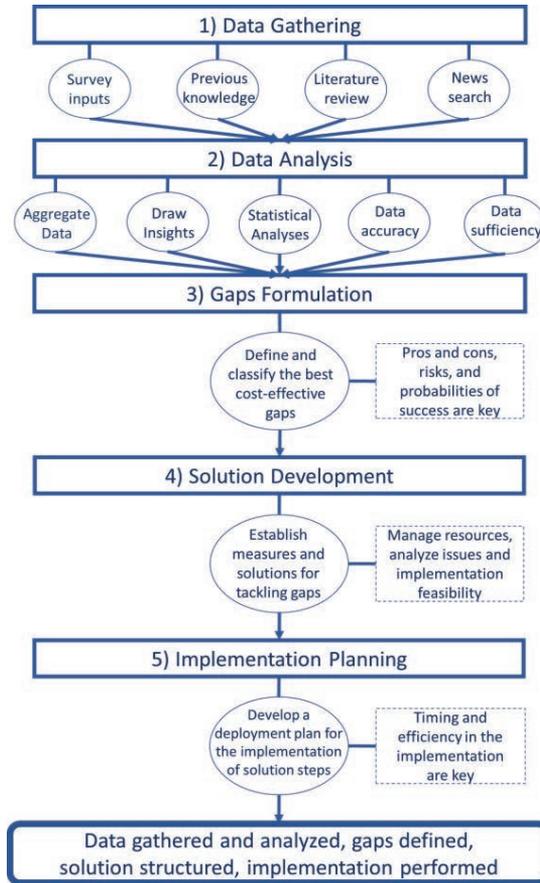


Figure 2: Educational resilience framework.

4. Illustrative example

It is known that the recent pandemic events have introduced several difficulties in the management of classes and resources of schools. Thus, a survey is launched for high school teachers over a significant number of public schools to collect information on the gaps, challenges, and improvements needed in the current educational system. The data gathered is analysed, whereby major gaps are found regarding the availability of materials and resources for laboratory classes (e.g., chemical/physics experimental classes). Those classes have to be in-person to provide a proper experience for the students, but the pandemic restrictions impose reduced classroom occupancy. Then, the laboratory classrooms are divided into multiple sub-groups, which leads to additional requirements of personnel and materials for carrying out the experiments, and hence, an overload of the system with consequent lack of resources. This includes achieving better allocation and distribution of resources to public schools to optimise such educational requirements. To properly address this concern, modelling and optimisation decision-making is employed, whereby building a mathematical model that represents the problem and utilising an optimisation algorithm to find an optimal solution to be implemented in the system. This would provide proper capabilities for systematically identifying and addressing the current and challenging educational needs.

5. Remarks and future guidelines

This work provides guidance for effectively addressing educational gaps through the development of an educational resilience framework assisted by decision-making capabilities. Guidelines are provided in a structured fashion which segments the methodology into five steps, namely, data collection, data analysis, gaps formulation, solution development, and implementation planning. They include identifying the most impactful educational gaps in times of crisis, developing solution methods and strategies that are appropriate and effective for each scenario, and providing practical guidance towards implementing these solutions to mitigate educational gaps.

The framework proposed herein is expected to: a) identify representative and impactful educational gaps and issues, which are insights drawn from the data gathered; b) provide cost-effectiveness or other similar analyses to systematically establish a criterion considering the importance to handle each gap, the cost and effort required, and the potential outcomes; and c) highlight quick and efficient methods, approaches, and solutions to be employed in the implementation steps.

From the proposed guidelines, educational resilience improvements can be achieved for multiple beneficiaries in a wide variety of problems and applications, and with multiple advantages, including better allocation of personnel, materials, and resources; improved teaching/learning; proper capabilities for the learning/development of students, teachers, and staff; more efficient investments from the government (e.g., allocation of financial resources for training, teaching, materials, infrastructure), enhanced resilience of the education system, among others. The results and conclusions derived from this work illustrate a decision-making framework can be effectively and interestingly employed towards easier and more efficient educational strategies, methodologies, and policies.

References

- M. Annarumma, L. Vitale, F. Sessa, I. Tedesco, 2020, The reorganization of the family system at the time of covid-19: educational resilience and assistive technologies, *Problems of psychology in the 21st century*.
- J.A. Downey, 2008, Recommendations for fostering educational resilience in the classroom, *Preventing School Failure: Alternative Education for Children and Youth*, 53(1), 56-64.
- R.E. Franzoi, B.C., Menezes, J.D. Kelly, J.A.W Gut, 2021, A moving horizon rescheduling framework for continuous nonlinear processes with disturbances, *Chemical Engineering Research and Design*, 174, 276-293.
- S.C. Goh, C. Elliott, T.K. Quon, 2012, The relationship between learning capability and organizational performance: A meta-analytic examination, *The learning organization*.
- H. Herrman, D.E. Stewart, N. Diaz-Granados, E.L. Berger, B. Jackson, T. Yuen, 2011, What is resilience?, *The Canadian Journal of Psychiatry*, 56(5), 258-265.
- K. Magis, 2010, Community resilience: An indicator of social sustainability, *Society and Natural Resources*, 23(5), 401-416.
- T.W. Malone, M.S. Bernstein, 2015, *Handbook of collective intelligence*, MIT Press, November.
- B.C., Menezes, R.E. Franzoi, M. Yaqot, A. Al-Banna, S. Ashkanani, M. Sawaly, J.D. Kelly, 2021, Complementarities of part-time and full-time postgraduates in process systems engineering (PSE): when industry and academia merges, *Computer Aided Process Engineering*, 50, 1559-1565.
- S. Naidu, 2021, Building resilience in education systems post-COVID-19, *Distance education*, 42, 1-4.
- B.M. Torsney, J.E. Symonds, 2019, The professional student program for educational resilience: Enhancing momentary engagement in classwork, *The Journal of Educational Research*, 112(6), 676-692.
- H.C. Waxman, J.P. Gray, Y.N. Padron, 2003, Review of research on educational resilience.
- A.M. Zamfir, 2020, Educational resilience in pandemic times and potential impacts on inequalities: The case of Romania, *Revista Românească pentru Educație Multidimensională*, 12(2sup1), 182-187.

