Vol. 4(1), pp. 1-11, 2025
eISSN: 2976-4009

ASEAN Journal of Process Control

Research Article

Process Simulation and Optimization of a Crude Distillation Unit Using Aspen Plus and MATLAB

Essi Mitcel Makele[1], Mohd Hardyianto Vai Bahrun[1], Zainal Ahmad[1,*]

[1]School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Seri Ampangan, Nibong Tebal, Pulau Pinang, MALAYSIA
*Corresponding Author: chzahmad@usm.my

Abstract: The crude distillation unit (CDU) is a primary refinery process where operating decisions impact both yield and energy requirements. This work presents an integrated Aspen Plus-MATLAB framework for CDU simulation and optimization. An Aspen Plus model was first established to simulate the base case, focusing on kerosene yield and energy consumption. Two surrogate modelling approaches, namely an autoencoder (AENN) and feed-forward (FFNN) neural networks were trained to predict CDU performance. FFNN outperformed AENN in predicting kerosene yield (R_{FFNN}^2 = 0.8467 *vs.* R_{AENN}^2 = 0.7834) and energy consumption (R_{FFNN}^2 = 0.9132 *vs.* R_{AENN}^2 = 0.8639), confirming its suitability as a reliable surrogate. The FFNN model was subsequently coupled with a genetic algorithm (GA) to identify optimal operating parameters for minimizing energy use. The FFNN-GA converged smoothly, recommending a CDU pressure of 2.58 bar and feed rate of 400 ton/h. Despite acceptable prediction error between FFNN/GA and Aspen Plus simulation below 20%, the optimized case yielded no net energy savings due to yield-energy trade-offs. The results demonstrate the viability of surrogate-based CDU optimization while emphasizing the need for multi-objective formulations and model refinement.

Keywords: Crude distillation unit; Process simulation; Neural networks; Genetic algorithm; Energy efficiency

1. Introduction

The crude distillation unit (CDU) is the primary separation process in petroleum refineries, fractionating crude oil into valuable products such as naphtha, kerosene, diesel, and residual fuels based on boiling point differences. Its efficient operation directly affects product yields, energy consumption, and both economic and environmental performance [1]. CDU energy usage typically represents 1-2% of the crude oil processed [2], contributing substantially to refinery CO_2 emissions. CDUs process highly complex feedstocks and are significantly more intricate than typical distillation columns, featuring multiple side draws, strippers, pump-arounds, and live steam injection. The design and operation of these systems must account for the interdependencies between numerous variables and constraints. Traditionally, CDU design and control relied on empirical correlations and heuristics derived from operational experience, which struggle with model non-linear coupling and variable feed streams [3].

With the increasing complexity of refinery operations and the tightening of energy and environmental regulations, process modelling has become essential for understanding and optimizing CDUs. Process modelling involves developing mathematical or data-driven representations of the

underlying physical and chemical processes. These models allow engineers to simulate operating conditions, test control strategies, and evaluate performance under different scenarios. The approach to modelling involves carefully selecting which aspects of the process to include or exclude, as well as determining how to represent these elements [4]. The optimization of the CDU process becomes increasingly important because of the high energy costs and ecological requirements for quality oil products. Given the energy-intensive nature of CDUs and the global push toward sustainable operations, optimizing their performance has become a strategic priority. High energy costs and increasingly stringent ecological requirements, such as reduced CO_2 emissions and stricter fuel specifications, have driven the need for more advanced optimization tools [5]. Conventionally, CDU optimizations were made using high-fidelity simulations, though they provide higher accuracy, but at a great computational expense, slow convergence in large-scale non-linear settings, and not all combinatorial nature of the optimization iteration within the search space are likely to lead to a converged simulation (a feasible design) [3, 6].

In recent years, data-driven surrogate modelling has gained prominence for its ability to emulate complex processes with drastically reduced computational cost. Regression-based surrogates, such as polynomial fitting, artificial neural networks (ANN), and support vector regression (SVR) have been applied to model CDUs based on plant data or simulations. For instance, Liau *et al.* [7] used an ANN-based surrogate model to optimize the yields of kerosene, diesel, and atmospheric gas oil, while Motlaghi *et al.* [8] developed a data-driven model to maximize product revenue based on market conditions. These surrogate models provide a flexible alternative to first-principles models, especially when integrated into optimization frameworks. Recent work also applies ANN for real-time monitoring and efficiency prediction in CDU operations [9].

When surrogate models are integrated into surrogate-assisted optimization frameworks, the number of expensive simulations needed is greatly reduced. The surrogate-assisted optimization frameworks have significantly reduced optimization time while achieving near-optimal process performance. Surrogate-assisted optimization frameworks aim to simplify originally expensive high-fidelity model and minimizing optimization expenses [10], as well as achieving design solution that is within a feasible region only [6]. Ye *et al.* [3] studied an intelligent optimization design of distillation columns using surrogate models, achieving an improved design solution with reduced cost and environmental impact with an advantages in convergence speed without sacrificing accuracy. Similarly, Xiong *et al.* [11] investigated an optimization design strategy for CDUs by proposing a novel bi-level surrogate column model aided constrained optimization design approach. The study integrates surrogate modelling, rigorous CDU simulations, support vector machine feasibility classification, and particle swarm optimization techniques to enhance the efficiency of solving optimization problems in CDU systems. These studies demonstrate that surrogate-assisted frameworks enable broader exploration of operating and design parameters, leading to more efficient and cost-effective CDU configurations.

