

A Neural Model of Pinch-Based Multicomponent Distillation for Applications in Flowsheet Synthesis

Alexander B. Wolf^a, Mirko Skiborowski^b, and Jakob Burger^{a*}

^a Technical University of Munich, TUM Campus Straubing for Biotechnology and Sustainability, Laboratory of Chemical Process Engineering, Straubing, Germany

^b Hamburg University of Technology, Institute of Process Systems Engineering, Hamburg, Germany

* Corresponding Author: burger@tum.de.

ABSTRACT

This work presents a data-driven surrogate modeling framework for predicting distillation behavior assuming an infinite number of stages and distillation limits informed by residue-curve topology and pinch-point feasibility analysis. The framework provides a direct mapping from feed composition and distillate-to-feed ratio (D/F) to distillate and bottom product compositions, making it suitable for flowsheet synthesis and optimization applications. The approach combines three components: a classifier that identifies feasible singular-point splits, a boundary regression model that predicts D/F limits separating pure- and mixed-product operating regimes, and a neural network that interpolates product compositions in the intermediate regime. The method is demonstrated for the ternary system ethanol, benzene, and water at 1 atm using data generated from rigorous vapor-liquid-liquid equilibrium analysis. Results show that the framework provides reliable predictions for pure splits while retaining smooth interpolation behavior in the mixed regime.

Keywords: Distillation, Surrogate Model, Machine Learning, Process Design, Modelling and Simulations

INTRODUCTION

Modeling distillation behavior of multicomponent systems is central to conceptual process design. Cumbersome process simulations based on equilibrium stage models are computationally demanding, prone to numerical failure, and require numerous specifications. They are therefore typically avoided in early design phases in favor of shortcut models that focus on the general feasibility of a split rather than detailed apparatus design. Shortcut models typically assume an infinite number of theoretical stages and target the minimum energy demand for pinched separations, e.g., the Underwood method for constant relative volatilities or the rectification body method for general nonideal mixtures [1]. Such shortcuts require prespecified products for which feasibility is determined. In many applications such as flowsheet synthesis, a mapping from feed and split specifications (e.g. distillate-to-feed ratio (D/F)) to product composition is required instead.

If such a mapping is required, one could rely on heuristics to suggest preferred column products [2] or wrap existing shortcut models inside iterative loops that adjust

product specifications until material balances and split specifications are met [3]. Heuristics often restrict the design space, for example by excluding non-sharp splits [2], while iterative loops can be slow and may fail to converge, particularly in nonideal and azeotropic settings [3]. Ryll et al. [4] propose direct mapping from feed specifications and D/F to product compositions under infinite-stages and total reflux (∞/∞) assumptions; however, these methods are geometric, work only for discrete pressures, and are limited to four components [5] [6].

In this work, we address these limitations by proposing a data-driven surrogate model for distillation columns assuming an infinite number of stages and operating under general distillation limits defined by residue-curve behavior at total reflux and pinch point-based feasibility limits. The key objective is a deterministic, direct mapping from the molar feed composition and molar D/F to the distillate and bottom molar compositions for *a priori* defined systems of any vapor-liquid equilibrium (VLE) topology. Accordingly, these models serve as tailor-made surrogates for feasibility studies in conceptual process design, intended for flowsheet synthesis and machine learning (ML)-based process optimization, where a

large number of process alternatives have to be screened before optimizing the column parameters (e.g., number of stages, reflux ratio) in detail.

Data-driven surrogate models have been used successfully for distillation column design and optimization, including neural network-based [7–9] and Gaussian process-based [10, 11] models for different column configurations. Further, Panofen et al. [12] have recently proposed a surrogate model of distillation boundaries in azeotropic multicomponent mixtures. However, data-driven surrogates are always tailor-made for, and thus limited to one specific task. To the best of our knowledge, no surrogate currently exists for predicting distillation column products at the limiting cases of an infinite number of stages and distillation boundaries informed by residue-curve topology and pinch-point feasibility analysis. We close this gap with the present work.

We present results for ternary systems, although the methodology is in principle not limited and can be extended to multicomponent systems when sufficient computational power is available during training. The resulting surrogate model can be used in any flowsheet synthesis problem. Because the applied ML-based models evaluate quickly and robustly, it is particularly suited for optimization methods that identify flowsheet concepts, such as superstructure optimization and ML-based flowsheet design [3], where numerous flowsheet evaluations are needed.