Teaching courses heavily dependent on computational resources to STEM students during Pandemics

Nelson Chibeles-Martins,^{a,b} Lourdes B. Afonso^{a,b}

^a*Departamento de Matemática, FCT-NOVA, 2829-516 Caparica, Portugal*

^b*Centro de Matemática e Aplicações, FCTNOVA, 2829-516 Caparica, Portugal*
npm@fct.unl.pt

Abstract

The current pandemic has blatantly affected teaching at all levels. Specifically, in Portugal, for STEM (Science, Technology, Engineering and Mathematics) students attending Universities and taking high learning degrees, it was particularly harsh. Most mandatory courses implied attendance to science laboratory or computational lab lessons that had to be postponed to after lockdown times, rescheduled to late night hours, or simply converted into videoconference sessions. Teachers and researchers had to be very creative and resourceful to keep their students interested and invested in their courses in spite of the disturbing and abnormal situations caused by the pandemic. Operational Research, Simulation, and Actuarial and Financial Mathematics courses were mostly converted to online versions with videoconference lectures supported by e-learning platforms. The courses' heavy usage of computational resources and the physical characteristics of the majority of existing computer labs prevented the possibility of face-to-face classes due to social distancing even after the end of the several lockdowns. Unfortunately, online versions of these very hands-on courses revealed to be really bland and having a low appealing to students, leading to pupils' increasingly dropping out of courses and even quitting their attendance for that particular semester.

In this work the authors will present the adaptations that they decided to implement on their courses during three long semesters to make the teaching and learning experiences engaging, interesting and effective and the assessment activities reliable and efficient. Lectures being recorded (previously or live), small group assignments during classes, the use of a professional video-conferencing software, each course having its Moodle page are some of the experimented adaptations, among others. The main goal was to meliorate and even to cease students' drop-outs. This work will explore the experiments and good practices that worked and the ones that did not correspond to the expected outcomes. Additionally, some framing will be made concerning how Portuguese Government and the authors' teaching institution have supported these experiments and good practices. Some comparisons between the students' performances from the affected semesters and from previous years will also be presented.

Keywords: Virtual Education, Teaching Practices, Pandemics, Active Learning, Peer Instruction

1. Introduction

FCT NOVA (NOVA School of Sciences and Technology) is one of the most important and renowned Sciences and Engineering schools in Portugal. It offers doctorate, master and bachelor degrees in Mathematics, Computer Sciences, Physics, Chemistry, Industrial Management, Geology, Environmental Sciences, Material Sciences, among others. All offered degree programmes are accredited by the national agency for assessment and accreditation of higher education and all Engineering programmes have the EUR-ACE certification. Around 8,000 students, 400 professors, 1,000 researchers and 200 staff (NOVA School of Sciences and Technology, 2021) study, work and live in its campus making it larger than several Portuguese villages. The current CoViD19 pandemic had a tremendous impact in this academic community like it happened across the Globe, slowing down research and teaching activities, leading to students dropping out of their hard earned higher learning opportunities.

From March, 2019 to June, 2020, FCT NOVA lived through total and partial lockdowns. Regular classes were converted into remote versions, whenever possible, affecting the lives of thousands of students and hundreds of teachers. The authors were the Coordinator Professors of a total of nine different courses offered to Mathematics and Engineering degrees' students. Almost 300 students attended to their courses during the mentioned period. Table 1 lists those courses and some of its characteristics.

Table 1 - List of Courses

Course	Year	# students	Type	Level
Decision Sciences	2020/21	2	Mandatory	Doctor.
Decision and Risk	2019/20	9	Optional	Master
Decision and Risk	2020/21	9		
Financial Mathematics	2019/20	77	Optional	Bach.
Financial Mathematics	2020/21	75		
Introduction to OR ¹	2020/21	42	Mandatory	Bach.
Non-Life Insurance	2019/20	17	Mandatory	Master
Non-Life Insurance	2020/21	17		
RM ² Non-Life Insurance	2020/21	11	Mandatory	Master
Simulation techniques in RM ²	2020/21	14	Mandatory	Bach.
Simulation	2020/21	16	Mandatory	Master
Social Security and Pension Funds	2020/21	11	Optional	Master

Decision Sciences is a mandatory course for students attending the doctorate degree in Mathematics – Operational Research. Decision and Risk is offered to the Master in Analysis and Engineering of Big Data Programme. Financial Mathematics is an optional but transversal course available to every FCT NOVA Bachelor students. Introduction to Operational Research is a mandatory course in both Mathematics and Risk Management Applied Mathematics degrees. The course “Simulation techniques in Risk Management” was new and the class of 2020/21 was the first one attending it. The remaining courses are offered to students attending the Applied Mathematics Master Programme.

¹ Operational Research

² RM - Risk Management

In the following the authors will address the various pedagogical approaches implemented on the listed courses.

2. What was tried

During the three mentioned semesters classes were held remotely through a videoconference platform. The authors had to adequate the courses' documents and other pedagogical materials to this format. Additionally several approaches were implemented in order to keep the students interested and invested. Namely:

- (1) Lectures were recorded live and the recordings were later made available to students;
- (2) Lectures were previously recorded and the recordings were available to students before the corresponding remote session;
- (3) Course was supported by a MOODLE³ course page;
- (4) The usual midterm tests were replaced by several smaller and more focused tests during class time;
- (5) Group assignments solved not only during classes but also as homework;
- (6) Students studied some small part of the course syllabus, or solved a more complex problem, and prepared and presented a lecture about it, individually or in group;
- (7) Students rated their peers' works;
- (8) Students were offered additional non-mandatory tasks that could be used to improve their final grade up to a limiting threshold;
- (9) Collections of solved exercises were made available before or after the corresponding classes. Although this was already a common practice during this period it was expanded. More exercises and with a higher variety were added to the collections and it was implemented in more courses.

Table 2 presents which of the listed approaches/techniques were used in each course.

Table 2 - Pedagogical approaches

Course	Years	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Decision Sciences	2020/21					X	X			
Decision and Risk	2019/21	X		X						X
Financial Mathematics	2019/21		X	X						X
Introduction to OR	2020/21	X		X	X				X	X
Non-Life Insurance	2019/21			X		X				X
RM in Non-Life Insurance	2020/21			X		X	X			
Simulation tech. in RM	2020/21	X		X	X	X	X	X	X	X
Simulation	2020/21			X		X	X			X
Social Security and Pension Funds	2019/20			X		X				

³ MOODLE is the acronym to Modular Object-Oriented Dynamic Learning Environment. It is a digital learning platform designed to provide an integrated, secure and highly flexible system for teachers, educators and students. (The Moodle Project, 2021). Often used by schools and universities as an e-learning platform.

A typical lesson included some lecturing time where the teacher presented the important theme concepts, deduced the relevant formulae and demonstrated the fundamental theoretical results using digital slides and/or writing on shared digital white boards. Later the teacher would explore practical examples supported by specific digital applications like spread sheets, Geometry software and an adequate programming language.

