



Optimization of a crude distillation unit using a combination of wavelet neural network and line-up competition algorithm☆



Bin Shi, Xu Yang, Liexiang Yan *

Department of Chemical Engineering, Wuhan University of Technology, Wuhan 430070, China

ARTICLE INFO

Article history:

Received 29 September 2016

Received in revised form 26 December 2016

Accepted 21 March 2017

Available online 21 April 2017

Keywords:

Crude oil distillation

Wavelet neural network

Line-up competition algorithm

Optimization

ABSTRACT

The modeling and optimization of an industrial-scale crude distillation unit (CDU) are addressed. The main specifications and base conditions of CDU are taken from a crude oil refinery in Wuhan, China. For modeling of a complicated CDU, an improved wavelet neural network (WNN) is presented to model the complicated CDU, in which novel parametric updating laws are developed to precisely capture the characteristics of CDU. To address CDU in an economically optimal manner, an economic optimization algorithm under prescribed constraints is presented. By using a combination of WNN-based optimization model and line-up competition algorithm (LCA), the superior performance of the proposed approach is verified. Compared with the base operating condition, it is validated that the increments of products including kerosene and diesel are up to 20% at least by increasing less than 5% duties of intermediate coolers such as second pump-around (PA2) and third pump-around (PA3).

© 2017 The Chemical Industry and Engineering Society of China, and Chemical Industry Press. All rights reserved.

1. Introduction

Distillation of crude oil is regarded to be one of the most fundamental processes in petroleum refining and petrochemical industries, where the crude oil is separated into different products each with specific boiling range. In response to the highly competitive market and stringent environmental laws, improving the operation level of crude oil distillation unit (CDU) is essential. In fact, the optimization of CDU system is beneficial to simultaneously achieve a well-controlled and stable system, high production rate and product quality as well as low operation cost for the economic consideration. Therefore, the engineering design, control strategy and process optimization of a CDU have been paid attention to improve product efficiency and quality assurance in petroleum industry in recent years [1].

As one of the most complicated operations in the field of chemical separation processes, the operation input and output variables among the CDU are highly interacted, which undoubtedly increase the difficulty in obtaining and maintaining the optimal operation condition for CDU. Moreover, the optimal manipulated parameters of CDU have to be frequently adjusted due to the variation of properties of crude oil supplied, and the production mode may be also changed over strategically from season to season. Besides, the oil supply can give rise to some severe problems in plant management or even lead to the

shutdown of the CDU on the condition that the specifications of oil products cannot be reached or the CDU operation is not stabilized. In brief, it is quite necessary to improve operation level of a complex CDU system.

In recent years, the research of CDU operation has been focused on the subject of process control and optimization. Inamdar *et al.* [2] proposed a steady state model based on $(C + 3)$ iteration variables to simulate an industrial CDU. The model was first tuned using industrial data. Then the elitist non-dominated sorting genetic algorithm (NSGA-II) was employed to solve a few meaningful multi-objective optimization problems. Case study has shown that the optimal operating conditions, where the profit could be increased keeping the product properties within acceptable limits were found by the proposed approach. More *et al.* [3] presented the optimization of a crude distillation unit using commercial Aspen Plus software. Optimization model constituted a rigorous simulation model supplemented with suitable objective functions with and without product flow rate constraints. Simulation study inferred that the product flow rate constraints sensitively affect atmospheric distillation column diameter and crude feed flow rate calculations. Based on all simulation studies, a generalized inference confirmed that it was difficult to judge upon the quality of the solutions obtained as far as their global optimality was concerned. Seo *et al.* [4] proposed the design optimization of CDU using a mixed integer non-linear programming (MINLP) method and realized the reduction in energy costs for an existing CDU system. As a meta-model, artificial neural networks (ANN) trained by historical data were also applied to the optimization aspect. With a method of design of experiment (DOE), Chen *et al.* [5] proposed a method using ANN models and information analysis

☆ Supported by the National Natural Science Foundation of China (No. 21376185).

* Corresponding author.

E-mail address: hgxtgc@163.com (L.X. Yan).

for design of experiment (AIDOE) which carried out the experimental or optimization process batch by batch. To maximize a CDU's valuable product yield under required product qualities, Liao *et al.* [6] developed an expert system where ANN trained with production data were used as the knowledge and the optimal conditions were solved with an optimization procedure. Motlaghi *et al.* [7] also designed an expert system for optimizing a crude oil distillation column where neural networks and genetic algorithm (GA) were used. Now, the method combining the data generated by a rigorous model with a meta-model has become a popular method to carry out the optimization of CDU. Yao and Chu [8] employed support vector regression (SVR) to optimize the CDU models constructed by Aspen Plus and revised DOE optimization procedure. Ochoa-Estapier *et al.* [9] simulated the distillation column using an ANN model and then the formulated optimization problem was solved using a simulated annealing (SA) algorithm.

In the terms of theoretical basis, rigorous models are more accurate than simplified and statistical models. Nevertheless, it is difficult to combine an optimization algorithm with rigorous models, because a large number of variables and non-linear equations need to be solved simultaneously [10,11] which will give rise to great computing burdens. In considering the industrial application of petrochemical process optimization, a meta-model with desirable fitting accuracy and generalization is more suitable for optimization calculation. As an emerging tool combining the strengths of discrete wavelet transform with neural network processing, wavelet neural network (WNN) models achieve strong nonlinear approximation ability, and thus have been successfully applied to forecasting [12], modeling and function approximations [13]. Therefore, WNN is proposed to model the CDU unit in this study, which is expected to simulate it accurately and efficiently. When the modeling of CDU is finished, operation optimization of CDU becomes the core problem. In general, for the complex non-linear optimization problem, evolutionary algorithms outperform DOE and SQP in finding global optimal solution. Among various evolutionary algorithms, the line-up competition algorithm (LCA) is a simple and effective stochastic global optimization technique primarily due to its attractive properties, such as parallel evolutionary strategy and asexual reproduction of individuals [14]. Based on these advantages, LCA is employed to find the best operation conditions of CDU in this work.

The aim of this study is to investigate the feasibility of combining WNN methodology with LCA in the area of optimizing operations in CDU.

2. WNN Model

The procedure for constructing the data-driven WNN model, as presented below, consists of three main steps. The first step is the construction of samples used as the knowledge database for the WNN model. The second step is the selection of the WNN structure and parameters; the third step is performing the training of the WNN model.

2.1. Basic WNN

Wavelet neural network, inspired by both the feed-forward neural networks and wavelet decompositions has received considerable attention and become a powerful tool for function approximation [15]. The main characteristic of WNN is that some kinds of wavelet functions are used as the nonlinear activation function in the hidden layer in place of the usual sigmoid function. Incorporating the time–frequency localization properties of wavelets and the learning abilities of general neural network, WNN has shown its advantages over other methods such as BPNN for complex nonlinear system modeling [16].

