Optimization-based design of distillation columns using surrogate models

Marc Caballero, Anton A. Kiss, Ana Somoza-Tornos\*

Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, The Netherlands

A.SomozaTornos@tudelft.nl

Abstract

Continuous process re-design and optimization are required to assess and implement emerging sustainable technologies. Being distillation one of the most energy intensive operations in the process industry, the design and optimization of such is crucial. This paper presents a novel approach to distillation column optimized design using a neural network surrogate model embedded in particle swarm optimization (PSO). The surrogate model resembles the input/output structure of shortcut models, yet it is trained with data from rigorous calculations carried out in Aspen Plus. The surrogate model is integrated into a PSO algorithm, enabling to obtain global optimum points while simultaneously including simplified constraints through cost function penalty. A case study illustrates how this novel methodology combines the simplicity of a shortcut method while offering design parameters closer to an already optimized rigorous model.

**Keywords**: distillation design, shortcut distillation, surrogate modelling, particle swarm optimization.

* 1. Introduction

The energy transition introduces new challenges for process design, from the shift to alternative feedstocks to the electrification of thermal demands (*Mallapragada et al., 2023; Lopez et al., 2023*). Continuous efforts are dedicated to exploring novel reactions, enhancing yields and selectivities, and transitioning to alternative fuels. Translating these efforts into fully-operable processes raises the critical need for constant process re-design and optimization, which poses a challenge for commonly used optimization approaches.

Traditional distillation column design often relies on heuristics and experience. Rigorous models, despite providing detailed insights, are computationally expensive and prone to convergence issues, challenging their use in deterministic optimization algorithms, and resulting in non-integer nonlinear problems. Meanwhile, simplified models lack the detail necessary for robust column design. Surrogate modeling, lying between these two alternatives, yields results close to rigorous simulations within the surrogate's training space, avoiding convergence issues and significantly reducing computational time.

Recent literature displays the use of surrogate modelling embedded in distillation optimization. It has been used for specific distillation systems, such as crude oil columns (*Ibrahim et al., 2018*) or vacuum distillation (*H’ng et al, 2021*). More generic approaches to this problem involve substituting the entire optimization for a surrogate model (*Quirante et al., 2015 and Keßler et al., 2018*).

Therefore, a novel optimization-based approach for designing distillation columns, using surrogate models, is proposed.

* 1. Methodology

The proposed methodology is depicted in Figure 1, formed by three main parts. First, gathering and pre-processing the data, followed by using it to train a surrogate model for distillation. Lastly, embedding the trained surrogate model into an optimization framework. This section provides a general overview of each part.



Figure 1. General workflow for the proposed methodology

* + 1. Data Gathering

The data collection approach for training the surrogate model involves several steps. First, the feed to the distillation column and the desired separation must be defined. Once the separation is chosen, the variability in the column feed is calculated to enable the model to handle a spectrum of compositions and operating conditions. This involves heuristically defined ranges and simplified simulations of the upstream.

After defining the ranges for the feed stream, the bounds for the column’s design parameters can be defined. The proposed method for that task is to run enough calculations with a distillation shortcut method (Fenske-Underwood-Gilliland-Kirkbride method) within the different feed conditions established.

To generate the data for training the surrogate model, Latin hypercube sampling is employed to generate input data for rigorous distillation calculations, within the bounds obtained previously. These calculations involve solving the MESH equations (Material balance, Phase Equilibrium, Mole fraction Summation, and Heat) using Aspen Plus. Both simulation results and corresponding inputs are stored in both .csv files with raw data and .json files with the object structure developed for distillation calculations.

As part of the data gathering process, raw data undergoes pre-processing. Non-converged simulations are filtered out, and normalization is applied to enhance machine learning model performance and prevent training issues. An essential pre-processing step is defining the input and output structure of the model. In this methodology, the model's inputs include feed compositions, pressure, temperature, and distillate flows, while the outputs encompass column design variables (number of trays, number of feed trays, reflux ratio, distillate to feed ratio), reboiler and condenser temperatures and duties, and the annualized fixed cost of production (FCOP) of the column. This input/output (I/O) structure resembles to the shortcut model, compensating for its lack of accuracy by training the surrogate model with data from rigorous simulations.