Building on these developments, the present study aims to simulate and optimize a CDU using an integrated Aspen Plus-MATLAB approach. First, a representative CDU flowsheet is constructed and simulated in Aspen Plus. Two neural network models are built to represent the CDU to facilitate operational optimization of the unit. Aspen Plus simulations are performed to build the dataset used to train the neural network models, so that accuracy is not compromised. Then, the best-perform neural network model is used as surrogate model for optimization using GA to explore and optimize key operating conditions with the goal of minimizing energy consumption.

2. Materials and Method

2.1. Process Description of CDU

The CDU simulation flowsheet in Aspen Plus is illustrated in Figure 1 [12]. The crude feed enters a preheater and furnace, and then to the main distillation column. The column has side draws to collect kerosene and other light cuts, pump-around loops for heat recovery, and a reboiler at the bottom. The condenser on top refluxes liquid back to the column. The pump-around circuits recycle heat from upper to lower stages. In this flowsheet, water and hydrogen sulfide are also included in the feed, as they often accompany crude. Key outputs of interest are the kerosene flowrate (as a measure of product yield) and the furnace duty, which indicates energy consumption.

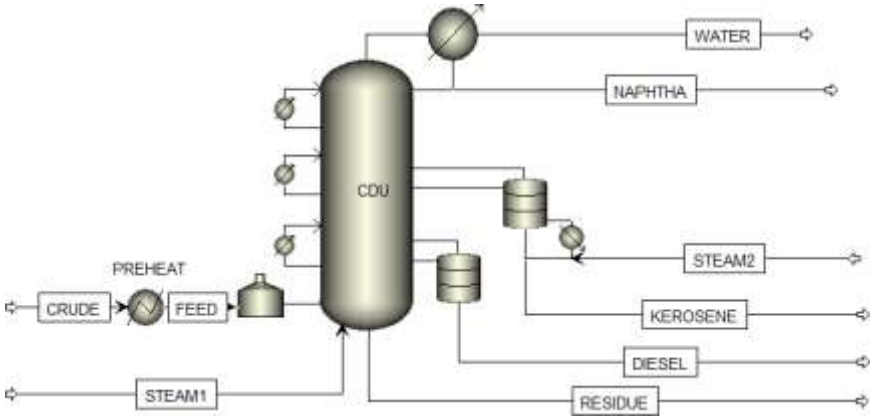

Figure 1. The simulation flowsheet of a CDU in Aspen Plus, showing a distillation column, preheater, furnace, pump-around loops, reboiler, and side draws for product fractions.

2.2.1. Components and thermodynamic properties

A set of ten components were defined to represent the crude and associated streams, including water, C1 – C5 alkanes, hydrogen sulfide, and the crude assay, as listed in Table 1. The Braun-K10 (BK10) thermodynamic method was used, as it is widely recognized to be suitable for petroleum applications al low pressures [12]. Figure 2 depicts the true boiling point (TBP) curve of the crude feed used to define its assay in the simulation.

Table 1. The list of components added from Aspen Properties databanks for CDU simulation.

Component	Chemical formula	Molecular weight (g/mol)
Water	H_2O	18.01
Methane	CH_4	16.04
Ethane	C_2H_6	30.06
Propane	C_3H_8	44.09
Isobutane	C_4H_{10}	58.12
n-butane	C_4H_{10}	58.12
2-methylbutane	C_5H_{12}	72.14
n-pentane	C_5H_{12}	72.14
Hydrogen sulfide	H_2S	34.08
Crude	Assay	N/A