Afterwards the class would be split into smaller groups and would go into separated videoconference “rooms” and would tackle with simple but consecutively more complex problems and practical case studies on their own under the supervision of the teacher. The lesson ended with the day’s theme being wrapped up with the problems solving or, at least, a possible resolution for the case study and the results’ discussion. The main goal was to emulate a face-to-face experience as close as possible using the videoconference platform and some additional digital means.

3. Some results and comments

At FCT NOVA, at the end of each semester students are obliged to fill an on-line anonymous questionnaire addressing the course and any out of the ordinary happenstance occurring during the semester. Additionally, the Coordinator Professor elaborates a report with his/hers perspectives about the course, discussing the students’ performances and commenting the students’ answers to the questionnaire.

This section presents the overall feeling of the authors concerning the implemented measures from the students’ point of view based on the information collected from the semester report of each course. No serious statistic treatment was made with that information and therefore no numbers will be presented.

3.1. Lectures’ recordings.

They were generally well received although were mostly used by students who missed some of the videoconference sessions. Some students also declared they reviewed the recordings if they were unsure of having fully understood some concept or explanation. Nevertheless, more timid students avoided asking questions knowing they were being recorded in spite of the guarantee that students’ identities were excluded from the recording. Live recordings were usually more messy and chaotic than previously recorded lectures but also realer because they included students’ questions during the lecture. Currently, even after classes went back to the classrooms and computational labs some of these last year’s recordings are still offered as additional studying material.

3.2. MOODLE

This was a no-brainer. Teaching this courses during lockdown without it or a similar platform would not be reasonably possible. The authors kept using it in all the courses they coordinate.

3.3. More evaluation moments

Students really enjoy them. Having short term objectives helped keeping them focused. Nevertheless this situation can easily become a heavy burden for the teachers. In the current school year, the courses referred in Table 2 keep using a diverse and more frequent

evaluation schedule although with some adaptations due to the lighten of the social distancing measures.

3.4. Group assignments

The general consensus is that they work really well even remotely. The current video and sound equipment present in most home computers and internet connections make it totally possible for students to follow their peers even from their bedrooms. Unfortunately, it can become a tricky situation when a student does not have a healthy family environment. Additionally, still not every family in Portugal has a good access to internet, particularly in rural areas.

Another issue arose when the groups were randomly assigned. Some students refused to work with colleagues they did not know. And simply turned off their microphones and cameras and worked alone for the rest of the session. In the current year group assignments are still a part of the courses evaluations.

3.5. Autonomous work with oral presentations

This was a hit or miss. When a student did a good job it was exciting to follow and motivating for the rest of the class. And the students enjoy it tremendously and praised it in the report. But when a student was shy, did not feel comfortable speaking in public or had some difficulties dealing with the videoconferencing technical issues the presentation could become a difficult situation to correct in the heat of the moment, a lost learning opportunity and very penalizing for the involved students. Currently, some of the courses mentioned in Table 2 still have some oral evaluation moment but with some teachers safeguard measures to ensure a higher evaluation equity.

3.6. Students marking their peers' work

It was a partial success. Some students really enjoyed it and discovered they would learn immensely from marking their colleagues' works. They became more aware of what was the teachers' expectations and therefore managed to achieve better results in the following tests. Other students never took it seriously. Additionally this kind of activities demand a tight control from the teacher otherwise it can lead to personal vendettas with a group of less mature students. This measure was kept for the current year.

3.7. Extra tasks for extra grade

As the previous measure, this was a partial success. The majority of students praised it and tried to execute those extra tasks as much as they could. Other students mentioned schedule issues and work overload and not even tried. Nevertheless, most students complained that the reward was sometimes too low for the extra work.

3.8. Collection of solved exercises.

This come to stay. The collections are already made. It is just a matter of adding more every other year.

From the authors' point of view all these implemented measures ensured that students had a good learning experience during the worse part of the pandemic. Not everything had the expected outcome but the general result is positive. Table 3 compares the average students' grades, in the [0 – 20] scale used in Portugal, where a 10 is needed for a student to successfully conclude a course. The grades in bold type and shaded in grey correspond

to the pandemic period. The excluded averages correspond to school years when the course was not being coordinated by one of the authors.

Table 3 - Average Grades

Course	Average Grades				
	2016/17	2017/18	2018/19	2019/20	2020/21
Decision Sciences	-	-	17.0	-	19.0
Decision and Risk	-	-	13.5	16.2	14.2
Financial Mathematics	14.9	-	14.8	13.1	14.6
Introduction to OR	-	-	16.1	15.8	16.9
Non-Life Insurance	-	13.8	13.9	14.0	15.5
RM in Non-Life Insurance	16.6	15.0	14.5	14.2	17.5
Simulation tech. in RM	-	-	-	-	16.3
Simulation	14.5	14.0	-	15.8	16.1
Social Security and Pension Funds	13.8	14.8	-	16.0	15.2

As can be seen, the grades did not suffer much during the lockdowns. On the contrary, in some courses an obvious improvement was noticed. And that led to the maintenance of some of the aforementioned measures even after classes returned, totally or partially, to the campus.

4. Conclusion

Both authors made important, creative and meaningful efforts during the lockdown semesters for keeping the courses they were coordinating engaging and interesting to students attending them. The additional load of work paid off because the main goal of those efforts was achieved. Students generally gave positive feedbacks on their semester questionnaire and the average grades were similar to the ones from recent years.

Additionally, in some courses the students' performance seemed to have improved leading to the maintenance, to some extents, of several of the teaching practices introduced during in the initial phase of the pandemic. Some lessons were learned and they were too good to be discarded!

Acknowledgments

This work was partially funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P. (National Foundation for Science and Technology), under the scope of the project UIDB/00297/2020.