METHODOLOGY

Overview

The objective is a direct and deterministic mapping from feed composition and D/F to the corresponding product compositions,

$$(\mathbf{z}, D/F) \rightarrow (\mathbf{x}, \mathbf{y}), \quad (1)$$

where $\mathbf{z} = (z_1, z_2, \dots, z_n)$ denotes the mole fractions of the feed an n -component mixture with $\sum_{i=1}^n z_i = 1$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$ the distillate's mole fractions, and $\mathbf{x} = (x_1, x_2, \dots, x_n)$ the bottom product's mole fractions.

The training data for this mapping are generated using the Automated Process Synthesis (APS) Tool v1.0 [2], which performs thermodynamically rigorous VLE analysis of multicomponent mixtures. The APS Tool identifies azeotropes, distillation regions, and feasible singular-point splits, that is, separations involving pure components or azeotropes, using general distillation limits defined by residue-curve topology and continuation-based pinch-branch calculations. Residue curves characterize separation behavior in the total reflux limit, whereas pinch lines describe the limiting behavior of reversible columns operated at finite reflux. A split is considered feasible if it is admissible under either limiting case. Consequently, the

feasible split set identified by the APS Tool represents the outer envelope of thermodynamically admissible separations for a given feed.

In the present work, these identified limiting singular-point splits are used as anchors for surrogate modeling. Once the limiting products are known, the separation behavior between them can be represented in terms of D/F , without explicitly distinguishing the underlying reflux mechanism.

A straightforward approach to learning the mapping in Eq. (1) would be to train a sufficiently large neural network directly on input–output data. While feasible, this approach was found to exhibit inconsistent accuracy in predicting azeotropic compositions and occasionally produced slightly impure product compositions in cases where a pure-product split is expected.

To address these limitations, we propose embedding the neural network into a structured hierarchical learning framework that leverages knowledge on distillation theory to guide the training process and enhance physical interpretability. The overall structure of the proposed framework is illustrated in Figure 1.

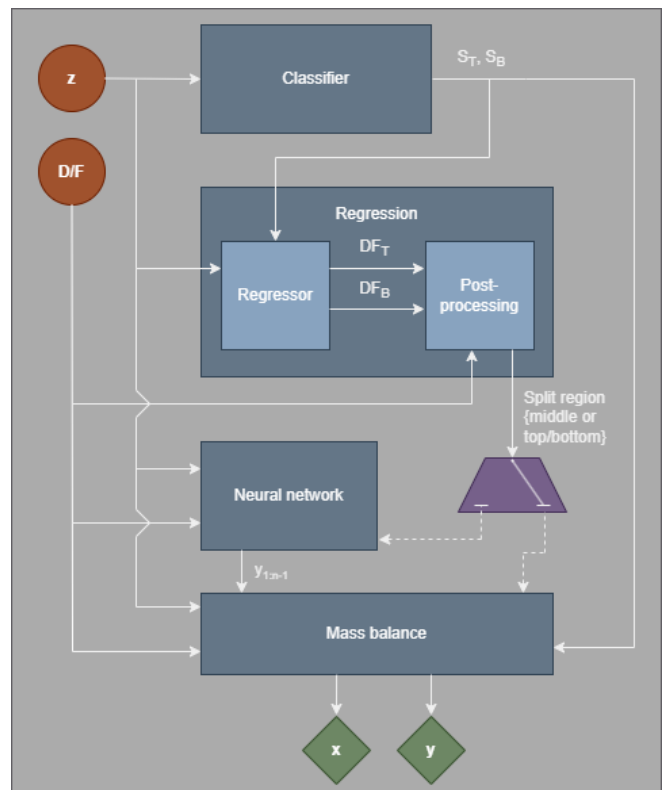


Figure 1. Hierarchical surrogate framework integrating data-driven models for predicting distillation behavior.

At the beginning of the pipeline, two learning tasks are introduced to extract physically interpretable information about the separation before predicting the compositions $(\mathbf{x}, \mathbf{y})$. The first task is a classification problem

that determines, based on $\mathbf{z}$, which singular points in the VLE act as unstable and stable nodes of the distillation lines governing the splits. The second task is a boundary regression that predicts the limiting D/F values associated with attainable singular-point separations for given feed composition $\mathbf{z}$ and singular points determined in the classifier.

In certain cases, the product compositions can be determined directly from the outcomes of these two tasks. For example, when separation to a pure singular-point product is feasible at a given D/F , the corresponding distillate or bottom compositions coincide with the composition of that singular point. Only for D/F values that do not admit a pure singular-point split, is the neural network used to predict the product compositions.

Both the classifier and the boundary regressor are entirely data-driven, with the required labels obtained and stored during the data-generation process. Overall, the proposed hierarchical framework effectively *distills* VLE knowledge into a data-driven surrogate model that is computationally efficient, numerically stable, and well-suited for flowsheet synthesis and ML-based process synthesis.

