12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Enhanced Decision-Making in Gas Lift Optimization through Deep Neural Network-based Multi-Objective Approaches and Feasible Operating Regions

Carine Menezes Rebello * **Johannes Jäschke** *
Idelfonso B. R. Nogueira *

* *Department of Chemical Engineering, Norwegian University of Science and Technology, Norway (e-mail: carine.m.rebello@ntnu.no, johannes.jaschke@ntnu.no, idelfonso.b.d.r.nogueira@ntnu.no).*

Abstract: Decision-making flexibility can be a key challenge in optimizing oil well production through a gas lift process. In this work, we introduce a multi-objective optimization strategy facilitated by deep neural networks (DNNs) as surrogate models to lessen the computational burden. Together with a likelihood test, we build a feasible operating region (FOR) using points from particle swarm optimization. Thus providing a tool for the refined process operation. We also subdivide the pareto region into constraint-compliant sub-regions, amplifying operational flexibility and identifying optimal settings. An optimality analysis is included to validate the results and assure the reliability of the surrogate-based optimization. This framework generates an operational map that can be instrumental for real-time process monitoring. Importantly, these computational tools can support the quality of real-time decisions in system operation by providing nuanced, data-driven insights into trade-offs and optimal conditions.

Keywords: Gas lift system, multi-objective optimization, surogate models, optimality, feasible operation region.

1. INTRODUCTION

A multi-objective approach can be used to address the complex challenges of allocating scarce resources, such as injected gas, in gas lift systems for oil wells (Ray and Sarker, 2006). Considering multiple objectives, such as maximizing daily oil production and minimizing gas consumption, enables operators to assess trade-offs between divergent goals and find solutions that effectively optimize oil production and available resources. It's important to note that multi-objective optimization can be computationally expensive if based on first-principle models (Guerra et al., 2022). Additionally, using a Pareto front can be limiting for operators who need to understand acceptable fluctuations in the system (Ranjan et al., 2015).

In the literature, some studies have explored how the decision-making process in multi-objective solutions can be significantly improved using the feasible operating regions (FOR) concept. The FORs have been employed to analyze operational variables and gain insights into the system's behavior (Rebello et al., 2022, 2021; Nogueira et al., 2019). To advance this work, our study employs a likelihood test to construct a FOR by recycling points provided by swarm particle optimization. This enhances unit operation flexibility and provides a new tool for efficient decision-making (Nogueira et al., 2019). Thus, a map of feasible operations for process operating variables is a valuable tool that provides significant support in decision-making, playing a crucial role in monitoring and effectively managing the process.

For gas lift systems specifically, multi-objective optimization remains a relatively small area in the literature. In this scenario, the operational mapping approach for multi-objective optimization of the gas lift system generates a refined operation map to tackle fluctuations in gas availability. This map furnishes valuable insights into the specific combinations of operational parameters that can maximize multiple objectives.

This study uses artificial neural networks (ANN) as surrogate models to capture the system's behavior, taking advantage of their ability to model complex and non-linear relationships between variables (Azlan Hussain, 1999; Lin et al., 1999; Ghiassi and Nangoy, 2009). The ANNs facilitate the optimization process by reducing computational effort and enabling the analysis of trade-offs between different objectives with low computational cost (Rashid et al., 2012; Ray and Sarker, 2007).

ANNs have become increasingly popular in optimization tasks across various domains (Santana et al., 2023; Rebello et al., 2021). However, these models can introduce artificial minima to the objective function landscape (Santana et al., 2023). Such spurious minima can distort the true topological features of the problem space. This emphasizes the need for a post-optimization evaluation to validate the optimality of results derived from ANNs. Here, we carry out a post-optimization test to ensure that the results ob-

⋆ SubPro project.

