Chemical Engineering & Processing: Process Intensification 194 (2023) 109603

Contents lists available at ScienceDirect

Chemical Engineering and Processing - Process Intensification

journal homepage: www.elsevier.com/locate/cep

Advancing neural network architectures for time series forecasting: A sustainable approach to intensified biobutanol production

Esteban A. Hernández-Vargas [a,b,*], Eduardo Sánchez-Ramírez [a], Abraham Rodarte de la Fuente [a], Martha Patricia Calderón-Alvarado [a], Juan Gabriel Segovia-Hernández [a,*]

[a] Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, Mexico
[b] Department of Mathematics and Statistical Science, University of Idaho, Moscow, ID, USA

ARTICLE INFO

Keywords:
Neural networks
Forecasting
Distillation columns
ABE purification

ABSTRACT

Sustainable production aligning with Industry 4.0 and the UN 2030 agenda is crucial for ecological balance and socio-economic progress. Accurate chemical process forecasting is vital in optimizing operations and cost reduction. We introduce a statistical methodology for determining Long Short Term Memory (LSTM) network architectures, focusing on distillation columns. Using Aspen Plus Dynamics, we generate datasets from a sustainable distillation column separating fermentation-derived effluents into acetone, butanol, and ethanol (ABE) for spark-ignition. Input variables (reflux ratio and reboiler heat duty) and output variables (acetone, butanol, ethanol purities) are analyzed. Results show a one-layer neural network effectively predicts ABE concentrations, with ADAM and RSMP as optimal training algorithms and five neurons to prevent overfitting. Linear activation outperforms hyperbolic tangent functions. Reflux ratio and reboiler duty alone sufficiently capture intensified column dynamics, eliminating the need for three features conventionally. This pioneering methodology demonstrates AI's potential in chemical engineering processes and beyond.

1. Introduction

The digitization and automation of chemical processes through the implementation of neural networks represent a significant technological advancement in the quest for sustainability in accordance with the United Nations' 2030 Agenda. These technologies facilitate more efficient resource management, waste reduction, and emissions mitigation, while simultaneously promoting process optimization and data-driven decision-making. In this context, it is imperative to underscore how these innovations are revolutionizing the chemical industry. First and foremost, the ability of neural networks to analyze and process vast volumes of real-time data is pivotal in enhancing efficiency in chemical processes. This not only leads to a more effective use of raw materials and energy but also minimizes wastage and detrimental emissions [1]. The chemical industry is well-known for its environmental footprint, and the implementation of these technologies directly contributes to the attainment of SDG 12 (Responsible Consumption and Production). Furthermore, process optimization via neural networks has a direct impact on the quality of chemical products, resulting in greater efficacy and uniformity in production. This benefits both the industry and consumers by ensuring safer and more reliable chemical products. By achieving operational efficiency, these technologies support SDG 9 (Industry, Innovation, and Infrastructure) by fostering technological innovation and competitiveness. Another critical aspect is safety in the chemical industry. The predictive and proactive problem prevention enabled by neural networks reduces the risk of serious accidents, chemical spills, and hazardous situations for workers. This is essential in compliance with safety standards and workplace well-being, aligning with SDG 8 (Decent Work and Economic Growth) by creating safer and more sustainable work environments. Finally, real-time data collection and analysis allow for more informed decision-making and agile adaptation to changing conditions, resulting in reduced operational costs and a diminished carbon footprint [2]. This, in turn, contributes to SDG 13 (Climate Action) by lowering greenhouse gas emissions and conserving valuable resources, benefiting both the company and the environment. The digitization and automation of chemical processes with neural networks play a crucial role in the chemical industry, in alignment with the UN's 2030 Agenda. By improving efficiency, safety, and sustainability, these technologies drive a significant transformation in the industry, comprehensively addressing economic, environmental, and

* Corresponding authors at: Departamento de Ingeniería Química, Universidad de Guanajuato, Noria Alta s/n, Guanajuato 36050, Mexico.
E-mail addresses: esteban@uidaho.edu (E.A. Hernández-Vargas), gsegovia@ugto.mx (J.G. Segovia-Hernández).

https://doi.org/10.1016/j.cep.2023.109603
Received 6 September 2023; Received in revised form 2 November 2023; Accepted 5 November 2023
Available online 7 November 2023

social challenges, paving the way toward a more sustainable future.

In the context of the chemical industry's digital transformation, it is imperative to highlight the particular relevance of these advancements in the production of biofuels. Biofuels, derived from renewable sources such as biomass and agricultural products, are integral to achieving sustainable energy production and reducing the carbon footprint. Neural networks and digital automation offer a comprehensive solution in optimizing the complex processes involved in biofuel production. By enabling precise control over feedstock conversion, reducing energy consumption, and ensuring consistent product quality, these technologies directly contribute to the advancement of Sustainable Development Goal 7 (Affordable and Clean Energy) and Sustainable Development Goal 15 (Life on Land) by promoting sustainable land use practices for biomass production [3]. Moreover, biofuel production often involves intricate enzymatic and microbial processes, where real-time data analysis and predictive maintenance provided by neural networks ensure process stability, minimizing the risk of contamination and increasing overall efficiency. Therefore, the integration of these technologies in the biofuel sector not only addresses industry-specific challenges but also aligns with broader sustainable development objectives.

Biofuels, such as ethanol and butanol, have been touted as a significant development because of their biomass fermentation origins and low emission levels [4]. Butanol is a feasible option owing to several qualities, including its high energy content, lower volatility than ethanol, high flash point, and compatibility with diesel engines without modifications [5]. ABE fermentation from biomass is one method of obtaining biofuels. However, the procedure results in a ternary combination of acetone, butanol, and ethanol, making it difficult to separate the components with little energy and high temperatures [6]. Intensified processes, which aim for a drastic shift in the unit operations used to meet current needs under the development of a sustainable strategy, have been made possible thanks to the use of new design approaches for the control, modeling, and simulation of chemical processes [7].

The production of biofuels through dividing wall distillation columns is an innovative and promising contribution to the United Nations 2030 agenda. In a world increasingly concerned about the scarcity of natural resources, climate change and the need to develop sustainable energy sources, biofuels have become an essential component to achieve the sustainable development goals established by the international community.

The dividing wall column (DWC) allows for a reduction of energy and capital costs due to the reduction of equipment since the separation of the mixture would be performed in one column and not in two, as in the case of the Petlyuk column, in addition to the fact that shorter piping and electrical currents are required [8]. Dynamic properties are an essential issue in DWC design because of a large amount of energy savings. For example, controllability can be significantly reduced, although it has been shown that control sequences can be better than conventional distillation sequences so that total annual cost savings can be considered as good dynamic performance [9]. Despite all the advantages offered by an intensified process, these are compromised due to the complexity of the control in these processes. Therefore, to achieve these advantages of DWC columns, appropriate control structures must be proposed and applied [10].

Dividing wall distillation columns are a technological advance that allows the efficient separation of bioproducts from biomass, such as vegetable oils and other organic materials, for their subsequent conversion into biofuels. This technology represents a double benefit, since it not only offers a clean and renewable alternative to fossil fuels, but also promotes the responsible use of natural resources and contributes to the mitigation of greenhouse gas emissions. In this context, it is essential to understand how the production of biofuels in dividing wall distillation columns perfectly aligns with the objectives of the United Nations 2030 Agenda, which seeks to address global challenges such as poverty reduction, energy security and protection enviroment. In this introduction, we will explore in detail how this promising technology can contribute to the achievement of the Sustainable Development Goals (SDGs) while opening new opportunities for a more sustainable and equitable future.