The basic wavelet theory is as follows.

For any function $\Psi(t) \in L^2(R)$ if it satisfies the admissibility condition [17]:

$$C_\Psi = \int_{-\infty}^{\infty} \frac{|\Psi(t)|^2}{|t|} dt < \infty \quad (1)$$

where $\Psi(t)$ is the mother wavelet, a double parameter family of wavelets created by translating and dilating this mother wavelet:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \left(\frac{t-b}{a} \right) \quad (2)$$

where a is the dilation parameter and b is the translation parameter. They can be used to control the magnitude and position of $\Psi(t)$.

In this work, a three-layer feed-forward wavelet neural network is designed, it has one input layer, one hidden layer and a linear output layer as shown in Fig. 1.

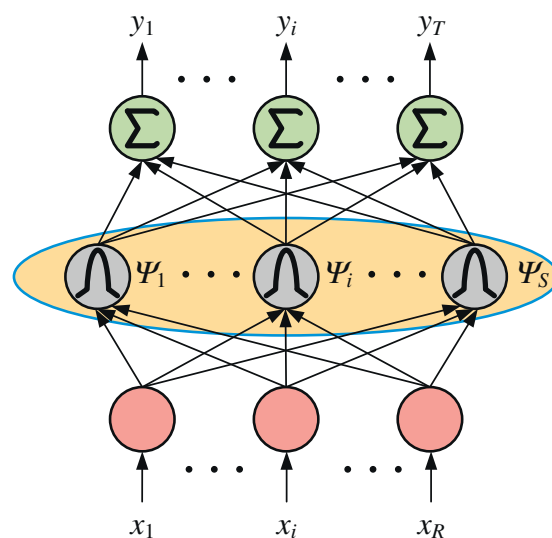


Fig. 1. WNN structure.

The hidden layer output is given by

$$H^q = \mathbf{W}^1 \mathbf{X}^q + t^1 \quad (3)$$

where $\mathbf{X}^q = [x_1^q, x_2^q, \dots, x_m^q]^T$ is the q th vector of input samples, and $q = 1, 2, \dots, Q$. The nodes of input and output layer are set in accordance with the training data. Hidden layer has h neurons. The connection weights from the input layer to the hidden layer are $h \times m$ matrix $\mathbf{W}^1$, and the threshold of hidden layer is a $h \times 1$ array t^1 . The WNN output is given by

$$\mathbf{y}^q = \mathbf{W}^2 \Psi \left(\frac{H^q - b}{a} \right) + t^2 \quad (4)$$

where $\mathbf{y}^q = [y_1^q, y_2^q, \dots, y_n^q]^T$ is the q th corresponding vector of expected output, and $q = 1, 2, \dots, Q$. The connection weights from the hidden layer to the output layer are $n \times h$ matrix $\mathbf{W}^2$, while $n \times 1$ array t^2 is the threshold of output layer. It should be noticed that the superscript in $\mathbf{W}^1$, $\mathbf{W}^2$, t^1 and t^2 note the layers of WNN. Note that the above WNN is a kind of basic neural network in the sense that the wavelets consist of the basis function, therefore, the scalar parameter and the translation parameter would be determined by a training algorithm.

In order to take full advantage of the local capacity of the wavelet basic functions, the performance of WNN which has one hidden layer

of neurons is measured by total error function, which is described as follows:

$$E = \sum_{q=1}^Q \sum_{k=1}^n (e_k^q)^2 \quad (5)$$

where $e_k^q = Y_k^q - D_k^q$, Y_k^q is the k th component in the q th network output and D_k^q is the k th component in the q th network expected output.

The training process aims to find a set of optimal network parameters. In the previous work, the training of WNN is achieved by the ordinary back propagation technique. According to the gradient method, the parameters are tuned by

$$\mathbf{w}_{i,j}^1(l+1) = \mathbf{w}_{i,j}^1(l) + \eta \frac{\partial E}{\partial \mathbf{w}_{i,j}^1} \quad (6)$$

$$t_i^1(l+1) = t_i^1(l) + \eta \frac{\partial E}{\partial t_i^1} \quad (7)$$

$$a_i(l+1) = a_i(l) + \eta \frac{\partial E}{\partial a_i} \quad (8)$$

$$b_i(l+1) = b_i(l) + \eta \frac{\partial E}{\partial b_i} \quad (9)$$

$$\mathbf{w}_{i,j}^2(l+1) = \mathbf{w}_{i,j}^2(l) + \eta \frac{\partial E}{\partial \mathbf{w}_{i,j}^2} \quad (10)$$

$$t_i^2(l+1) = t_i^2(l) + \eta \frac{\partial E}{\partial t_i^2} \quad (11)$$

where η is the learning rate and l is the current iteration numbers.

Remark 1. The Morlet wavelet function $\Psi_M(x)$ is often considered as a “mother wavelet” in the hidden nodes of the WNN,

$$\Psi_M(x) = \cos(1.5x)e^{-x^2/2} \quad (12)$$

We consider different activation functions, e.g. the Sigmoid shown as follows:

$$\Psi_S(x) = \frac{1}{1 + e^{-x}} \quad (13)$$

and the Gauss function is defined as

$$\Psi_G(x) = e^{-x^2} \quad (14)$$

that are often selected in the hidden nodes of some neural network (NN) frameworks. Moreover, the profiles of activation functions by using Morlet, Sigmoid and Gauss are depicted in Fig. 2(a), (b) and (c), respectively.

2.2. Modified WNN

The WNN is trained comparatively slowly while calculating the samples one by one without any numerical optimization method. Recently, some work was devoted by introducing evolutionary algorithm, such as particle swarm optimization (PSO) [18,19], to initialize the parameters of WNN for accelerating the training process of WNN. However, the number of parameters in a practical WNN is up to dozens or even hundreds, which is difficult for evolutionary algorithm to carry out the optimization. In our approach, the numerical optimization algorithm, namely Levenberg–Marquardt (LM) algorithm, is introduced into the training process to accelerate convergence of WNN parameters. At the same time the training mode is changed into batch mode, which adjusts