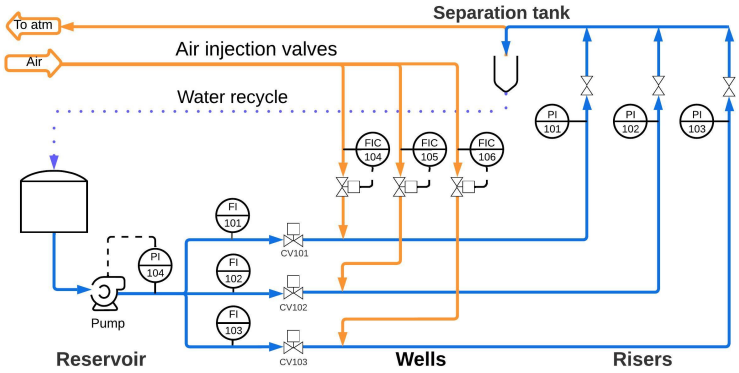


Fig. 1. Experiment scheme of small-scale gas lift pilot unit.

tained from the surrogate-based optimization are reliable. This can be listed as another contribution of this work.

The objective of this work is to perform multi-objective optimization of a gas lift system to maximize oil production and minimize gas injection at steady state (SS). The main contributions consist of the following:

- Artificial neural networks are used as surrogate models to reduce the computational burden in solving multi-objective optimization problems.
- An operational map is developed to assist decision-making on operational parameters amid fluctuations in gas availability.
- A post-optimization test is conducted to confirm the reliability of the results of the ANN-based optimization process.

2. METHODOLOGY

The methodological approach presented in this work begins with synthetic data generated from the phenomenological process model. These data were employed to identify the neural network's (NN) architecture. Multi-objective optimization was carried out using surrogate models to construct the Pareto Front. The optimal operating variables from the Pareto Front were utilized to query the phenomenological model and assess the consistency of the developed optimization. Subsequently, the Fisher-Snedecor test was employed to define a feasible operational region close to the Pareto Front, which was later subdivided into three subgroups to construct the operational map, see Figure 2. The details are described below.

2.1 Process model

This work used a phenomenological model to simulate underwater oil wells in a gas lift lab unit. The system includes reservoir, well, and riser sections using water and air instead of oil and gas. Valve adjustments mimic reservoir behaviors, and flow meters are placed before the reservoir valves; see Figure 1

The experimental prototype simulates wells using three flexible hoses and a gas lift system injecting air after control valves CV101, CV102, and CV103. A control system (FIC104, FIC105, and FIC106) regulates the injected gas flow. The risers are represented by three vertical tubes with pressure measured by gauges PI101, PI102, and PI103. The liquid is recirculated to the reservoir, while air is expelled. This study employs a validated first-principle model of the gas lift process. The model used for this system is based on Matias et al. (2022) adaptable model to the system. It includes algebraic and differential equations and considers hydrostatic pressure and friction-induced pressure loss. In calculations, only two pressures are considered: one at the bottom and one at the top of the riser, (Matias et al., 2022). This algorithm was developed in MATLAB with CASADI to solve the system's differential algebraic equation (DAE).

2.2 Neural network identification

The development of artificial neural networks entails gathering comprehensive and varied data that includes exceptions and boundary conditions. This study used synthetic data from a validated gas lift model as the virtual plant in a software-in-the-loop (SIL) approach.

Careful selection of input variables is important, e.g., to avoid data discrepancies and overfitting. We use latin hypercube sampling (LHS) to efficiently generate representative data (Stein, 1987). The LHS was used to generate the sequence of experiments to be performed in the model, thereby producing the database for training neural networks. In this work, we chose the multiple-input single-output (MISO) strategy due to its straightforward implementation and real-time neural network training capabilities. The input variables of the neural network model are the gas flow rates for three different wells ($Q_{g,1}, Q_{g,2}, Q_{g,3}$) and the output variables are the total liquid production ($w_{l,t}$) and the gas total injected ($w_{g,t}$).

The data set was divided into three subsets according to the following proportion: 75% for training data, 15% for validation data, and 15% for test data.

The next step consists of defining the hyperparameters of the neural network model. Traditional methods, such as random and grid search for hyperparameter tuning, are computationally expensive. An efficient alternative is HYPERBAND, which minimizes the evaluation of hyperparameter combinations through random sampling and early stopping (Li et al., 2018).