Data generation and representation

The training data are derived from rigorous VLE analyses performed with the MATLAB-based APS Tool. The proposed framework approximates these results while preserving physical interpretability and numerical robustness. Rather than relying on repeated evaluations with the APS Tool, physically meaningful separations are extracted from the tool's thermodynamic analyses and combined with material-balance calculations to determine the D/F value corresponding to a particular product split. This information can then be embedded into predictive models that accept D/F as an explicit input and can be evaluated efficiently in a Python-based environment.

First, a separation system is defined by specifying pure-component properties, such as Antoine equation parameters, together with a binary activity coefficient model. Based on this thermodynamic description, the singular points of the system are identified. The finite set of thermodynamically admissible singular points as top products is denoted as $\mathbf{S}_T = \{s_{T,1}, s_{T,2}, \dots, s_{T,K}, \emptyset_T\}$, and the corresponding set for bottom products by $\mathbf{S}_B = \{s_{B,1}, s_{B,2}, \dots, s_{B,K}, \emptyset_B\}$, where $\emptyset_T$ and $\emptyset_B$ denote the absence of a feasible pure distillate or pure bottom split, respectively.

Feed compositions $\mathbf{z}$ are then systematically sampled over the composition simplex. The evaluated feeds evenly cover the simplex, as studies on surrogate models show that such coverage reduces error and delivers large speedups when embedded later in flowsheet optimization [12]. For each feed composition, a one-hot encoding is assigned to indicate which elements of $\mathbf{S}_T$ and $\mathbf{S}_B$

correspond to feasible top and bottom splits, respectively. In addition, the corresponding product compositions $\mathbf{y}$ and $\mathbf{x}$ for the pure split at the maximum achievable yield are provided.

The APS Tool may identify multiple competing singular points for a given feed composition, for example near distillation boundaries where more than one pure top or bottom product is theoretically feasible. In such cases, expert judgment is applied to select the most appropriate split.

Once the singular-points splits have been identified, two critical values of D/F are defined for each feed composition: DF_T , the maximum D/F at which a pure distillate is achievable, and DF_B , the minimum D/F at which a pure bottom product is achievable. These values are calculated from a component-wise molar balance,

$$z_i = y_i(D/F) + x_i(1 - D/F). \quad (2)$$

The resulting values partition the operating space into three regimes:

Top: $0 < D/F \leq DF_T$,

Middle: $DF_T < D/F < DF_B$

Bottom: $DF_B \leq D/F < 1$.

If no pure distillate split exists for a given feed, DF_T is set to zero, and if no pure bottom split DF_B is set to one. This information constitutes the complete dataset required to train the classifier and boundary regressor.

To train the neural network, additional data are generated to capture the relationship outlined in Eq. (1). For singular-point splits, product compositions are computed for pure-product yields Y_i ranging from $Y_i = 0.01$ to the maximum feasible yield. An example of such data points for a single feed composition in a ternary mixture of ethanol, benzene, and water at 1 atm is illustrated in Figure 2.

The feed selected in Figure 2 corresponds to a separation governed by residue-curve topology, such that the intermediate operating points remain consistent with the distillation boundary associated with the total-reflux limit. Data points corresponding to a pure bottom split (orange circles) represent the bottom operating regime at high D/F values ($D/F = 0.78$ – 0.99), whereas points associated with a top split with a minimum ternary azeotrope (red squares) represent the top operating regime at low D/F values ($D/F = 0.003$ – 0.33).

Data for the middle operating regime are generated by interpolating between the limiting pure splits. When a feed admits both a pure distillate and bottom product, intermediate compositions are sampled along the discretized binary line connecting the corresponding pure-split product compositions, as indicated by the green diamonds in Figure 2. For each sampled composition, the corresponding D/F is computed using Eq. (2).

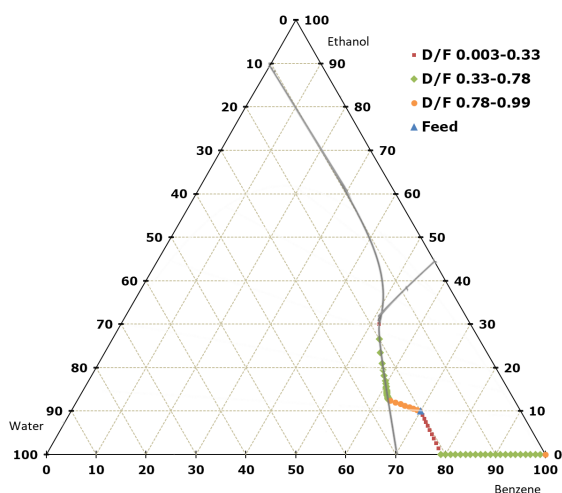


Figure 2. Example data points for ternary system collected using APS Tool [2]. Separation behavior falls within limits of residue-curve topology.