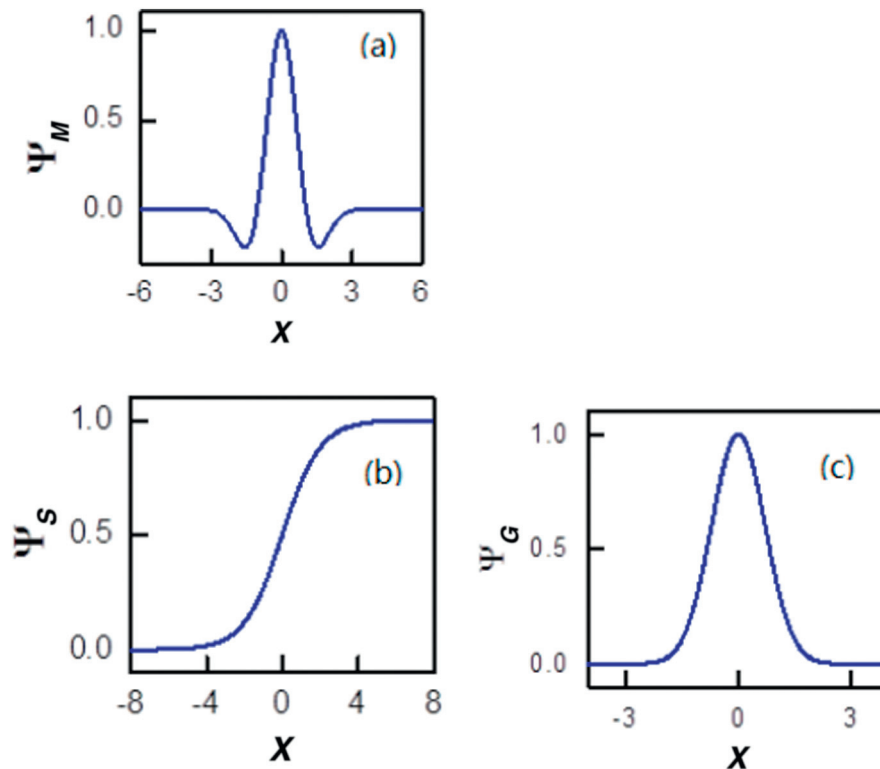


Fig. 2. Profiles of (a) Morlet, (b) Sigmoid, and (c) Gauss function.

the parameters of WNN by calculating all the samples. Referring to the Levenberg–Marquardt algorithm, the updating law for the matrix $\mathbf{W}^1$ and $\mathbf{W}^2$, the arrays t^1 , t^2 , a , and b are shown as follows:

$$\mathbf{s}_{l+1} = \mathbf{s}_l - [\mathbf{J}^T(\mathbf{s}_l)\mathbf{J}(\mathbf{s}_l) + \mu_l \mathbf{I}^{-1}]^{-1} \mathbf{J}^T(\mathbf{s}_l)\mathbf{v}(\mathbf{s}_l) \quad (15)$$

with

$$\mu_{l+1} = \theta \mu_l \text{ if } E_{l+1} < E_l \quad (16)$$

$$\mu_{l+1} = \mu_l / \theta \text{ if } E_{l+1} \geq E_l \quad (17)$$

where $\theta > 1$ is an adjusted parameter and $\mathbf{I}$ is a unit matrix. The vectors $\mathbf{s}$ and $\mathbf{v}$ are defined as

$$\mathbf{s}^T = [w_{1,1}^1, \dots, w_{h,m}^1, t_1^1, \dots, t_h^1, a_1, \dots, a_h, b_1, \dots, b_h, w_{1,1}^2, \dots, w_{n,h}^2, t_1^2, \dots, t_n^2] \quad (18)$$

$$\mathbf{v}^T = [e_1^1, \dots, e_n^1, e_1^2, \dots, e_n^2, \dots, e_1^Q, \dots, e_n^Q] \quad (19)$$

and the Jacobian matrix of the network is given as

$$\mathbf{J}(\mathbf{s}) = \begin{bmatrix} \frac{\partial e_1^1}{\partial w_{1,1}^1} & \frac{\partial e_1^1}{\partial w_{1,2}^1} & \dots & \frac{\partial e_1^1}{\partial w_{h,m}^1} & \frac{\partial e_1^1}{\partial t_1^1} & \dots \\ \frac{\partial e_2^1}{\partial w_{1,1}^1} & \frac{\partial e_2^1}{\partial w_{1,2}^1} & \dots & \frac{\partial e_2^1}{\partial w_{h,m}^1} & \frac{\partial e_2^1}{\partial t_1^1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_n^1}{\partial w_{1,1}^1} & \frac{\partial e_n^1}{\partial w_{1,2}^1} & \dots & \frac{\partial e_n^1}{\partial w_{h,m}^1} & \frac{\partial e_n^1}{\partial t_1^1} & \dots \\ \frac{\partial e_1^2}{\partial w_{1,1}^1} & \frac{\partial e_1^2}{\partial w_{1,2}^1} & \dots & \frac{\partial e_1^2}{\partial w_{h,m}^1} & \frac{\partial e_1^2}{\partial t_1^1} & \dots \\ \frac{\partial e_1^2}{\partial w_{1,1}^1} & \frac{\partial e_1^2}{\partial w_{1,2}^1} & \dots & \frac{\partial e_1^2}{\partial w_{h,m}^1} & \frac{\partial e_1^2}{\partial t_1^1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (20)$$

It is noted that $\mathbf{s}$ is solved by Eq. (15), elements of $\mathbf{s}$ are allocated to the $\mathbf{W}^1$, $\mathbf{W}^2$, t^1 , t^2 , a and b which are used to recalculate the total error E . The parameter μ is updated by Eqs. (16) and (17).

Remark 2. Comparing with the back propagation technique for training the network, the parametric updating law by Eqs. (15)–(17) is used to enhance the convergence rate of the iteration while the parameter μ is updated by adjusting θ . Regarding the BPNN model using Sigmoid function and the Gauss function in RBFNN, they usually use the gradient method to update the weights.

3. Process Optimization

If the proposed WNN model can precisely predict the operational model of real CDU, then it can assist to simplify the complicated process model and improve the solvability of the constrained optimization problem.

3.1. Optimization model

Referring the optimization issue for the CDU, the constrained optimization model for the CDU process is formulated as

$$\max_x J = \sum_{j=1}^{N_p} C_{\text{prod},j} F_{\text{prod},j} - C_s F_s \quad (21)$$

subject to

$$y = \Phi(x, \mathbf{s}) \quad (22)$$

$$\begin{aligned} mv_{lb} &\leq x \leq mv_{ub} \\ ps_{lb} &\leq y \leq ps_{ub} \end{aligned} \quad (23)$$

where the objective function J represents the net revenue. $C_{\text{prod},j}$ and $F_{\text{prod},j}$ are the price and flow rate of product j respectively. The amount of the steam, F_s used in the process operation multiple its price, C_s , represents the energy cost. Φ is the WNN model of CDU modeled by our approach, and this model is a combination of nonlinear algebraic equations where the vector parameter $\mathbf{s}$ is estimated by the updating law by Eqs. (15)–(20). mv_{lb} and mv_{ub} are the lower and upper bounds of process inputs x . ps_{lb} and ps_{ub} are the lower and upper bounds of process outputs y . Constraints by Eq. (23) are specified according to the real process operating conditions.

