processes

MDPI

Article

Intelligent Optimization Design of Distillation Columns Using Surrogate Models Based on GA-BP

Lixiao Ye [1], Nan Zhang [2,*], Guanghui Li [1,*], Dungang Gu [1], Jiaqi Lu [1] and Yuhang Lou [1]

[1] Innovation Centre for Environment and Resources, School of Chemistry and Chemical Engineering, Shanghai University of Engineering Science, No.333 Longteng Road, Songjiang District, Shanghai 201620, China; lixiaoye9526@163.com (L.Y.); dgu@sues.edu.cn (D.G.); wilsherelu@foxmail.com (J.L.); yuhang.lou@canalis-tech.com (Y.L.)

[2] Centre for Process Integration, Department of Chemical Engineering and Analytical Science, The University of Manchester, Manchester M13 9PL, UK

* Correspondence: nan.zhang@manchester.ac.uk (N.Z.); ghli@sues.edu.cn (G.L.); Tel.: +86-21-67795965 (G.L.)

Abstract: The design of distillation columns significantly impacts the economy, energy consumption, and environment of chemical processes. However, optimizing the design of distillation columns is a very challenging problem. In order to develop an intelligent technique to obtain the best design solution, improve design efficiency, and minimize reliance on experience in the design process, a design methodology based on the GA-BP model is proposed in this paper. Firstly, a distillation column surrogate model is established using the back propagation neural network technique based on the training data from the rigorous simulation, which covers all possible changes in feed conditions, operating conditions, and design parameters. The essence of this step is to turn the distillation design process from model-driven to data-driven. Secondly, the model takes the minimum TAC as the objective function and performs the optimization search using a Genetic Algorithm to obtain the design solution with the minimum TAC, in which a life-cycle assessment (LCA) model is incorporated to evaluate the obtained optimized design solution from both economic and environmental aspects. Finally, the feasibility of the proposed method is verified with a propylene distillation column as an example. The results show that the method has advantages in convergence speed without sacrificing accuracy and can obtain an improved design solution with reduced cost and environmental impact. Compared with the original design using rigorous simulation, the TAC is reduced by 6.1% and carbon emission by 27.13 kgCO_2/t.

Keywords: distillation column; BP neural network; genetic algorithm; surrogate modeling; intelligent design; life cycle assessment

Citation: Ye, L.; Zhang, N.; Li, G.; Gu, D.; Lu, J.; Lou, Y. Intelligent Optimization Design of Distillation Columns Using Surrogate Models Based on GA-BP. *Processes* **2023**, *11*, 2386. https://doi.org/10.3390/pr11082386

Academic Editor: Raul D. S. G. Campilho

Received: 16 July 2023
Revised: 2 August 2023
Accepted: 3 August 2023
Published: 8 August 2023

1. Introduction

Distillation is a separation technique that has been used for thousands of years, with evidence of its use dating back to ancient civilizations such as Greece and Egypt. Distillation remains a widely used separation technique, accounting for over 90% of separation processes. Due to the use of large equipment and low energy efficiency, distillation processes' economic costs and energy consumption are typically high. It is estimated that distillation energy consumption accounts for about 3% of global energy use and more than 40% of the chemical industry's energy use, and a large amount of energy consumption also results in significant CO_2 emissions [1]. As the main equipment of the distillation process, the optimal design of distillation columns significantly impacts economic investment, energy consumption, and the environment.

Systematic procedures have been developed for distillation column design, which often requires design understanding and experience to obtain a good design solution for the same distillation column design task. The current distillation design procedure still largely relies on engineers' experience. Different designers will provide different design solutions.

Due to the existence of a large number of degrees of freedom within distillation systems and complex interactions between variables, the optimization of distillation column design presents a high degree of nonlinearity and discretization; therefore, realizing the optimal design of distillation columns has always been a challenge.

Distillation systems include column equipment and auxiliary equipment such as reboilers, condensers, and pumps, and this work focuses on the design of the column. The optimal design of distillation columns involves structural and operational parameters, each with complex interactions. The key variables for optimization include discrete variables, such as the total number of plates (N) and the feed position (N_F), as well as continuous variables, such as the reflux ratio (R) and the operating pressure (P). The combination of discrete and continuous variables leads to complex mixed-integer nonlinear problems (MINLP) [2]. Scholars have proposed various distillation column design methods with different objectives in the past decades, mainly shortcut design, rigorous simulation, process optimization, artificial intelligence, and hybrid methods [3].

In the era of Industry 4.0, the realization of intelligent manufacturing has become one of the main goals in the chemical industry [4], which should cover the entire life cycle of chemical processes, including the design stage. The use of intelligent methods at the design stage in the life cycle of distillation processes can realize the simultaneous optimization of structural and operating parameters of distillation towers; improve the consistency and efficiency of process design; and obtain more economic design solutions with reduced capital cost, energy consumption, and carbon emissions.

The use of artificial intelligence or machine learning techniques such as data-driven, surrogate models, neural networks, and stochastic optimization algorithms in the field of optimal design of distillation columns has the potential to overcome the difficulties of achieving intelligent design of distillation columns. Therefore, in order to overcome the challenges of process complexity and difficulties in achieving intelligent distillation column design, this work aims to develop an intelligent method for optimizing distillation column design, which firstly establishes an alternative neural network model to replace the equation based rigorous model for distillation columns, and then combines it with a stochastic optimization algorithm to obtain the optimal design.

2. Previous Research on Distillation Column Design

Simulation-based modeling optimization approaches have been commonly applied in the design of distillation columns. Process optimization methods use mathematical optimization techniques to find a distillation column's optimal design and operating conditions. Examples of process optimization methods include the use of linear programming, mixed-integer programming, nonlinear programming (NLP), and heuristic optimization techniques. These are some references for process optimization methods in distillation design and operation [5].

Shortcut design methods are mainly based on simple mass and energy balance equations, empirical correlations, and procedures such as iteration and trial and error [6]. For example, the shortcut models based on the FUG [7–9] (Fenske–Underwood–Gilliland) equation and McCabe–Thiele graphical method can determine the minimum N and R for designing a distillation column. The early design methods also extend FUG and McCabe–Thiele method [10,11]. The predictions from simplified models are rather poor. Consequently, the application of the shortcut models in practice is limited. However, they can provide a good initial design for the rigorous model [12].

Due to the discrete and continuous variables, the rigorous distillation column model is usually expressed in the form of MINLP and generalized disjunctive programming [13]. Discrete variables are usually expressed as binary in the rigorous model based on the material–equilibrium–summation–heat (MESH) equation. Viswanathan and Grossman [14] first formulated the optimization problem of distillation columns as a MINLP problem. They used the 0–1 variable to represent the N_F and then constructed a superstructure model that can optimize the N of the distillation column. However, the MINLP mod-

eling involves the solution of redundant equations, such as phase equilibrium on each tray, which will affect the robustness of problem solving. Yeomans and Grossmann [15] proposed a Generalized Disjunctive Programming (GDP) model. The modeling method is applied to the optimization of the distillation sequence, thermal coupling distillation, and reactive distillation processes. The GDP model models the optimization of discrete decision variables as disjunctive variables. Under the framework of disjunctive logic conditions, the solution of redundant equations, such as phase equilibrium on the plate, is avoided. Jackson and Grossmann [16] established an optimization model for the reactive distillation column based on the GDP model and optimized the number of plates and feeding position. Barttfeld et al. [2] used both MINLP and GDP to model a single multicomponent distillation column; the MINLP model took longer than the GDP model to solve. Caballero et al. [17] combined the GDP model with the process simulator HYSYS and optimized the N, N_F, and *R* with the minimum total annualized cost (TAC) as the objective function. However, he made some specific settings during the process. Although this method yielded good results for the studied problems, it is not universally applicable. Tsatse et al. [18] optimized the design and operating parameters of complex reactive distillation processes based on the superstructure method. Compared with the MINLP model, the GDP model is easier to handle. However, both MINLP and GDP models are difficult to converge when dealing with complex nonlinearity. This drawback often leads to infeasible solutions. Therefore, the initialization and delimitation of variables are important to obtain valid solutions for these models.

Pattison and Baldea [19] proposed the pseudo-transient continuation (PTC) method. The main idea is to convert the original algebraic equation (AE) that is difficult to solve into the differential algebraic equation (DAE) that is easy to initialize. This model significantly improves the convergence of equation solving and the robustness of the initial solution. Later, this method optimized the dividing wall column (DWC) [20]. Ma et al. [21] established rigorous models for distillation columns, considering the P simultaneously. For optimization, the PTC model guarantees robustness, and a bypass efficiency method is used to optimize the number of plates efficiently. Based on the concept of bypass efficiency, Yeoh et al. [22] proposed two NLP models with a strict plate-to-plate MESH calculation. The models can be used to simultaneously optimize the N, N_F, and operating parameters such as the R and P.

Compared with the shortcut models, rigorous simulation methods based on the MESH equation are more consistent with distillation processes because of the display temperature, pressure, flow rate, flow composition, and other details in each plate of a distillation process, as well as the balanced relationship of each plate. Since proposed, modeling and design of distillation processes have been primarily based on these methods, typically via software such as Aspen Plus, HYSYS, and PRO/II. However, rigorous simulation methods require a lot of iterative calculations, which significantly slows down the speed of distillation column design. Besides, to obtain a valid solution, a good initial solution is required to ensure the convergence of a model [23].