Figure 3 illustrates a case in which the selected feed admits feasible operating points associated with pinch-line feasibility. In this case, some middle-regime data points extend beyond the residue-curve-based distillation boundary, reflecting admissible separations under reversible-column limits. Both types of behavior are retained in the training dataset and treated as physically admissible by the surrogate framework.

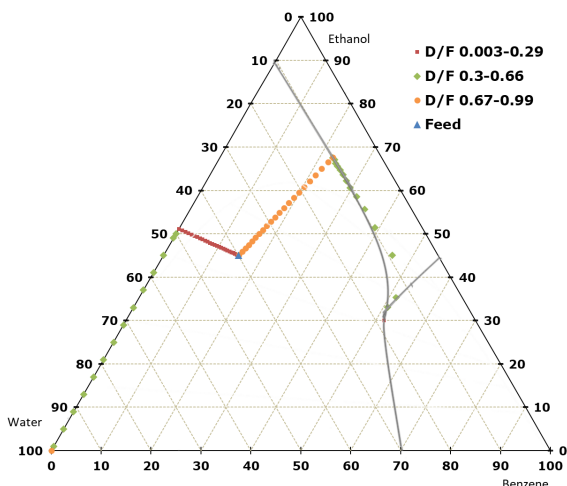


Figure 3. Example data points for ternary system collected using APS Tool [2]. Separation behavior falls within limits of pinch-line feasibility analysis.

Moreover, due to the geometry of the distillation boundaries and pinch-points limits, the resulting D/F values along these interpolation paths do not necessarily vary monotonically. This further motivates the use of a neural network to learn the relationship between D/F and product composition in the middle regime. In addition,

numerical convergence issues may arise in rigorous VLE calculations for certain intermediate operating points, motivating the development of a surrogate model capable of delivering robust predictions across the full operating range.

Singular point split classification

The first ML task addresses the identification of singular-points for a given feed composition. The input to this task is the reduced feed composition vector $\tilde{z} = (z_1, z_2, \dots, z_{n-1})$ for an n -component system. The output is a one-hot encoded vector over the predefined sets S_T and S_B , indicating which singular points correspond to the top and bottom splits, respectively, for the given feed compositions.

The model employs a histogram-based gradient boosting classifier implemented in the *scikit-learn* library in Python. This method constructs an ensemble of decision trees sequentially, where each tree is trained to correct misclassifications of the current ensemble. The histogram-based discretization of the input space improves numerical robustness and computational efficiency while preserving the ability to represent nonlinear decision boundaries in the composition simplex.

The trained classifier yields a discrete and interpretable mapping from $\tilde{z}$ to S_T and S_B within the training domain.

Operating regime prediction via boundary regression

The regression task is formulated as a boundary regression problem. Two continuous regression models are trained to predict the functions $\widehat{DF}_T = f_T(\tilde{z}, S_T, S_B)$ and $\widehat{DF}_B = f_B(\tilde{z}, S_T, S_B)$. The operating regime for a queried D/F value is then obtained by comparison with these predicted boundaries.

Explicit constraints are enforced to ensure physically consistent predictions, including clamping predicted ratios to the interval $[0, 1]$ and enforcing the condition $\widehat{DF}_T \geq \widehat{DF}_B$. The regression models are implemented using histogram-based gradient boosting regressors from the *scikit-learn* library, which efficiently capture nonlinear boundary surfaces over the composition simplex while remaining numerically robust.

Neural network for impure product splits

For operating points in the middle regime, product compositions are predicted using a feed-forward neural network implemented in Python with *PyTorch*. The network is trained as a regression model that maps $\tilde{z}$ and D/F to the distillate compositions $\hat{y} = (y_1, y_2, \dots, y_{n-1})$. The remaining y_n and corresponding bottom compositions x are computed from an overall material balance.

Product composition reconstruction

The final product compositions are predicted using a hybrid reconstruction strategy that combines analytical material balances and the neural network prediction. In the top and bottom regimes, product compositions are computed using Eq. (2) consistent with the corresponding limiting pure split. For operating points in the middle regime, the neural network is used to predict product compositions.