Finally, the data used for surrogate model training is divided into 3 subsets: a training set, a validation set, and a test set. These splits are standard in machine learning training procedures, with the training set used for model training, the validation set for adjusting hyperparameters, and the test set for evaluating model performance with "unseen data."

* + 1. Model Training

Distillation, inherently nonlinear due to phase equilibrium between liquid and vapor in the column, requires nonlinear machine learning models. Once the machine learning model is chosen, a dedicated implementation is performed to unlock hyperparameter tuning and architecture modification capabilities.

Regarding the performance, the r2 factor is the metric chosen over other regression metrics since its scale ranges from 0 to 1. The learning curves are useful for assessing the model’s performance, as well as for hyperparameter tuning. Learning curves are valuable for assessing model performance and hyperparameter tuning. The divergence between the training and validation curves suggests overfitting, while a significant gap between the curves indicates the model may not be effectively learning the relationships between inputs and outputs.

* + 1. Optimization framework embedding

Given the complexity of implementing gradient-based methods with nonlinear machine learning models, a stochastic (derivative-free) optimization method is recommended. In this work, the chosen method is particle swarm optimization (PSO).

* 1. Case Study

The general case study (regarding compositions and conditions for streams) is derived from *Somoza et al. 2020*. The target product is ethylene, with an alternative sustainable upstream process based on polyethylene pyrolysis; hence the aim to train the surrogate model toward the C2 fractionator column, which is the main product of the case study.

To capture compositions and design parameters resembling the C2 column, the distillation train feed stream from the case study is taken and fed into a cascade of two shortcut models in Aspen HYSYS V12. The first models the C1 column upstream, while the second estimates the design parameters for the C2 column; number of trays, feed tray, and reflux ratio. The first shortcut specifies molar fractions of light and heavy key components in bottoms and distillate, uniformly distributed from 0.1 to 0.001. Both shortcuts are set with a reflux ratio spec of 1.5 times the minimum reflux.

Rigorous simulations in Aspen Plus V12 are carried out using a RadFrac model for column simulations, with bounds from the shortcut simulations. Compositions are derived from the shortcut simulations, and the remaining parameters have their ranges defined as shown in Table 1. Pressure, feed temperature, and distillate to feed ratio (DFR) values are within commonly seen ranges in literature, such as in the work of *Spallina et al., 2017*. Combining 1,000 normally distributed values from Table 1 with 100 different compositions from the shortcut simulations results in 100,000 rigorous simulations. These simulations yield distillate flows and parameters necessary to calculate the total annualized cost (TAC) of the column, following the equations proposed by Sinnot & Towler, 2020.

Table 1. Bounds for design parameters in data gathering rigorous simulations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Pressure[kPa] | Feed Temp.[ºC] | DFR[-] | Reflux Ratio[-] | Number of trays[-] | Feed Tray[-] |
| Upper Bound | 3200 | 92 | 0.7 | 4.54 | 40 | 16 |
| Lower Bound | 1600 | 25 | 0.3 | 0.72 | 16 | 6 |
|  |  |  |  |  |  |  |

Following the proposed methodology, a multilayer perceptron (MLP) is trained using the Keras API of Tensorflow 2.14; with the following “shortcut-like” input/output structure described in section 2.2. The model architecture is comprised of an input layer matching the input dimension, 8 hidden layers with 256 neurons each, and an output layer with the same size as the output dimension; all densely connected. The training was conducted for 10,000 epochs.