For optimization algorithms, deterministic optimization algorithms search for the optimum according to the determined direction and step size. The quality of the solution depends more on the initial point, and the local optimum rather than the global optimum is obtained. Common deterministic algorithms include branch and bound (BB), outer approximation (OA) [16], generalized Benders decomposition (GBD), and successive quadratic programming (SQP) [24]. These algorithms have been implemented in GAMS, MATLAB, and other modeling and optimization software [25]. Caballero [26] used OA to optimize R, N, and N_F, but the solutions were just local optimum. The deterministic algorithms have good applicability to convex problems. However, when solving the complex MINLP problems representing the distillation process, the deterministic algorithms are not quite efficient in finding the optimal solution within an appropriate time range.

Unlike deterministic optimization algorithms, stochastic optimization algorithms perform adaptive optimization based on the solutions found. They do not rely on derivative

information of the objective function and constraints while searching for the best solution. By introducing random parameters, the stochastic optimization algorithm can jump out of the local optimal solution and obtain a solution infinitely close to the solution when time is sufficient. It can deal with complex nonlinear and large-scale numerical problems, including NLP, Mixed-Integer Linear Programming (MILP), and MINLP. Common stochastic optimization algorithms include simulated annealing (SA) [27], genetic algorithm (GA), ant colony optimization (ACO) [28], particle swarm optimization (PSO) [29], and artificial immune algorithm (IA) [30].

Javaloyes-Antón et al. [25] conducted a rigorous design of distillation columns by combining HYSYS with PSO, considering variables such as R, recovery rate, N, and N_F. Li et al. [31] realized the optimization of benzene-isopropanol-water ternary extractive dividing wall distillation process by integrating SA and Aspen Plus. Christopher et al. [32] used the method of combining HYSYS with PSO to optimize mechanical steam recompression (MVR) and self-heat recuperation (SHR) devices. Ibrahim [33] combined GA and SQP algorithms to carry out flexible optimization designs for crude oil distillation units (CDU), and the combination of the two improved the calculation efficiency. It should be noted that although stochastic optimization algorithms have many advantages, the final results may reveal small fluctuations due to the randomness of the algorithm itself. These research findings demonstrate that problems related to the optimal design of distillation columns can be solved using rigorous models. However, they all involve complex modeling and iterative calculations.

Artificial intelligence methods use artificial intelligence and machine learning techniques to optimize the design and operation of a distillation column based on historical process data and real-time process monitoring. Examples of artificial intelligence methods include neural networks, deep learning, fuzzy logic, and expert systems [34]. Recently, surrogate models have received increasing attention in the optimization design of distillation columns. As a type of model based on accessible data, surrogate models are often easier to establish and optimize compared with a rigorous model [35]. Moreover, the continuous emergence of new optimization algorithms greatly promotes the development of surrogate models [36].

Surrogate models treat the whole system as a black box, regardless of the actual internal processes of distillation. This kind of method is completed by data learning that obtains mathematical relationships between specific inputs and outputs of the system. Standard methods for building surrogate models include kriging [37,38], support vector regression (SVR), support vector machine (SVM), artificial neural networks (ANN), and radial basis functions (RBF).

Nentwich et al. [39] introduced the applications of surrogate models in the optimization design of chemical processes. Yao and Chu [40] used the SVR method to model and optimize crude oil distillation plants, with optimization variables including feed temperature, reflux ratio, product flow rate, pump cycle temperature drop, product flow rate, and steam flow rate. López C. et al. [41] established a meta-model with a second-order polynomial function to optimize the CDU in order to maximize profit. Based on the data from an actual plant, Liau et al. [42] established an ANN model to optimize the distillation unit's operation to improve product quality. TGARGUIFA et al. [43] use the principles of central composite design (CCD), the ANN method, and the desirability function. To optimize the operating conditions, such as the R, P, and N_F, to reduce the distillation of bioethanol production's energy consumption and cost. Ochoa-Estopier and Jobson (2015a) [36] proposed establishing the ANN surrogate model of CDU to promote the operation optimization of the CDU system. The model considers the operational variables, such as coil outlet temperature, pump-around temperatures, flow rates, and flow rates of distillation products and stripping steam. Osuolale et al. [24] proposed a modeling strategy for distillation columns to maximize the distillation system's energy performance. This strategy was based on bootstrap-ANN and the second law of thermodynamics, with decision variables includ-

ing product and steam flow rate, pump cycle temperature drop and load, and product quality constraints.

Ibrahim [33,44,45] pointed out that the design variables (such as the numbers of N, Feed position of N_F, etc.), which were assumed to be fixed in the previous models, need to be optimized. Ibrahim established an ANN surrogate model specifically for the design of CDU, considering discrete (the number of plates in each tower section) and continuous (feed temperature, load around the pump, temperature drop, and steam flow rate, R) design variables. The ANN model of CDU was established by associating all independent variables with a set of specific dependent variables. The infeasible design was removed from the solution space by SVM. The output variables include boiling temperature, product flow rate, supply temperature of the steam to be heated and cooled, target temperature, corresponding enthalpy change, and diameter of each section. Peng et al. [46] established the improved Back Propagation neural network (BP-NN) models for the reactor and distillation column but did consider the cost of capital. Moreover, some use surrogate models to achieve multi-objective optimization of more complex distillation columns [47]. Some researchers have realized the deterministic global optimization of the artificial neural network, and based on this, the cumene process has been optimized for operation points [48].

Hybrid methods combine elements of the other design methods, such as using rigorous simulation to generate training data for an artificial intelligence model or using a heuristic optimization method to refine the results of a mathematical optimization method [49]. Tables 1 and 2 summarize the advantages and disadvantages of different modeling techniques and algorithms commonly used to optimize distillation design.

Table 1. Modeling techniques and their advantages and disadvantages.

Model		Advantages	Disadvantages
Simplified models	FUG McCabe-Thiele	• Provide initial solutions	• Not accurate enough for rigorous design
Rigorous models	MINLP/NLP GDP PTC Aspen	• More detailed information can be obtained • More accurate	• Could be time-consuming for computation • Potential convergence issues
Surrogate models	ANN Kriging degree polynomial	• Easy to establish, optimize, and solve	• Could be time-consuming to establish accurate and complete data samples

Table 2. Algorithms and their advantages and disadvantages.

Algorithm		Advantages	Disadvantages
Deterministic algorithm	SQP LBOA BB GBD	• Generally faster	• Solution depends on the initial point • Local optimum • Possible convergence issues
Stochastic algorithm	GA SA PSO IA ACO	• Can handle MINLP problems more reliably • Robust computing ability • Global optimization	• Slower computation • Stability of solutions

In summary, there are two main issues with shortcut-rigorous iterative optimization or automatic optimization by linking rigorous simulation (such as Aspen Plus) and deterministic algorithms (such as the ones built in MATLAB or GAMS): the long computational time required in the iterative optimization procedure, and potential failure of convergence due to the MINLP nature of the problem. On the other hand, switching to data-driven

approaches with surrogate models means that the major computational effort of many simulation runs is made up front rather than within the optimization process, which speeds up the whole procedure. In addition, it removes the possibility of convergence failure in rigorous simulation, which makes the optimization procedure more robust.

Upon reviewing the previous research on distillation column design, surrogate models have good applicability in the optimization design process of distillation columns, and combining it with the stochastic optimization algorithm for optimization design greatly improves the efficiency of obtaining the optimal design scheme. Therefore, this research aims to adapt artificial intelligence techniques in optimizing distillation design by turning a model-driven design process into a data-driven one, built upon an extensive training dataset from rigorous simulation.

In this work, a back propagation neural network surrogate model of the distillation column is established. The sample data are obtained through rigorous simulation, considering the design and operating variables. Combining the random optimization algorithm with the genetic algorithm for optimization design to find the optimal design and operating parameters with the goal of minimizing total annual cost, this method greatly reduces the dependence on experience and manpower, improves the efficiency of obtaining the optimal design, and demonstrates the feasibility of this method through the common separation system of propane-propylene separation.

3. Modeling

The proposed model building process is mainly composed of three main steps:

(1) Data generation. Analyze the impact of each variable on operational and economic costs, determine the upper and lower limits of the input and output variables of the optimization variables, generate sampling points for each input variable within its range using the Latin hypercube sampling method, and conduct multiple rigorous simulations to establish data samples.

(2) Establish a surrogate model. Using sample data to establish a surrogate model, compared with other surrogate models, ANN has great flexibility and good learning ability and is suitable for dealing with complex problems and large amounts of data [50]. BP neural network (BP-NN) is a multilayer feedforward neural network based on an error back propagation algorithm (BP algorithm). It modifies the weight coefficient according to the gradient, decreasing the direction of the error function to reduce the error. Finally, it reflects the nonlinear mapping relationship between input and output data, characterized by simple structure and effective calculation [51]. Therefore, BP-NN is selected to establish this research's surrogate model for distillation columns.