To ensure physical feasibility, all predicted compositions are postprocessed to satisfy component-wise material constraints. For a given D/F , the distillate mole fractions must satisfy

$$y_i \leq \min\left(1, \frac{z_i}{D/F}\right), \quad (3)$$

reflecting that a distillate stream cannot be in excess of its availability in the feed. If necessary, the predicted distillate vector is projected onto the feasible composition simplex under these caps, ensuring nonnegativity and a unit sum. Once a feasible distillate composition is obtained, the bottom composition is computed from the overall material balance (see Eq. (2)).

EXAMPLE CASE STUDY

The proposed framework is demonstrated using the ternary mixture of ethanol, benzene, and water at 1 atm, modeled using the Non-Random Two-Liquid (NRTL) g^E -model [13] with parameters taken from Aspen Plus [14]. The mixture exhibits a liquid-liquid phase split and four minimum azeotropes (three binary, one ternary; two homogeneous, two heterogeneous). The set S_B comprises pure bottom splits yielding ethanol, benzene, or water, as well as the absence of a pure bottom split. Since all feasible compositions admit a top split with the ternary minimum azeotrope, the set S_T contains only a single option and yielding a trivial classification task.

Data for split classification and operating regime prediction are generated with approximately 124,000 feed compositions samples across the VLE simplex. The samples are drawn from a structured grid with component mole fractions in the range $0.002 \leq z_i \leq 0.996$. To avoid artificial grid-induced artifacts and improve generalization, small Gaussian perturbations with standard deviations of $\sigma = 0.0003$ are added to the sampled compositions. This stochastic smoothing prevents the classifier and regression models from memorizing grid-aligned decision boundaries and encourages learning of physically meaningful, continuous separation surfaces in the composition space.

A comparatively high sampling density is used for classification and boundary regression tasks because these models must resolve sharp transitions between distillation regions and accurately capture singular-point behavior near curved distillation boundaries. In addition, generating training data for pure splits is computationally

efficient, as it requires only the identification of limiting separation behavior, rather than singular points across the entire operating range.

In contrast, training data for the neural network are generated using feed compositions in the range of $0.02 \leq z_i \leq 0.96$. For middle-regime splits, product compositions are generated along the binary line connecting the pure bottom split and the bottom composition associated with the pure distillate split, as described above and outlined in Figures 2 and 3. These compositions are sampled in increments of $\Delta z_i = \pm 0.01$. In total, approximately 293,000 data points are generated for neural network training.

Results

Classifier

The pure-split classification model was evaluated on 4,851 validation feed compositions that were not used in the training dataset, with component mole fractions in the range $0.01 \leq z_i \leq 0.98$. The results of the evaluation are shown in Figure 4, which shows the predicted pure bottom split for each feed composition z selected from the admissible set S_B .

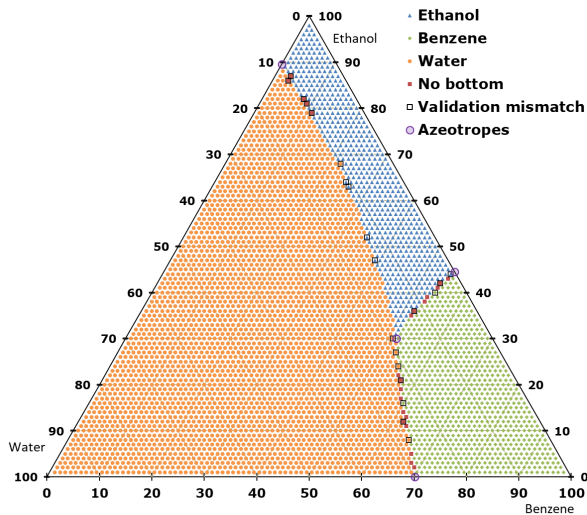


Figure 4. Validation run for singular-point split classification.

Of the analyzed feed compositions, only 21 showed mismatches with the expected splits. Notably, all of these cases occur at distillation boundaries separating feasible pure splits. In eight cases, the model predicted an incorrect pure bottom component, while in the remaining 13 cases it incorrectly classified whether a pure bottom split exists at the boundary point. Further improvements could be achieved by increasing the training density or applying additional weighting for feed compositions near distillation boundaries. In addition, incorporating confidence

measures based on the consistency of predictions in a local neighborhood of the composition space could enable the use of fallback strategies or alternative predictors when the model encounters uncertain splits.