A PSO library from L.J.V Miranda, 2018, *Pyswarms 1.3,* is used to optimize a column with a fixed feed flow (by specifying the input bounds). The optimization problem is:

|  |  |
| --- | --- |
|   | (1) |
|   | (2) |
|    | (3) |
|   | (4) |

Where *“FCOP”* in Eq. (2), corresponds to the (annualized) Fixed Cost of Production; Qr, Tr, Qc and Tc are the reboiler and condenser duties and temperatures, which are taken as an output of the neural network. The cost for the utilities (*Ut*) in Eq. (3) comes from the work of *Ulrich and Vasudevan, 2006*, as well as the CS,f parameter, used to relate energy with fuel cost. The *CEPCI* corresponds to the Chemical Engineering Plant Cost Index, which value is the one corresponding to 2022 from *ChemEngOnline.com*. The recovery constraint is implemented as a penalty constraint. The optimizing algorithm used was the global optimizer, with cognitive, social and inertia parameters being respectively 1, 1 and 0.0001. The number of particles used in the swarm is 500.

* 1. Results and discussion

The results of the MLP training are shown in Figure 2, Figure 3 and Table 2. Figure 2 displays the learning curves for the training, showing consistent trends with minor differences. This indicates that the MLP model effectively learns the nonlinear relationships between inputs and outputs without overfitting the data. While the majority of learning is achieved within 2,000 epochs, extending the process to 10,000 epochs enhances model performance. The parity plot in Figure 3 illustrates the correlation between the data and the outputs generated from the MLP model for the FCOP, the variable with the lowest r2 factor, yet it can be appreciated how the deviations resemble more outliers rather than a deviation trend.

In addition, the calculated R2 factors with the test dataset, shown in Table 2, reflect the performance of the MLP model in predicting the output values. Moreover, the MLP model outperforms in terms of computational time the rigorous models, in single simulations by a factor of x17, and in performing multiple simulations, thanks to the parallelization capabilities of the neural network, by a factor of x1500.

|  |  |
| --- | --- |
| Figure 2. Learning curves of MLP | Figure 3. FCOP parity plot |

This computational time advantage plays a crucial role in the optimization step, where multiple function evaluations are performed. For the optimization, values in Table 3 are used as a case study, retrieved from the test data slice, with the lowest cost that meets the recovery constraint rc ≥ 0.95.

The data from Table 3 is processed by the surrogate-based optimization methodology proposed, as well as in the “traditional” workflow for column design (first use a shortcut model to obtain the design variables, and use those results in a rigorous model). For a more meaningful comparison, the values are processed in a different instance with the pressure and temperature values from the optimization. Table 4 lists the design parameters and results. This reveals that using a rigorous simulation based on the parameters from a shortcut model yields a case with low cost, but does not meet the C2 recovery constraint. Using the design parameters from the shortcut model and the parameters obtained from the optimization, the constraint is still not met (despite better recovery), in addition to a significant cost increase. From the optimization output, the cost is much lower and the recovery constraint is met. However, when using the feed conditions and design parameters in a rigorous simulation, although the C2 recovery is increased significantly, the cost suffered more than a twofold increase. Nevertheless, this case meets the constraints and has less cost than the case where the shortcut design parameters and optimized pressure and temperature for the feed were used.

Table 2. R2 factors of the trained MLP

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Number of trays | Feed Tray | Reflux Ratio | DFR | Reb. Temp | Cond.Temp. | Reb. Duty | CondDuty | FCOP |
| R2  | 0.978 | 0.981 | 0.988 | 0.992 | 0.999 | 0.998 | 0.993 | 0.994 | 0.939 |

Table 3. Feed stream summary

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Methane | Ethylene | Propylene | Propyne | 1-Butene | 1,3-Butadiene | Benzene | Temp.[ºC] | Pres.[kPa] |
| Flow [kg/s] | 0.00778 | 2.31663 | 0.91118 | 0.05499 | 0.09434 | 0.58862 | 0.69578 | 49.89 | 1603 |

Table 4. Parameters and results of different approaches

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Case | Temp.[ºC] | Pres.[kPa] | NT[-] | FT[-] | RR[-] | DFR[-] | C2 Recov.[ - ] | TAC [$] |
| Shortcut | 49.89 | 1603 | 13.58 | 10.49 | 1.17 | 0.4845 | 0.971 | N/A |
| Shortcut + Rigorous | 49.89 | 1603 | 14 | 11 | 1.17 | 0.4845 | 0.882 | 1,557,575 |
| Shortcut + Rig. + Opt. P,T | 45.13 | 2230 | 14 | 11 | 1.17 | 0.4845 | 0.929 | 3,587,728 |
| Optimization | 45.13 | 2230 | 32.14 | 11.71 | 2.84 | 0.6380 | 0.974 | 1,051,084 |
| Optimization + Rigorous | 45.13 | 2230 | 32 | 12 | 2.84 | 0.6380 | 0.999 | 2,227,929 |