3.2. Line-up competition algorithm

Regarding the above WNN-based optimization model, the LCA algorithm [20,21] is adopted to solve this constrained optimization problem. In the LCA, independent and parallel evolutionary families are always kept during evolution, each family producing offspring only by asexual reproduction. There are the two levels of competition in the algorithm. One is the survival competition inside a family. The best one of each family survives in each generation. The other level is the competition between families. According to their values of objective function, families are ranked to form a line-up. The best family is located in the first position in the line-up, while the worst family is put in the final position. The families of different positions have different driving forces of competition. The driving force of competition may be understood as the power of impelling family mutation. By the above two levels of the competitions, the first family in the line-up is replaced continually by other families, accordingly the value of its objective function is continually updated. As a result, the optimal solution can be approached rapidly.

The above two levels of competition in the algorithm can be illustrated in Fig. 3. It is seen that a two-dimensional search space is occupied by four families, each consisting of five members. Afterwards, all the members in each family compete with each other. The member having the best objective value is chosen as the candidate of this family to strive for a better position in the next line-up.

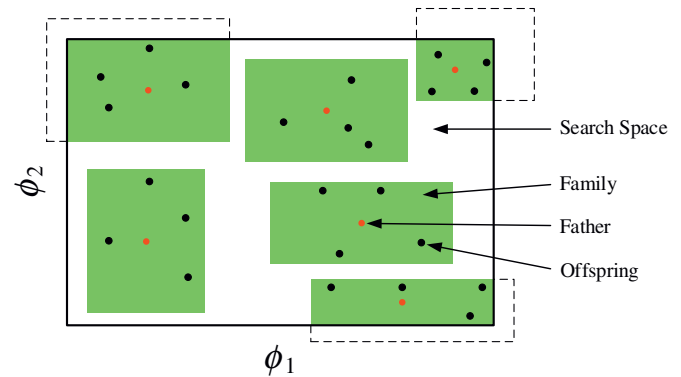


Fig. 3. Mapping diagram of LCA.

The LCA includes mainly the four operating processes: reproduction, ordering, allocation of the search space and contraction of the search space. The calculation steps are detailed as follows:

Step 1. Assign the numbers of evolutionary generation, individual and family, N_g , N_i and N_f , respectively. Initialize the starting evolutionary generation counter g as 1.

Step 2. Uniformly and dispersedly generate N_f individuals, so-called families, to form the initial population.

$$\{P_{g,1}, P_{g,2}, \dots, P_{g,N_f}\}, \quad g = 1 \quad (24)$$

The f th individual in the g th generation consists of N_c decision variables as follows:

$$P_{g,f} = [p_{g,f}^1, p_{g,f}^2, \dots, p_{g,f}^{N_c}] \quad (25)$$

Therein, the value of $P_{g,f}$ can be assigned randomly between L_0 and U_0

$$p_{g,f}^c = L_{g,f}^c + \lambda \Delta_{g,f}^c, \quad g = 1, c = 1, \dots, N_c, f = 1, \dots, N_f \quad (26)$$

where $\Delta_{g,f}^c \equiv mv_{ub}^c - mv_{lb}^c$ ($g=1$) represents the initial scale of the search interval for c th decision variable; λ is a random number that ranges from 0 to 1. The scale vector of the f th individual is thus defined as $\Delta_{g,f} \equiv [\Delta_{g,f}^1, \dots, \Delta_{g,f}^{N_c}]$.

Step 3. Compute the corresponding fitness value for each individual.

$$Y_{g,f} = \Theta(P_{g,f}), \quad f = 1, \dots, N_f \quad (27)$$

Step 4. According to the fitness values, the individuals are ranked to form a line-up. For the problems of global minimum, the line-up is an ascending sequence. Otherwise, for the problems of global maximum, the line-up is a descending sequence. Sort the individuals. For the minimization problem, the individuals will be sorted in ascending order. Conversely, the individuals will be sorted in descending order for the maximization problem. The sorted individuals are expressed below:

$$(\bar{P}_{g,f}, \bar{Y}_{g,f}), \quad f = 1, \dots, N_f \quad (28)$$

Step 5. Allocate the associated search space proportionally for each individual according to their position in the line-up. The first one in the line-up will be allocated the smallest sub-space, while the

last one will be allocated the largest sub-space. The lower bound $L_{g,f}^c$ and upper bound $U_{g,f}^c$ of the c th decision variable in the sub-space is calculated by

$$L_{g,f}^c = \max\left(p_{g,f}^c - \frac{f \Delta_{g,f}^c}{2N_f}, L_0^c\right) \quad (29)$$

$$U_{g,f}^c = \min\left(p_{g,f}^c + \frac{f \Delta_{g,f}^c}{2N_f}, U_0^c\right) \quad (30)$$

where $\bar{\Delta}_{g,f} \equiv [\bar{\Delta}_{g,f}^1, \bar{\Delta}_{g,f}^2, \dots, \bar{\Delta}_{g,f}^{N_c}]$ is the new size variable after ordering.

Step 6. Through asexual reproduction based on the most diversity, each individual, so-called father, reproduces N_i offspring within its search space. The manner that the offspring are produced is same with the way in Step 3.

$$(\bar{P}_{g,f,i}, \bar{Y}_{g,f,i}), \quad i = 1, \dots, N_i \quad (31)$$

Step 7. For the f th individual, the N_i offspring together with their father compete with each other, and the best one survives as father in the next generation.

$$(P_{g,f}, Y_{g,f}) = \max\{\bar{Y}_{g,f}, \bar{Y}_{g,f,i}\}, \quad f = 1, 2, \dots, N_f, i = 1, 2, \dots, N_i \quad (32)$$

Step 8. Contract the largest sub-space

$$\Delta_{g+1} \leftarrow \beta \Delta_g \quad (33)$$

where β is the contraction factor which can be set between 0 and 1. If $g < N_g$, then go back to Step 6, or else stop the iteration.