Regression

The operating regime regression model was evaluated using the same validation set by comparing queried D/F values against the predicted DF_T and DF_B boundaries. The predicted DF_T values exhibited a mean absolute error of $|\varepsilon_T| = 0.0039$ with a standard deviation of $\sigma_T = 0.0056$, with a maximum observed error of $\varepsilon_T = 0.0178$. The prediction error for DF_B was slightly greater, with a mean absolute error of $|\varepsilon_B| = 0.0062$ and a standard deviation of $\sigma_B = 0.0129$. The two largest errors, $\varepsilon_B = 0.277$ and $\varepsilon_B = 0.194$, occurred for feed compositions where the bottom split was misclassified by the singular-point classifier, with ethanol predicted as the bottom product instead of water. Both cases lie near distillation boundaries, underscoring the sensitivity of the regression task to accurate split classification near these regions. While this suggests that additional fine-tuning of the classifier could reduce such errors, alternative strategies for improving the regression model could also be explored, including the use of different regression architectures or uncertainty-aware regression formulations.

Neural Network

The neural network is trained using mini-batch gradient descent with the Adam optimizer and a mean-squared-error (MSE) loss function. Key modeling choices, including data normalization, loss formulation, validation strategy, optimization settings, and stopping criteria, were systematically evaluated. A randomized train-validation split with a fixed random seed and early stopping based on validation loss were selected to balance accuracy and generalization. Among the tested configurations, a network with three hidden layers of 32 neurons each achieved the best validation performance, with an MSE of 2.36×10^{-5} .

Overall model

The full predictive framework was evaluated using representative feed compositions by first generating reference product compositions and the corresponding D/F values using the APS Tool. Figure 5 illustrates four representative cases. The first three feed compositions (cases (a)-(c)) were selected for their location within distinct distillation regions, for which a pure bottom split is expected. The fourth (case (d)) was chosen because it represents a composition for which the singular-point classifier produced a mismatch. The APS Tool-derived D/F values were provided as inputs to the hybrid surrogate and neural network models, and the resulting product compositions were compared against the APS reference data.

Although the hybrid surrogate and standalone neural network show similar overall agreement with the APS Tool data in cases (a)-(c), closer inspection reveals qualitative and quantitative differences in performance. In particular, for cases (a) and (b), the standalone network fails to correctly reproduce the expected pure-bottom product splits at high D/F values where thermodynamics dictates a pure split. The hybrid surrogate successfully predicts the compositions of these pure splits, as pure-split compositions are enforced by construction.

These discrepancies are reflected in the error statistics summarized in Table 1, which report the mean absolute error of the predicted product composition vectors relative to the APS Tool reference, evaluated across all components simultaneously and averaged over all considered D/F values for each split.

Table 1: Comparison of prediction errors between hybrid surrogate and neural network for representative feed compositions.

Case study	Prediction error (mean $\pm$ std. dev.)	
	Surrogate model	Neural network
(a)	4.17×10^{-4} $\pm 2.82 \times 10^{-3}$	3.57×10^{-3} $\pm 2.54 \times 10^{-2}$
(b)	3.32×10^{-5} $\pm 1.58 \times 10^{-4}$	1.69×10^{-3} $\pm 1.15 \times 10^{-2}$
(c)	8.56×10^{-5} $\pm 7.11 \times 10^{-4}$	3.23×10^{-4} $\pm 7.59 \times 10^{-4}$
(d)	1.26×10^{-1} $\pm 2.83 \times 10^{-1}$	1.10×10^{-2} $\pm 2.53 \times 10^{-2}$

For cases (a)-(c), the hybrid surrogate achieves lower mean errors, on the order of 10^{-5} - 10^{-4} , compared with 10^{-4} - 10^{-3} for the neural network, and shows substantially lower variability. In particular, the standard deviations of the surrogate predictions remain less than 3×10^{-3} , whereas the neural network exhibits standard deviations exceeding 1×10^{-2} for all splits. These differences also arise from the tendency of the neural network to produce inconsistent predictions of the ternary azeotrope composition, whereas the hybrid surrogate enforces exact singular-point compositions.

For the feed corresponding to the classifier mismatch (case (d)), errors show the opposite trend. The neural network exhibits a lower mean error than the hybrid surrogate, whose errors increase substantially due to the incorrect pure-split classification of ethanol instead of water, as seen in Figure 5. In this case, the surrogate predictions trend toward the misclassified bottom split, whereas the neural network captures a smooth transition toward the correct split. This behavior suggests that additional uncertainty-aware handling near classification boundaries could further improve the hybrid framework.