* 1. Conclusion

Through the case study, the methodology proposed in this work has been proven successful. The use of particle swarm optimization provides the capability to obtain the design parameters of distillation columns through the use of surrogate modelling while meeting recovery constraints. The neural network is capable of providing the design parameters for a distillation column, resembling the shortcut model input/output structure, with the reliability of having been trained with data from rigorous models, bridging the gap between these two models for distillation. The application on a case study for the distillation of light hydrocarbon shows that not only the recovery constraints can be met with this method, but also improved up to an additional 4%. Additionally, the resulting design parameters and feed conditions lead to cost reduction up to a 62%. The TAC obtained from the optimization provides an estimate for its order of magnitude. Therefore, this methodology holds great promise for speeding up the design process of distillation columns compared to the traditional workflow, while meeting constraints.

References

D. S. Mallapragada, Y. Dvokin, M. A. Modestino, D. V. Esposito, W. A. Smith, B. Hodge, M. P. Harold, V. M. Donnelly, A. Nuz, C. Bloomquist, K. Baker, L. C. Grabow, Y. Yan, N. N. Rajput, R. L. Hartman, E. J. Biddinger, W. S. Aydil, A. D. Taylor, 2023, Decarbonization of the chemical industry through electrification: Barriers and opportunities, Joule, 7, 1, 23-41.

G. Lopez, D. Keiner, M. Fasihi, T. Koiranen, C. Breyer, 2023, From fossil to green chemicals: sustainable pathways and new carbon feedstocks for the global chemical industry, Energy & Environ. Sci., 26, 2879.

N. Quirante, J. Javaloyes, J. A. Caballero, 2015, Rigorous Design of Distillation Columns Using Surrogate Models Based on Kriging Interpolation, AIChE Journal, 61, 7, 2169 - 2187.

D. Ibrahim, M. Jobson, J. Li, G. Guillén-Gosálbez, 2018, Optimization-based design of crude oil distillation units using surrogate column models and a support vector machine, Chemical Engineering Research and Design, 134, 212-225.

S. X. H’ng, L. Y. Ng, D. K. S. Ng, V. Andiappan, 2021, Optimising of vacuum distillation units using surrogate models, IOP Conf, Ser.: Mater. Sci. Eng., 1195, 012050.

T. Keßler, C. Kunde, N. Mertens, D. Michaels, A. Kienle, 2018, Global optimization of distillation columns using surrogate models, SN Appl. Sci. 1, 11(2019).

A. Somoza-Tornos, A. Gonzalez-Garay, C. Pozo, M. Graells, A. Espuña, G. Guillén-Gosálbez, 2020, Realizing the Potential High Benefits of Circular Economy in the Chemical Industry: Ethylene Monomer Recovery via Polyethylene Pyrolysis, ACS Sustainable Chem. Eng., 8,9, 3561 – 3572.

V. Spallina, I. Campos Velarde, J.A. Medrano Jimenez, H. Reza Godini, F. Gallucci, M. van Sint Annaland, 2017, Techno-economic assessment of different routes for olefins production through the oxidative coupling of methane (OCM): Advances in benchmark technologies, Energy Conservation and Management, 154, 244 – 261.

R. Sinnot and G. Towler, Chemical Engineering Design (Sixth Edition), 2020.

G. D. Ulrich, P. T. Vasudevan, How to Estimate Utility Costs, 2006, Chemical Engineering, 113, 66-69.

L. J. V. Miranda, 2018, PySwarms: a research toolkit for Particle Swarm Optimization, The Journal of Open Source Software, 3(21) 433.

The Chemical Engineering Plant Cost Index. Chemical Engineering, http://www.chemengonline.com