Before using HYPERBAND, it is necessary to define the hyperparameters and their search spaces, including param-

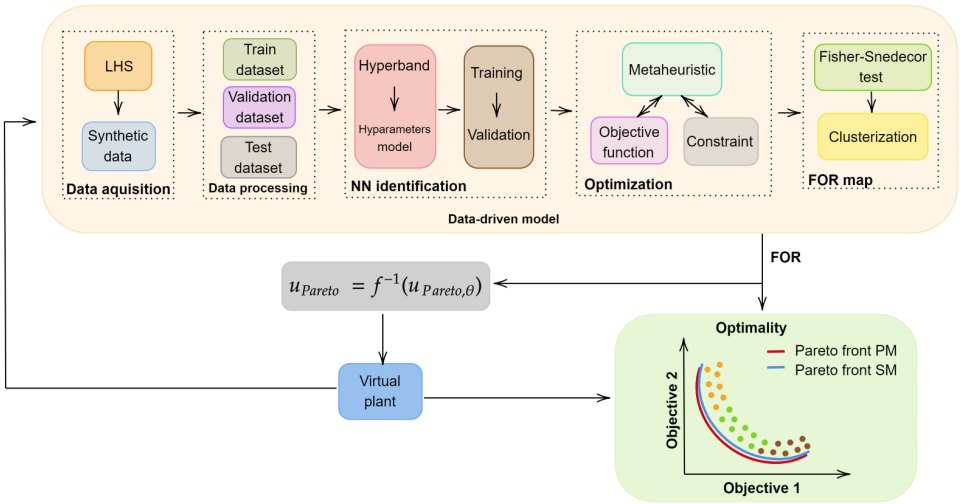


Fig. 2. Schematic diagram of the methodology proposed in this work.

eters such as learning rate and network architecture. This study selected hyperparameters such as initial learning rate, dense layer count, activation functions, and neuron count per layer. Once the most suitable architecture for training neural networks through hyperband was defined, these were trained and validated with a batch size of 100 and an epoch number of 200. The model's performance was evaluated using the test data. The identification of neural networks was carried out using the TensorFlow library, implemented in Python, and the high-level library Keras.

2.3 Optimization problem

In this study, a multi-objective optimization approach is employed to achieve a balance between the $w_{l,t}$ and the $w_{g,t}$ at steady state.

The decision variable vector ($\boldsymbol{u}$) comprises three key process variables, specifically, the gas flow rates for three distinct wells, $Q_{g,1}$, $Q_{g,2}$, $Q_{g,3}$. The search range for the decision variables is defined by the minimum limits $\boldsymbol{u}_{min} = [0, 0, 0]$ and maximum limits $\boldsymbol{u}_{max} = [10, 10, 10]$. To simulate various production conditions for each well, CV101, CV102, and CV103, the valve openings characterizing the reservoirs were defined as 80%, 60%, and 40%. This necessitates that the optimizer can adjust the gas flow rates to optimal for varying conditions in each well.

Mathematically, the multi-objective optimization problem is formulated as follows:

$$\begin{cases} min\ w_{g,t}, \\ max\ w_{l,t}. \end{cases} \tag{1}$$

Subject to

$$\boldsymbol{u} = [\boldsymbol{u}_{\min}, \boldsymbol{u}_{\max}] \tag{2}$$

Particle swarm optimization (PSO) addressed the multi-objective optimization challenge. PSO stands out for its ease of implementation and efficiency, presenting a linear correlation between population size and solution time, which makes it more computationally efficient than genetic algorithms (GA) in similar scenarios (Kachitvichyanukul, 2012; Nogueira et al., 2019). Its simplicity, combined with its effectiveness, makes it a suitable choice for optimization tasks.

2.4 Feasible operation map

The Fisher-Snedecor extended test was applied to solve the multiobjective optimization problem addressed in this work. This test evaluates a population generated through optimization, producing a feasible operational region map close to the optimal Pareto front. This approach allows for a more complete evaluation of the set of optimal solutions and a better understanding of the variability in near-optimal operating variables. The Equation 3 shows the Fisher-Snedecor test applied to this study, with 95% confidence level (α) (Rebello et al., 2022, 2021).