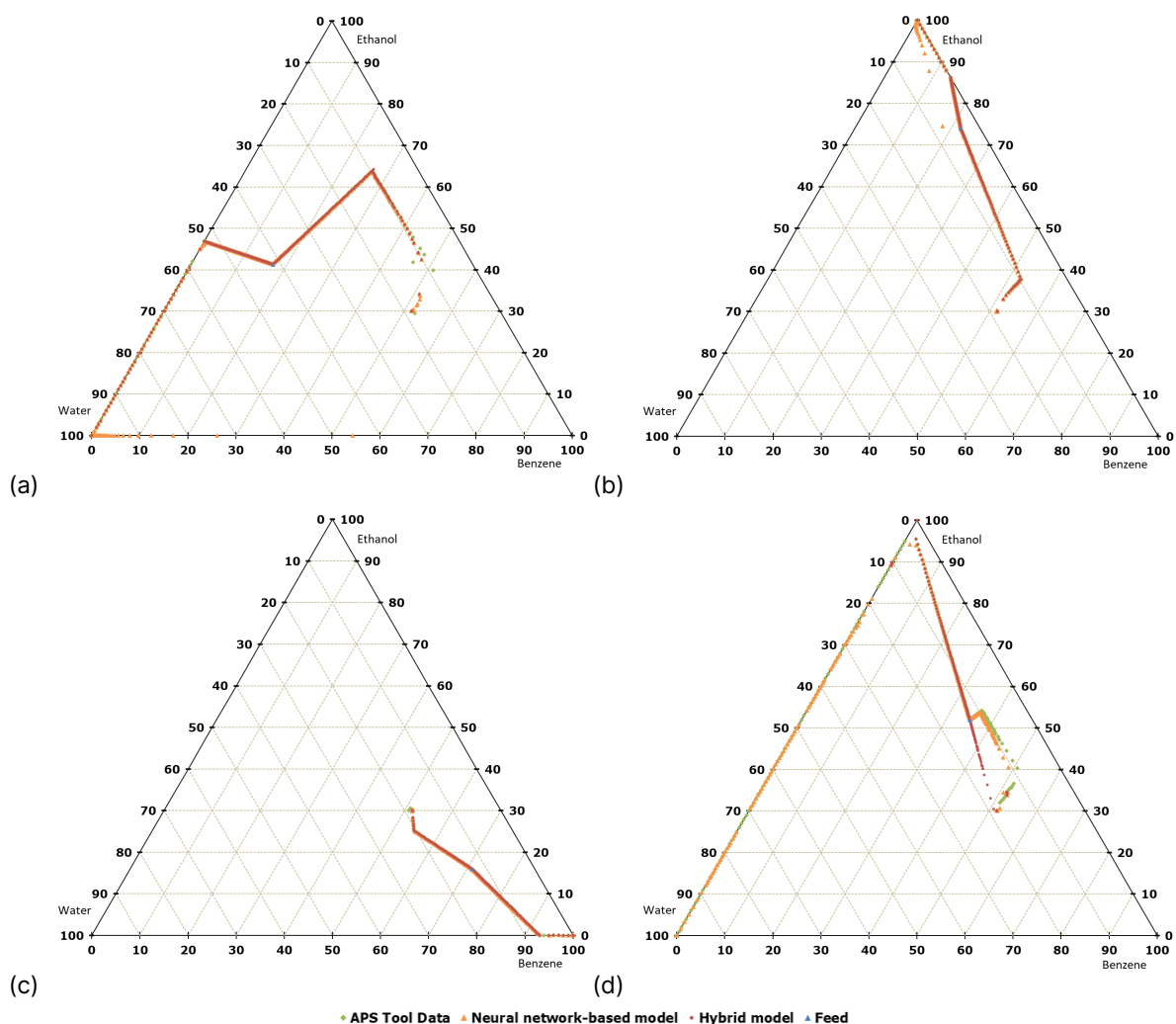


Figure 5. Hybrid surrogate vs. neural network-based product predictions for four feed compositions from different distillation regions

Nonetheless, due to its consistency and accuracy in predicting pure splits and its physically constrained behavior across operating regimes, the proposed hybrid surrogate provides a more reliable overall predictor than a purely neural network-based approach.

CONCLUSION

This work presents a structured, data-driven surrogate modeling framework for predicting distillation behavior under infinite-stage assumptions, operating within distillation limits informed by residue-curve behavior at total reflux and pinch point-based feasibility analysis. The framework provides a direct and deterministic mapping from feed composition and distillate-to-feed ratio to product compositions, making it well-suited for flowsheet synthesis and optimization applications in early conceptual design that require repeated and robust column

evaluations. By decomposing the overall problem into a singular-point split classification task, an operating regime boundary regression, and a neural network-based predictor for impure splits, the approach combines physical interpretability with numerical robustness and computational efficiency.