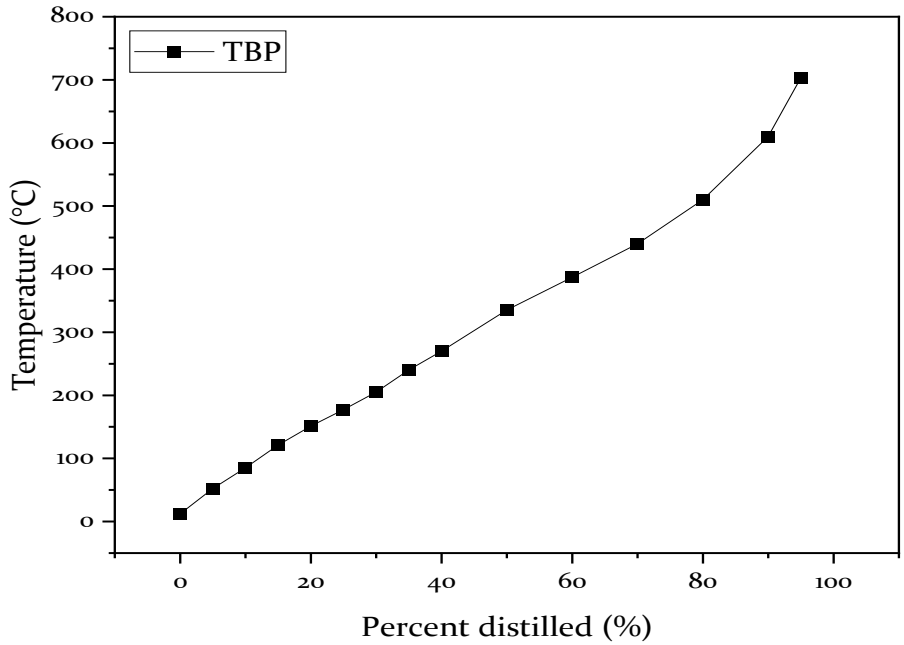

Figure 2. The crude assay feed true boiling point (TBP) curve.

2.2. Modelling of the CDU using

2.2.1. Data collection and preprocessing

The accuracy and reliability of neural network models are intrinsically linked to the datasets used for model training and testing [13]. In this study, a dataset was obtained from a validated simulation model and was generated by varying the CDU operating variables around the base case conditions. Five input features, identified as having a significant influence on the CDU were selected as inputs during neural network models training. Then, the corresponding output vectors were determined from the validated simulation model. Here, two performance indicators of the CDU were selected as the output variables, including kerosene yield and energy consumption. Table 2 lists the input and output variables used for modelling of the CDU, along with their corresponding lower and upper bounds of the input variables dataset.

A single-output strategy was implemented in this study, whereby one neural network model predicts one output variable. A data normalization based on the min-max normalization, scaling datasets to a range of [0, 1] was performed during pre-processing to address the large variations in variables magnitudes and their differing units in the datasets. This normalization techniques accelerates the training process and helps to eliminate bias in the data [14]. Then, a systematic training-validation-testing splitting strategy was followed, using a default ratio in MATLAB environment.

Table 2. The input and output variables used to develop neural network models for CDU.

Parameter	Boundary value
Input variables	
Feed flowrate (ton/h)	[400, 420]
Pump-around flowrate (ton/h)	[9, 14]
CDU pressure (bar)	[1, 5]
Reflux ratio	[2.0, 2.5]
Kerosene flowrate (ton/h)	[200, 230]
Output variables	
Kerosene yield (%)	-
Energy consumption (GJ/h)	-

2.2.2. Autoencoder neural network model

An autoencoder network (AENN) was developed in MATLAB using the deep learning toolbox to model the CDU. The AENN is an unsupervised learning architecture designed to compress the input vector into a lower-dimensional latent space, and subsequently reconstruct an output vector that closely approximates the original input [15]. This encoding-decoding process forces the network to learn and retain the most relevant non-linear features of the data while discarding redundant information. In this study, the AENN architecture comprised an encoder with 20 neurons and a matching decoder structure. The encoder employed a saturating linear activation function, selected for its ability to approximate identity mapping while still enabling effective data compression. The network was trained over 200 epochs, with the objective of minimizing the reconstruction error between the input and the output vectors. The final tuned parameters and architectural settings for the AENN are summarized in Table 3. Because the AENN operates in an unsupervised mode, no explicit target features are required. Instead, the same input dataset is passed through both the encoder and decoder, enabling the network to learn underlying patterns directly through self-reconstruction.

Table 3. Configuration and tuning parameters of the AENN used for CDU surrogate modelling.

Parameter	Value	Description
hiddenLayerSize	20	Number of neurons in the encoder
EncoderTransferFunction	“satlin”	Activation function used in the encoder for saturating linear
MaxEpochs	200	Maximum number of training epochs for the autoencoder

The rationale for adopting the AENN in this study was to obtain a compressed yet information-rich representation of CDU operation, which could then be used in surrogate modelling and optimization. By constraining the network to encode and decode the same input features, the model inherently identifies and retains the essential structure of the operating data.

2.2.3. Feed-forward neural network model

A conventional feed-forward neural network (FFNN) was also developed in MATLAB for performance comparison with the AENN model. Unlike the unsupervised AENN, the FFNN employs supervised learning, which requires paired input-output training data to establish the mapping between operating conditions and process responses. This direct mapping directly makes FFNNs particularly suitable for predictive modelling tasks where accurate output estimation is crucial.

The FFNN architecture consisted of two hidden layers containing 10 and 5 neurons, respectively. The Levenberg-Marquardt algorithm was employed as the training method to optimize the network's weights and biases, aiming to minimize the error between predicted outputs and corresponding target values. This algorithm was selected for its robustness and fast convergence when applied to medium-sized networks, making it well suited for process modelling applications. In MATLAB, the FFNN model was configured using the built-in "feedforwardnet" function. During model configuration, the input features were explicitly defined along with their corresponding target outputs to ensure proper supervised training. The complete set of architectural and tuning parameters used for the FFNN is presented in Table 4.