$$L_i(\boldsymbol{\theta}_k, \boldsymbol{\lambda_k}) \leq L_i(\boldsymbol{\theta}_j^*, \boldsymbol{\lambda_j^*}) + \frac{(n_k n_j)}{(n_k - n_{\boldsymbol{\theta}} + n_i)} F_\alpha(n_j, n_k - n_{\boldsymbol{\theta}} + 1) \tag{3}$$

where L denotes the Lagrangian, $\boldsymbol{\theta_k}$ represents the particle vector in the Fisher test, $\boldsymbol{\theta_j^*}$ signifies the optimal condition vector, $\boldsymbol{\lambda_j^*}$ indicates the Lagrange multipliers for optimal conditions, $\boldsymbol{\lambda_k}$ indicates the Lagrange multipliers for particles vetor, n_k is the particle count, $n_{\boldsymbol{\theta}}$ is the count of decision variables, n_i represents the number of objective functions, n_j is the quantity of points on the Pareto front, and F stands for the Fisher test.

This work used a clustering tool to create the operation map to identify and group groups with similar characteristics into categories or clusters. Thus, the clustering technique classified the points that passed Fisher's test into three sub-regions, and each sub-region prioritized different objectives in the multi-objective optimization problem. It is worth highlighting that the feasible operation map is a tool capable of assisting in decision-making and can play an important role in monitoring the process, providing relevant information about the behavior of the system's optimal conditions. Additionally, clusters of operating regions provide flexibility for operating processes.

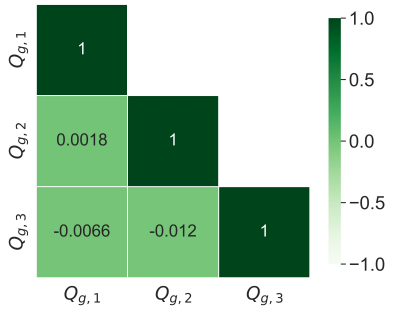


Fig. 3. Correlation heatmap for the input variables of the surrogate models.

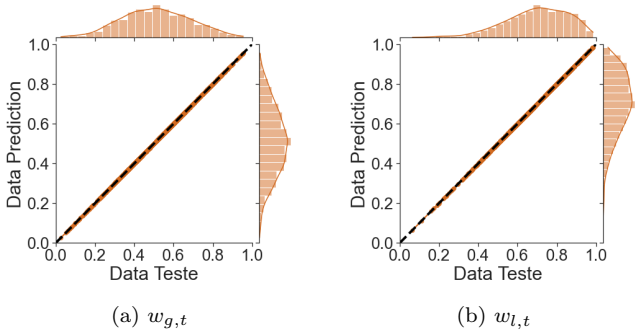
3. RESULTS

For data acquisition, a design of experiments (DoE) tool, latin hypercube sampling (LHS), was used to promote representative data sampling in a multidimensional space. LHS generated a matrix of inputs to be used as experimental inputs (given to a virtual plant in the present work) and generated output synthetic data in the steady state. The correlation between the input variables that compose the final dataset is then evaluated. Figure 3 displays a heatmap depicting the low correlations, approaching zero. This figure indicates a well-designed input space by LHS, preventing data bias and undesirable training biases. Subsequently, these perturbations were injected into the phenomenological model, creating datasets constituting the model's output matrix.

After the DoE step, we proceeded to a data preparation and segmentation phase into training, validation, and test sets. This phase plays an essential role in ensuring the effectiveness and reliability of neural network training. The training data served as the basis for the neural networks to learn the complex relationships between system variables. Meanwhile, the validation set was crucial in selecting and adjusting neural network architectures. HYPERBAND was employed to find the optimal architecture for the neural networks. Table 1 presents the search space this optimizer uses to define the network structure. In the same Table 1, it is possible to find the best configuration identified for the surrogate model, representing the most effective and suitable result for the system's needs.

After defining the optimal neural network architecture, the trained and validated. Figure 4 visually presents the model predictions against the actual test data, showing its accuracy in capturing the system. The randomly distributed points along the diagonal line indicate the model's precision. The graph reveals close alignment between model predictions and test data, with random residuals, validating the model's ability to predict system behavior accurately.

Continuing the discussion regarding the predictive capacity of the surrogate models, Table 2 displays performance metrics, including mean absolute error (MAE) and mean squared error (MSE), indicating low error values for test data. These results confirm successful model identification and accurate predictions, showcasing model reliability in tracking data.