(3) Optimize the process. The surrogate model is combined with the stochastic optimization algorithm, and the TAC minimization is taken as the objective function. Under the constraints of producing qualified products, the optimal design is searched by considering the simultaneous optimization of design variables (N and N_F) and operating variables (R, P, and fraction feed ratio). GA is a common random optimization algorithm with the advantage of simple operation. Therefore, this study chose this algorithm for optimization operations, and the final model is the GA-BP optimization design model. Figure 1 shows the detailed process of the GA-BP optimization model.

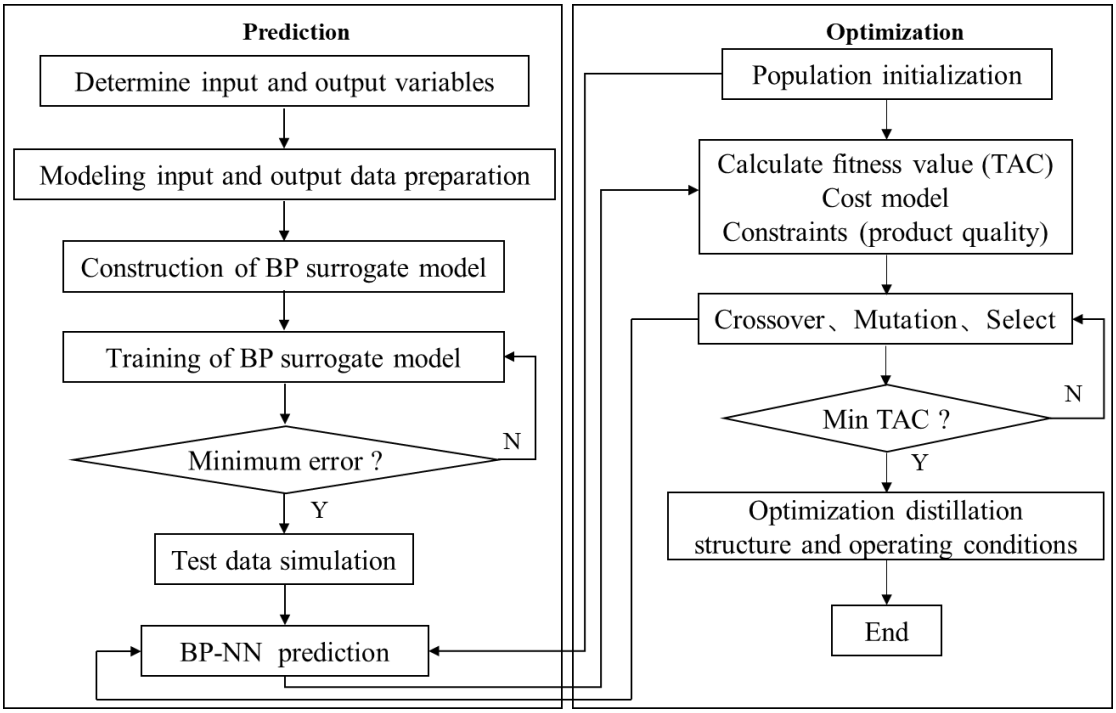


Figure 1. GA-BP intelligent design optimization modeling.

3.1. Data Generation

The first step in building the ANN surrogate model is to collect network-training datasets. This work uses the results of extensive, rigorous simulations (such as Aspen Plus®) as the sample data needed to build a surrogate model. The set-up of sample data includes: (1) selecting input and output and determining the upper and lower bounds for each input variable; (2) generating random sample points for each input variable; (3) obtaining the output corresponding to the input by rigorous simulation. This is accomplished through the integration of Aspen Plus with MATLAB through the following interface:

```
handles.aspen = actxserver('Apwn. Document');

handles.filepathname = strcat(filepath,filename);

handles.aspen.InitFromFile2(handles.filepathname);

handles.aspen.Visible = 0/1;
```

where filepath is the path and filename is the file name, and this achieves the interconnection between Aspen Plus and MATLAB platforms. By finding the FindNode corresponding to the variable, you can rewrite the data of Aspen Plus in MATLAB.

For the design of the distillation column for a given separation task, key variables considered in this research include feed temperature (T), feed flow (F), the mole fraction of feed components, N, NF, R, feed ratio of distillate (D/F), product quality, tower top temperatures (T1) and bottom temperatures (T2), reboiler duties (Qr), and condenser duties (Qc), as well as the calculated reflux ratio (Rc). Under the constraints of producing qualified products and minimizing the loss of column bottom products to the greatest extent possible, these variables are the minimum requirements for the design of a distillation column and the main variables affecting the objective function. The range of variables is determined by sensitivity analysis to ensure that it converges to a meaningful, feasible solution space satisfying specified constraints such as the MESH equation and product specifications.

The next step is to generate random sample points for each input variable. There are many methods available for sampling, such as Latin Hypercube Sampling (LHS) and Markov Chain Monte Carlo sampling (MCMC), both of which are randomized sampling techniques. LHS is adopted in this research because the sampling method can be applied to uniformly distributed sample points [52]. Each set of samples is made up of different combinations of variables in their respective design range spaces.

Data transmission is realized through the integration of data sampling and rigorous simulations. Only converged samples are collected because not all rigorous simulations will converge. The purpose of removing non-converged samples is to avoid generating optimal solutions that do not meet constraints and improve model accuracy. Figure 2 shows the process of obtaining sample data.

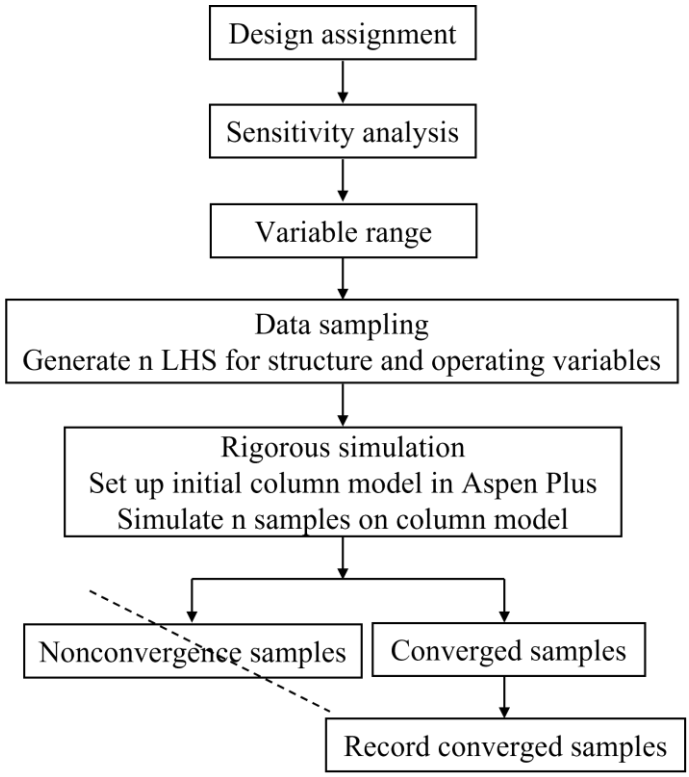


Figure 2. LHS sampling process (The dotted line indicates that non-convergent samples are excluded from modeling).

3.2. Surrogate Modeling of Distillation Columns

After obtaining the data samples, a distillation column surrogate model can be established. BP-NN is a typical forward network consisting of input, hidden, and output layers. Full interconnection is among the layers (the connection between each unit in the upper layer and each unit in the lower layer), as illustrated in Figure 3, where there is a BP-NN with a hidden layer, where X_1, $X_2 \ldots$, X_N are inputs and Y_1, $Y_2, \ldots$, Y_N are predicted outputs. w_{hi}, w_{jh}, and b_h, b_j are weighting and threshold values. Unlike other ANNs, the activation function of BP-NN usually adopts sigmoid differentiable functions such as logsig and tansig functions because such nonlinear functions can characterize both linear and nonlinear problems. The output layer selects a linear purelin function to output any value. Other critical parameters are the number of hidden layers and hidden layer nodes, which have major impacts on the model's accuracy. Some literature shows that neural networks with a hidden layer can approximate any continuous function with a closed domain [53]. On the other hand, too few hidden layer nodes will lead to poor accuracy, while too many will lead to increased complexity or overfitting. The empirical

formula shown in Equation (1) is typically used to calculate the number of hidden layer nodes [44], which is adopted in this research and calibrated with error detection.

$$m = \sqrt{n + o} + z \;\; z \in [1,\ 10] \tag{1}$$

where m is the number of hidden layer nodes, n is the number of variables at the input layer, o is the number of variables in the output layer, and z is a constant of [1, 10].

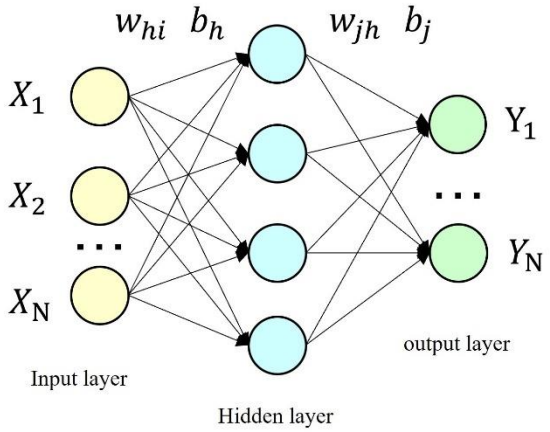


Figure 3. The structure of a simple BP-NN with one hidden layer.