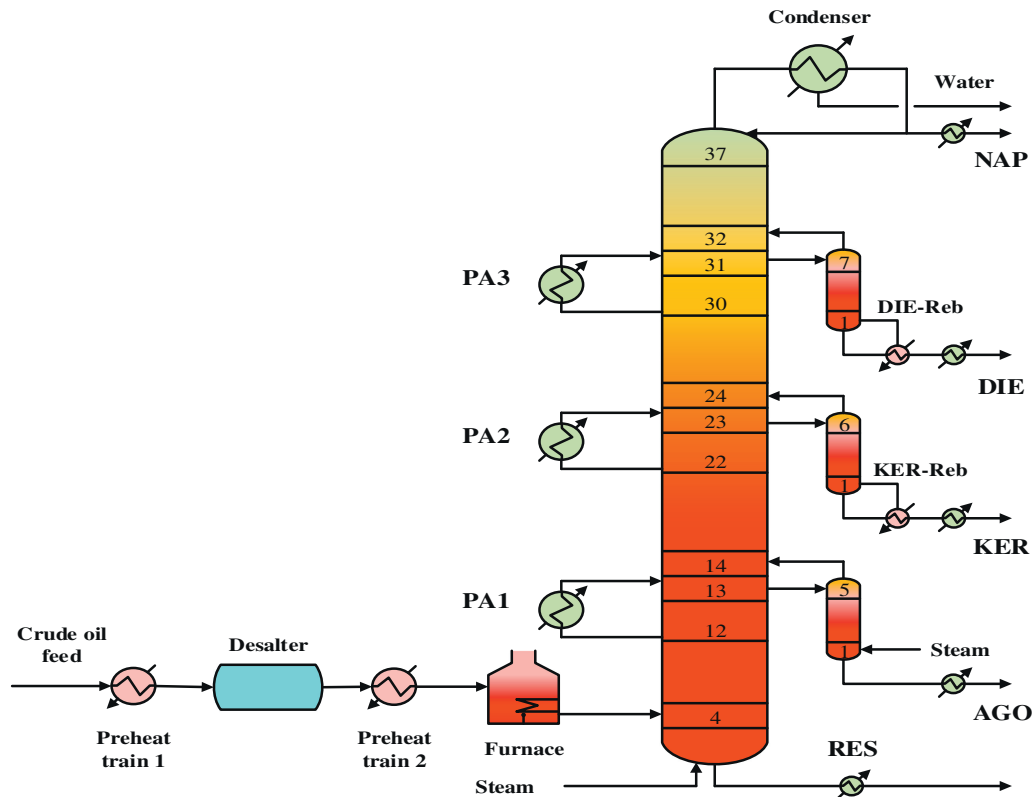


Fig. 4. Crude distillation unit (CDU).

Table 1
Input specifications of the CDU process

Item	Lower bound	Upper bound	Base case	Units
Column bottom temperature	359	366	365.03	°C
Column bottom steam	1.5	3.5	2.79	t·h ⁻¹
DIE flow rate	20	36	24.46	t·h ⁻¹
KER flow rate	22	36	29.25	t·h ⁻¹
AGO flow rate	90	110	94.91	t·h ⁻¹
PA1 flow rate	230	270	241.90	t·h ⁻¹
PA2 flow rate	220	260	246.02	t·h ⁻¹
PA3 flow rate	320	360	355.66	t·h ⁻¹
RES flow rate	–	–	355.07	t·h ⁻¹
Furnace duty	–	–	31.75	MW
PA1 duty	–	–	4.83	MW
PA2 duty	–	–	6.10	MW
PA3 duty	–	–	11.56	MW

It is very important to choose a set of appropriate control parameters for decreasing computing time and increasing quality of solution. The LCA includes three parameters in all: population size (N_f), number

of reproduction (N_i) for each family in each generation and contraction factor (β).

Larger N_f and N_i provide generally high quality solution, but may result in a longer computing time. Small ones can speed up the convergence rate, but may result in trapping in a local minimum. We have to trade between the computing time and solution quality.

Contraction factor influences strongly on solution quality and computation time. Based on our computing experiences, for a difficult problem, the global optimal solution can be obtained only when $0.9 < \beta < 0.99$.

4. Case Study

4.1. Process description

Referring to the specifications of a CDU system in a real refinery in Wuhan, China, the crude oil at 40 °C and 300 kPa with flow rate of 569.6 t·h⁻¹ (702.2 m³·h⁻¹) is fed into the CDU, which consists of the preheat train, the main tower, one condenser, three pumparounds (PA1, PA2, PA3) and three side strippers. Steam at 300 kPa and 400 °C

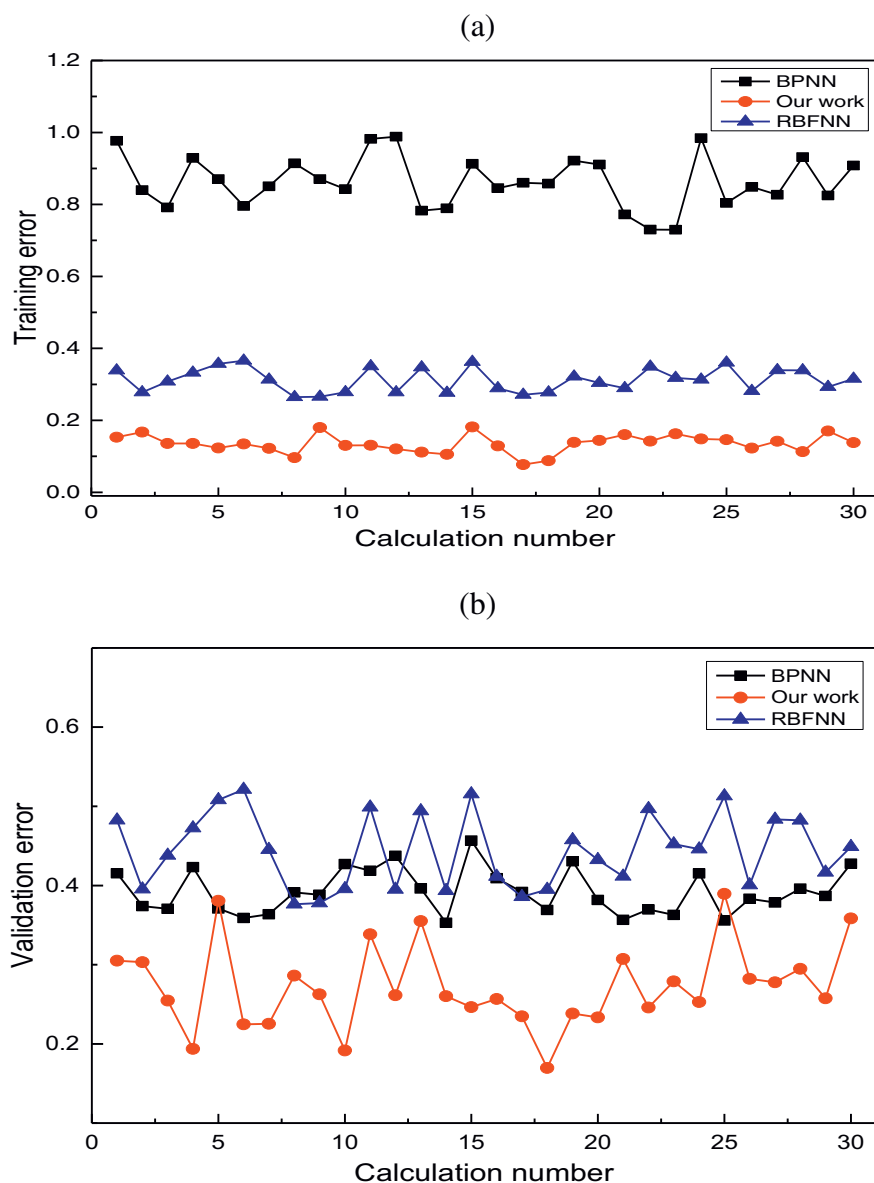


Fig. 5. Identification of CDU using BPNN, modified WNN and RBFNN: (a) training errors, (b) validation errors.