Table 4. Configuration and tuning parameters of the FFNN used for CDU surrogate modelling.

Parameter	Value	Description
hiddenLayerSize	[10 5]	Two hidden layers with 10 and 5 neurons respectively. They control the model's capacity to learn complex non-linear patterns.
Normalization	"range"	The normalization of inputs and outputs range to ensure consistent scale for better training.
trainFcn	"trainlm"	The training algorithm for a fast, second order optimization method.

2.1. Optimization of the CDU Performance using Genetic Algorithm

Genetic algorithms (GAs) have been widely used in the literature to optimize chemical processes. These algorithms work by recombining and reproducing members of the initial population [13], and therefore works well in solving complex optimization problems as its algorithm repeatedly modifies a population of individual solutions [16]. The GA was used in this study to find the optimal CDU performance metric, due to its suitability to handle problems that are discontinuous, non-differentiable, stochastic, and highly non-linear [17]. In this study, the GA was integrated with the best-performing neural network model, which served as a surrogate model to replace computationally expensive high-fidelity simulations.

The GA interacts with the neural network in an iterative loop; it generates a new set of independent variables (operating parameters), passes them to the neural network, and receives the predicted output, in this case is the CDU's energy consumption. This predicted value is then evaluated using a fitness function that quantifies how close the solution is to the optimum. The objective function for the optimization, defined in Eq. 1, is to minimize CDU energy consumption as represented by furnace duty. Furnace duty was selected because it accounts for the largest portion of energy usage in CDU that accounts to almost 70% and, thus, has a significant impact on fuel cost and CO_2 emissions. By focusing on furnace duty minimization, the optimization directly addresses a key operational and environmental performance driver in CDU operations.

$$\begin{cases} \text{Obj: min (Energy consumption)} \\ \text{Subject to: } Y_i = f_{NN_i}(X_i) \\ \\ \text{Boundary: } X_i \in [\text{LB}, \text{UB}] \end{cases} \quad \text{Eq. 1}$$

where X_i represents the decision variables, Y_i is the objective function, f_{NNi} is the best-perform neural network model and LB and UB stand for lower and upper bounds of the input (decision) variables, respectively.

3. Results and Discussion

3.1. The base case CDU simulation

The CDU simulation was successfully developed in Aspen Plus using process data and specifications reported by Gary *et al.* [18]. Among the various simulated output variables, particular emphasis was placed on kerosene flowrate and furnace duty, as these represent critical performance indicators for CDU operation. The kerosene flowrate serves as a direct measure of product yield, reflecting the unit's separation efficiency and production capability. In contrast, the furnace duty quantifies the CDU's energy demand, accounting for the largest portion of its total energy consumption. These two parameters were selected as the primary evaluation metrics for both benchmarking the base-case simulation and comparing subsequent optimization results. By jointly considering kerosene yield and furnace duty, it is possible to assess the trade-offs between production performance and energy efficiency. The detailed performance results for the base-case CDU configuration are summarized in Table 5, which serves as the reference point for analyzing the effects of the proposed optimization strategies.

Table 5. Base case operating parameters and CDU performance from Aspen Plus simulation.

Parameter	Value
Feed flowrate (ton/h)	390.827
Pump-around flowrate (ton/h)	190.000
CDU pressure (bar)	3.24
Reflux ratio	2.136
Kerosene flowrate (ton/h)	38.144
Furnace duty (MW)	290.061
Energy consumption (GJ/h)	0.737

The base-case CDU operates at a feed flowrate of 390.827 ton/h under a column pressure of 3.24 bar with a reflux ratio of 2.136. The pump-around flowrate is set to 190 ton/h, which contributes to internal heat recovery and influences both product yield and energy demand. From a product perspective, the kerosene flowrate of 38.144 ton/h represents approximately 9.8% of the total feed. In terms of energy usage, the furnace duty is 290.061 MW, corresponding to an overall energy consumption of 0.737 GJ/h.

3.1. Neural network-based models' performance for CDU prediction

The predictive performance of both the AENN and the FFNN was evaluated for kerosene yield and energy consumption in the CDU process. Figures 3 and 4 compare the actual values from Aspen Plus simulations and predicted outputs for both kerosene yield and energy consumption across all sample datasets using the AENN and FFNN, respectively.

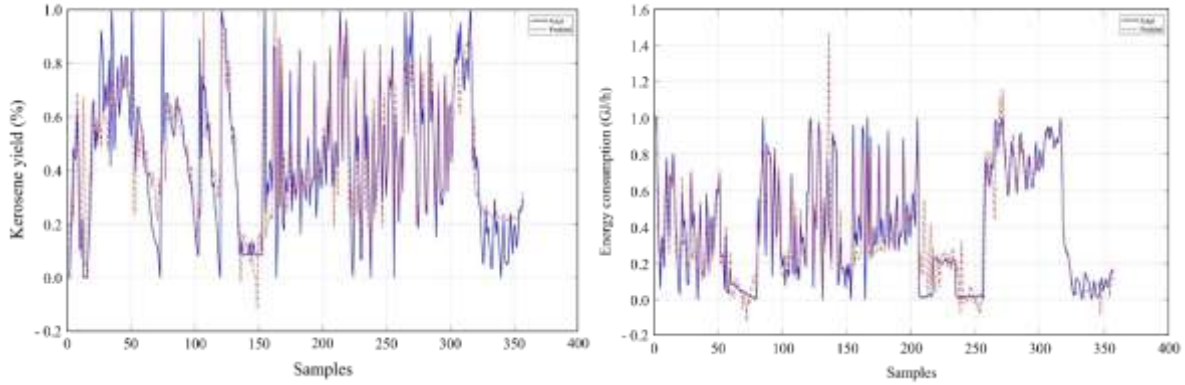

Figure 3. Kerosene yield and energy consumption (AENN model).