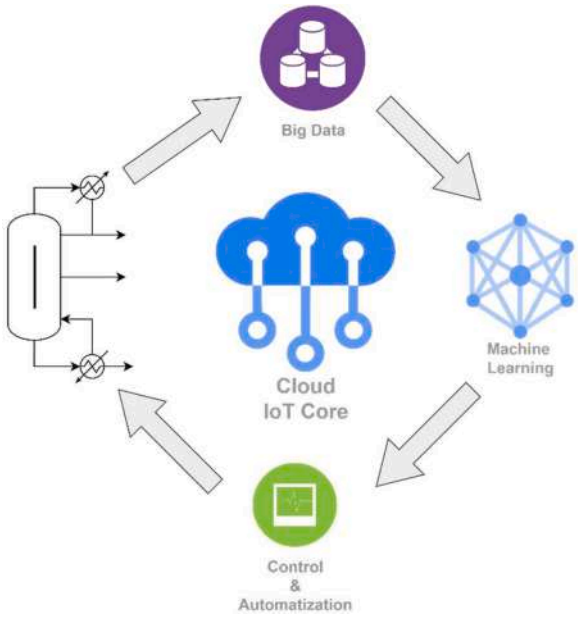

Fig. 1. Driving digital transformation in complex distillation columns.

The integration of Industry 4.0 technologies, coupled with the strategic deployment of machine learning algorithms, assumes a pivotal role in the advancement of the objectives delineated within the UN 2030 Agenda. Industry 4.0, characterized by the confluence of digitalization, automation, and data-centric decision-making, empowers industrial sectors to optimize resource allocation, diminish environmental footprints, and amplify overall operational efficiency. Machine learning, a subset of artificial intelligence, augments this transformative potential by facilitating predictive analytics, pattern recognition, and adaptive optimization across a diverse array of domains, spanning from healthcare and agriculture to energy.

In concert, these technologies foster sustainable development by reinforcing economic expansion, stimulating innovation, and confronting global challenges, harmonizing seamlessly with the UN's mission to eradicate poverty, combat climate change, and ensure equitable and inclusive prosperity for all by 2030. The symbiotic relationship between Industry 4.0 and machine learning furnishes a potent toolkit for policymakers and organizations as they navigate the intricate terrain of sustainable development, paving the way for a brighter and more sustainable future for our planet.

The progress of computational capacities, digitization, and collection of large volumes of data are triggering the implementation of Artificial intelligence (AI) algorithms for knowledge extraction in chemical engineering. Machine learning will be the strategic catalyzer of the chemical industry toward a market that is more decentralized, flexible, and customizable [11], see Fig. 1.

Data of complex distillation columns will be stored in the cloud. Machine learning algorithms can use the collected data to develop predictors controlling the distillation column to foresee potential production problems and keep quality requirements.

Among different machine learning algorithms, Artificial Neural Networks (ANNs) are the most promising computational approach because of their ability to model highly nonlinear systems [12]. For instance, control in distillation columns has been studied by implementing ANNs in conjunction with model predictive control [13]. ANNs were used by [14] to predict new setpoints after disturbances and proved to be a fast and feasible solution for extractive distillation columns.

However, chemical engineers have used ANNs to capture complex dynamics with a limited impact so far [15]. A methodology to fine-tune neural networks in the chemical process engineering community is still fragmented and based on trial and error. In particular, selecting the ANN architecture is a critical obstacle to obtaining accurate predictions of the process, *e.g.,* the number of hidden layers, neurons, features, training algorithms, and activation functions, among others. This pioneering work proposes a methodology to design NN architectures with predictive value. To exemplify our methods, we focus on biofuel production, one of the most relevant processes for society.

The use of neural networks is not new in the field of process system engineering. There have even been some approaches to model and predict the behavior of distillation columns using neural networks. For example, [16] performed the modeling of a binary distillation column for the Toluene-Methylcyclohexane system. For the neural network modeling, he used the NARMAX computational package. On the other hand, [17] performed an optimization process of a binary distillation column (methanol-water and benzene-toluene) using the MATLAB computational package. Likewise, [18] performed an optimization process for a distillation column to putify hydrocarbons using the MATLAB computational package.

On the other hand, both [13,20] performed a control study approach both using the NARX computational package. Y Shin et al. [13] presented their work aimed at a purification process of a hydrocarbon system and Taqvi et al. [20] for an Ethanol-Water binary mixture.

It is evident that neural networks have been used to model, optimize or control processes whose model is highly nonlinear. However, it is important to emphasize that in all the cases presented here, black box type computational packages have been used, and they have been oriented to address the purification of relatively simple mixtures. The use of software packages with preloaded neural network libraries presents an important practicality; however, their use does not allow modifying, knowing and/or modifying the modeling of the neural networks, which has a direct impact on their performance. In that sense, this work uses a free and open-source software that allows a direct interaction between the user and the code in order to generate a better performance in the use of this tool, as well as an adequate design of the neural network.

The innovation of digitizing and automating through neural networks in a biofuel purification separation process brings a significant and novel contribution to the field of industry and sustainable energy production. This technology enables dynamic open-loop responses, signifying real-time adaptability to changing conditions, including variations in feedstock composition, temperature, pressure, and other environmental variables. The capacity of these neural networks to anticipate and automatically adjust separation parameters based on real-time data is revolutionary. This results in a more efficient purification of biofuels, ensuring compliance with required quality and sustainability standards [21]. Furthermore, by minimizing waste and reducing resource consumption, this technology aligns directly with the United Nations' Sustainable Development Goals (SDGs), particularly SDG 7 (Affordable and Clean Energy) and SDG 12 (Responsible Consumption and Production). In the current context, where the transition to cleaner and renewable energy sources is essential to combat climate change, efficient biofuel purification plays a pivotal role. Integrating digitization and automation with neural networks into this process promotes a more sustainable and competitive biofuel production, reducing reliance on fossil fuels and fostering a low-carbon economy. In summary, the application of neural networks in biofuel purification separation processes represents an innovation that not only enhances operational efficiency but also significantly contributes to global efforts towards a more sustainable future with reduced carbon emissions. This novel contribution has the potential to transform the energy industry and support sustainability goals on a global scale. Moreover, it is crucial to underscore that the reliability of these neural networks hinges on the careful design of their architecture, including the selection of the activation function, the number of layers, and the quantity of neurons in each layer. The choice of the activation function plays a pivotal role in shaping how information is processed within each neuron, influencing the network's ability to capture complex relationships in the data. Similarly, the number of layers and neurons within the network directly affects its capacity to learn and generalize patterns, making it vital to strike the right balance between model complexity and overfitting. In the context of biofuel purification, the implications of this architectural design are profound. It enables the network to adapt and optimize its performance for the specific challenges and nuances of biofuel separation processes. This adaptability, combined with the power of deep learning, allows for improved real-time decision-making and process control. As a result, the energy industry stands to benefit from enhanced