The results demonstrate that the proposed framework reliably predicts feasible pure splits and operating regimes across the example ternary composition space of ethanol, benzene, and water. In particular, the combination of the classifier and regime regression outperforms a purely neural network-based predictor for pure splits, where exact reproduction of singular-point compositions is essential. The neural network, on the other hand, is well-suited for interpolating product compositions in the middle regime between limiting pure splits, as well as for feed compositions associated with greater uncertainty in the feasible bottom split. This hybrid

structure allows each modeling component to be applied where it is most effective, resulting in improved consistency and physical fidelity.

While the current implementation represents a significant step toward practical surrogate models for nonideal distillation systems, several opportunities for further improvement remain. The accuracy of the regime boundary predictions could be further refined, and the data generation workflow should be streamlined to reduce manual intervention and computational effort. In addition, the present framework is limited to fixed-pressure operation. Extending the model to include pressure as an additional input would require expanded training data and, in the case of azeotropic singular points, prediction of pressure-dependent azeotrope compositions. A longer-term objective is the development of a more general surrogate model that incorporates thermodynamic descriptors as inputs, enabling application to previously unseen chemical systems.

ACKNOWLEDGEMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the Priority Programme “SPP 2331: Machine Learning in Chemical Engineering.”

AUTHOR IDENTIFIERS

Author ORCIDs:

Wolf AB: 0009-0007-0259-4518

Skiborowski M: 0000-0001-9694-963X

Burger J: 0000-0002-9376-9438

REFERENCES

1. Bausa J, Watzdorf RV, Marquardt W. Shortcut methods for nonideal multicomponent distillation: I. simple columns. *AIChE Journal* 44:2181-2198 (2004). <https://doi.org/10.1002/aic.690441008>
2. Sasi T, Skiborowski M. Automatic synthesis of distillation processes for the separation of homogeneous azeotropic multicomponent systems. *Ind. Eng. Chem. Res.* 59:20816-20835 (2020). <https://doi.org/10.1021/acs.iecr.0c04555>
3. Göttl Q, Pirnay J, Burger J, Grimm DG. Deep reinforcement learning enables conceptual design of processes for separating azeotropic mixtures without prior knowledge. *Computers & Chemical Engineering* 194:108975 (2025). <https://doi.org/10.1016/j.compchemeng.2024.108975>
4. Ryll O, Blagov S, Hasse H. ?/?-analysis of homogeneous distillation processes. *Chemical Engineering Science* 84:315-332 (2012).

5. Bekiaris N, Morari M. Multiple steady states in distillation: ?/? predictions, extensions, and implications for design, synthesis, and simulation. *Ind. Eng. Chem. Res.* 35:4264-4280 (1996). <https://doi.org/10.1016/j.ces.2012.08.018>
6. O. Ryll, Thermodynamische Analyse gekoppelter Reaktions-Destillations-Prozesse: konzeptioneller Entwurf, Modellierung, Simulation und experimentelle Validierung, Ph.D. Thesis, University of Stuttgart 2009.
7. Gutiérrez-Antonio C. Multiobjective stochastic optimization of dividing-wall distillation columns using a surrogate model based on neural networks. *Chem. Biochem. Eng. Q.* 29:491-504 (2016). <https://doi.org/10.15255/cabecq.2014.2132>
8. 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. *Chemical Engineering Research and Design* 134:212-225 (2018). <https://doi.org/10.1016/j.cherd.2018.03.006>
9. Wang Z, Zhou T, Sundmacher K. Data?driven integrated design of solvents and extractive distillation processes. *AIChE Journal* 69: (2023). <https://doi.org/10.1002/aic.18236>
10. Quirante N, Javaloyes J, Caballero JA. Rigorous design of distillation columns using surrogate models based on kriging interpolation. *AIChE Journal* 61:2169-2187 (2015). <https://doi.org/10.1002/aic.14798>
11. Keßler T, Kunde C, McBride K, Mertens N, Michaels D, Sundmacher K, Kienle A. Global optimization of distillation columns using explicit and implicit surrogate models. *Chemical Engineering Science* 197:235-245 (2019). <https://doi.org/10.1016/j.ces.2018.12.002>
12. Panofen M, Rolland T, Pela J, Skiborowski M. Surrogate models for distillation boundaries in azeotropic multicomponent mixtures. *Ind. Eng. Chem. Res.* 63:13723-13739 (2024). <https://doi.org/10.1021/acs.iecr.4c00475>
13. Prausnitz, J., Lichtenthaler, R., Azevedo, E. *Molecular Thermodynamics of Fluid-Phase Equilibria*, Prentice Hall PTR (1999).
14. Aspen Plus 8.8, Aspen Technology Inc.

© 2026 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

