Asian Journal of Chemical Sciences

12(1): 54-67, 2022; Article no.AJOCS.89365
ISSN: 2456-7795

Application of Artificial Neural Network in Chemical Separation Process

Mengyu Wang [a*] and Zhu Zhu [a]

[a] *East China University of Science and Technology, Shanghai-200237, China.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJOCS/2022/v12i119140

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/89365

Review Article

Received 10 May 2022
Accepted 14 July 2022
Published 21 July 2022

ABSTRACT

The chemical separation process is a necessary unit operation to obtain qualified purity products. The design and operation control of the separation device are very important. Due to the numerous parameters that affect the separation efficiency and product purity, it is very difficult to optimize by traditional methods. Artificial neural network (ANN) has strong fault tolerance because of its self-learning, self-organization, self-adaptive and strong nonlinear function approximation ability. ANN can be used to map the complex nonlinear relationship between dependent variables and independent variables, and can be used for the design and control of chemical separation devices. This paper summarizes the characteristics of artificial neural network and several important artificial neural network models, and discusses the application of artificial neural network in different chemical separation processes.

Keywords: Artificial neural network; chemical separation; distillation column; seawater desalination; wastewater treatment.

1. INTRODUCTION

Separation devices often account for about 80% of the total investment of a chemical plant [1,2]. The process optimization and control of the separation process are crucial to the utilization efficiency of equipment and the improvement of product quality [3-5]. Due to the numerous parameters that affect the separation efficiency and product purity, it is very difficult to optimize

**Corresponding author: E-mail: Y82210069@mail.ecust.edu.cn;*

by traditional methods[6]. “Artificial neural network (ANN) has been a research hotspot in the field of artificial intelligence since 1980s. It can fully approximate complex nonlinear relations, has strong robustness and fault tolerance, can process a large number of complex data” [7]. It can realize simulation, binary image recognition[8-10], prediction and fuzzy control [11-13]. It is a powerful tool for dealing with nonlinear systems. Therefore, it has great application in chemical separation process [14,15]. In this paper, the characteristics of ANN are introduced and several important artificial neural network models are summarized, and the application of ANN in different chemical separation processes are analyzed and discussed.

2. ARTIFICIAL NEURAL NETWORK

2.1 Overview of Artificial Neural Networks

“Artificial neural network is the abstraction and simulation of some basic characteristics of human brain or natural neural network, and it is based on the physiological research results of the brain. The research of artificial neural network can be traced back to the McCulloch psychologist and Pitts mathematician who considered to find the basic principle behind neurons in 1943” [16]. “Threshold function was regarded as the main characteristic of computational neurons, logical calculus was expressed as neural computing architecture, and the concept of "neural network" and M-P model were proposed. It marks the germination of artificial neural network ANN. In the 1980s, after obtaining feasible algorithms for artificial neural networks and the traditional algorithms based on Von Neumann system showed their weakness in knowledge processing, people began to show interest in artificial neural networks again, leading to the revival of neural networks” [17]. At present, there are many schools of neural network research methods, and the most fruitful research work includes: BP algorithm of multi-layer network [18,19], Hopfield network model [20,21], RBE network model [22,23], nonparallel training approaches [24,25], adaptive resonance theory [26], self-organizing feature mapping theory [27] and so on. “These theories have greatly promoted the development of artificial neural networks. In recent years, with the continuous deepening of the research on artificial neural network, not only the software technology is relatively mature, but also the hardware has been greatly developed, improving the information processing ability of artificial neural network system” [28].

2.2 Characteristics of Artificial Neural Network

“Artificial neural network combines many advantages of biological neural network, so it has its inherent characteristics: high parallelism and nonlinearity, good fault tolerance and associative memory function, and very strong self-organization, self-adaptation and self-learning function” [29].