(a) $w_{g,t}$ (b) $w_{l,t}$

Fig. 4. Parity plots for the test dataset by training the neural networks.

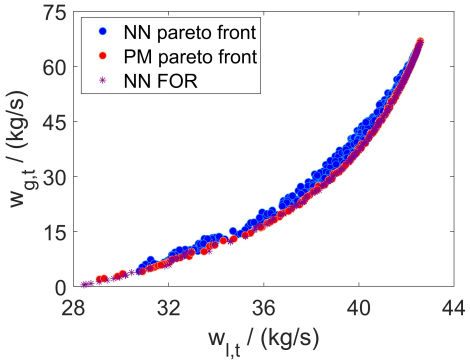


Fig. 5. Overlap of the Pareto front and region of feasible operation for surrogate models and phenomenological models (PM).

The solution for this optimization problem is represented by a set of solutions known as the Pareto optimal front. This Pareto front describes the best solutions simultaneously satisfying multiple objectives, highlighting the relationship and possible compromises between these objectives. Using the methodology proposed in this work, Figure 5, it was possible to obtain a Pareto front composed of 214 points, covering the liquid-produced flow and injected gas, using 50 particles and 50 iterations defined using the PSO algorithm. Furthermore, with the results obtained in the optimization, we used the respective optimal operating conditions at each point on the Pareto front to consult the phenomenological model and validate the consistency of optimization, as described in Figure 2. The overlap of the optimization strategy solutions and the query of optimal points from the phenomenological model is illustrated in Figure 5. The overlapping demonstrates the consistency of the surrogate-based solutions. Hence assuring the reliability of the results obtained.

With the reliability ensured, we could proceed with the next step, to establish a feasible region near the Pareto optimal front. The main idea was to create a more flexible region to assess potential fluctuations the process could undergo without deviating from the desired operating conditions. This FOR was constructed based on the Fisher-Snedecor test. Figure 5 illustrates the feasible operating region for surrogate model optimization. This approach enables a comprehensive understanding of optimal operating conditions and the available margins of flexibility for the process.

Table 1. Search hyperspace configurations and best parameters.

Variable	Hyperparameters	Search space	Best hyperparameters
$w_{g,t}$	Initial learning rate	0.0001, 0.001, 0.01	0.001
	Number of dense layers	1 - 4	3
	Layer activation function	relu, tanh, linear	tanh, tanh, linear
	Number of neurons per layer	20 to 200 with 10 step	170, 190, 1
	Number of parameters per layer	-	680,32490,191
$w_{l,t}$	Initial learning rate	0.0001, 0.001, 0.01	0.001
	Number of dense layers	1 - 4	3
	Layer activation function	relu, tanh, linear	tanh, tanh, linear
	Number of neurons per layer	20 to 200 with 10 step	170, 190, 1
	Number of parameters per layer	-	680,32490,191

Table 2. Final MAE and MSE to test dataset.

Variable	MSE	MAE
$w_{g,t}$	$4.26 \cdot 10^{-7}$	$5.83 \cdot 10^{-4}$
$w_{l,t}$	$5.99 \cdot 10^{-7}$	$5.88 \cdot 10^{-4}$

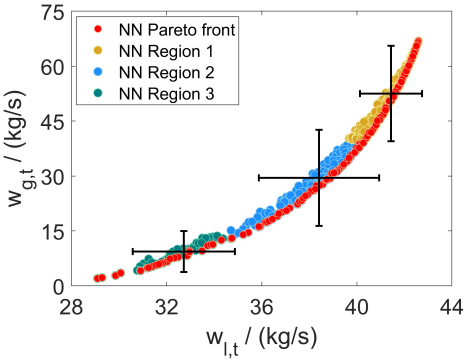


Fig. 6. Operation map obtained by clustering and Pareto optimal front.

Given that transitioning between different points on the Pareto front and the feasible operating region entails making different trade-offs with the objectives addressed in the optimization problem, we adopted clustering techniques to subdivide the feasible region into three subregions. The region 3 prioritizes the commitment to the injected gas flow rate, thereby minimizing resource consumption. Meanwhile, region 1 focuses more on production, and region 2 represents a balance between the two more extreme regions. Figure 6 displays the Pareto optimal front and the operating subregions derived from surrogate models. These subregions were identified using clustering techniques, where the Euclidean distance between the centroids of each group was employed as a metric to assign points to their respective subregions. The error bars calculated according to the average of each sub-region and 95% confidence level show that the clusters are independent, without overlapping.