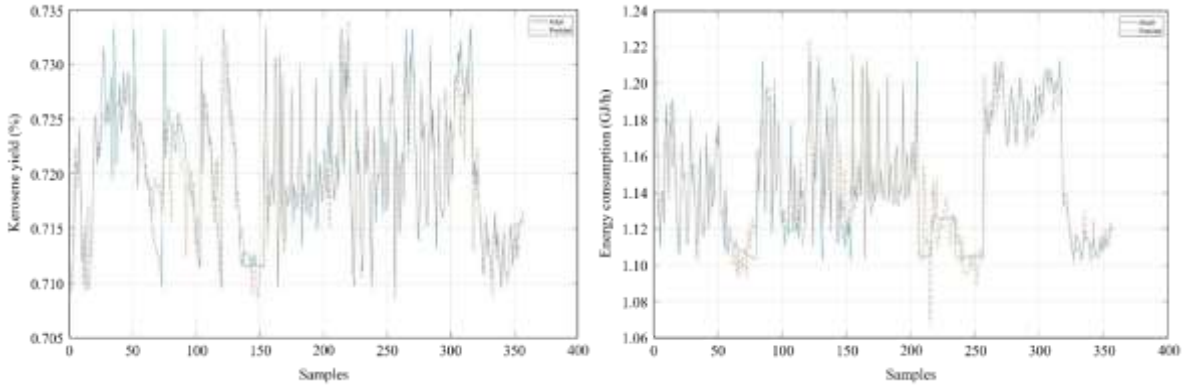

Figure 4. Kerosene yield and energy consumption (FFNN model).

Both models successfully capture the key non-linear dynamics of the CDU process, however, their predictive accuracy and error characteristics differ, as summarized in Table 6. For kerosene yield, the FFNN achieved an R^2 of 0.8467 and an MSE of 1.70×10^{-5}, substantially outperforming the AENN (R^2 = 0.7834, MSE = 0.0155). This indicates that the FFNN's direct mapping from raw input features to outputs is more effective than the AENN's latent-space reconstruction in predicting yield. For energy consumption, both models achieved higher accuracy compared with yield prediction. The FFNN again led with an R^2 of 0.9132 and MSE of 2.98×10^{-4}, outperforming the AENN (R^2 = 0.8639, MSE = 0.0132). The AENN's improved performance in energy prediction compared with yield prediction suggest that energy-related features (e.g., reflux ratio, pump-around flowrates) are more linearly separable in the latent space it generates. Nonetheless, the FFNN's simple architecture yielded faster convergence, lower error, and greater predictive stability across both outputs.

Table 6. Comparison of AENN and FFNN performance on kerosene yield and energy consumption prediction in CDU.

Model	Output	R^2	MSE
AENN	Kerosene yield (%)	0.7834	0.0155
AENN	Energy consumption (GJ/h)	0.8639	0.0132
FFNN	Kerosene yield (%)	0.8467	1.70×10^{-5}
FFNN	Energy consumption (GJ/h)	0.9132	2.98×10^{-4}

These observations are reinforced by the regression plots in Figure 5, where the R values for the FFNN model during training, validation, testing, and overall evaluation were 0.9724, 0.8917, 0.8611, and 0.9416, respectively. The scatter points cluster tightly along the 45° ideal line for both kerosene yield and energy consumption, indicating minimal systematic bias and strong generalization capability.