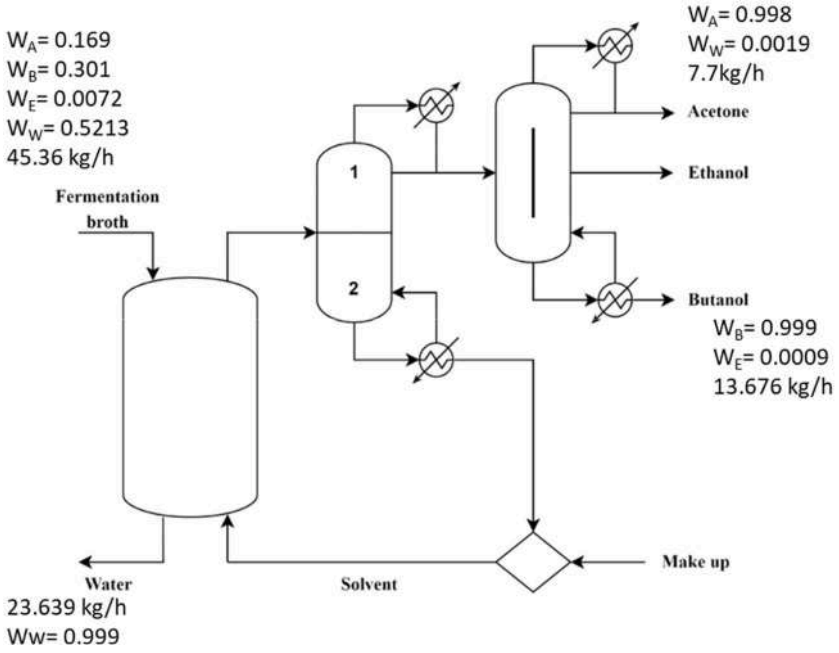

Fig. 2. Case study of DWC.

Table 1
Design parameters for the sequence considered as a case study [22].

	Fig 2(a)			Fig 2(b)		
	Extractor	C_1	DWC	Extractor	DWC	C_3
Total number of stages	5	23	48	5	49	31
Number of stages across the wall	—	—	18	—	8	—
Feed location	—	12	31	—	21	16
Side stream location	—	—	33	—	20	—
Reflux ratio	—	0.894	20.314	—	39.053	8.284
Distillate flowrate [kg h^{-1}]	—	21.685	7.697	—	8.035	7.705
Side stream flowrate [kg h^{-1}]	—		0.327	—	13.663	—
Liquid split flowrate [kg h^{-1}]	—	—	2.183	—	39.295	—
Vapor split flowrate [kg h^{-1}]	—		9.821	—	481.178	—
Extract flowrate [kg h^{-1}]	733.873	—	—	733.873	—	—
Solvent flowrate [kg h^{-1}]	712.147	–	–	712.147	—	—
Diameter [m]	0.335	0.288	0.302	0.335	0.413	0.287
Pressure [kPa]	101.3	101.3	101.3	101.3	101.3	101.3
Condenser duty [kW]	—	7.241	23.452	—	47.00	10.216
Reboiler duty [kW]	—	66.218	23.818	—	106.357	10.211
TAC [k$ yr^{-1}]	111.86			122.99		
EI99 [kpoints yr^{-1}]	17.50			19.50		
CN	10.35			10183.72		

yield and resource efficiency, reduced waste, and ultimately, a smaller environmental footprint. Additionally, the lessons learned from the biofuel industry can be extrapolated to other sectors, further promoting the application of neural networks in sustainable practices across various domains. By harnessing the potential of neural networks while ensuring their architecture is carefully tailored to the task at hand, we are not only advancing the field of artificial intelligence but also contributing to a greener and more sustainable future on a global scale.

2. Methodology to determine neural network architectures

Here, we aim to forecast the dynamics of complex distillation columns using neural networks. Before establishing the methodology, it is essential to understand some general concepts and then develop the computational approach. In this sense, we presented a revision concept in Supplementary Material.

The main objective of this study is to propose a methodology based on statistical foundations that optimizes the ANN architecture to represent the dynamic behavior of complex distillation columns by analyzing different aspects of ANNS such as the activation function, training algorithms, number of hidden layers as well as number of neurons.

STEP 1. DATA COLLECTION

To underscore the practical viability and demonstrable efficacy inherent in our advanced methodology, we have harnessed the comprehensive capabilities of the Aspen Dynamics software platform. This enabled us to intricately instantiate a sophisticated modeling framework tailored to the intricate dynamics of a divided-wall distillation column (DWC), a conceptual schema elucidated in the meticulous study conducted by [22]. Through the judicious assimilation of real-time operational data, meticulously curating salient variables, and pertinent features of the system, we have painstakingly assembled a corpus for exhaustive scrutiny and analysis. Of particular focus are the variables imbued with paramount significance: the mass fractions of Acetone, Ethanol, and Butanol – pivotal constituents that collectively govern the intricacies of the separation process.

Our research endeavors have been markedly concentrated on an elevated echelon of separation mechanism, epitomized by the intrinsic role that the distillation columns, themselves tantamount to partitioning barriers, play in sculpting the divided-wall distillation column configuration (DWC). This innovative configuration seamlessly affords the astute segregation of a ternary amalgam renowned as ABE (Acetone-Butanol-Ethanol). The quintessential crux of our approach resides in our adroit manipulation of cardinal determinants encompassing the reflux ratio, reboiler duty, and side flowrate. This orchestration grants us dominion over the trajectory of the separation, culminating in the distinct extraction of Acetone from the summit, Ethanol siphoned through the side conduit, and Butanol judiciously harvested from the nether regions of the column, a visual portrayal meticulously etched in Fig. 2 and more parameters are presented in Table 1.

Within the ambit of the liquid-liquid extraction column, the introductory feed stream manifests as a composite fusion comprising the constituent elements of acetone, butanol, ethanol, and water, their respective volumetric ratios standing as a testament to their relative prevalence: 0.3018, 0.1695, 0.0073, and 0.5214 wt%. To effectuate the targeted separation of water, a task of paramount import, the strategic deployment of n-hexyl acetate transpires as the solvent of choice, its efficacious interaction within the liquid-liquid extraction matrix serving as the crux of our methodology. In due course, water recuperation assumes prominence through the established conventions of a customary distillation column.

The salient variables, often denominated manipulable parameters or features within the expansive sphere of machine learning, encompass the trinity of the reflux ratio, side flowrate, and reboiler duty. This triumvirate of modifiable parameters, once finessed to articulate the zenith of design configuration, precipitated the ensuing phase of open-loop dynamic simulations, deftly realized through the robust framework of Aspen Dynamics. Our quest for comprehensiveness and resilience prompted us to subject these manipulable variables to an array of external influences, spanning both affirmative and negative perturbations, and spanning the spectrum from 1% to 5%, doing so on a continuous basis. Initially, positive perturbations were made up to 5%, then negative perturbations to reach 0% again and then -0.5%, and again positive perturbations to reach 5%. This envelope of perturbations, meticulously sourced from the well-regarded insights imparted by [23], meticulously aligns with the canonical thresholds encapsulating the characteristic perturbations endemic to industrial processes.

The intricate choreography of data acquisition bore semblance to a symphony orchestrated at five-hour intervals, spanning an aggregate timeline of 100 hours. This methodical cadence lay at the heart of our endeavor to capture the nuances of dynamic responses embedded within the Acetone-Butanol-Ethanol composition under the duress of orchestrated perturbations. The culmination of these meticulous data-gathering efforts is vividly encapsulated in the graphical rendition furnished in Fig. 3.