It became possible to conduct the same analysis from the perspective of operating variables by clustering to define subgroups with different trade-offs concerning optimization objectives. In Figure 7, the location of the optimal points comprising the Pareto front is visible, allowing for the simultaneous identification, for each well, of the operating conditions that would result in higher oil production or lower resource consumption. Note that clusters related to higher oil flow rates will require larger quantities of injected gas, which, in turn, entail associated energy

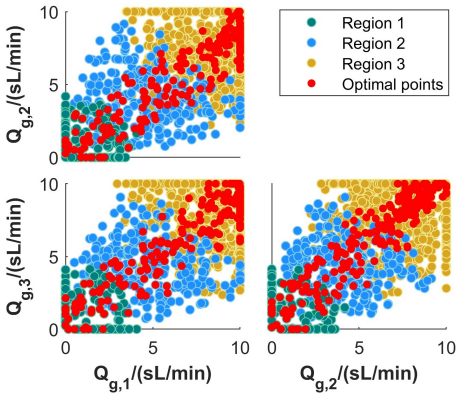


Fig. 7. Operation map obtained by clustering and Pareto optimal front for decision variables.

consumption. Conversely, lower resource usage, associated with reduced gas injections, may result in lower liquid production. This operating map constitutes a valuable tool for processes experiencing continuous fluctuations in operating conditions, demanding real-time decision-making..

By mapping a process's feasible operating region, it becomes possible to understand the interactions among the involved variables, which play an essential role in guiding operational and monitoring decisions. The FOR assessment and the corresponding operational variables represent a valuable tool for exploring the process's operational behavior, especially near the desired optimal conditions. Through this analysis, we can identify which operational variables significantly impact established objectives and use them efficiently during execution. In particular, within the context of this work, an increase in gas injection rates is directly linked to higher well production, considering the close correlation between well flow rates. Therefore, operations focusing on Region 1 may be particularly interesting for process optimization. Applying the proposed methodology has securely identified the operational variables that genuinely influence the optimal operational conditions of the studied process.

Based on the FOR, the operator can select a point within a given region, for example, operation region 1, and then operate the process in the corresponding operating conditions. As a feasible region exists, the process monitoring can tolerate any process fluctuations within that region. Hence, it is an alternative to the traditional optimization strategies that consider a single point, and any deviations from it are treated as a disturbance that should be rejected.

4. CONCLUSION

In this work, we contribute to the field of multi-objective optimization in gas lift systems, a critical component in oil and gas production, by addressing the inherent trade-offs between maximizing oil output and minimizing gas usage. We reduced the computational burden commonly associated with first-principle models in optimization tasks by utilizing deep artificial neural networks as surrogate models. The particle swarm optimization technique was employed for its effectiveness in population-based optimization scenarios.

To validate the optimization's reliability, we engaged in a post-optimization evaluation using a phenomenological model, thereby confirming the consistency of the generated solutions. The extensive data points obtained through PSO enabled the construction of feasible operating regions, determined through a likelihood test using the Fisher-Snedecor criterion. These FORs were further dissected into sub-regions via clustering techniques, each focusing on different trade-offs between oil production and gas consumption.

The proposed framework excels in mapping a nuanced pareto front segmented into these operational sub-regions, thus offering an advanced tool for real-time decision-making in process monitoring. This multi-layered approach provides operators with a rich operating map that delineates not just optimal conditions but also acceptable fluctuations within those conditions. This transforms the Pareto front into more adaptable, robust decision-making tools vital for navigating the complexities and variabilities inherent in oil and gas production.

ACKNOWLEDGEMENTS

The present work contributes to completion of a subproject at SUBPRO, a research-based innovation center within Subsea Production and Processing at the Norwegian University of Science and Technology. The authors would like to express their gratitude for the financial support received from SUBPRO, funded by the Research Council of Norway through grant number 237893, major industry partners, and NTNU.