References

- Nova School of Sciences and Technology, 2021, *Overview*. <https://www.fct.unl.pt/en/about-fct/overview>
- The Moodle Project, 2021, *About Moodle*. https://docs.moodle.org/311/en/About_Moodle

Author Index

Name	Page		Page
A. Alibrahim, Hussain	703	Al-Mohannadi, Dhabia M.	907
A. Matos, Henrique	397	Alnouri, Sabla	673
Abdelouahed, Lokmane	163		907
	571	Alnouss, Ahmed	115
Abildskov, Jens	1357		379
Abou El Qassime, Mehdi	1171		1339
Acevedo, Ariana Camille	1081	Al-Obaidi, Mudhar	133
Addis, Bernardetta	685		697
Adjiman, Claire S.	601	Al-Obaidli, Houd	1231
	607	Alpy, Nicolas	67
Afonso, Lourdes B	1687	Al-Qadri, Ali Abdullah	421
Aguilar Laserre, Alberto A.	391	Alqashouti, Noof	1681
Aguilar-Murguía, Dulce María	901	Al-Rumaihi, Aisha	805
Ahmednooh, Mahmoud	979	Alrwashdeh, Saad	697
Akhmetova, Sabira	577	Alsarayreh, Alanood	697
Akisque, Rafael Akira	139	Al-Slaihat, Ferdaous	439
Al-Ansari, Tareq	85	Alvarado-Rodríguez, Carlos Enrique	559
	115	Alves, Ediane Sá	553
	367	Amao, Khalid	1621
	379	Ambye-Jensen, Morten	877
	493	Amorim, Ana Sofia	883
	805	Anandan, Paul Danny	1093
	1231	Anantharaman, Rahul	1441
	1243	Andersson, Leif Erik	1441
	1339	Andersson, Niklas	619
	1573		1183
Albasi, Claire	1519	Andrade, Thalles Allan	877
Alcantara Avila, J. Rafael	853	Andrade Neto, Ataíde Souza	1387
Alcántara Avila, Jesús Rafael	475	Angarita-Martinez, Javier David	997
Alcántara Maciel, Francisco Daniel	475	Ansart, Renaud	511
Alcocer-García, Heriberto	811	Antxustegi, Mirari	547
Alherbawi, Mohammad	85	Anye Cho, Bovinille	19
	805	Aouichaoui, Adem Rosenkvist Nielsen	1357
	1243		1393
Al-Hotmani, Omer	133	Araujo, Mariana	1081
Allen, David T.	937	Ardakani, Mohammad Hamed	1033
	943	Arellano-Garcia, Harvey	271
Allen, Louis	1399		277
Allen, R. Cory	835		505
	919		769
Almeida Benalcázar, Eduardo Francisco	1303		1621

Arges, Christopher G.	1537	Bausa, Irina	151
Arjun, Ramprasad	1429	Belaidi, Chanez	529
Armstrong, Alan	601	Benjelloun, Saad	415
	607		445
Aroniada, Magdalini	31		1291
Arrieta-Escobar, Javier A.	847	Benyahia, Brahim	181
Arsuaga, José Ignacio	1513		775
Arvanitidis, Achilleas L.	7		1093
Aryan, Venkat	829	Bernardi, Andrea	955
Asante, Samuel	349	Beykal, Burcu	205
Ataman, Nazlı	991		835
Attarakih, Menwer Mohammed	439	Bezzo, Fabrizio	25
Attramadad, Kari J. K.	1495		1567
Avraamidou, Styliani	835	Biggins, Flora	79
	1573	Bisotti, Filippo	55
Aydin, Erdal	661	Bispo, Artur	1645
Azais, Antonin	1519	Blakely, Logan	583
Azzaro-Pantel, Catherine	67	Bolt, Cameron Ernst	145
	391	Boly, Vincent	847
	895	Bouchkira, Ilias	445
B. B. Costa, Caliane	781	Boukharfane, Radouan	415
Babi, Deenesh K.	925	Boukouvala, Fani	583
Bachawala, Sunidhi	1087	Bozzano, Giulia	55
	1129	Braeken, Leen	175
Bähner, Franz D.	217	Brand Rihm, Gerardo Martin	385
Baillon, Fabien	1459		1489
Bajcinca, Naim	481	Braz, Ana Catarina	73
Bakošová, Monika	1021		1123
Baldan, Maurício Ribeiro	199	Brener, Arnold	577
Baldea, Michael	937	Briceno-Mena, Luis A.	1537
	943	Brignole, Nelida Beatriz	1405
Bano, Gabriele	25	Brown, Solomon	79
	31		157
Bansal, Abhishek	1153		427
Barbieri, Massimo	1321	Brown, Solomon F.	1399
Barbosa-Póvoa, Ana	1651	Bruns, Bastian	859
Barbosa-Póvoa, Ana P.	967	Bugryniec, Peter	157
Bardow, André	1201	Burghardt, Célia	1219
	1441	Bynum, Michael	583
Bareschino, Piero	1	Caballero Suarez, Jose Antonio	649
Barolo, Massimiliano	25		781
Barz, Tilman	1069		793
Basso, Thiago Olitta	211		973
Batool, Iqra	481	Cabrera Ruiz, Julián	475
Baumgärtner, Nils	1201		853

Cafaro, Diego C.	949	Choubert, Jean-Marc	1519
Cafaro, Vanina G.	949	Christodoulou, Charalampos	25
Calderon-Alvarado, Martha Patricia	49	Claussen, Heiko	1507
Calixto, Ewerton	1645	Coelho, Pedro	397
	355	Colet-Lagrille, Melanie	223
Camarda, Kyle V.	229	Continillo, Gaetano	1
Cameron, Ian T.	1663	Contreras-Zarazua, Gabriel	613
	1675		655
Campbell, Timothy Joel Samuel	775	Cordiner, Joan	1399
Campos, Jean Carlos	271	Cormos, Ana-Maria	631
Cansino-Loeza, Brenda	1579	Cormos, Calin-Cristian	631
Capron, Bruno Didier Olivier	1387	Cortez-González, Jazmín	559
	1423		1555
Carabineiro, Hugo M.D.	817	Cos, Ralph	37
Cardenas, Cristian	457	Costa, Alexandra	331
	463	Costa, Caliane	649
Carnero, Mercedes	1291		781
Carrera, Eduardo	895	Couenne, Françoise	187
	1531	Coutu, Arnaud	529
Carvalho, Ana	1651	Cremaschi, Selen	709
Casillas Céspedes, Victor Enrique	475	Cristea, Vasile Mircea	1375
Casoni, Andres	715	Cristofari, Christian	895
Castel, Christophe	679	Cruz Bournazou, Mariano Nicolas	1069
	685		1111
Castro Oliveira, Miguel	397	Cuttillo, Enrico Alberto	1
Castro-Amoedo, Rafael	871	Da Silva, Liangrid Lutiani	199
Castro-Rodriguez, Daniel	331	Dachselt, Raimund	1141
Cegla, Maximilian	1189	D'Amore, Federico	1567
Cenci, Francesca	25	Danesh Alagheband, Tina Sadat	1501
Cha, Sujin	241	D'Angelo, Sebastiano Carlo	799
Chachuat, Benoit	931	De Beer, Julien	1345
	1633	De Beer, Thomas	259
Chakraborty, Arijit	1507		289
Chalkis, Apostolos	1297	De Blasio, Cataldo	811
Chavez-Salas, Jorge	1225	De Clercq, Dries	175
Chibeles-Martins, Nelson	1687	De Lima, Lionete Nunes	139
Chico-Proano, Andres	343	De Meio Reggiani, Martin Carlos	1405
Chkifa, Abdellah	1291	De Mel, Ishanki	1207
Cho, Bovinille Anye	109	De Miranda, Gabriel Fiúza Moreira	1423
Cho, Hyungtae	241	De Moraes, Marcellus Guedes Fernandes	1423
	1369	De Oliveira Júnior, Silvio	1315
Cho, Jae Hyun	283	De Prada, Cesar	151
Cho, Seolhee	841		985
Chokparova, Zarina	1267	De Sousa Júnior, Ruy	139