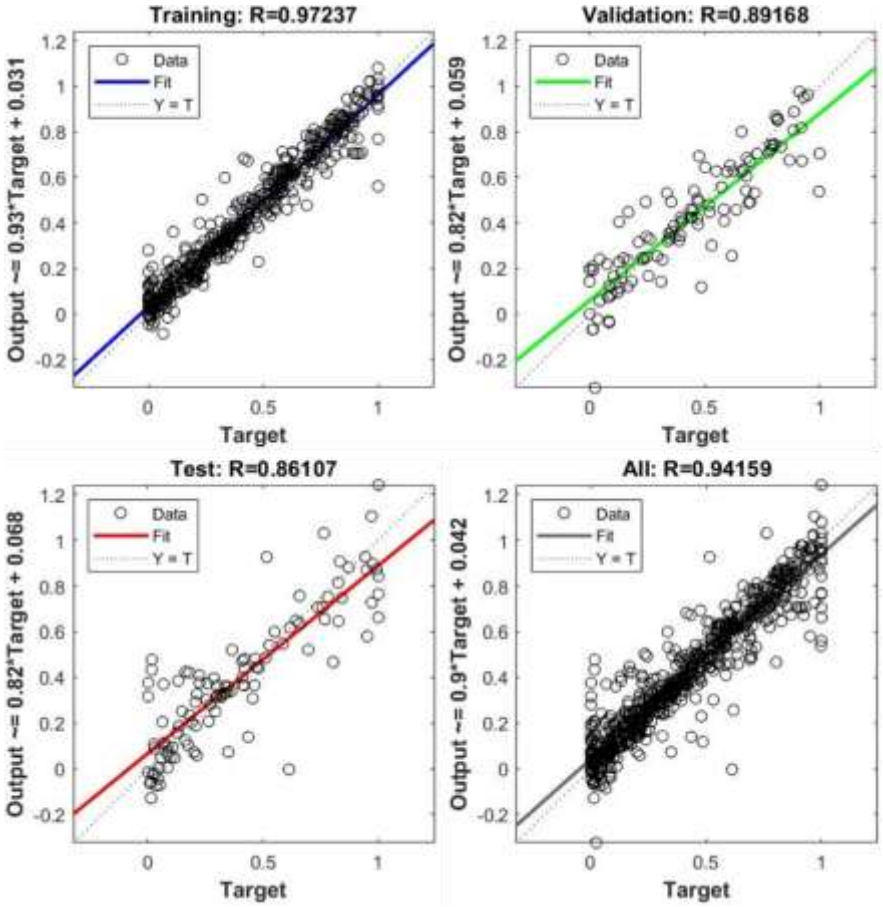

Figure 5. Parity plot showing predicted versus actual values for the FFNN model.

From an architectural perspective, the AENN's encoder-decoder design effectively reduces data dimensionality and can help mitigate overfitting in high-dimensional process datasets. However, the reconstruction phase introduces additional sources of bias and requires careful balancing between compression and predictive accuracy. Improper tuning during compression can lead to the loss of critical process features, contributing to the deviations observed in both kerosene yield and energy consumption predictions. In contrast, the FFNN propagates data directly through its hidden layers without a reconstruction step, focusing solely on the mapping between inputs and outputs. This enables the model to capture complex non-linear relationships more robustly, with lower computational overhead and reduced susceptibility to error propagation. Training the FFNN is generally less computationally demanding compared with the AENN, as it does not need to optimized for dual objectives (reconstruction accuracy and output prediction) [19]. The superior performance suggests that raw feature inputs provide more direct predictive criterial for both yield and energy compared to the AENN's latent space representation.

3.1. Genetic algorithm performance in optimizing CDU energy consumption

Since the FFNN surrogate demonstrated superior predictive accuracy, it was selected for GA-based optimization. The optimization objective was to minimize CDU energy consumption by adjusting key operating parameters within practical bounds. For this single-objective task, the GA used the FFNN model to estimate energy consumption for each candidate solution. Figure 6 illustrates the optimization progress. The FFNN/GA converged smoothly, reaching a minimum fitness value of 1.0356 GJ/h by the 65th generation, with the average fitness value close at 1.0357 GJ/h. Beyond this point, fluctuations were minimal, indicating convergence to a stable solution. The relatively steady convergence pattern suggests that the FFNN provided a smooth and consistent fitness landscape for the GA search process.

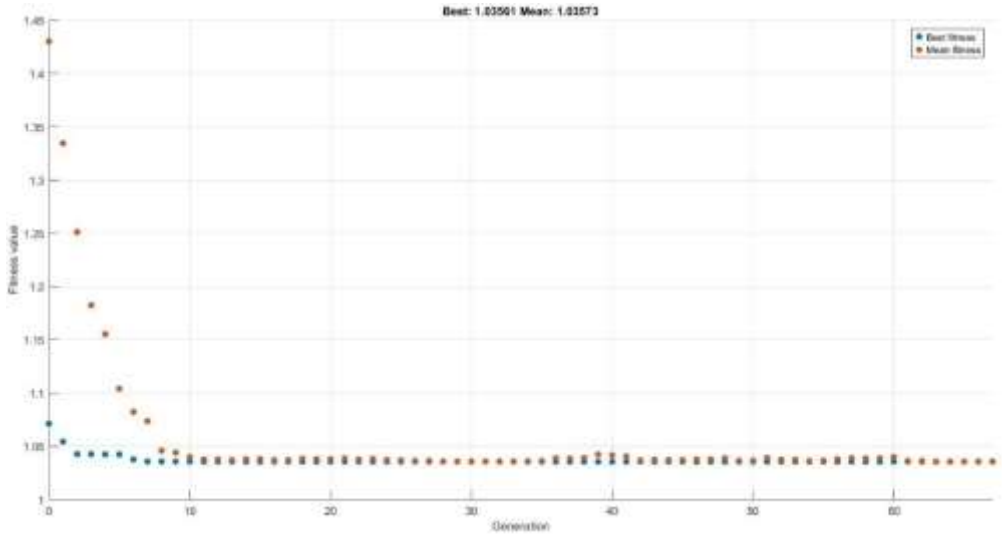

Figure 6. Convergence evolution of the GA optimization using FFNN as the surrogate model.