Remark. To train a neural network, the signals interacting with the parameter estimations should not decay too rapidly. "*Persistent excitation*" is a concept of identification systems [24] that can be understood as constant network weight stimulation to effectively estimate model

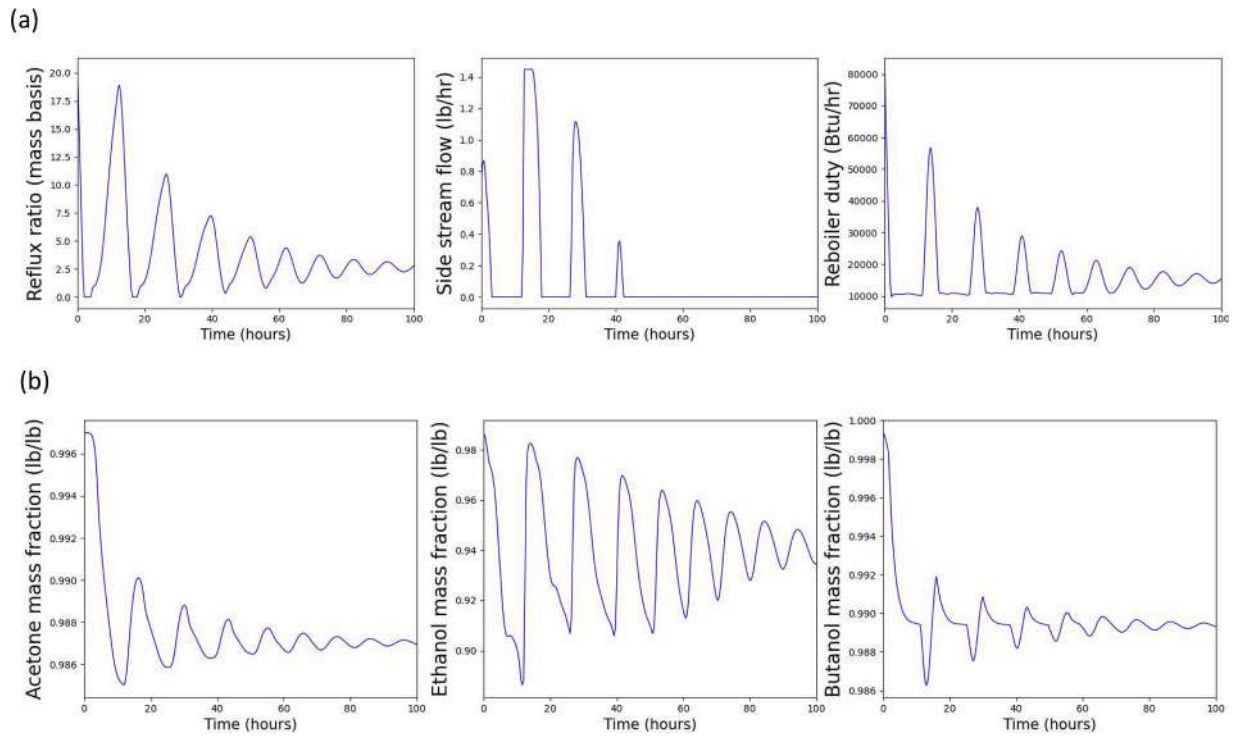

Fig. 3. Collected data sets. Panel (a) shows the manipulated variables. Panel (b) shows the target variables.

parameters either in a mechanistic or black box model [25]. To this end, the collected data was obtained in a closed loop, using a PI controller in the composition of each mixture component to be separated (Acetone-Butanol-Ethanol). Using this approach, three setpoint changes were performed and tuned simultaneously, producing time-varying operational data sets. ■

STEP 2. DATA VISUALIZATION

Before using any machine learning tool, a good data analysis practice is visualizing the data and finding potential pitfalls. For instance, if we want to predict the target y using three features x_1, x_2, and x_3, but during visualization, we find these three features are highly correlated. Therefore, it will be a red flag before providing the three features to our neural network. To this end, it is recommended to use a Correlation Matrix or Scatter Plots, which will uncover potential correlations between all the different possible pairings of data. It is a potent tool for summing up enormous datasets and finding and visualizing patterns within the data provided.

In our case study, the data set obtained in ASPEN Dynamics can be visualized by the Seaborn library of Python. We can observe in Fig. 4a that some of the manipulatable variables, we call them features, have more impact on the target variable, the butanol composition profile. In this case, the reflux ratio and the reboiler duty are the ones that have similar behavior when related to the compound since the lateral output creates a more complex behavior. Fig. 4b shows how the two target variables, butanol, and acetone, behave significantly similarly to the reflux ratio and reboiler. Ethanol is the most distant dynamic of the target variables, strongly correlated with the side stream flow.

STEP 3. STATISTICAL CRITERIA TO DETERMINE NEURAL NETWORK ARCHITECTURES

In previous work [26,27,13], neural network model architectures are designed by minimizing the MSE (Mean Squared Error) concerning the variation of the number of hidden neurons. This criterion is calculated by Eq. (1) using the total number of data n, the value obtained from the simulation y_i and the predicted value $\overline{y_i}$

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \overline{y_i})^2}{n} \tag{1}$$

where n is the number of total data, y_i is the actual value, and $\overline{y_i}$ is the predicted value. The *MSE* is the most used mean error to estimate the quality of a quantitative prediction model, which is applied to the data set on which the quality of the quantitative model is being calculated. Focusing only on minimizing MSE would result in overfitting as the neural network complexity would increase to minimize the criteria. To avoid this problem, we propose here, for a given collection of models for a data set, the AIC (Akaike's information criterion) that reckons the quality of each model concerning other proposed models [28]. Thus, the AIC provides a means for model selection by addressing the trade-off between model goodness of fit and complexity. AIC does not offer a test of a model in the sense of testing a null hypothesis, so it cannot say anything about model quality in an absolute sense. If all candidate models fit poorly, the AIC will not warn about it. A lower AIC value means a given model describes the data better than other models with higher values. The AIC is calculated by Eq. (2) using the number of data points N, the number of unknown parameters M, in our case, the number of neural network parameters, and the mean square error (MSE).

$$AIC = Nlog(MSE) + \frac{2MN}{N - M - 1} \tag{2}$$

STEP 4. DATA NORMALIZATION

Data normalization is an important step in training a neural network and was performed by employing Eq. (3) to transform all feature values into an interval of $[0, 1]$, mathematically speaking y_i refers to the data values, $\min(y_i)$ is the minimum value and $\max(y_i)$ refers to the maximum value. Increase the efficiency of the algorithm by reducing fluctuations.