De Souza Jr., Mauricio Bezerra	1423	El-Halwagi, Mahmoud M.	1591
De Vaal, Philip	145	Elmaz, Furkan	175
Del Rio Chanona, Antonio	19	Elmisaoui, Sanae	1291
	1039	Elomri, Adel	85
Delage, Thibault	535	Engell, Sebastian	499
Delgado, Pedro	1381		1165
Deligianni, Vasiliki	1297		1189
Deliismail, Ozgun	661		1483
Demirhan, C. Doga	919	Eränen, Kari	73
Dering, Daniela	1099	Errington, Ethan	1615
Deutz, Sarah	1201	Erturk, Emrullah	661
Di Caprio, Ulderico	175	Esche, Erik	385
	1411		1489
Di Martino, Marcello	835	Espinosa-Cardenas, Juan Pablo	997
	1573	Espinoza, Daniel	1183
Di Pretoro, Alessandro	91	Esposito, Flora	463
Diangelakis, Nikolaos A.	205	Espuña, Antonio	787
	1159	Estel, Lionel	163
Diaz, Soledad	715	Estrada, Carlos A.	505
Dimitriou, Konstantinos	889	Estrada, Vanina	715
Dirion, Jean-Louis	1459	Eydenberg, Michael	583
Dirza, Risvan	1249	F. Santos, Lucas	781
Do, Thai Ngan	763	Facco, Pierantonio	25
Dologlu, Pelin	991	Falk, Véronique	847
Dorneanu, Bogdan	271	Fan, Fan	1393
	277	Favre, Eric	679
	505		685
	769	Fedeli, Matteo	55
	1621	Fichera, Alberto	547
Dos Santos, Allyne M.	1495	Fikar, Miroslav	1045
Dou, Yi	1585	Filipe, Rui	883
Doubek, Gustavo	361	Floquet, Pascal	1501
Drofenik, Jan	487	Flores-Perez, Jose Manuel	391
Du Toi, Jan-Pierre	19	Flórez-Orrego, Daniel	1315
Duarte, Alexandra	997	Foo, Dominic C Y	1603
Duarte, Inês	1009	Foppoli, Marco	451
Dürr, Robert	691	Fraga, Eric S.	343
Earp, Dave	37	Franke, Meik	637
Eden, Mario R.	709	Franke, Meik B.	595
Ejeh, Jude	79	Franzoi, Robert	979
	1399		1003
El Hamra, Fatima Ez-Zahra	415		1609
Eleff, Sam	835		1681
Elekidis, Apostolos P.	1561	Fricke, Armin	439

Friis, Jesper	1237	Govindan, Rajesh	367
Frohmann, Julia	1201		1231
Fujii, Shoma	469	Graells, Moisès	373
Furlan, Felipe Fernando	307		787
	553	Granjo, José	217
Gabriel, Alex	847		517
Gagliano, Antonio	547		1123
Galeazzi, Andrea	1321	Grisales Diaz, Victor Hugo	1075
	1543	Groß, Sebastian	1069
Galindo, Amparo	601	Grossmann, Ignacio E.	841
	607	Grünewald, Marcus	859
Gallo, Francesco	1543	Guerras, Lidia	625
Gando-Ferreira, Licínio M.	517	Gui, Lingfeng	607
Garcia Castro, Florencia Lujan	973	Guillén Gosálbez, Gonzalo	169
García-Saravia, Rafael Cristobal	997		799
García-Trejo, Juan Fernando	901	Guimarães, Renato Jose	553
Geissler, Caleb H.	913	Guo, Miao	1615
Gencer, Emre	1597	Gutiérrez-Antonio, Claudia	193
Gentimisc, Thanassis	565		901
Georgiadis, Michael C.	17		1195
	1561	Gutiérrez-Guerra, Roberto	559
Gerard, Bruno	1531	Haddad, Joshua	583
German, Boris G.	343	Hagelien, Thomas F	1237
Gernaey, Krist V.	925	Hahne, Mihaela	1345
	1627	Hall, Richard	61
Gerogiorgis, Dimitrios	331	Hamedi, Nazanin	1141
	1117	Hamroun, Boussad	187
Ghebredngl, Salmon	1135	Han, Chengyu	1057
Ghiat, Ikhlas	493	Hanich, Ronny	829
Ghidaglia, Jean-Michel	415	Hans-Jörg, Bart	439
Giannikopoulos, Ioannis	937	Hasebe, Shinji	319
	943	Haubensack, David	67
Giordano, Roberto De Campos	307	Hayashi, Yusuke	721
	553	Heiho, Aya	1585
Gómez-Castro, Fernando Israel	193	Heinzelmann, Norbert	277
	1015	Hellinckx, Peter	175
	1195		1411
	1225	Hemmati, Hadiseh	1033
Gomis-Fons, Joaquín	619	Hemrle, Jaroslav	871
Gonçalves, Bruno S.	967	Heng, Jerry Yong Yew	1615
Gonzalez, Marcial	1087	Herbinet, Olivier	1471
	1129	Hernández, José	1285
Gonzalez Alriols, Maria	547	Hernández, Salvador	193
González-Núñez, Sofia	625		901