is used as a stripping agent in the main column and strippers. Five products including naphtha (NAP), diesel (DIE), kerosene (KER), atmospheric gas oil (AGO), and residue (RES) are exhausted at different stages. Fig. 4 shows the CDU system, which can be simulated in Aspen Plus environment. The process inputs (x) include the steam flow rate and temperature of steam at the bottom of the column, flow rates of DIE, KER, AGO, PA1 (first pump-around), PA2 (second pump-around), and PA3 (third pump-around). The process outputs (y) are ASTM D86 100% point of DIE, ASTM D86 95% point of AGO, RES of CDU, furnace duty, and duties of PA1, PA2 and PA3.

4.2. Identification

The independent variables are randomly varied between their upper and lower bounds to ensure the full exploration of the search space. Table 1 shows the upper and lower bounds of the independent variables. The bounds of each variable is specified according to the real process operating conditions of the CDU. 500 feasible scenarios, in the sense of leading converged simulation, were generated to build WNN distillation column model. The purpose of this case study is mainly to enhance the profitability of the CDU process via optimizing its operation. WNN structure is created to finish the modeling and identification of the CDU. To validate the approximation ability of WNN, BPNN and RBFNN are also constructed to model the same CDU system for comparison. The structure of NN comprises 30 neurons in the hidden layer. 350 scenarios of 500 converged simulation scenarios are used to train the three networks, and the rest 150 scenarios are used to validate the trained networks. A comparison of the identification performance (training and validation) of the modified WNN using Morlet function (our work), the BPNN using Sigmoid function and the RBFNN using Gauss function is depicted in Fig. 5(a) and (b), respectively. Apparently, the training and validation errors of the modified WNN are smaller than other approaches. The modified WNN, BPNN, and RBFNN use the similar network structure, which contains one input layer, one hidden layer and one output layer. It is verified that the activation function embedded in the hidden layer, which extraordinarily affects the ability of approximating the complex nonlinear system and extracting the nonlinear characteristics of it. In this case, the sigmoid function used in BPNN is not orthogonal which may lead the slow convergence rate. However, the Morlet function in WNN is orthogonal which reduces the redundant part. Moreover, Fig. 6 shows that the identified WNN model provides

Table 2

Output specifications of the CDU process

Property	Lower specification	Upper specification	Base case	Units
NAP D86_5	70	85	78.30	°C
DIE D86_5	150	170	159.08	°C
KER D86_5	180	200	188.37	°C
AGO D86_5	240	260	246.1	°C
NAP D86_100	160	180	174.23	°C
DIE D86_100	200	220	211.94	°C
KER D86_100	–	300	273.77	°C
AGO D86_95	–	360	333.50	°C

Table 3

Feed, product and utility prices

Item	Price	Units
Naphtha	6800.0	CNY · t ⁻¹
Diesel	6900.0	CNY · t ⁻¹
Kerosene	6300.0	CNY · t ⁻¹
Atmospheric gas oil	6100.0	CNY · t ⁻¹
Residue	4000.0	CNY · t ⁻¹
Industrial power	0.989	CNY · kW ⁻¹ · h
Stripping steam (300 °C, 400 kPa)	752.27	CNY · t ⁻¹

the high accuracy to predict the outputs of the CDU process as compared with the rigorous model in Aspen Plus.

4.3. Optimization

Based on the identified WNN model, the input/output specifications of the CDU process in Tables 1 and 2 and the prices of the products and the operating cost list in Table 3 are taken into account for the constrained optimization problem. A comparison of the LCA, GA and PSO for solving the same problem has been done here. For fair comparison, function evaluations in each iteration of the three algorithms are set to 150, the corresponding other detailed parameter settings for the algorithms are given in Table 4. Fig. 7 shows the results of the three algorithms. It is clear that the values of the objective (J) by using LCA, GA and PSO increase very fast in beginning few generations. At the same generations, the profit predictions reveal that the value of LCA is higher

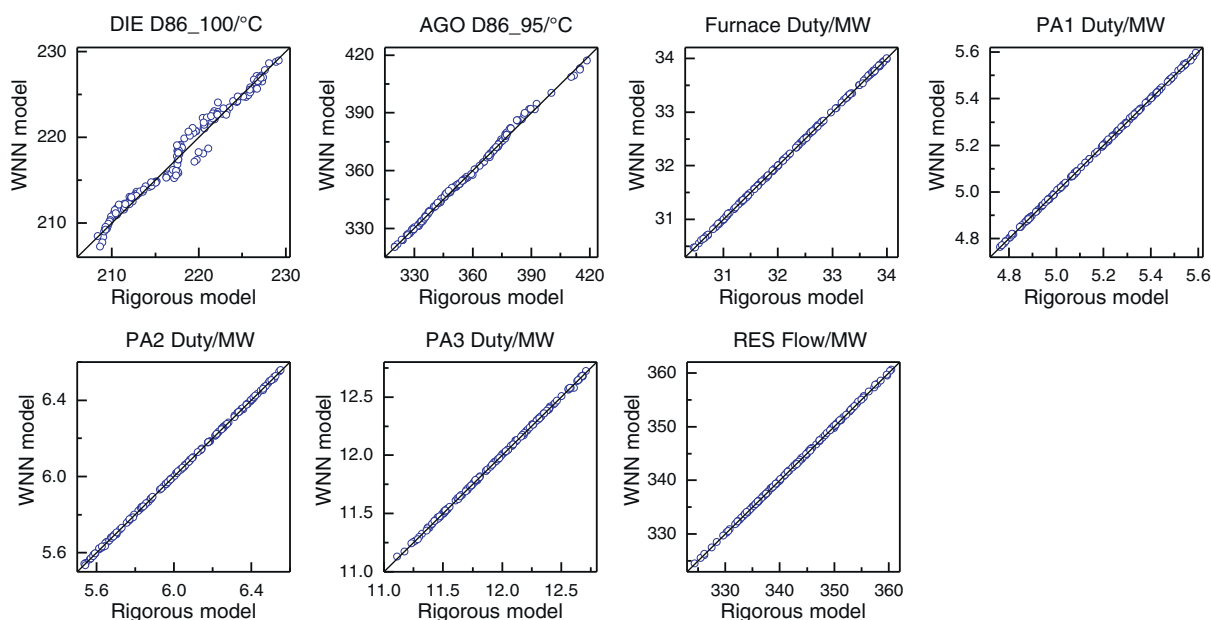


Fig. 6. Validation of CDU using the modified WNN.