$$y_{scaled} = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)} \tag{3}$$

STEP 5. MODEL TRAINING

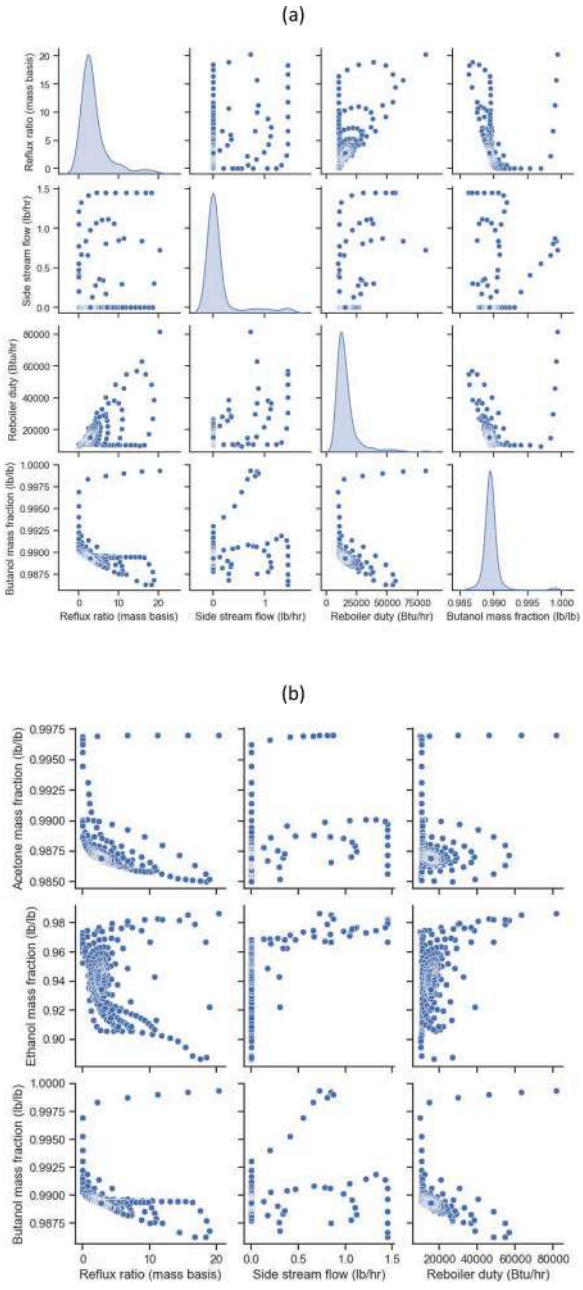

Fig. 4. Data visualization. Panel (a) shows the scatter plots for the key target variable (Butanol) and the three features of interest. Panel (b) shows the three target variables and the three features.

The data obtained from Aspen Dynamics are normalized and fed to a single hidden layer neural network to determine the number of epochs. The data were subdivided into 76% for training and 24% for testing the systems analyzed. Fig. 5 shows that as the number of epochs increases, the MSE value decreases to a minimum value.

STEP 6. EXPLORING PARAMETERS AND ARCHITECTURES OF NEURAL NETWORKS

Based on the AIC, we will determine a minimalistic configuration that predicts the data, see Fig. 6. To this end, we explore different features and hyperparameters, such as the number of layers, hidden neurons of a single hidden layer, type of activation function, and training algorithms.

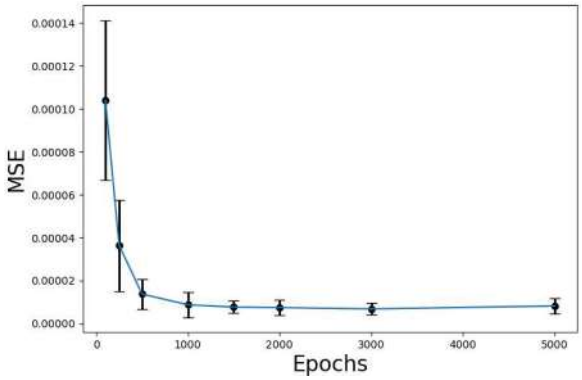

Fig. 5. Training process.

4. Case study in biobutanol purification

Obtaining biobutanol can occur through the ABE fermentation process from biomass such as lignocellulose. In this process, a mixture of acetone, ethanol, and butanol is obtained, which creates a challenge when trying to reduce the energy needed in the separation process. When used as fuel (in the form of blends), this compound is called biobutanol, emphasizing its vegetable origin since butanol can also be produced from fossil fuels with the same chemical properties. The methodologies proposed by [22] aim to generate intensified alternatives for separating butanol from fermentation (ABE) considering the mixed design of liquid-liquid extraction columns and conventional distillation columns.

In a conventional column, there are respective condensers and reboilers so that in the first column, the acetone is obtained as the product of the dome, and at the bottom, a mixture of ethanol and butanol are separated in the second column, separating ethanol as the product of the dome and butanol at the bottom. An alternative to conventional columns is the divided wall columns (DWC). For an intensified system where the distillation columns are divided wall (DWC), separating a ternary mixture (ABE) is possible by manipulating the reflux ratio, reboiler duty, and lateral outflow. Acetone is obtained from the dome, ethanol from the side outlet, and butanol from the bottom.

The data referring to the ethanol and butanol mass fraction, reflux ratio, and reboiler duty were obtained in Aspen Dynamics. The process data are obtained, considering NRTL-HOC as the thermodynamic model describing the phase equilibrium. The simulations are performed in a closed test loop where a setpoint was implemented in the composition of each mixture component to be separated. In this way, three setpoint changes were performed and tuned simultaneously, thus producing time-varying operational data sets. The feed stream to the liquid-liquid extraction column considers a mixture of acetone, butanol, ethanol, and water with the following ratios of 0.3018, 0.1695, 0.0073, and 0.5214 wt%, respectively. The solvent used in the liquid-liquid extraction is n-hexyl acetate, and is used to separate the water, which is finally recovered in a conventional distillation column.

Consequently, the data obtained from the column simulations in Aspen Dynamics were used to feed the ANN. The number of neurons for each neural network configuration implemented to model the system does not exceed 15 neurons, as chosen by the AIC criterion. In the case of the ADAM algorithm, the Tanh function needs 10 neurons, the Linear 5 neurons, and the ReLU function 15 neurons. With the RMSPROP optimizer, the Tanh function needs 15 neurons, and the Linear and ReLU functions need 5 neurons. Finally, to implement the SGD optimizer, all the activation functions need 5 neurons to model the conventional system.

With the data from the conventional dynamic state separation process, a comparison of three activation functions, Linear, ReLU, and Tanh, was performed using ADAM, RMSPROP, and SGD algorithms (see

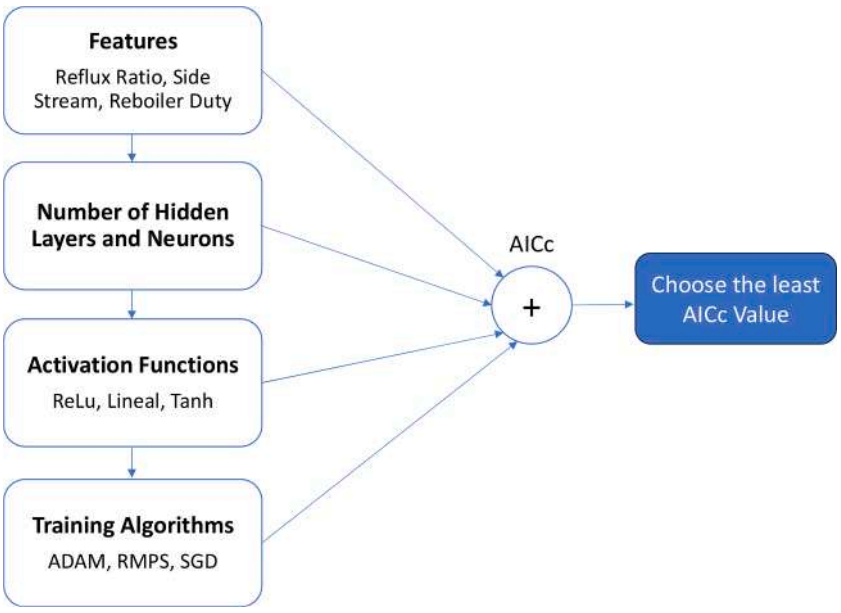

Fig. 6. Methodology for determining a neural network architecture.