2.2.1 High degree of parallelism and nonlinearity

Artificial neural network (Ann) is a highly parallel nonlinear system composed of a large number of simple neurons. Although the processing function of each neuron is very limited, the parallel activities of a large number of neurons make the network present rich functions and have high speed. The parallelism of the structure makes the information storage of the network adopt the distributed mode, that is, the information is not stored in a part of the network, but distributed in all the connection weights of the network. A neural network can store many kinds of information, in which the connection weight of each neuron stores only a part of many kinds of information. The inherent parallelism and distribution of neural network is reflected in the parallel storage and processing of information in space distribution and time. In the network, neurons are in two different states of activation or suppression. This behavior shows a nonlinear relationship in mathematics. The extensive interconnection and parallel work of neurons will inevitably make the whole network show highly nonlinear characteristics.

2.2.2 Good fault tolerance and associative memory function

The structural characteristics of distributed storage of information make neural network show good fault tolerance in two aspects: on the one hand, due to the distributed storage of information, the damage of some neurons in the network will not affect the overall performance of the system; On the other hand, when fuzzy, incomplete or deformed information is input, neural network can recover the memory through association, so as to realize the correct recognition of incomplete input information. When the stored knowledge needs to be acquired, the neural network uses the method of "association" to recall under the excitation of input information, so it has the function of association memory. Neural network can process continuous analog signals and inaccurate and incomplete fuzzy information, so it gives the optimal solution rather than the exact one. This kind of association can be realized by the feedback network of artificial neural network.

2.2.3 Self-organization, self-adaptation and self-learning functions

Under the stimulation of the external environment, the nervous system can adjust the strength of synaptic connections between neurons according to certain rules, and gradually build the neural network. This construction process is called self-organization of the network (or reconstruction).The self-organization ability of neural network is related to self-adaptability, which is realized by self-organization. Adaptability refers to the ability of a system to change its own performance to adapt to environmental changes, which is an important characteristic of neural network. Adaptability includes self-learning and self-organization. Self-learning of neural network refers to that when the external environment changes, after a period of training or perception, neural network can automatically adjust the network structure parameters to produce the desired output for a given input. Training is the way of neural network learning, so people often use the words learning and training interchangeably.

3. ARTIFICIAL NEURAL NETWORK MODEL

3.1 MLP Neural Network

Single layerperceptron (neuron model) is the most basic model in neural network, but perceptron cannot solve complex nonlinear problems. Therefore, multi-layer perceptron (MLP) neural network emerged [30,31]. MLP neural network is also known as multi-layer feed-forward neural network. In the network, neurons at two adjacent layers cascade with each other, while nodes at the same layer are not connected with each other. The first layer is the input layer, the last layer is the output layer, and one or more hidden layers can be included in the middle to form a fully connected neural network. MLP network is the most basic neural network, and its activation function usually selects Sigmoid function. The multilayer perceptron model uses enough hidden layers to express more complex nonlinear functions to achieve any pattern classification. However, the multi-layer perceptron model did not reasonably solve how to determine the weight (learning algorithm), which hindered the development of neural network until the appearance of back-propagation (BP) neural network and applied to the training of multi-layer perceptron [32].

3.2 BP Neural Network

“The full name of BP neural network is back-propagation network.It is a kind of multi-layer forward feedback neural network. Its name comes from the backward propagation learning algorithm, namely BP learning algorithm, which is used to adjust the network connection weights. A typical BP network is a three-layer network, including input layer, hidden layer and output layer” [33]. Its structure is shown in Fig.1.

The main characteristic of BP neural network is that the signal propagates forward while the error propagates backward. The process of BP neural network is mainly divided into two stages. The first stage is the forward propagation of signal, from the input layer through the hidden layer, and finally to the output layer. The second stage is the back propagation of errors, from the output layer to the hidden layer, and finally to the input layer, adjusting the weight and bias of the hidden layer to the output layer, and the weight and bias of the input layer to the hidden layer in turn. The basic content of BP algorithm is to get the minimum value of error function. Its basic idea is gradient descent method, which uses gradient search technology to minimize the mean square error of the actual output value and the expected output value of the network. This kind of algorithm has the advantages of low computation, simple and easy operation and good parallelism. This is a relatively mature algorithm used in the current neural network training.