Table 4
Detailed parameter settings for LCA, GA and PSO

Algorithm	Parameter	Value	Evaluations/generation
LCA	Family number	10	150
	Offspring number	15	
GA	Population number	150	150
PSO	Particle number	150	150

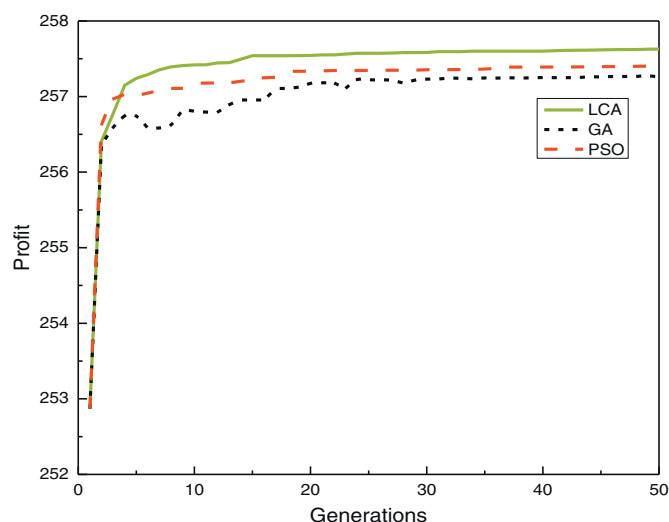


Fig. 7. Profit predictions of CDU using LCA, GA and PSO.

than those by GA and PSO. Moreover, Table 5 indicates that all input/output differences between the WNN model and the model in Aspen Plus are less than 0.54%. It is verified that the optimal operating conditions obtained by the WNN-based optimization approach are reliable. The input and output patterns of the CDU with regard to base and optimal conditions are shown in Fig. 8(a) and (b), respectively. As compared with the base conditions, the optimal operation increases the production of diesel, kerosene, and atmospheric gas oil by 22%, 25% and 10%, respectively. The corresponding duties of furnace, PA1, PA2 and PA3 increase 10%, 17%, 8%, and 3%, respectively. Apparently, the performance of the CDU is improved by increasing a few duties of coolers. Consequently, the proposed approach based on WNN and LCA can reduce energy consumption in regard to the increments of oil products. In addition, by introducing different operation and property constraints into the optimization model, the new operational scheme with different product distributions can be obtained easily.

Table 5
Comparisons of the WNN model and process model at optimal operating condition

Item	Optimal results	Verified results	Error/%
DIE D86_100/°C	219.98	219.26	0.32
AGO D86_95/°C	359.99	361.95	0.54
RES flow rate/t·h ⁻¹	335.21	334.76	0.13
Furnace duty/MW	33.79	33.79	0.00
PA1 duty/MW	5.56	5.57	0.18
PA2 duty/MW	6.53	6.51	0.31
PA3 duty/MW	12.06	12.08	0.17
Profit × 10 ⁻⁴ /CNY	257.63	257.45	0.07

Optimal results: the optimal operating conditions predicted by our WNN-based optimization model.

Verified results: the actual simulation results obtained by the CDU model built in Aspen Plus for verifying optimal operating conditions obtained by the proposed WNN-based optimization model.

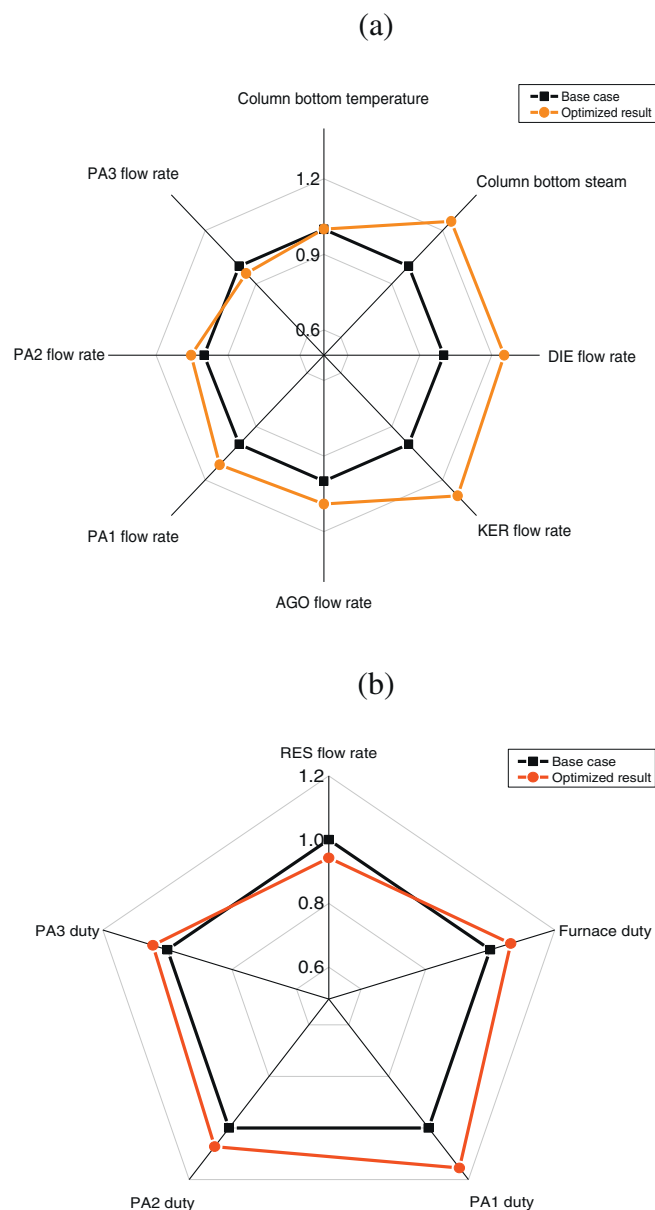


Fig. 8. Radar plots for comparisons of base and optimal conditions of CDU: (a) output flow rates of coolers and products, (b) duties of coolers and furnace.