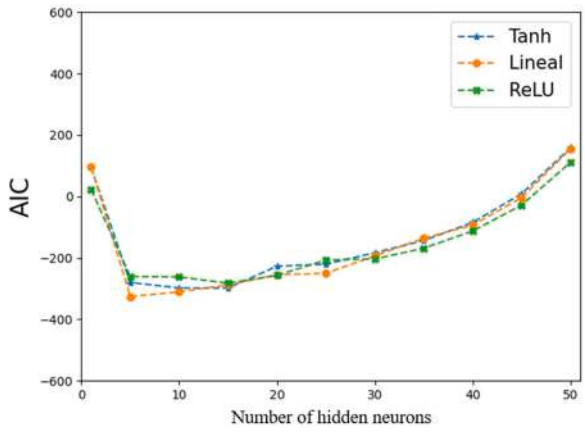

Fig. 7. Comparison of different activation functions with the ADAM algorithm.

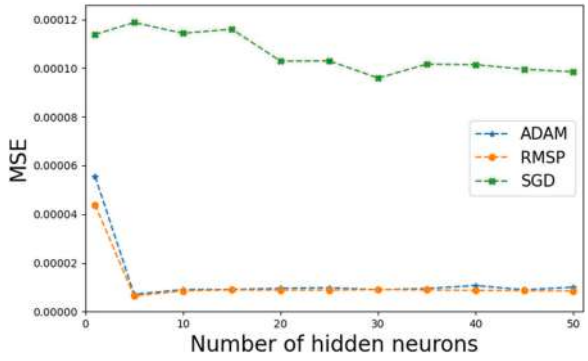

Fig. 8. MSE value comparison for different algorithms.

Fig. 7). Results were obtained by running the ANN simulation 100 times. According to this behavior, it is also concluded that the models with more than 5 hidden neurons do not show an appreciable increase in the reduction of the MSE. Therefore, it is possible to work with 5 neurons to save computational time. ReLU activation function showed the highest MSE value concerning the 5 neurons compared to the other two

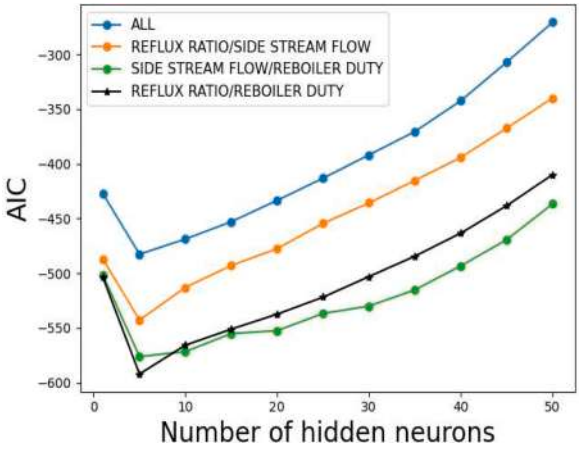

Fig. 9. Effects of the input manipulable variables on the AIC.

functions employed. However, with this methodology, it is challenging to obtain a conclusive result about the lowest number of hidden neurons needed for a good prediction. Therefore, an AIC indicator analysis is performed.

The learning algorithms ADAM, RMSP, and SGD were used to compare the impact of loss function minimization, and by changing the ANN optimizer, it is observed that ADAM and RMSP practically merge in the graph shown in Fig. 8. This could be because the ADAM optimization algorithm is based on RMSP. The SGD optimizer is the one that presents the highest error values and moves away from the behavior obtained with the other algorithms.

Next, we worked with the features (manipulable variables) such as reflux ratio, lateral outflow, and reboiler duty. In the linear activation function and the RMSP algorithm, an analysis of the impact of these manipulatable variables was performed in Fig. 9. According to these results, it can be observed that the worst behavior was the one that omitted the reboiler duty as data feed, and the ones that showed the best results were the ones that fed the three manipulatable variables and the ones that overlooked the sidestream flow feed and the reflux ratio. The one that showed the best result for 5 neurons was the one that fed the reflux ratio and the reboiler duty.

In this context, it becomes evident that the influential factors

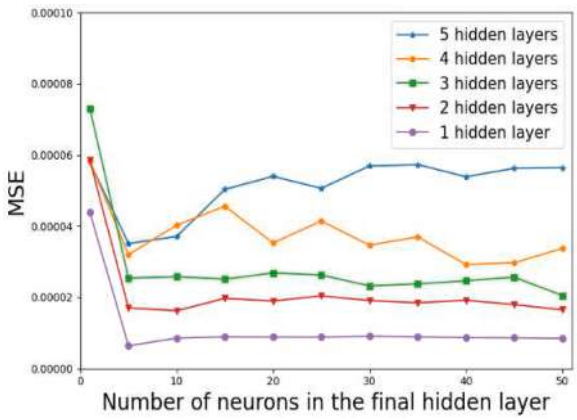

Fig. 10. Comparison of the number of hidden layers.

impacting the process dynamics possess distinct significance levels. Specifically, the reboiler duty and reflux ratio emerge as pivotal design variables, whereas the influence of the lateral flowrate appears comparatively subdued within the confines of this case study. This behavior can be attributed to the relatively low magnitude of flow obtained through the lateral outlet. Consequently, both the reboiler duty and reflux ratio exhibit the capacity to mitigate disturbances that might otherwise affect the lateral outlet, thereby underscoring their pronounced effects. These findings align with the outcomes elucidated by [29], which assert that in scenarios where the intermediate component's quantity is limited within a Dividing Wall Column (DWC), the interplay involving the lateral flowrate variable remains subdued. Instead, the preeminent dynamics are driven by the reboiler duty and reflux ratio.

The observed phenomenon concerning these two variables can also be attributed to the inherent system topology. As documented in the research undertaken by Errico et al. [22], the investigated process system epitomizes an intensified configuration, wherein the internal recirculation of steam and liquid within the confines of the dividing wall column attenuates disturbances within the process scheme. Additionally, the nature of a dividing wall configuration inherently entails a larger diameter than a conventionally designed counterpart tasked with the same function. This larger diameter operationally augments the process's controllability.

An error analysis was performed with the RMSP algorithm, from 1 to 50 neurons, a reasonable prediction using 5 neurons in a single layer, as shown in Fig. 10. Therefore, using a single layer with minimal error is sufficient for prediction; if the number of layers increases, the error also increases. This behavior is associated with the increase in the number of layers, the complexity of the ANN increases, and with it, the error as well. The best architecture is shown in Fig. 11. The results obtained by ANN show a good prediction of the independent validation data simulated by Aspen Dynamics for the mass fraction of Acetone, Ethanol, and Butanol, as is shown in Fig. 12.