REFERENCES

Azlan Hussain, M. (1999). Review of the applications of neural networks in chemical process control simulation and online implementation. *Artificial Intelligence in Engineering*, 13(1), 55–68. doi: https://doi.org/10.1016/S0954-1810(98)00011-9.

Ghiassi, M. and Nangoy, S. (2009). A dynamic artificial neural network model for forecasting nonlinear processes. *Computers & Industrial Engineering*, 57(1), 287–297. doi:https://doi.org/10.1016/j.cie.2008.11.027. Collaborative e-Work Networks in Industrial Engineering.

Guerra, L., Temer, B., Loureiro, J., and Silva Freire, A. (2022). Experimental study of gas-lift systems with inclined gas jets. *Journal of Petroleum Science and Engineering*, 216, 110749. doi:https://doi.org/10.1016/j.petrol.2022.110749.

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., and Talwalkar, A. (2018). Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185), 1–52. URL `http://jmlr.org/papers/v18/16-558.html`.

Lin, J.S., Jang, S.S., Shieh, S.S., and Subramaniam, M. (1999). Generalized multivariable dynamic artificial neural network modeling for chemical processes. *Industrial & Engineering Chemistry Research*, 38(12), 4700–4711. doi:10.1021/ie990312e.

Matias, J., Oliveira, J.P., Le Roux, G.A., and Jäschke, J. (2022). Steady-state real-time optimization using transient measurements on an experimental rig. *Journal of Process Control*, 115, 181–196. doi: 10.1016/j.jprocont.2022.04.015.

Nogueira, I.B., Martins, M.A., Requio, R., Oliveira, A.R., Viena, V., Koivisto, H., Rodrigues, A.E., Loureiro, J.M., and Ribeiro, A.M. (2019). Optimization of a true moving bed unit and determination of its feasible operating region using a novel sliding particle swarm optimization. *Computers & Industrial Engineering*, 135, 368–381. doi: https://doi.org/10.1016/j.cie.2019.06.020.

Ranjan, A., Verma, S., and Singh, Y. (2015). Gas lift optimization using artificial neural network. In *SPE Middle East Oil & Gas Show and Conference*. OnePetro.

Rashid, K., Bailey, W., and Couët, B. (2012). A survey of methods for gas-lift optimization. *Modelling and Simulation in Engineering*, 2012, 24–24.

Ray, T. and Sarker, R. (2006). Multiobjective evolutionary approach to the solution of gas lift optimization problems. In *2006 IEEE International Conference on Evolutionary Computation*, 3182–3188. doi: 10.1109/CEC.2006.1688712.

Ray, T. and Sarker, R. (2007). Genetic algorithm for solving a gas lift optimization problem. *Journal of Petroleum Science and Engineering*, 59(1), 84–96. doi: https://doi.org/10.1016/j.petrol.2007.03.004.

Rebello, C.M., Martins, M.A.F., Santana, D.D., Rodrigues, A.E., Loureiro, J.M., Ribeiro, A.M., and Nogueira, I.B.R. (2021). From a pareto front to pareto regions: A novel standpoint for multiobjective optimization. *Mathematics*, 9(24). doi:10.3390/math9243152.

Rebello, C.M., Martins, M.A., Rodrigues, A.E., Loureiro, J.M., Ribeiro, A.M., and Nogueira, I.B. (2022). A novel standpoint of pressure swing adsorption processes multi-objective optimization: An approach based on feasible operation region mapping. *Chemical Engineering Research and Design*, 178, 590–601. doi: https://doi.org/10.1016/j.cherd.2021.12.047.

Santana, V.V., Martins, M.A.F., Loureiro, J.M., Ribeiro, A.M., Queiroz, L.P., Rebello, C.M., Rodrigues, A.E., and Nogueira, I.B.R. (2023). Novel framework for simulated moving bed reactor optimization based on deep neural network models and metaheuristic optimizers: An approach with optimality guarantee. *ACS Omega*, 8(7), 6463–6475. doi:10.1021/acsomega.2c06737. URL `https://doi.org/10.1021/acsomega.2c06737`.

Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2), 143–151. doi:10.1080/00401706.1987.10488205.