The optimal operating parameters determined by the FFNN/GA are summarized in Table 7, alongside the base-case values. The optimized solution lowered the column pressure from 3.24 bar to 2.58 bar and increased the feed flowrate from 390.827 ton/h to 400 ton/h. These adjustments suggest that the GA favoured higher throughput and reduced column pressure, which could lead to milder separations and lower specific energy consumption per unit feed. The reflux ration was slightly increased, while the pump-around flowrate decreased substantially, indicating reduced external liquid circulation. Despite these adjustments, the FFNN-predicted optimal configuration resulted in an energy consumption of 1.036 GJ/h, which is higher than the base-case Aspen Plus value of 0.737 GJ/h. This indicates that while the optimization targeted energy minimization, the FFNN/GA did not outperform the base-case energy performance. The lower pressure recommendation implies an attempt to improve energy efficiency by reducing system resistance, yet the increased feed rate appears to have shifted the optimization bias towards yield improvement, potentially offsetting the energy savings.

Table 7. The optimal operating conditions obtained through FFNN/GA optimization in comparison with the base case Aspen Plus simulation.

Operating parameters	Base case	FFNN/GA
Feed flowrate (ton/h)	390.827	400.000
Pump-around flowrate (ton/h)	190.000	13.979
CDU pressure (bar)	3.24	2.58
Reflux ratio (-)	2.136	2.483
Kerosene flowrate (ton/h)	38.144	204.930
Furnace duty (MW)	290.061	-
Energy consumption (GJ/h)	0.737	1.036

Table 8 presents energy consumption recalculated in Aspen Plus using the FFNN/GA-optimized parameters. The FFNN/GA predicted an energy consumption of 1.036 GJ/h, whereas the recalculated Aspen Plus value for the same conditions was higher, at 1.22 GJ/h. This difference corresponds to a prediction error of approximately 15%, indicating that although the FFNN provided a consistent and convergent optimization path, it tended to underestimate the actual energy consumption when tested in the rigorous process simulator.

Table 8. Comparison of the optimized energy consumption obtained via Aspen Plus and FFNN/GA optimization.

Operating parameters	Aspen Plus	FFNN/GA	Percentage error (%)
Energy consumption (GJ/h)	1.226	1.036	15.5

Despite achieving convergence and maintaining a prediction error well below 20%, the final optimized energy consumption remained higher than the Aspen Plus base-case value of 0.737 GJ/h. This result demonstrates that the objective of reducing energy consumption was not effectively met. While the FFNN/GA optimization recommended operational adjustments that aligned with energy-saving strategies like lowering column pressure and reducing pump-around flow rate, the other adjustments, particularly the increase in feed flowrate, may have shifted the optimization toward yield improvement rather than pure energy minimization. This outcome suggests that trade-offs embedded in the neural network training process or the GA fitness function configuration influenced the optimizer's search priorities.

4. Conclusions

This study presented an integrated modelling and optimization framework for a CDU. An Aspen Plus simulation of the CDU was first developed, incorporating detailed component specifications and appropriate thermodynamic models. Two neural network architectures, including an AENN and a FFNN were trained using the simulation dataset to predict key performance indicators of the CDU. Both models achieved promising predictive accuracy, with FFNN demonstrated superior performance accuracies for both outputs. Leveraging this accuracy, the FFNN was embedded within a GA to identify operating parameters that minimize energy consumption. The FFNN/GA yielding an operational configuration characterized by reduced column pressure and increased feed flowrate. The FFNN/GA optimal energy consumption did not outperform the base-case value, highlighting a gap between surrogate-predicted and simulation-verified performance. The findings underscore that while surrogate-assisted GA optimization can provide rapid, computationally efficient process screening, its effectiveness depends critically on surrogate fidelity across the entire search space and on the formulation of the fitness function to capture competing objectives. Future work should address these aspects by incorporating multi-objective optimization frameworks and hybrid surrogate strategies to enhance robustness and predictive capability across diverse search space.

Acknowledgments: The authors would like to thank the Ministry of Higher Education Malaysia for financial support through Fundamental Research Grant Scheme (FRGS) number FRGS/1/2022/TK05/USM/01/5.