5. Conclusions

In this work, a statistical methodology was proposed to find a minimalistic architecture of the ANN. We consider a sustainable and intensified chemical process, such as the dividing wall column. Numerical results show that a good process prediction can be obtained by only using two features: the reflux ratio and the reboiler duty. In contrast to previous literature, we show that only one layer with 5 neurons is enough to predict complex distillation columns. This highlights the importance of our methodology to avoid overfitting. This methodology will impact future work in chemical engineering by measuring only two features and the three perturbed output variables of the DWC system.

The harmonious fusion of machine learning techniques with advanced process intensification methodologies, exemplified by the divided wall column, has delivered a systematic and efficient framework for the generation of neural network architectures. This pioneering methodology not only heightens the precision and efficacy of neural network design but also makes a substantial contribution to the realization of the ambitious objectives outlined in the UN 2030 Agenda.

By harnessing the potential of machine learning, we empower industries and organizations to optimize their operations, curtail environmental impacts, and grapple with global challenges, in perfect alignment with the agenda's overarching mission to promote sustainable economic growth, combat climate change, and ensure inclusivity and prosperity for all. As we look ahead, the integration of these methodologies across diverse sectors holds the promise of accelerating progress towards the UN's sustainable development goals, paving the way for a more equitable and sustainable future for our planet.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

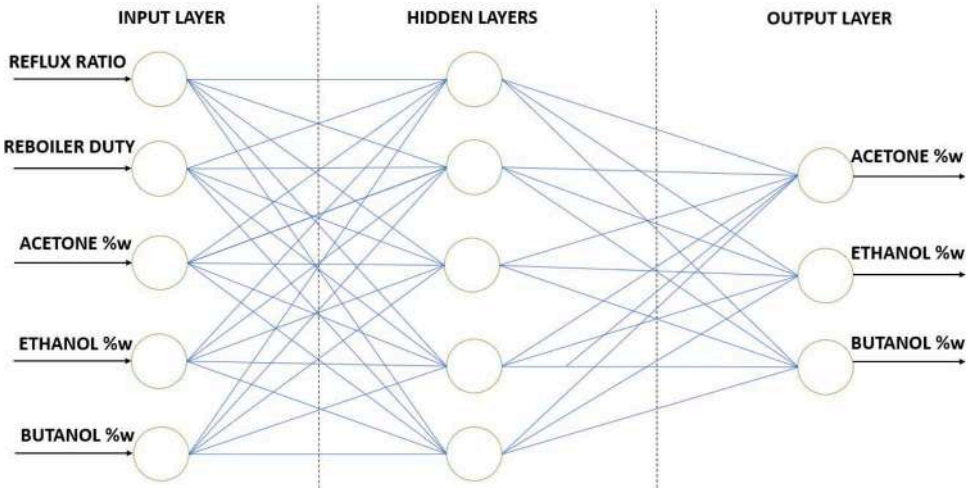

Fig. 11. Best ANN architecture.

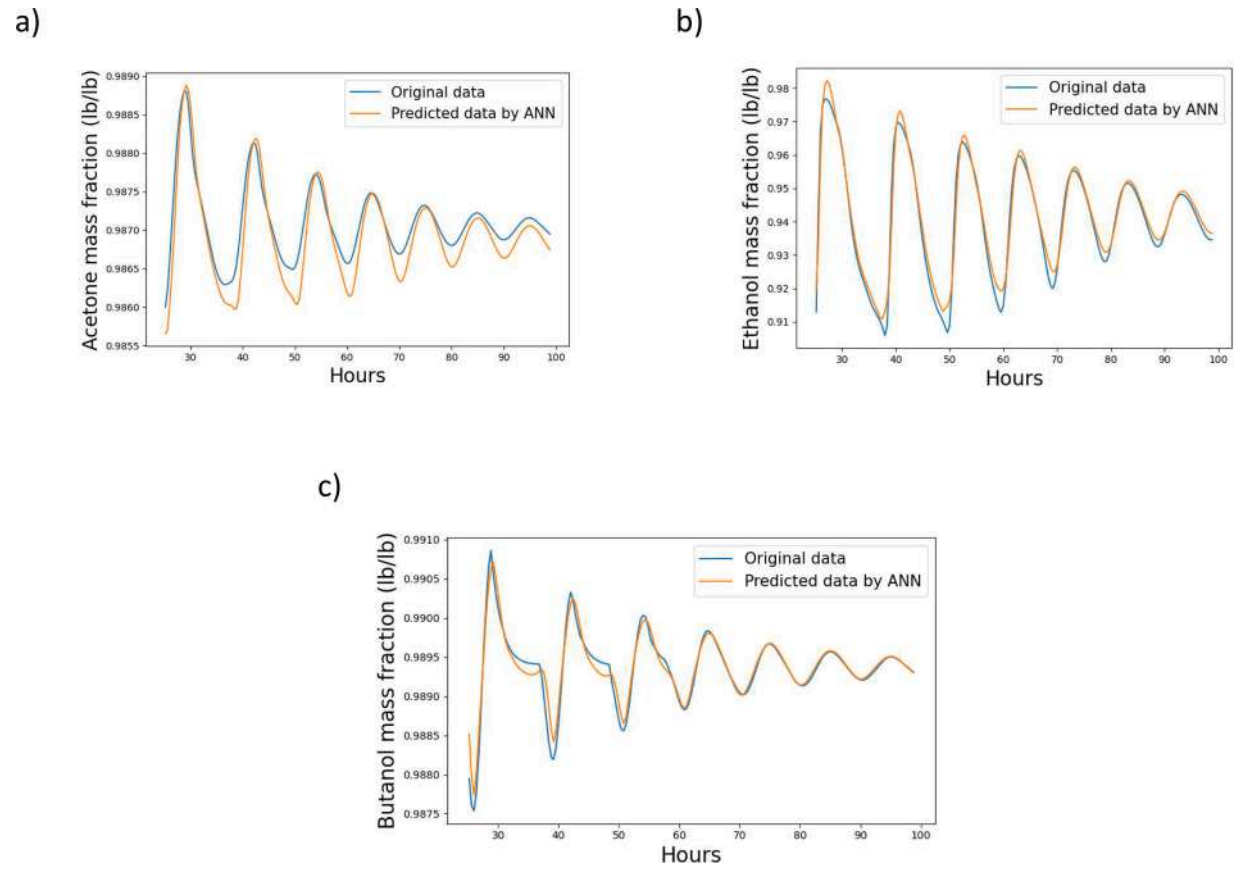

Fig. 12. Prediction of the chosen ANN architecture. Comparison between the predicted and actual (a) acetone, (b) ethanol, and (c) butanol mass fraction.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cep.2023.109603.

References

[1] F. Biermann, N. Kanie, R.E. Kim, Global governance by goal-setting: the novel approach of the UN sustainable development goals, Curr. Opin. Environ. Sustain. 26–27 (2017) 26–31, https://doi.org/10.1016/J.COSUST.2017.01.010.

[2] Y. Lu, N. Nakicenovic, M. Visbeck, A.S. Stevance, Policy: five priorities for the UN sustainable development goals, Nature 520 (7548) (2015) 432–433, https://doi.org/10.1038/520432a. *2015 520:7548*.