3.3. Training of Surrogate Model

After obtaining the input and output sample data and determining the structure and parameters of the model, a BP-NN surrogate model can be trained, verified, and tested. MATLAB neural network toolbox supplies a function newff for neural network establishment. The grammar of it is as follows:

$$\text{net = newff (PR,[S1 S2 SN],\{TF1 TF2...TFN\},BTF)} \tag{2}$$

where *PR* is an input vector, which contains two lines that denote the minimum and the maximum of the input vector, respectively; each element in [S1 S2 ... SN] represents the number of neurons in each layer; each element in {TF1 TF2 ... TFN} [10] means the transfer function neurons used in each layer, and BTF is the training function used during training. In the process of network initialization, the newff function will automatically call the default function 'init' to initialize the weights and thresholds and return to 'net'—that is, the trained network.

The first step in training is to normalize the data, the purpose of which is to limit preprocessed data to $[-1, 1]$ in order to eliminate the adverse effects of excessive or small sample data, avoid problems such as slow convergence and long training time due to an extensive range of data changes, and improve model accuracy. A BP-NN model can then be trained with the determined structure of BP-NN, including the number of hidden layers, the number of hidden layer nodes, the transfer function between layers, the training function, and the network parameters (training times, learning rate, and training objectives). The numbers of input and output layer nodes are the same as the numbers of input and output variables. This process also includes the calculation of weighting and threshold values. The train stops when the network reaches the specified maximum error or maximum training times. The error can be evaluated by comparing the actual values and the predicted values with mean squared errors (MSE) defined in Equation (3) and mean absolute error (MAE) in Equation (4). A desired value of R^2 is close to 1, indicating a more accurate prediction of BP-NN.

$$\text{MSE} = \sum_{i=1}^{S} \frac{(t_i - y_i)^2}{S} \tag{3}$$

$$\mathrm{MAE} = \frac{1}{\mathrm{S}}\sum_{i=1}^{\mathrm{S}}|t_i - y_i| \tag{4}$$

where MSE is the mean square error, MAE is the mean absolute error, S is the total sample points, and t and y are the prediction output and actual output, respectively.

3.4. Optimization Process

In this work, the BP-NN surrogate model is integrated with GA in the same framework for optimization. GA is one of the most widely used stochastic optimization algorithms. The optimization principle is to continuously realize 'survival of the fittest' based on changes in fitness and pass on better genes to the next generation until the best individual is found. It mainly includes five processes: coding, setting the initial population, formulating a fitness function, determining selection-crossover-mutation operators, and setting operation parameters. The process is easy to build, which is adopted in this work as the algorithm for optimization.

The objective function is the only criterion for the continuous evolution of a GA. Different objective functions such as energy cost, CO_2 emissions [54], product revenue, maximum net profit, and product output can be selected according to the design purpose. This paper takes the minimum TAC as the objective function, as defined in Equation (5), which can be referred to in the work of Luyben [55].

$$\mathrm{TAC} = \mathrm{C_{OPE}} + \frac{\mathrm{C_{CAP}}}{\mathrm{P_t}} \tag{5}$$

where $\mathrm{C_{OPE}}$ is the operating cost, including utility costs; C_{CAP} is the capital cost determined by N, column height ($\mathrm{H_c}$), column diameter ($\mathrm{D_c}$) [56], heat exchanger area ($\mathrm{A_c}$ + $\mathrm{A_r}$), etc.; and $\mathrm{P_t}$ is the payback period of investment.

Relevant constraints for process design should also be specified, such as the feasible range (lower and upper bounds) of design and operating variables, product specifications, hydraulic restrictions, etc.

Generally, the GA-BP optimization design framework mainly includes: (1) using trained BP-NN to predict outputs; (2) GA takes the trained BP-NN prediction results as individual fitness values and searches for the optimal global value of the objective function, with corresponding variables through the iterative operation of selection, crossover, and mutation.

Unlike the usual GA-BP optimization framework, in Step (2) of this study, the results of the BP prediction outputs are not directly used as the fitness values of GA optimization. They are first used to calculate the TAC in GA through the BP prediction outputs, which is the fitness function. In the optimization framework shown in Figure 2, GA randomly generates an initial population within the range of parameters—that is, a group of design schemes, and the BP-NN proxy model performs the prediction outputs. Then, the fitness value of TAC is calculated in GA, and the process continues until the convergence criterion or a set number of iterations is reached. The optimal solution corresponds to the design scheme with the minimum TAC.

For the setting of GA parameters for distillation design optimization, a floating-point method is adopted for coding because standard binary coding has mapping errors when continuous functions are discretized. The length of individuals will affect the accuracy of the solution [57], while the floating-point method is convenient for solving complex problems and has high accuracy. The setting of population size, evolutionary algebra, crossover, and mutation probability was determined by an orthogonal test method, discussed by Katoch, S., etc., in detail [58].

3.5. Life Cycle Assessment

The environmental evaluation of distillation processes is extremely important, and its evaluation index can be quantified as carbon emissions. The life cycle assessment method is a common theory that can systematically evaluate a product or activity's carbon emissions. In this work, the carbon emissions of distillation processes are estimated with the emission factor method, considering the carbon emissions of both the manufacturing and operation phases of the equipment, and the system boundary is shown in Figure 4. The plant life is assumed at 20 years, and the environmental impact assessment is conducted on the design scheme obtained from the model based on the functional unit of carbon emissions generated by processing 1 ton of raw materials during the total operating time. In the case study of this work, the main units of the distillation column, reboiler, and condenser are made of carbon steel. The consumption of carbon steel for the distillation column can be estimated based on the column height and diameter. The consumption of carbon steel for the reboiler and the condenser can be estimated based on the heat exchange area. The consumption of steam, cooling water, and electricity during operation can be calculated based on energy consumption [59]. The emission factors (f) and calculation process of specific carbon steel and process mass are shown in Table 3. The emission factor for electricity is set at 0.84 $kgCO_2/(kW{\cdot}h)$.

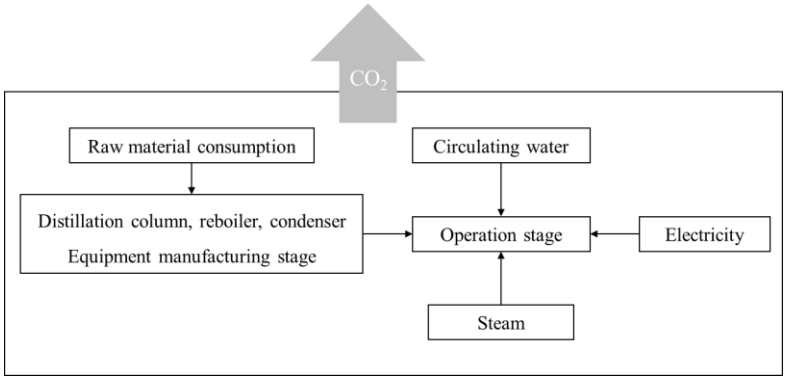


Figure 4. System boundary.

Table 3. Emission factors of different substances.

Stage	Material	f $kgCO_2/t$
Equipment Manufacturing	Carbon steel	720 [60]
Operation	Circulating water	0.30
	3.5 MPa steam	262.64
	1.0 MPa steam	226.82
	0.35 MPa steam	196.98

Carbon emission calculation process:

1. Heat exchanger mass calculation:

$$m_{HEX} \approx \rho_{steel} \times V_{plate} = \rho_{steel} \times b \times A \quad (6)$$

$$A = A_{co} + A_{re} \quad (7)$$

where m_{HEX} is for the heat exchanger steel consumption, t; ρ_{steel} and ρ_s are the density of steel, 7.85 g/cm^3; b indicates the thickness of the heat exchanger plate, 0.5 mm; A is the total heat transfer area, m^2; A_{co} is the condenser heat transfer area, m^2; and A_{re} is the reboiler heat transfer area, m^2.

2. Electricity consumption of the pump:

$$N_{pump} = \frac{N_e}{\eta_{pump}} \tag{8}$$

$$N_e = hq_l\rho_c g \tag{9}$$

where N_{pump} is the shaft power, W; N_e is the effective power, W; η_{pump} is the efficiency of the pump; q_l is the flow rate, m^3/s; ρ_c is the density of the product, kg/m^3; h is the delivery head, m; and g is gravity acceleration, 9.8 m/s^2.