BP neural network is one of the most widely used networks at present. In terms of water quality detection, Zhao[34] et al. constructed salinity and ammonia nitrogen models of water plants in the Tidal reach of Minjiang River based on BP artificial neural network model and scenario analysis method, and quantitatively simulated the influence mechanism and countermeasures of salt tide invasion in water plants in the tidal reach of Minjiang River. On the whole, BP neural

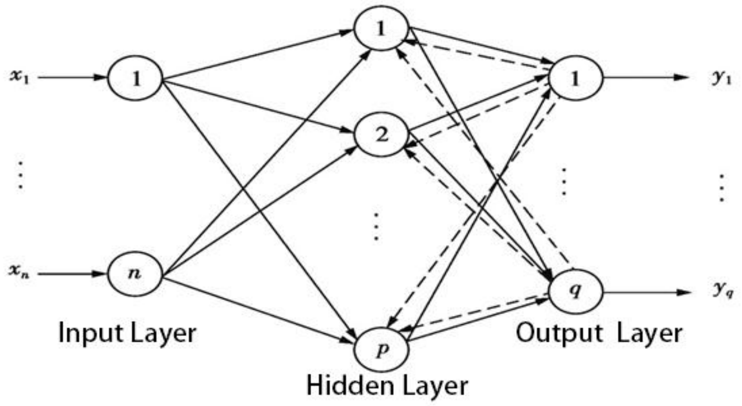

Fig. 1. BP neural network structure

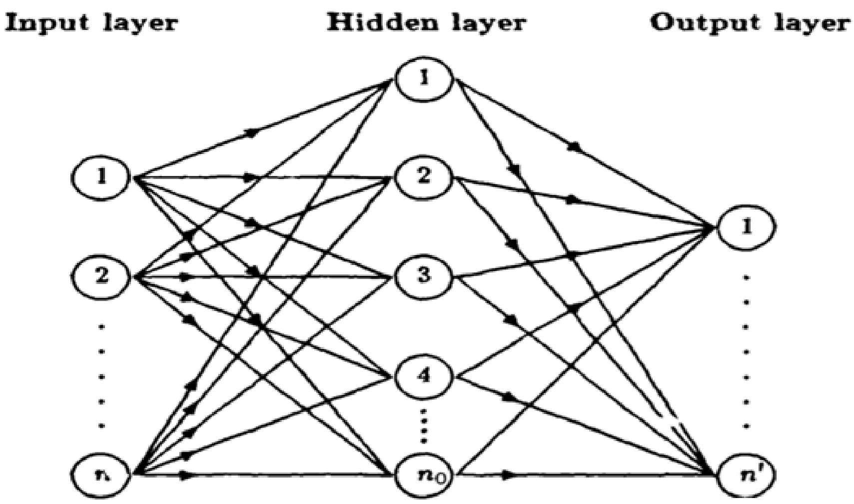

Fig. 2. RBF neural network structure

network model has a good effect on salinity simulation of tidal reach of Minjiang River basin, and can better simulate the change trend of ammonia nitrogen. This shows that the neural network model has good applicability for salinity simulation of tidal river reach in Minjiang River Basin. Similarly, Wang et al.[35] found in the study of Qiantang River that the neural network model can effectively simulate water quality changes in the estuary area. In terms of stock prediction, the three-layer BP neural network established by Chu et al. [36] has good data processing ability, and its convergence speed is faster than previous algorithms with higher accuracy. When the selected data are reasonable and have good properties, the fitting effect will be more accurate, so as to get more accurate prediction of stock price in a short time. It shows that BP neural network is feasible and reasonable to predict stock price. Therefore, it can not only get the law of stock price to nonlinear, but also show that BP artificial neural network has a strong fitting effect in practical life application.