5. Conclusions

This study proposed a methodology using the combination of WNN-based optimization model and LCA to model and optimize the operation of crude distillation unit. The main results of this article are summarized as follows:

- (1) A WNN model of CDU is constructed, where the *Levenberg–Marquardt* algorithm is introduced into the WNN to speed up the training procedure.
- (2) Based on the WNN model of CDU, an economic optimization model for crude oil distillation process is built under prescribed constraints.
- (3) A practical framework combined with WNN-based optimization model and LCA is presented for optimizing the complex operation of non-linear CDU.

Case study result has shown that the optimal operating condition obtained by the proposed approach can increase the yield of high valuable

products and reduce the energy consumption as compared with those in base operating conditions, therefore, increasing the total profits of the CDU.

Nomenclature

a	dilation parameter
b	translation parameter
$C_{\text{prod},j}$	prices of products, $\text{CNY} \cdot \text{t}^{-1}$
C_s	prices of stripping steam, $\text{CNY} \cdot \text{t}^{-1}$
D_k^q	k th component in the q th network expected output
$\mathbf{D}^q$	vector of q th corresponding vector of expected output
E	error function
$\mathbf{e}^q$	vector of q th sample error
$F_{\text{prod},j}$	flow rates of products, $\text{t} \cdot \text{h}^{-1}$
F_s	flow rates of stripping steam, $\text{t} \cdot \text{h}^{-1}$
H^q	hidden layer output
h	number of neurons in the hidden layer
$\mathbf{I}$	unit matrix
$\mathbf{J}(\mathbf{s})$	Jacobian matrix of the network
L_0	lower bounds of variables in LCA
$L^2(R)$	Lebesgue square integrable function
m	input number of neural network
mv_{lb}	lower bounds of manipulated parameters
mv_{ub}	upper bounds of manipulated parameters
N_g	numbers of evolutionary generation
N_i	numbers of individual in each family
N_f	numbers of family in each evolutionary generation
n	output number of neural network
obj	constrained objective function, 1×10^4 CNY
P	population in LCA
$\tilde{P}$	newly generated population
ps_{lb}	lower bounds of product specifications, $^{\circ}\text{C}$
ps_{ub}	upper bounds of product specifications, $^{\circ}\text{C}$
Q	number of input samples
RBFNN	radial basis function neural network
$\mathbf{S}$	reshaped vector of network parameters
t	threshold of neurons
U_0	upper bounds of variables in WNN
W	weight of connected neurons
$\mathbf{X}^q$	q th vector of input samples
x	manipulated parameters
$\mathbf{Y}$	fitness value of individual
$\mathbf{Y}^q$	vector of q th network output
Y_k^q	k th component in the q th network output
$\tilde{Y}$	newly calculated individual fitness
y	ASTM D86 point of specified products, $^{\circ}\text{C}$
β	contraction factor in LCA
Δ	scale of the search interval
η	learning rate
θ	factor in Levenberg–Marquardt algorithm
λ	random number that ranges from 0 to 1

μ	parameter in Levenberg–Marquardt algorithm
Φ	formulation of WNN model
$\Psi(t)$	wavelet base function

Subscripts

g	evolutionary generation counter
l	current iteration number
lb	lower bound
prod	product
S	steam
ub	upper bound

References

- [1] A. Mizoguchi, T.E. Marlin, A.N. Hrymak, Operations optimization and control design for a petroleum distillation process, *Can. J. Chem. Eng.* 73 (1995) 896–907.
- [2] S.V. Inamdar, S.K. Gupta, D.N. Saraf, Multi-objective optimization of an industrial crude distillation unit using the elitist non-dominated sorting genetic algorithm, *Chem. Eng. Res. Des.* 82 (2004) 611–623.
- [3] R.K. More, V.K. Bulasara, R. Uppaluri, V.R. Banjara, Optimization of crude distillation system using aspen plus: effect of binary feed selection on grass-root design, *Chem. Eng. Res. Des.* 88 (2010) 121–134.
- [4] J.W. Seo, M. Oh, T.H. Lee, Design optimization of a crude oil distillation process, *Chem. Eng. Technol.* 23 (2000) 157–164.
- [5] J. Chen, D.S.H. Wong, S.S. Jang, S.L. Yang, Product and process development using artificial neural-network model and information analysis, *AIChE J.* 44 (1998) 876–887.
- [6] L.C.K. Liau, T.C.K. Yang, M.T. Tsai, Expert system of a crude oil distillation unit for process optimization using neural networks, *Expert Syst. Appl.* 26 (2004) 247–255.
- [7] 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 (2008) 1540–1545.
- [8] H. Yao, J. Chu, Operational optimization of a simulated atmospheric distillation column using support vector regression models and information analysis, *Chem. Eng. Res. Des.* 90 (2012) 2247–2261.
- [9] L.M. Ochoa-Estopier, M. Jobson, R. Smith, Operational optimization of crude oil distillation systems using artificial neural networks, *Comput. Chem. Eng.* 59 (2013) 178–185.
- [10] K. Basak, K.S. Abhilash, S. Ganguly, D.N. Saraf, On-line optimization of a crude distillation unit with constraints on product properties, *Ind. Eng. Chem. Res.* 41 (2002) 1557–1568.
- [11] J.C.M. Hartmann, Determine the optimum crude intake level: A case history, *Hydrocarb. Process.* 80 (2001) 77–84.
- [12] H. Chitsaz, N. Amjadi, H. Zareipour, Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm, *Energy Convers. Manag.* 89 (2015) 588–598.
- [13] Q. Zhang, A. Benveniste, Wavelet networks, *IEEE Trans. Neural Netw.* 3 (1992) 889–898.
- [14] L.X. Yan, D.X. Ma, Global optimization of non-convex nonlinear programs using line-up competition algorithm, *Comput. Chem. Eng.* 25 (2001) 1601–1610.
- [15] J. Zhang, G.G. Walter, Y. Miao, W.N.W. Lee, Wavelet neural networks for function learning, *IEEE Trans. Signal Process.* 43 (1995) 1485–1497.
- [16] S. Billings, H.L. Wei, A new class of wavelet networks for nonlinear system identification, *IEEE Trans. Neural Netw.* 16 (2005) 862–874.
- [17] I. Daubechies, Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [18] S. Chi, Character recognition based on wavelet neural network optimized with PSO algorithm, *Appl. Mech. Mater.* 602 (2014) 1834–1837.
- [19] Y. Lu, N. Zeng, Y. Liu, N. Zhang, A hybrid Wavelet Neural Network and Switching Particle Swarm Optimization algorithm for face direction recognition, *Neurocomputing* 155 (2015) 219–224.
- [20] L.X. Yan, Solving combinatorial optimization problems with line-up competition algorithm, *Comput. Chem. Eng.* 27 (2003) 251–258.
- [21] L.X. Yan, K. Shen, S. Hu, Solving mixed integer nonlinear programming problems with line-up competition algorithm, *Comput. Chem. Eng.* 28 (2004) 2647–2657.