3. Carbon emissions generated per 1 t of raw material processed:

$$GHG_{P-CO_2} = f_{steel} \times m_c + f_{steel} \times m_{re} + f_{steel} \times m_{co} \tag{10}$$

$$GHG_{R-CO_2} = \left(f_{water} \times m_{water} + f_{steam} \times m_{steam} + f_{electricity} \times N_p\right) \times 8000 \times 20 \tag{11}$$

$$ECP = \frac{GHG_{P-CO_2} + GHG_{R-CO_2}}{m_s \times 8000 \times 20} \tag{12}$$

where GHG_{P-CO_2} is the carbon emissions from the production of the plant; GHG_{R-CO_2} is the carbon emissions during the total operating time of the plant, $kgCO_2$; m_c, m_{re}, and m_{co} are the amount of raw materials consumed by the distillation column, reboiler, and condenser, respectively, t; m_{water} and m_{steam} are the consumption of circulating water and steam, respectively, t/h; f_{steel}, f_{water}, and f_{steam} are the carbon emission factors of carbon steel, circulating water, and steam, respectively, $kgCO_2/t$; $f_{electricity}$ is the electricity carbon emission factor of electricity, $kgCO_2/(kW{\cdot}h)$; ECP denotes the carbon emission per ton of raw material treated, $kgCO_2/t$; and m_s is the amount of raw material treated per hour, t/h.

4. Case Study

This section demonstrates the feasibility of the developed design approach via the design of a propane–propylene distillation system.

4.1. Problem Description

Propane–propylene is a typical near-boiling system (at atmospheric conditions, the boiling points of propylene and propane are −47.6 °C and −42.1 °C, respectively). The number of plates (N) required for propylene distillation columns is usually very high to obtain high-purity propylene products. R is also large, resulting in high capital and energy costs. Therefore, the design optimization of propylene distillation columns has been studied. There are many systems like propane–propylene distillation. Therefore, the application of the proposed GA-BP optimization modeling for intelligent design optimization of propylene distillation columns has a wider implication. The propylene column's initial operating conditions are based on the work by Cui [56]. The specific composition and initial operating conditions are shown in Table 4. The plate efficiency is set at 100%; P_t is at 3 years, with 8000 operating hours per year; the cooling water (25 °C to 35 °C) cost is 0.354 USD/GJ; and the low-pressure steam (6 bars at 160 °C) cost is 7.78 USD/GJ.

The optimization variables of the propylene distillation system are selected as N, N_F, R, and P. The mathematical process can be described as follows:

$$Min\ TAC = f(N, N_F, R, P) \tag{13}$$

Subject to

$$X_{C_3H_6} \geq 99.5\%, \quad X_{C_3H_8} \leq 1.1\% \tag{14}$$

Table 4. Operating conditions of the propylene distillation unit.

Parameter	Value
Feed composition (mol %)	
Propane	0.5
Propylene	0.5
Feed flow rate (kmol/h)	100
Distillate molar flow rate(kmol/h)	50
Feed temperature (°C)	44.59
Distillate specification (mol %) propylene	99.5
Bottom specification (mol %) propylene	1.1
Feed pressure (MPa)	1.7
N	201
N_F	137
Overhead pressure (MPa)	1.6789
Bottom pressure (MPa)	1.6989

4.2. GA-BP Modeling

First, a rigorous simulation model is built in Aspen Plus, using PENG-ROB [61] as the thermodynamic property package for the propylene distillation column. Sensitivity analyses are carried out to identify bounds for each independent variable, as shown in Table 5.

Table 5. Range of variables.

Input Variables	lower Bound	Upper Bound
Feed temperature (°C)	—	
Feed propylene and propane component (mol%)	—	
Feed flow rate (kmol/h)	—	
D/F	—	
N	150	230
N_F	80	170
R	15	22
P	1.65	2.2

Next, using LHS, 2000 samples are generated, each sample consisting of different combinations of the independent variables within the bounded region. Through an interface established between Aspen Plus and MATLAB, all the samples are simulated. Of these samples, 59.3% (1186) simulations are converged; for the remaining 40.7% (814), the simulations do not converge. The sampling is carried out on an HP desktop PC with an Intel(R) Core i5 processor running at 3.0 GHz and 8 GB of RAM.

The converged sample data are used to train and establish the BP-NN surrogate model, in which 85% (1007) are used as the training set, and 15% (179) are used as the verification set. According to the empirical formula, about 5–14 neurons could be used in the hidden layer. The MSE of the prediction results for different node numbers is calculated and compared, and the optimal number of nodes in the hidden layer of BP-NN is determined to be 11. Therefore, the structure of BP-NN of the propylene distillation column under this design is set at 9-11-6. The detailed structure and parameter operations of the BP-NN agent model are shown in Table 6.

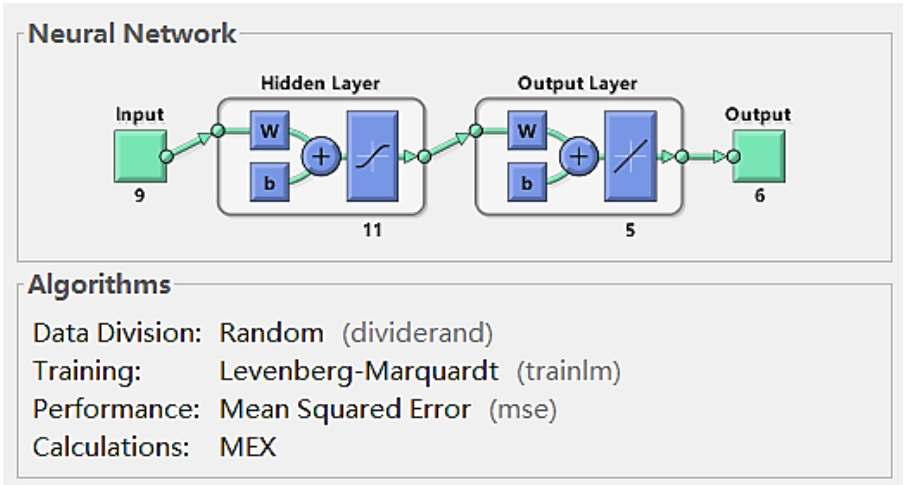
Table 6. BP-NN surrogate model parameter.

Parameter	
Input layer	9
Hidden layer	11
Output layer	7
Hidden layer transfer function	tansig
Output layer transfer function	purelin
Training function	trainlm
Training error	0.0001
Training times	2000
Learning rate	0.1
Training algorithm	Levenberg–Marquardt

Figure 5a shows the training structure of BP-NN. Figure 5b shows the BP-NN training error versus the number of iterations. It can be seen that the error change of the number of BP-NN iterations reached the minimum value after 284 and stabilized afterward. Figure 5c shows the results of R^2, indicating that the surrogate model has good fitness. According to the calculation of the test sample set, the MSE is at 0.000122, and the MAE is at 0.011. It can be seen that the prediction results of the BP model are almost consistent with the results of the rigorous model calculation. These results demonstrated the feasibility and accuracy of the established BP-NN surrogate model.

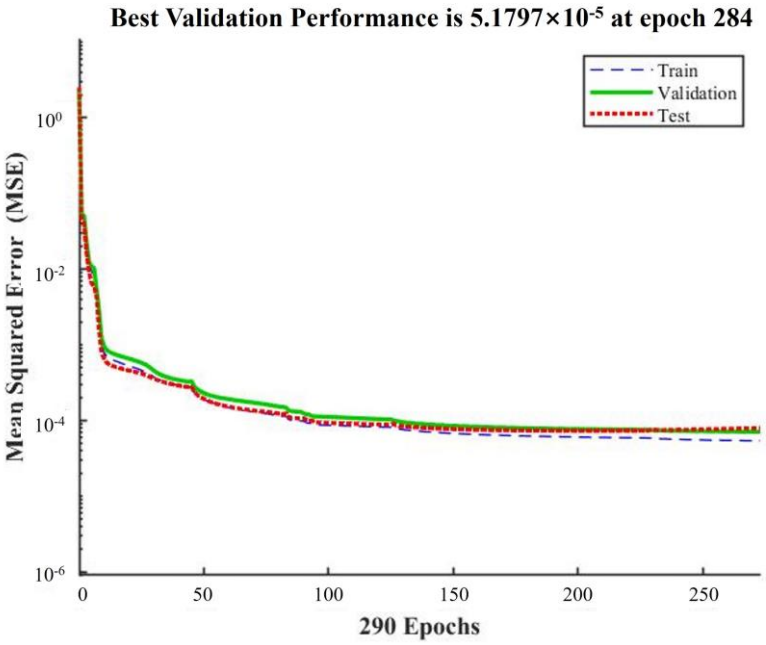
4.3. GA-BP Optimization Results

Orthogonal experiments provide GA operation parameters. The TAC values are compared under different parameters, with the population number set at 200, the crossover probability at 0.5, and the mutation probability at 0.03. With 1000 iterations, good results can be obtained. Encoding with actual numbers was used, with binary crossing. The selection rules follow roulette. Because GA is a random optimization algorithm, it has a certain amount of randomness. In order to obtain a reliable and consistent solution, five runs are needed to obtain each optimization result. Figure 6a and b show, respectively, the results of a rigorous model and the results of optimization by establishing a shortcut model in the literature [51].