[3] O.E. Ogunmakinde, T. Egbelakin, W. Sher, Contributions of the circular economy to the UN sustainable development goals through sustainable construction, Resour. Conserv. Recycl. 178 (2022), 106023, https://doi.org/10.1016/J.RESCONREC.2021.106023.

[4] N. Ribeiro, A.C. Pinto, C.M. Quintella, G.O. da Rocha, L.S.G. Teixeira, L.L. N. Guarieiro, M.do C. Rangel, M.C.C. Veloso, M.J.C. Rezende, R.S. da Cruz, A.M. de Oliveira, E.A. Torres, J.B de Andrade, The role of additives for diesel and diesel blended (ethanol or biodiesel) fuels: a review, Energy Fuels 21 (4) (2007) 2433–2445, https://doi.org/10.1021/EF070060R/ASSET/IMAGES/LARGE/EF070060RF00003.JPEG.

[5] S. Maiti, G. Gallastegui, S.J. Sarma, S.K. Brar, Y. le Bihan, P. Drogui, G. Buelna, M. Verma, A re-look at the biochemical strategies to enhance butanol production, Biomass Bioenergy 94 (2016) 187–200, https://doi.org/10.1016/J.BIOMBIOE.2016.09.001.

[6] R. González-Bravo, E. Sánchez-Ramírez, J.J. Quiroz-Ramírez, J.G. Segovia-Hernández, L.F. Lira-Barragán, J.M. Ponce-Ortega, Total heat integration in the biobutanol separation process, Ind. Eng. Chem. Res. 55 (11) (2016) 3000–3012, https://doi.org/10.1021/ACS.IECR.5B03168/ASSET/IMAGES/LARGE/IE-2015-031689_0019.JPEG.

[7] E. Sánchez-Ramírez, J.J. Quiroz-Ramírez, S. Hernández, J.G. Segovia-Hernández, A.A. Kiss, Optimal hybrid separations for intensified downstream processing of biobutanol, Sep. Purif. Technol. 185 (2017) 149–159, https://doi.org/10.1016/J.SEPPUR.2017.05.011.

[8] F.I. Gómez-Castro, J.G. Segovia-Hernández, S. Hernández, C. Gutiérrez-Antonio, A. Briones-Ramírez, Dividing wall distillation columns: optimization and control properties, Chem. Eng. Technol. 31 (9) (2008) 1246–1260, https://doi.org/10.1002/CEAT.200800116.

[9] R.R. Rewagad, A.A. Kiss, Dynamic optimization of a dividing-wall column using model predictive control, Chem. Eng. Sci. 68 (1) (2012) 132–142, https://doi.org/10.1016/J.CES.2011.09.022.

[10] S. Tututi-Avila, A. Jiménez-Gutiérrez, J. Hahn, Analysis of multi-loop control structures of dividing-wall distillation columns using a fundamental model, Processes 2 (1) (2014) 180–199, https://doi.org/10.3390/PR2010180. *2014*, Vol. 2, Pages *180-199*.

[11] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L.H. Chiang, M.S. Reis, Recent trends on hybrid modeling for Industry 4.0, Comput. Chem. Eng. 151 (2021), 107365, https://doi.org/10.1016/J.COMPCHEMENG.2021.107365.

[12] A. Carranza-Abaid, J.P. Jakobsen, Neural network programming: integrating first principles into machine learning models, Comput. Chem. Eng. 163 (2022), 107858, https://doi.org/10.1016/J.COMPCHEMENG.2022.107858.

[13] Y. Shin, R. Smith, S. Hwang, Development of model predictive control system using an artificial neural network: a case study with a distillation column, J. Cleaner Prod. 277 (2020), 124124, https://doi.org/10.1016/J.JCLEPRO.2020.124124.

[14] T.G. das Neves, W.B. Ramos, G.W. de Farias Neto, R.P. Brito, Intelligent control system for extractive distillation columns, Korean J. Chem. Eng. 35 (4) (2018) 826–834, https://doi.org/10.1007/S11814-017-0346-0/METRICS.

[15] V. Venkatasubramanian, The promise of artificial intelligence in chemical engineering: is it here, finally? AlChE J. 65 (2) (2019) 466–478, https://doi.org/10.1002/aic.16489.

[16] Y. Chetouani, Using Artificial Neural networks for the modelling of a distillation column, Int. J. Comput. Sci. Appl. © 4 (2007) 119–133.

[17] F.N. Osuolale, J. Zhang, Energy efficiency optimisation for distillation column using artificial neural network models, Energy 106 (2016) 562–578, https://doi.org/10.1016/J.ENERGY.2016.03.051.

[18] S. Motlaghi, F. Jalali, M.N. Ahmadabadi, 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 (4) (2008) 1540–1545, https://doi.org/10.1016/J.ESWA.2007.08.105.

[19] S.A. Taqvi, • Lemma, D. Tufa, H. Zabiri, • Abdulhalim, S. Maulud, • F. Uddin, Fault detection in distillation column using NARX neural network, Neural. Comput. Appl. 32 (2020) 3503–3519, https://doi.org/10.1007/s00521-018-3658-z.

[20] S. Sparviero, M. Ragnedda, Towards digital sustainability: the long journey to the sustainable development goals 2030, Digit. Policy Regul. Govern. 23 (3) (2021) 216–228, https://doi.org/10.1108/DPRG-01-2021-0015/FULL/XML.

[21] M. Errico, E. Sanchez-Ramirez, J.J. Quiroz-Ramìrez, B.G. Rong, J.G. Segovia-Hernandez, Multiobjective optimal acetone-butanol-ethanol separation systems using liquid-liquid extraction-assisted divided wall columns, Ind. Eng. Chem. Res. 56 (40) (2017) 11575–11583, https://doi.org/10.1021/ACS.IECR.7B03078/SUPPL_FILE/IE7B03078_SI_001.PDF.

[22] W.L. Luyben, Practical distillation control. IEE Colloquium (Digest), 1st ed., Van Nostrand Reinhold, 1992 https://doi.org/10.1049/ic:20000114. Issue 23.
[23] Narendra, K. S., & Annaswamy, A. M. (2007). Persistent excitation in adaptive systems. 10.1080/00207178708933715, 45(1), 127–160. doi:10.1080/00207178708933715.
[24] Nar, K., & Sastry, S. S. (2019). Persistency of excitation for robustness of neural networks. *ArXiv*. http://arxiv.org/abs/1911.01043.
[25] A.B. Bulsari, S. Palosaafi, Neural computing & applications application of neural networks for system identification of an adsorption column, Neural Comput. Appl. 1 (1993) 160–165.
[26] T.R. Biyanto, B.L. Widjiantoro, A. Abo Jabal, T. Budiati, Artificial neural network based modeling and controlling of distillation column system, Int. J. Eng. Sci. Technol. 2 (6) (2010). www.ijest-ng.com.
[27] J.E. Cavanaugh, A.A. Neath, The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements, Wiley Interdiscip. Rev. Comput. Stat. 11 (3) (2019), https://doi.org/10.1002/wics.1460.
[28] J.G. Segovia-Hernández, S. Hernández, V. Rico-Ramírez, A. Jiménez, A comparison of the feedback control behavior between thermally coupled and conventional distillation schemes, Comput. Chem. Eng. 28 (5) (2004) 811–819, https://doi.org/10.1016/J.COMPCHEMENG.2004.02.019.