Hernández-Vargas, Esteban Abelardo	49	Kandlbauer, Lisa	121
Herwig, Christoph	1177	Kanematsu, Yuichiro	469
Heuberger-Austin, Clara F.	919	Kanniche, Mohamed	679
Hicks, Aaron	97	Karlsen, Espen	1495
Hirono, Keita	721	Kato, Kazuki	325
Hlawitschka, Mark	349	Kay, Harry	97
Hoffmann, Christian	1489		1525
Hong, Seokyoung	1369	Kay, Sam	97
Horsch, Martin	1237		1525
Hortua Orjuela, Hector Javier	103	Kayahan, Emine	1411
Hu, Guoqing	1051	Kedawat, Ayush	1153
Huang, Yan-Shu	1081	Kenefake, Dustin Ryan	1273
	1087	Kenig, Eugeny	127
Huerta-Rosas, Brenda	247	Khalifa, Ridab	85
Husfeld, Russell R.	835	Khamar, Lhachmi	445
Huusom, Jakob Kjøbsted	1075	Khaydarov, Valentin	1363
Huynh, Thien An	595	Khodier, Karim	121
	637	Kienle, Achim	691
Ibrahim, Dauda	1633		727
Ioannou, Iasonas	169	Kikuchi, Yasunori	469
Ipanaque, William	271		1585
	505	Kilwein, Zachary	583
Işeri, Funda	991	Kim, Changsu	763
Iten, Muriel	397	Kim, Jiyong	763
Ja, Ahmed	415		1447
Jaber, Hadi	913		1453
Jain, Aarjav	157	Kim, Jong Woo	1069
Jalving, Jordan	583	Kim, Junghwan	241
Jang, Kyojin	1447		1369
Janiga, Gábor	13	Kim, So-mang	265
Jäschke, Johannes	1135	Kim, Yeonghyun	283
	1147	Kim, Youngjin	283
Javaloyes-Anton, Juan	649	Kim, Yurim	241
Ji, Cheng	1057	Kino-Oka, Masahiro	721
Jing, Keju	109	Kis, Zoltán	1633
John, Yakubu	133	Kitagawa, Akihiro	325
Jones, Wil	1117	Klaimi, Rachid	673
Jonuzaj, Suela	601	Klein, Peter	1237
Joo, Chonghyo	1369	Klymenko, Oleksiy V.	1207
Juárez-García, Maricruz	613	Kokossis, Antonis	565
Jurischka, Constantin	769		889
Kabak, Özgür	991		1297
Kager, Julian	1177		1351
Kaldellis, John K.	1309		1657

Kolis, Stanley P.	607	Leonard, Gregoire	1663
Konchakova, Natalia	1237	Leonard, Grégoire	265
Kondili, Emelia	1309		301
	1663		1675
	1675	Lewin, Daniel R.	1663
Kontoravdi, Cleo	733		1669
	1633		1675
Kostoglou, Margaritis	7	Li, Dan	961
Kourdourli, Fatma	163	Li, Jie	823
Krausch, Niels	1069		961
Kravanja, Zdravko	487		1279
Krenn, Tobias	121	Li, Yang	1063
Krishna M, Aswin	1153	Lim, Jonghun	241
Kröger, Corinna	1363	Lim, Jongkoo	1369
Kruber, Kai Fabian	829	Lima, Fernando Arrais Romero Dias	1423
	1261	Lima Fernandes, Marcelino Artur	337
Kumar, Ashish	5	Lindahl, Simon Brædder	925
	289	Linke, Patrick	1573
Kunde, Christian	727	Liu, Jiayu	181
Kurban, Sena	991	Lizama-Valenzuela, Franco	223
Kuşoğlu Kaya, Gizem	991	Lizcano-Prada, Javier Oswaldo	997
L. E. Swartz, Christopher	1099	Llano-Ponte, Rodrigo	547
Lagare, Rexonni	1081	Longati, Andreza Aparecida	307
Lagoda, Konrad	1177	Lopes, Braulio Haruo Kondo	199
Laird, Carl	583	López García, Jesús Armando	475
Lakelin, Matthew	955	López-Ramírez, Varinia	1555
Lane, Amanda	97	Loureiro, Luciana Neves	1387
	1525	Loweski Feliz, Minda	571
Lara-Montaño, Oscar Daniel	1195	Lubušký, Karol	1045
Lasala, Silvia	1471	Lucia, Sergio	1069
Latifi, Abderrazak M	445	Lun, Denis	1531
	457	Luna, Martin Francisco	1111
	463	Ma, Fangyuan	1063
	1291	Ma, Jiaying	1063
Le Roux, Galo Antonio Carrillo	211	Ma, Yingjie	1279
	235		
	1105	Macchietto, Sandro	523
Leal, Adriano Galindo	1549	Machado, Ricardo Antonio Francisco	199
Leal Silva, Jean Felipe	361	Mache, Julian	799
Leblebici, M. Enis	175	Maciel, Maria Regina Wolf	313
	1411	Maciel Filho, Rubens	313
Lee, Seongju	763		361
	1453		1303
Lehr, Annemarie	13	Mackey, Hamish R.	805

Mädler, Jonathan	1255	Martínez-Sánchez, Osiris	1015
	1435	Martins, Cristina	1381
Maenaka, Kaori	319	Martins, Fernando G.	1663
Majgaonkar, Pranav	829		1675
Maki, Taisuke	325	Martin, Irem	991
Manca, Davide	451	Massari, Dario	451
	1639	Matias, Jose	1135
Mancino, Gianluca	745	Matos, Henrique A.	73
Mancusi, Erasmo	1		883
Manenti, Flavio	55		1123
	91	Matos, Henrique A.S.	817
	595	Maxwell, Peter	1147
	1321	Mckay, Gordon	115
	1543		805
Manno, Andrea	1213		1243
Manos, George	343		1339
Manrique-Silupu, Jose	271	Medonza, Cesar	97
	505	Meimaroglou, Dimitrios	1471
Mansouri, Seyed Soheil	1357	Melampianakis, Evripidis	1657
	1663	Méndez, Ronald	727
	1675	Mendoza, Cesar	1525
Maravelias, Christos T.	913	Menezes, Brenno	979
Marcato, Gianmarco	1567		1003
Marchesan, Andressa Neves	313		1609
Maréchal, François	865		1681
	871	Mercandetti, Alex	169
	1315	Mercelis, Siegfried	175
Marias, Frédéric	1105		1411
Marinos, Ioannis	565	Mészáros, Alajos	1021
Marousi, Asimina	1297	Metten, Bert	175
Marques, Luíza	883	Mihály, Norbert-Botond	1375
Marr, Stephen	61	Milhé, Mathieu	1459
Marsteau, Stéphanie	457	Minetti, Gabriela	1285
	463	Minta, Ali Badara	1519
Martelli, Emanuele	1213	Miranda, Everson Alves	307
Martelo Schmalbach, Nezly	547	Miranda, Letícia Passos	553
Martes Hernández, Leonela	571	Mirlekar, Gaurav	1027
Martin, Philip	97	Miroschnitschenko, Anna	1261
	1525	Mocellin, Paolo	1567
Martín, Mariano	625	Moffatt, Scott G.	835
Martin Salvador, Pedro	259	Mohamed, Abdulkarim	493
Martinez, Ernesto Carlos	1111	Mohamed, Ismail	883
Martinez Triana, Alvaro Andres	679	Mojto, Martin	1045
Martínez-Guido, Sergio Iván	901	Montastruc, Ludovic	91