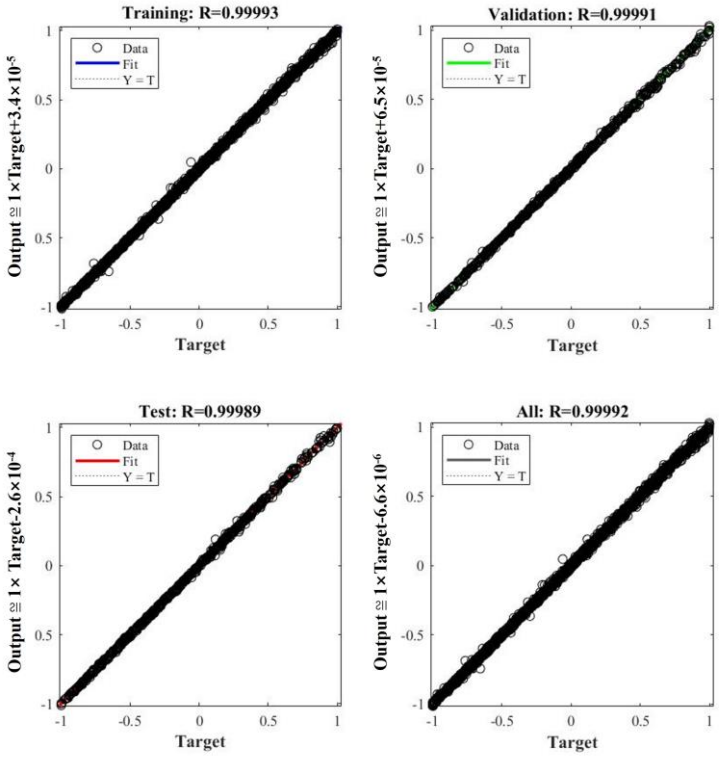


(**a**)

Figure 5. *Cont.*



(**b**)



(**c**)

Figure 5. (**a**) BP-NN surrogate model network structure diagram; (**b**) Change in BP-NN training error with training times; (**c**) BP-NN network regression diagram (training, verification, test set, and all datasets).

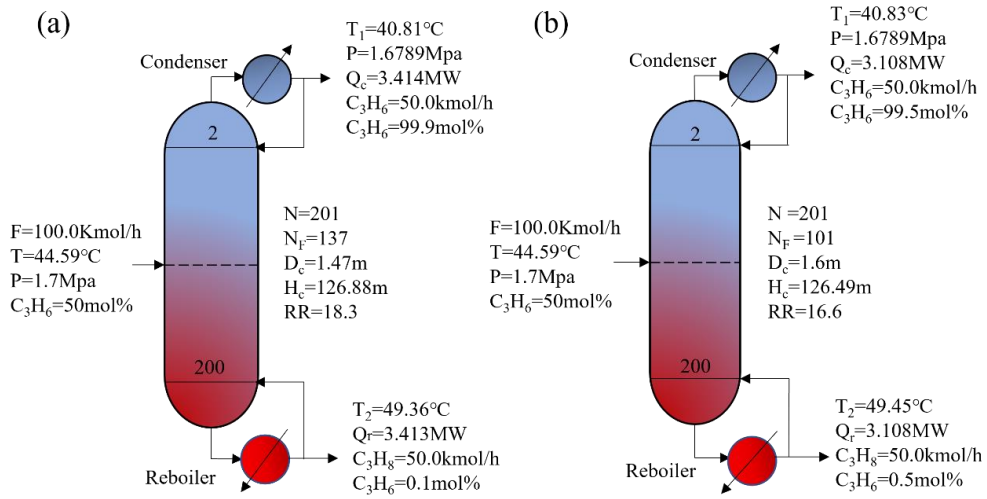


Figure 6. (**a**,**b**) Represents rigorous design and simplified model design in the literature, respectively.

Optimization of the propylene distillation column design by the GA-BP model covers the number of plates, feed position, reflux ratio, and operating pressure simultaneously. Due to the stochastic nature of GA, the optimization process was run five times, with the results shown in Table 7. The results of the five runs did not differ much, and the optimal TAC was found in the third run. Figure 7 shows the GA search process for optimal results. The TAC of the optimal solution fluctuates greatly during the first 600 generations of decline, and the TAC value tends to be minimized and gradually stabilized after 800 generations. Therefore, taking 1000 generations as the maximum iteration of evolutionary generations is reasonable. The global optimal solution can be approximated to the optimal solution of 1000 generations. Figure 8 shows the results of the specific optimization design.

Table 7. Optimization results for multiple runs.

Run	TAC (USD/a)
1	1,331,609.71
2	1,325,233.37
3	1,324,611.42
4	1,326,403.67
5	1,328,464.11

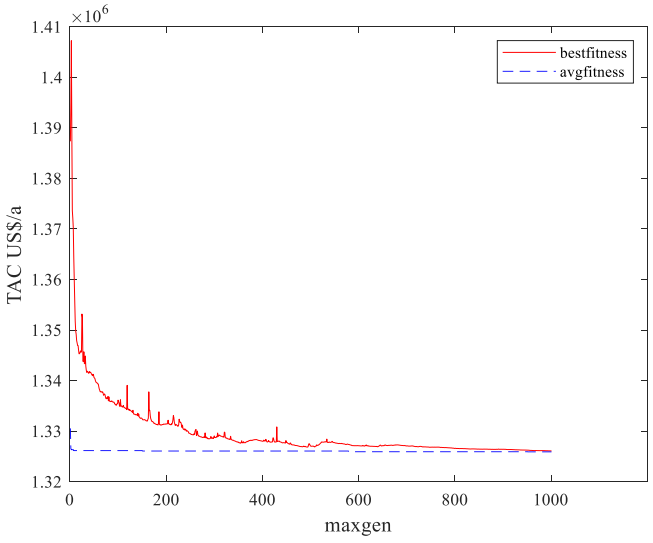


Figure 7. GA optimization process of the optimal result.

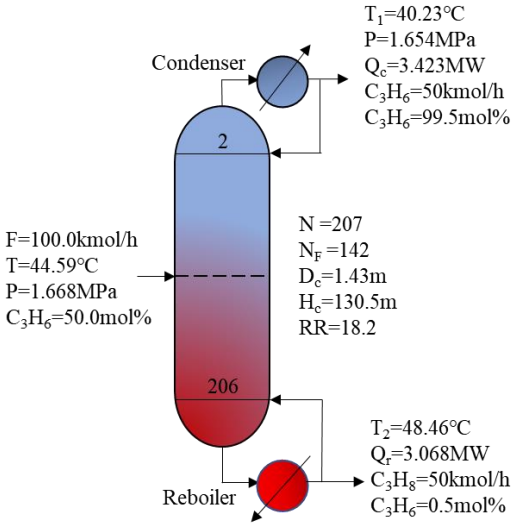


Figure 8. Detailed design of GA-BP optimization results.

Comparison of Figures 6a,b and 8 shows that compared to the original design, the design solution obtained in this work has an increase in the number of plates N, a change in the feed position N_F, and a decrease in the operating pressure P, as well as a change in the reflux ratio R, which has led to a decrease in the energy consumption and a decrease in the column diameter, leading to a decrease in the TAC. Table 8 shows a detailed comparison of the results. Under the same propylene product specifications, the optimal design scheme from the GA-BP optimization framework in this work is better than that reported in the literature. Considering the simultaneous optimization of the column plate number N, feed position N_F, reflux ratio R, and operating pressure P, the optimized design solution TAC obtained by GA-BP saves 6.1% compared with the original, rigorous model and 3.7% compared with the shortcut model.

Table 8. Result comparison.

Cost	Figure 6a	Figure 6b	GA-BP
Utility requirements			
Hot utility (MW)	3.10837	3.10837	3.068
Cold utility (MW)	3.414	3.10837	3.423
Annualized operating cost (USD/a)	799,462	728,165	718,649
Annualized capital cost (USD/a)	611,307	646,932	605,962
Total annualized cost (USD/a)	1,410,834	1,375,097	1,324,611

The results of the GA-BP optimization design framework are compared with the results of the Aspen Plus simulation. The specific results shown in Table 9 are the Aspen Plus simulation results. It can be seen from the relative error that the error between the two results is small. We can see that the maximum relative error is 6.94×10^{-3}, which further verifies the accuracy of the GA-BP optimization framework.

Although they have the same order of magnitude in the results, a slight reduction may bring substantial economic benefits to the wide context of the process industry. The error comparison with the output results of rigorous simulation shows the accuracy of the established surrogate model. Moreover, establishing a BP-NN surrogate model using the proposed approach is also relatively straightforward, which can reduce the dependence on designers' engineering experience and improve the design quality as well as job efficiency; the average CPU running time for obtaining the best design solution is 44.5 min.

Table 9. Comparison with results of Aspen Plus.

Output	GA-BP	Aspen Plus	Relative Error
T_1 (°C)	40.23	40.15	1.99×10^{-3}
T_2 (°C)	48.46	48.7	4.93×10^{-3}
Q_c (MW)	3.423	3.432	2.62×10^{-3}
Q_r (MW)	3.068	3.078	-3.25×10^{-3}
C_3H_6 (mol/%)	99.5	99.5	0
D_c (m)	1.43	1.44	-6.94×10^{-3}
Rc	18.2	18.2	0

4.4. Environmental Assessment

Based on the carbon emission calculation process in Section 3.5, carbon emissions calculations were conducted on the design schemes obtained from the literature and GA-BP models, and environmental impacts were compared. The literature and GA-BP model correspond to the consumption of raw materials in the design scheme equipment manufacturing phase. The consumption of each utility in the operation phase is shown in Table 10, and the carbon emissions generated are shown in Table 11.