References

1. Ochoa-Estopier, L., Enriquez-Gutierrez, V.M., Lu, C., Fernandez-Ortiz, J.M., Herrero-Soriano, L., Jobson, M.: Industrial Application of Surrogate Models to Optimize Crude Oil Distillation Units. Chem. Eng. Trans. 69, 289–294 (2018). https://doi.org/10.3303/CET1869049.
2. Ibrahim, D., Jobson, M., Guillén-Gosálbez, G.: Optimization-Based Design of Crude Oil Distillation Units Using Rigorous Simulation Models. Ind. Eng. Chem. Res. 56, 6728–6740 (2017). https://doi.org/10.1021/acs.iecr.7b01014.
3. Ye, L., Zhang, N., Li, G., Gu, D., Lu, J., Lou, Y.: Intelligent Optimization Design of Distillation Columns Using Surrogate Models Based on GA-BP. Processes. 11, 2386 (2023). https://doi.org/10.3390/pr11082386.
4. Ochoa-Estopier, L.M., Jobson, M., Smith, R.: Operational optimization of crude oil distillation systems using artificial neural networks. Comput. Chem. Eng. 59, 178–185 (2013). https://doi.org/10.1016/j.compchemeng.2013.05.030.
5. Shi, B., Yang, X., Yan, L.: Optimization of a crude distillation unit using a combination of wavelet neural network and line-up competition algorithm. Chin. J. Chem. Eng. 25, 1013–1021 (2017). https://doi.org/10.1016/j.cjche.2017.03.035.
6. Ibrahim, D., Jobson, M., Li, J., Guillén-Gosálbez, G.: Optimization-based design of crude oil distillation units using surrogate column models and a support vector machine. Chem. Eng. Res. Des. 134, 212–225 (2018). https://doi.org/10.1016/j.cherd.2018.03.006.
7. Liau, L.C.-K., Yang, T.C.-K., Tsai, M.-T.: Expert system of a crude oil distillation unit for process optimization using neural networks. Expert Syst. Appl. 26, 247–255 (2004). https://doi.org/10.1016/S0957-4174(03)00139-8.
8. Motlaghi, S., Jalali, F., Ahmadabadi, M.N.: An expert system design for a crude oil distillation column with the neural networks model and the process optimization using genetic algorithm framework. Expert Syst. Appl. 35, 1540–1545 (2008). https://doi.org/10.1016/j.eswa.2007.08.105.

9. Popoola, L.T., Al-Saidi, A., Asmara, Y.P., Taura, U., Obende, B.A., Osibuamhe, M.O.: Artificial neural network model for process monitoring of crude oil distillation column in a petroleum refinery. Concurr. Eng. 0, 1–14 (2025). https://doi.org/10.1177/1063293X251329460.
10. Streb, A., Mazzotti, M.: Performance limits of neural networks for optimizing an adsorption process for hydrogen purification and CO2 capture. Comput. Chem. Eng. 166, 107974 (2022). https://doi.org/10.1016/j.compchemeng.2022.107974.
11. Xiong, Y., Shi, X., Ma, Y., Chen, Y.: Optimization design of crude oil distillation unit using bi-level surrogate model. Front. Control Eng. 4, 1162318 (2023). https://doi.org/10.3389/fcteg.2023.1162318.
12. Doust, A.M., Sadeghi, J., Shahraki, F.: Steady-state and dynamic simulation and control of crude oil distillation unit: results investigation in real processes. Pet. Coal. 53, 167–179 (2011).
13. Bahrun, M.H.V., Bono, A., Othman, N., Zaini, M.A.A.: Neural network-assisted multi-objective optimization of dual-reflux pressure swing adsorption in biogas upgrading. J. Chem. Technol. Biotechnol. (2025). https://doi.org/10.1002/jctb.7879.
14. Wang, Z., Shen, Y., Zhang, D., Tang, Z., Li, W.: A comparative study of multi-objective optimization with ANN-based VPSA model for CO2 capture from dry flue gas. J. Environ. Chem. Eng. 10, 108031 (2022). https://doi.org/10.1016/j.jece.2022.108031.
15. Li, P., Pei, Y., Li, J.: A comprehensive survey on design and application of autoencoder in deep learning. Appl. Soft Comput. 138, 110176 (2023). https://doi.org/10.1016/j.asoc.2023.110176.
16. Katoch, S., Chauhan, S.S., Kumar, V.: A review on genetic algorithm: past, present, and future. Multimed. Tools Appl. 80, 8091–8126 (2021). https://doi.org/10.1007/s11042-020-10139-6.
17. Hannan, M.A., Al-Shetwi, AliQ., Begum, R.A., Young, S.E., Hoque, M.M., Ker, P., Mansur, M., Alzaareer, K.: The value of thermal management control strategies for battery energy storage in grid decarbonization: Issues and recommendations. J. Clean. Prod. 276, 124223 (2020). https://doi.org/10.1016/j.jclepro.2020.124223.
18. Gary, J.H., Handwerk, G.E., Kaiser, M.J.: Petroleum Refining: Technology and Economics. CRC Press, Boca Raton (2007).
19. Berahmand, K., Daneshfar, F., Salehi, E.S., Li, Y., Xu, Y.: Autoencoders and their applications in machine learning: a survey. Artif. Intell. Rev. 57, 28 (2024). https://doi.org/10.1007/s10462-023-10662-6.

Citation: Makele, E.M.; Bahrun, M.H.V; Ahmad, Z. Process Simulation and Optimization of a Crude Distillation Unit Using Aspen Plus and MATLAB. *ASEAN Journal of Process Control* **2025**, *vol*, issue. https://doi.org/10.3390/xxxxx