Montes, Daniel	985	Nilges, Benedikt	1219
Moon, Il	241	Nilsson, Bernt	619
	283		1183
	1369	Nolzen, Niklas	1201
	1447	Noorman, Henk	1303
Morales, Alejandro	301	Nopens, Ingmar	289
Morales-Rodriguez, Ricardo	757	Nord, Lars O.	1027
	1225	Novak Pintarič, Zorka	487
Motta, Ingrid Lopes	319	Oh, Seunghyeon	1447
Moulines, Eric	1171	Ojo, Gbemisola	229
Moustapha Mai, Tchougoune	895	Oliveira, Nuno M.C.	817
Mowbray, Max Richard	97	Oliveira, Rafael David De	211
	1039	O'Neill, Eric G.	913
	1525	O'Neill, Sam	61
Muhieddine, Mohamad H.	601	Orozco Alvarado, Gustavo Adolfo	103
Mujtaba, Iqbal	133	Othman, Mohamad Rizza	433
	697	Otto, Eric	691
Müller, Don Fabian	1177	Ouaret, Rachid	1501
Munguia-Lopez, Aurora Del Carmen	1579		1519
	1591	Oudon, Baptiste	529
Muñoz, Raúl	151	Pacheco-López, Adrián	787
Murrieta-Dueñas, Rodolfo	559	Pahor, Bojan	487
	1555	Papadokostantakis, Stavros	1465
Musabekova, Leila	577	Papageorgiou, Lazaros G.	343
Muthukkumaran, Arun	1429	Papathanasiou, Maria M.	733
Nagy, Zoltan	1081		931
	1087		955
Nair, Purusothmn Nair S Bhasker	1603		1633
Najarnezhadmashhadi, Ali	73	Pappas, Iosif	835
Nakama, Caroline Satye Martins	1147	Paredes, Rodrigo	1417
Nakamura, Tomoki	853	Park, Hyundo	1369
Namany, Sarah	1573	Park, Minseong	763
Naranjo, Luz Nadiezda	289	Park, Song Won	1333
Narváez C., Ricardo A.	343		1549
Narváez Rincón, Paulo César	847	Parthasarathy, Prakash	805
Nascu, Ioana	1159		1243
Nasir Al-Rowaili, Fayez	703	Patel, Raj	133
Nasr Esfahani, Kourosh	373		697
Negny, Stephane	1501	Patrascioiu, Cristian	253
Neubauer, Peter	1069	Patrocínio, Francisco J.G.	817
Neveux, Thibaut	535	Paulen, Radoslav	1045
	679	Paulo, Helena	967
	685	Pavao, Leandro	649
Nielsen, Adem A.R.	1597	Pedrozo, Alejandro	715

Peña-Torres, Daniel	223	Pruvost, Jérémy	403
Pepe, Francesco	1	Pujan, Robert	295
Père, Valentin	1459	Pyladarinos, Alexandros	565
Pérez-Moya, Montserrat	373	Quina, Margarida M. J.	517
Pesenti, Giuseppe	451	Quiroz-Ramirez, Juan Jose	247
Pessoa, Fernando. L. P.	355		643
	1645		655
Peters, Greg	1465	R. Mackey, Hamish	1243
Petrescu, Letitia	631	Raeisi, Maryam	595
Petsagkourakis, Panagiotis	1039	Raeisi, Maryam	637
Piccialli, Veronica	685	Rahm, Julian	1435
Pico, Davide	829	Rahman, Nawwar A.	253
Picón-Núñez, Martín	757	Ramírez Márquez, César	475
Pinheiro, Carla I.C.	1123	Ramírez-Corona, Nelly	811
Pinheiro, Carolina T.	517		1015
Pinho, Manuel L.	1123	Ramlow, Heloisa	199
Pinto-Varela, Tânia	967	Ramos, Fernando	715
Pirola, Carlo	55	Rato, Tiago J.	1381
Pistikopoulos, Efstratios N	205		1417
	409	Ravagnani, Mauro	649
	835		781
	919	Ravichandran, Arjun	1429
	1159	Raynal, Ludovic	745
	1273	Reinert, Christiane	1201
	1573		1219
Pitarch, José Luis	985	Reis, Marco S.	1381
Pitault, Isabelle	187		1417
Podar, Smaranda	151	Reklaitis, Gintaras	1081
Ponce-Ortega, Jose Maria	613	Reklaitis, Gintaras V.	1087
	1579	Rendall, Ricardo	1381
	1591	Reneaume, Jean-Michel	1105
Ponche-Rocha, Jaime D.	757	Rengaswamy, Raghunathan	1429
Ponsich, Antonin	391	Repke, Jens-Uwe	385
Popescu, Marian	253		1489
Posada, John	1303	Reyes Alonzo, Luis Cesar	571
Pott, Robert William McClelland	19	Ribeiro Domingos, Meire Ellen Gorete	1315
Prado-Rubio, Oscar Andrés	1075	Riboldi, Luca	1441
Preisig, Heinz A.	295	Ricardez-Sandoval, Luis	1663
	1237		1675
Presser, Demian J.	949	Rielly, Chris D	775
Price, Nicholas	157		1093
Prifti, Kristiano	1321	Riese, Julia	859
	1543	Rivadeneira, Daniel	343
Procópio, Dielle Pierotti	211	Rivas-Interian, Raul Mauricio	643