Table 10. Design results for material and utility consumption of the literature and GA-BP optimization.

	Equipment Manufacturing Stage			Operation Stage		
	Column t	Condenser t	Reboiler t	Steam t/h	Circulating Water t/h	Electricity kW·h
Figure 6a	208.09	0.37	2.08	5.89	10.37	5.87
Figure 6b	226.49	0.34	1.89	5.36	9.45	6.91
GA-BP	208.28	0.33	2.08	5.29	10.40	6.29

Table 11. Carbon emissions from the literature and GA-BP optimization.

	Equipment Manufacturing Stage $kgCO_2/t$	Operation Stage $kgCO_2/t$	Carbon Emissions Generated per 1 t of Raw Material Processed $kgCO_2/t$
Figure 6a	0.221	271.13	273.35
Figure 6b	0.240	247.21	247.45
GA-BP	0.222	244.00	244.22

As can be seen in Table 11, the design solution obtained from the GA-BP model produces lower carbon emissions, with a reduction of 27.13 $kgCO_2/t$ per ton of propylene-propane feedstock processed compared to the results based on the rigorous model in the literature and a reduction of 3.23 $kgCO_2/t$ compared to that from the shortcut model. The number of plates in the optimized design is increased. However, the change in the feed position and the reduction in operating pressure reduce the energy consumption and carbon emissions during operation to a greater extent. It can also be seen in Table 11 that the carbon emissions from the operation phase account for the majority of the carbon emissions from the distillation process, and reducing the energy consumption during operation can significantly reduce the environmental impacts.

5. Conclusions

The traditional design process for distillation columns is iterative, which heavily relies on engineers' experience as inputs, resulting in inconsistent design results and quality. In this work, an optimization framework of BP-NN combined with GA, the GA-BP optimization framework, is developed to achieve intelligent distillation column design. Compared with the conventional design process, a data-driven model is utilized to replicate

a rigorous simulation model, with the advantages of a simple modeling process and fast convergence for optimization. At the same time, GA can take TAC as the objective function for optimization from a global perspective. N, N_F, R, and P of the distillation column can be optimized simultaneously under specified quality and operating constraints. A case study for propylene distillation column design shows that the proposed method in this work can reach the optimal design scheme faster with sufficient accuracy. The design example of a propylene distillation column shows that the proposed method can achieve the economically optimal design solution with a reduced environmental impact faster and with sufficient accuracy, which is useful for the optimal design of distillation columns.

In the future, this method could be further extended to consider the detailed design of column internals and other ancillaries, such as overhead condensers and bottom reboilers, as well as azeotropic distillation columns and complex distillation systems.

Author Contributions: Conceptualization, N.Z.; methodology, L.Y. and N.Z.; software, L.Y., N.Z. and J.L.; validation, L.Y., N.Z. and Y.L.; formal analysis, L.Y. and N.Z.; investigation, L.Y.; resources, N.Z.; data curation, L.Y.; writing—original draft preparation, L.Y.; writing—review and editing, D.G.; visualization, L.Y.; supervision, N.Z.; project administration, G.L.; funding acquisition, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Capacity Building Project of Some Local Colleges and Universities in Shanghai (No. 21010501400).

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

D_c	Column diameter	m
F	Feed flow rate	kmol/h
H_c	Column height	m
N	Total number of plates	
N_1	Distillation section	
N_2	Stripping section	
N_F	Feed position	
P	Operating pressure	MPa
Q_c	Condenser duty	MW
Q_r	Reboiler duty	MW
R	Specified reflux ratio	
Rc	Calculated reflux ratio	
T_1	Column top temperature	°C
T_2	Bottom temperature	°C
g	Gravitational acceleration	m/s^2
h	Head of delivery	m
N_e	Effective power	
Np	Shaft power	
P_t	Payback period	a
q_l	Flow	m^3/s
ρ_l	Product density	kg/m^3
ρ_{still}	Carbon steel density	g/cm^3
Superscript		
A	Heat exchange area	m^2
C	Cost	USD
f	Emission factors	$kgCO_2$
ECP	Carbon emissions per 1 ton of raw material processed	$kgCO_2/t$
GHG	Carbon emission	$kgCO_2$
m	Consumption	t

Subscripts	
c	Column
CAP	Equipment investment
co	Condenser
HEX	Heat exchanger
OPE	Operate
P-co_2	Equipment manufacturing stage
R-co_2	Operate stage
re	Reboiler
s	Raw material

References

1. Zhu, J.; Chen, L.; Liu, Z.; Hao, L.; Wei, H. Synergy of electrification and energy efficiency improvement via vapor recompression heat pump and heat exchanger network to achieve decarbonization of extractive distillation. *Sep. Purif. Technol.* **2022**, *293*, 121065. [CrossRef]
2. Barttfeld, M.; Aguirre, P.A.; Grossmann, I.E. Alternative representations and formulations for the economic optimization of multicomponent distillation columns. *Comput. Chem. Eng.* **2003**, *27*, 363–383. [CrossRef]
3. Costa, A.L.H.; Bagajewicz, M.J. 110th Anniversary: On the departure from heuristics and simplified models toward globally optimal design of process equipment. *Ind. Eng. Chem. Res.* **2019**, *58*, 18684–18702. [CrossRef]
4. Chiang, L.H.; Braun, B.; Wang, Z.; Castillo, I. Towards artificial intelligence at scale in the chemical industry. *AlChE J.* **2022**, *68*, e17644. [CrossRef]
5. Sahinidis, N.V. Optimization under uncertainty: State-of-the-art and opportunities. *Comput. Chem. Eng.* **2004**, *28*, 971–983. [CrossRef]
6. Sinnott, R.; Towler, G. Separation columns (distillation, absorption and extraction). In *Chemical Engineering Design*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 645–772.
7. Gilliland, E.R. Multicomponent rectification estimation of the number of theoreticalplates as a function of the reflux ratio. *Ind. Eng. Chem. Res.* **1940**, *32*, 1220–1223. [CrossRef]
8. Underwood, A. Fractional Distillation of Multicomponent Mixtures. *Ind. Eng. Chem. Res.* **2002**, *41*, 2844–2847. [CrossRef]
9. Fenske, M.; Quiggle, D.; Tongberg, C. Composition of straight-run pennsylvania gasoline. *Ind. Eng. Chem. Res.* **1949**, *24*, 408–418. [CrossRef]
10. Dragomir, R.M.; Jobson, M. Conceptual design of single-feed hybrid reactive distillation columns. *Chem. Eng. Sci.* **2005**, *60*, 4377–4395. [CrossRef]
11. Adiche, C.; Vogelpohl, A. Short-cut methods for the optimal design of simple and complex distillation columns. *Chem. Eng. Res. Des.* **2011**, *89*, 1321–1332. [CrossRef]
12. Uwitonze, H.; Han, S.; Kim, S.; Hwang, K.S. Structural design of fully thermally coupled distillation column using approximate group methods. *Chem. Eng. Process. Process Intensif.* **2014**, *85*, 155–167. [CrossRef]
13. Zhang, X.; Song, Z.; Zhou, T. Rigorous design of reaction-separation processes using disjunctive programming models. *Comput. Chem. Eng.* **2018**, *111*, 16–26. [CrossRef]
14. Viswanathan, J.; Grossmann, I.E. A combined penalty function and outer-approximation method for MINLP optimization. *Comput. Chem. Eng.* **1990**, *14*, 769–782. [CrossRef]
15. Yeomans, H.; Grossmann, I.E. Optimal design of complex distillation columns using rigorous tray-by-tray disjunctive programming models. *Ind. Eng. Chem. Res.* **2000**, *39*, 4326–4335. [CrossRef]
16. Jackson, J.R.; Grossmann, I.E. A disjunctive programming approach for the optimal design of reactive distillation columns. *Comput. Chem. Eng.* **2001**, *25*, 1661–1673. [CrossRef]
17. Caballero, J.A.; Milán-Yañez, D.; Grossmann, I.E. Rigorous design of distillation columns: Integration of disjunctive programming and process simulators. *Ind. Eng. Chem. Res.* **2005**, *44*, 6760–6775. [CrossRef]
18. Tsatse, A.; Oudenhoven, S.R.G.; ten Kate, A.J.B.; Sorensen, E. Optimal design and operation of reactive distillation systems based on a superstructure methodology. *Chem. Eng. Res. Des.* **2021**, *170*, 107–133. [CrossRef]
19. Pattison, R.C.; Baldea, M. Equation-oriented flowsheet simulation and optimization using pseudo-transient models. *AlChE J.* **2014**, *60*, 4104–4123. [CrossRef]
20. Pattison, R.C.; Gupta, A.M.; Baldea, M. Equation-oriented optimization of process flowsheets with dividing-wall columns. *AlChE J.* **2015**, *62*, 704–716. [CrossRef]
21. Ma, Y.; Luo, Y.; Yuan, X. Towards the really optimal design of distillation systems: Simultaneous pressures optimization of distillation systems based on rigorous models. *Comput. Chem. Eng.* **2019**, *126*, 54–67. [CrossRef]
22. Yeoh, K.P.; Hui, C.W. Rigorous NLP distillation models for simultaneous optimization to reduce utility and capital costs. *Cleaner Eng. Technol.* **2021**, *2*, 100066. [CrossRef]
23. Grossmann, I.E.; Aguirre, P.A.; Barttfeld, M. Optimal synthesis of complex distillation columns using rigorous models. *Comput. Chem. Eng.* **2005**, *29*, 1203–1215. [CrossRef]
24. Osuolale, F.N.; Zhang, J. Energy efficiency optimisation for distillation column using artificial neural network models. *Energy* **2016**, *106*, 562–578. [CrossRef]

25. Javaloyes-Antón, J.; Ruiz-Femenia, R.; Caballero, J.A. Rigorous design of complex distillation columns using process simulators and the particle swarm optimization algorithm. *Ind. Eng. Chem. Res.* **2013**, *52*, 15621–15634. [CrossRef]
26. Caballero, J.A. Logic hybrid simulation-optimization algorithm for distillation design. *Comput. Chem. Eng.* **2015**, *72*, 284–299. [CrossRef]
27. Yusup, N.; Zain, A.M.; Hashim, S.Z.M. Evolutionary techniques in optimizing machining parameters: Review and recent applications. *Expert Syst. Appl.* **2012**, *39*, 9909–9927. [CrossRef]
28. Mavrovouniotis, M.; Li, C.; Yang, S. A survey of swarm intelligence for dynamic optimization: Algorithms and applications. *Swarm Evol. Comput.* **2017**, *33*, 1–17. [CrossRef]
29. Escamilla-Salazar, I.G.; Torres-Trevi no, L.; Gonzalez-Ortiz, B. Intelligent parameter identification of machining Ti64 alloy. *Int. J. Adv. Manuf. Technol.* **2016**, *86*, 1997–2009. [CrossRef]
30. Aydin, I.; Karakose, M.; Akin, E. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Appl. Soft Comput.* **2011**, *11*, 120–129. [CrossRef]
31. Li, M.; Cui, Y.; Shi, X.; Zhang, Z.; Zhao, X.; Zhu, X.; Gao, J. Simulated annealing-based optimal design of energy efficient ternary extractive dividing wall distillation process for separating benzene-isopropanol-water mixtures. *Chin. J. Chem. Eng.* **2021**, *33*, 203–210. [CrossRef]
32. Christopher, C.C.E.; Dutta, A.; Farooq, S.; Karimi, I.A. Process synthesis and optimization of propylene/propane separation using vapor recompression and self-heat recuperation. *Ind. Eng. Chem. Res.* **2017**, *56*, 14557–14564. [CrossRef]
33. Ibrahim, D.; Jobson, M.; Li, J.; Guillén-Gosálbez, G. Surrogate models combined with a support vector machine for the optimized design of a crude oil distillation unit using genetic algorithms. In Proceedings of the 27th European Symposium on Computer Aided Process Engineering, Computer Aided Chemical Engineering, Barcelona, Spain, 1–5 October 2017; Volume 3, pp. 481–486.
34. Salehi, H.; Burgueño, R. Emerging artificial intelligence methods in structural engineering. *Eng. Struct.* **2018**, *171*, 170–189. [CrossRef]
35. Ochoa-Estopier, L.M.; Jobson, M.; Smith, R. The use of reduced models for design and optimisation of heat-integrated crude oil distillation systems. *Energy* **2014**, *75*, 5–13. [CrossRef]
36. Ochoa-Estopier, L.M.; Jobson, M. Optimization of heat-integrated crude oil distillation systems. Part I The distillation model. *Ind. Eng. Chem. Res.* **2015**, *54*, 4988–5000. [CrossRef]
37. 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. *Chem. Eng. Sci.* **2019**, *197*, 235–245. [CrossRef]
38. Quirante, N.; Javaloyes, J.; Caballero, J.A. Rigorous design of distillation columns using surrogate models based on Kriging interpolation. *AlChE J.* **2015**, *61*, 2169–2187. [CrossRef]
39. Nentwich, C.; Engell, S. Application of surrogate models for the optimization and design of chemical processes. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 1291–1296.
40. Yao, H.; Chu, J. Operational optimization of a simulated atmospheric distillation column using support vector regression models and information analysis. *Chem. Eng. Res. Des.* **2012**, *90*, 2247–2261. [CrossRef]
41. López, C.D.C.; Hoyos, L.J.; Mahecha, C.A.; Arellano-Garcia, H.; Wozny, G. Optimization model of crude oil distillation units for optimal crude oil blending and operating conditions. *Ind. Eng. Chem. Res.* **2013**, *52*, 12993–13005. [CrossRef]
42. Liau, L.C.-K.; Yang, T.C.-K.; Tsai, M.-T. Expert system of a crude oil distillation unit for process optimization using neural networks. *Expert Syst. Appl.* **2004**, *26*, 247–255. [CrossRef]
43. Tgarguifa, A.; Bounahmidi, T.; Fellaou, S. Optimal design of the distillation process using the artificial neural networks method. In Proceedings of the 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 16–19 April 2020.
44. Ibrahim, D.; Jobson, M.; Li, J.; Guillén-Gosálbez, G. Optimization-based design of crude oil distillation units using surrogate column models and a support vector machine. *Chem. Eng. Res. Des.* **2018**, *134*, 212–225. [CrossRef]
45. Ibrahim, D.; Jobson, M.; Li, J.; Guillén-Gosálbez, G. Optimal design of flexible heat-integrated crude oil distillation units using surrogate models. *Chem. Eng. Res. Des.* **2021**, *165*, 280–297. [CrossRef]
46. Peng, H.; Zhou, L.; Liu, G. Thermodynamics-based neural network and the optimization of ethylbenzene production process. *J. Cleaner Prod.* **2021**, *296*, 126615. [CrossRef]
47. Gutiérrez-Antonioa, C.; Briones-Ramírez, A. Multiobjective stochastic optimization of dividing-wall distillation columns using a surrogate model based on neural networks. *Chem. Biochem. Eng. Q.* **2016**, *29*, 491–504. [CrossRef]
48. Schweidtmann, A.M.; Mitsos, A. Deterministic Global Optimization with Artificial Neural Networks Embedded. *J. Optim. Theory Appl.* **2018**, *180*, 925–948. [CrossRef]
49. Schöneberger, J.C.; Aker, B.; Fricke, A. Explaining and integrating machine learning models with rigorous simulation. *Chem. Ing. Tech.* **2021**, *93*, 1998–2009. [CrossRef]
50. Abdolrasol, M.G.M.; Hussain, S.M.S.; Ustun, T.S.; Sarker, M.R.; Hannan, M.A.; Mohamed, R.; Ali, J.A.; Mekhilef, S.; Milad, A. Artificial neural networks based optimization techniques: A review. *Electronics* **2021**, *10*, 2689. [CrossRef]
51. Liu, L.; Chen, J.; Xu, L. Realization and application research of BP neural network based on MATLAB. In Proceedings of the 2008 International Seminar on Future BioMedical Information Engineering 2008, Wuhan, China, 18–18 December 2008; pp. 130–133. [CrossRef]

52. Subramanyan, K.; Diwekar, U.; Zitney, S.E. Stochastic modeling and multi-objective optimization for the APECS system. *Comput. Chem. Eng.* **2011**, *35*, 2667–2679. [CrossRef]
53. Kůrková, V. Kolmogorov's theorem and multilayer neural networks. *Neural Networks* **1992**, *5*, 501–506. [CrossRef]
54. Brahim, A.O.; Abderafi, S.; Bounahmidi, T. Optimization of the distillation column of qetroleum gractions using ANN method. In Proceedings of the 2018 6th International Renewable and Sustainable Energy Conference (IRSEC), Rabat, Morocco, 5–8 December 2018.
55. Luyben, W.L. Design and control of the cumene process. *Ind. Eng. Chem. Res.* **2010**, *49*, 719–734. [CrossRef]
56. Cui, C.; Liu, S.; Sun, J. Optimal selection of operating pressure for distillation columns. *Chem. Eng. Res. Des.* **2018**, *137*, 291–307. [CrossRef]
57. Abdul-Rahman, O.A.; Munetomo, M.; Akama, K. An adaptive resolution hybrid binary-real coded genetic algorithm. *Artif Life Robotics* **2011**, *16*, 121–124. [CrossRef]
58. Katoch, S.; Chauhan, S.S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimedia Tools Appl.* **2021**, *80*, 8091–8126. [CrossRef]
59. Wang, S.; Liu, C.; Ren, J.; Liu, L.; Li, Q.; Huo, E. Carbon footprint analysis of organic rankine cycle system using zeotropic mixtures considering leak of fluid. *J. Clean. Prod.* **2019**, *239*, 118095. [CrossRef]
60. Yang, A.; Su, Y.; Teng, L.; Jin, S.; Zhou, T.; Shen, W. Investigation of energy-efficient and sustainable reactive/pressure-swing distillation processes to recover tetrahydrofuran and ethanol from the industrial effluent. *Sep. Purif. Technol.* **2020**, *250*, 117210. [CrossRef]
61. Chen, D.; Yuan, X.; Xu, L.; Yu, K.T. Comparison between different configurations of internally and externally heat-integrated distillation by numerical simulation. *Ind. Eng. Chem. Res.* **2013**, *52*, 5781–5790. [CrossRef]