Rivera-Gil, Jose-Luis	847	Sawaly, Razan	493
Roberts, Diarmid	79	Sayyed, Fareed Bhasha	607
Robertson, Daniel	61	Schellhas, Lars	1201
Rocha, Antônio	1387	Schilling, Johannes	1441
Rodarte De La Fuente, Abraham	49	Schlesinger, Robert	349
Rodriguez Hernandez, Manuel	1033	Schnitzlein, Klaus	277
Rogers, Alexander	109	Sébile-Meilleroux, Joris	403
Roh, Kosan	1219	Secchi, Argimiro Resende	1387
Romagnoli, Jose A.	1537	Segovia-Domínguez, Ignacio De Jesus	1555
Romero, Michelle	343	Segovia-Hernandez, Juan Gabriel	49
Romero-García, Ana Gabriela	811		247
Romero-Izquierdo, Araceli Guadalupe	193		613
Ross, Brandon Sean	19		643
Rossi, Mattia	595		655
Ruiz Femenia, Ruben	973		811
Russo, Vincenzo	73	Seidel-Morgenstern, Andreas	13
S. Khalafalla, Siddig	703	Semrau, Robin	1165
S. S. Ravagnani, Mauro A.	781	Serna, Juliana	847
Sa, Tongxin	667	Serneels, Sven	1507
Sachio, Steven	733	Serra, Sylvain	1105
Sailau, Assel	577	Serrano-Rubio, Juan Pablo	1555
Salaun, Jessica	529	Shah, Nilay	931
Salcedo-Diaz, Raquel	973		955
Salmi, Tapio	73		1633
Salto, Carolina	1285	Shahbaz, Muhammad	115
Sami Abushaikha, Ahmad	493	Shaikh, Abdul Aziz	379
Sánchez, Mabel	1285	Shamaki, Patience Bello	235
Sánchez-Bautista, Aurora De Fátima	1591	Shanbhag, Sai	1429
Sanchez-Ramirez, Eduardo	49	Sharma, Shivom	865
	247		871
	643	Shavaliyeva, Gulnara	1465
	655	Sheriff, M. Ziyen	1081
	811		1087
Santana, Leonardo O. S.	355	Shi, Songjie	1597
	1645	Shokry, Ahmed	1171
Santos, Alessandro Santiago Dos	1549	Short, Michael	1207
Santos, Andreia	1651		1561
Santos, Lino O.	517	Sidek, Norul Malakiah	433
	1417	Sildir, Hasan	661
Santos, Lucas	781	Simonin, Olivier	511
Sarc, Renato	121	Sin, Gürkan	925
Sarigol, Ilkan	907		1357
Sarkis, Miriam	931		1393
Sarrafi, Sahin	661		1597
Sasi, Thulasi	751		1627

Siqueira Lopes Ritta, Almir Guilherme	511	Teles Dos Santos, Moisés	1315
Skiborowski, Mirko	751	Thévenin, Dominique	13
	829	Tian, Xueyu	639
	1261	Tillmanns, Dominik	1201
Skogestad, Sigurd	1249	Titica, Mariana	403
	1495	Tokoro, Chiharu	1585
Skouteris, Alkiviadis	937	Tomaselli, Andrea	91
	943	Tonomura, Osamu	319
Sochard, Sabine	1105		325
Somoza-Tornos, Ana	787	Topcu, Y. Ilker	991
Song, Ziqi	109	Torres, Alice De Carvalho Lima	139
Sotowa, Ken-Ichiro	325	Torres, Ana Inés	1513
Sousa, Ana M.	517	Toth, Andras Jozsef	43
Sousa, Renata Samara	1333	Toth, Eszter	745
Souza, Ana	1645	Tran, Anh	835
Spang, Bernhard	127	Tregambi, Claudio	1
Stadtherr, Mark A.	937	Triantafyllou, Niki	955
	943	Trinh, Cindy	1471
Stijepovic, Mirko	673	Tsakona, Dimitra	1351
Stollberg, Christian	769	Tseng, Wei-Fu	523
Stouffs, Pascal	67	Tula, Anjan K.	709
Stylianopoulou, Konstantina	1309	Tzanakopoulou, Vasiliki E.	331
Suárez-Toriello, Víctor Alejandro	247	Udugama, Isuru A.	721
Su-Feher, Denis	409	Untrau, Alix	1105
Sugiyama, Hirokazu	721	Urbas, Leon	1141
	1663		1255
	1675		1267
Sun, Wei	1057		1363
	1063		1435
Sundmacher, Kai	1477	Uribe-Rodriguez, Ariel	997
Suresh, Resmi	1153	Usama, Ahmed	421
Suwa, Izuru	1585	Valenzuela-Venegas, Guillermo	223
Taira, Gustavo Ryuji	1549	Vallejo-Blancas, David	655
Tak, Kyungjae	931	Vallières, Cécile	457
	1633	Van Der Ha, Olivier	175
Tallvod, Simon	619	Van Gerven, Tom	1411
Tamagnini, Filippo	499	Vasičkaninová, Anna	1021
Tan, Raymond R	1603	Vechetti, Aldo	1327
Tan, Ruomu	217	Veldandi, Pavan Kumar	37
Tao, Tingting	1057	Veloso, Madalena	1651
Taouk, Bechara	163	Venkatasubramanian, Venkat	1507
	571	Ventura, Patrícia Rodrigues	1387
Tardioli, Paulo Waldir	139	Vergara Espinosa, Laura Alejandra	103
	553	Vernuccio, Sergio	157
Tatschl, Lisa	121	Verrengia, María De Los Milagros	1327

Vescovi, Guilherme	67	You, Chanhee	763
Vicente, Margarida L.C.	217	You, Fengqi	639
Viedt, Isabell	1255		1051
	1435	Young, Katherine	1081
Vieira, Miguel	967	Yu, Haoshui	1597
Vigier, Hernan Pedro	1405	Yusuf, Noor	367
Villar, Luciana Belen	1405	Zahid, Umer	703
Viswanath, Shekhar K.	601	Zahidova, Aysel	661
Vollmer, Nikolaus I.	1627	Zannane, Sanaa	535
Volpe, Rosaria	547	Zaryab, Sied Ali	1213
Von Der Aßen, Niklas	1219	Zhang, Aibo	1597
Vueva, Yuliya	25	Zhang, Dongda	19
	31		97
Waldherr, Steffen	1411		109
Wang, Jiayuan	667		961
Wang, Jingde	1057		1039
	1063		1525
Wang, Wanrong	823	Zhang, Liping	961
Wang, Zihao	1477	Zhang, Nan	823
Weigert, Joris	1489		961
Wibbing, Daniel	1177		1279
Wilkes, Mathew	427	Zhao, Lu	1531
Winz, Joschka	1483	Zhou, Teng	1477
Wood, Paul	61	Zhu, Lingyu	667
Wrigley, Paul	61	Zibart, Alexander	127
Wu, Min	175	Zirngast, Klavdija	487
	1411	Zitte, Bertrand	187
Xenios, Stefanos	541	Zomer, Simeone	25
Xu, Shuang	709	Zondervan, Edwin	595
Yamane-Nolin, Mikael	619		637
Yang, Jiadi	1165		1663
Yaqot, Mohammed	1609		1675
Yeadley, Aaron	79		
Yeadley, Aaron Steven	157		
	1399		