3.3 RBF Neural Network

"Radial basis function (RBF) neural network is a feedforward neural network. It is a kind of 3-layer forward network: the first layer is the input layer, which is composed of signal source nodes. The second layer is the hidden layer. The transformation function of the hidden element is a locally distributed nonnegative nonlinear function, which is radial symmetric and attenuated to the central point. The number of hidden layers is determined by the needs of the problem described. The third layer is the output layer, and the output of the network is the linear weighting of the implicit unit output. The transformation from the input space to the hidden layer space of RBF neural network is nonlinear, while the transformation from the hidden layer space to the output layer space is linear" [37]. Its structure is shown in Fig. 2.

Compared with BP artificial neural network, RBF neural network is superior to BP neural network in approaching ability, classification ability and learning speed. RBF neural network has simple structure, simple training, fast learning convergence speed, and can approach any nonlinear function and overcome the local minimum problem. Therefore, RBF artificial neural network has been widely used in recent years. BP network is widely used in enterprise credit evaluation. However, when BP network is used for function approximation, the weight adjustment adopts negative gradient descent method. This method of weight adjustment has its limitations, such as slow convergence speed and local minimum. RBF network is superior to BP network in approaching ability, classification ability and learning speed. Therefore, people began to use RBF network for credit evaluation, and combined with "5C evaluation method" to create a network model. Through network training and testing with a small sample size, it is shown that RBF artificial neural network has good performance in enterprise credit evaluation [38].

In terms of housing construction engineering prediction [39], "people seek the non-linear relationship between the quota index and the actual construction and installation engineering cost of final settlement, so as to establish the RBF artificial neural network estimation model, and verify and analyze the results, providing a new idea and method for engineering cost estimation". In the MATLAB environment, the selected cost factors are taken as the input vector of the model training sample set, and the main cost indicators are taken as the output vector. After the sample set data is normalized, the RBF artificial neural network is called to train the data, and the generalized RBF artificial neural network model of project cost prediction is constructed.Finally, the model is verified by the actual sample. The comparison between the predicted value and the actual value of the cost index shows that the prediction accuracy of the model is reliable. Cai et al. [40] in sitting position evaluation application, based on RBF neural network position description model was applied, through building sitting pressure acquisition experiment platform, using MATLAB for data normalization processing, calculation, model simulation, training a set of good learning ability and fault tolerance of sitting position state neural network. Compared with the conventional sitting posture evaluation method, the application method can not only reduce the difficulty of data acquisition, but also improve the calculation efficiency of sitting posture evaluation by 5%, which provides a theoretical basis for the design of sitting posture monitoring system.

4. APPLICATION OF ARTIFICIAL NEURAL NETWORK IN CHEMICAL SEPARATION PROCESS

4.1 Energy Efficiency Optimization of Distillation Column

Distillation towers continue to grow in importance in the traditional petrochemical industry. Their critical role in the petrochemical industry and the quest to make them more energy efficient has made the separation process a top priority for the industry. Distillation units pose great challenges to process engineers because of their complexity. Studies on the thermal efficiency of crude oil distillation column show that the energy and exergic loss of the column is very high[41], and the total efficiency of the column ranges from 5% to 23%[42].In order to achieve energy saving while meeting product quality constraints, the optimization of distillation column operation is very important.

At present, domestic and foreign research teams begin to use neural network to optimize the energy efficiency of distillation system.Neural network models have been proved to be able to approximate arbitrary continuous nonlinear functions. Osuolale et al. [43] used neural network to model exergic efficiency and product composition of distillation system. Then, the exergic efficiency was optimized under product quality constraints by using neural network model. Osuolale built a bootstrap aggregation neural network (BANN) consisting of 30 neural networks to predict exergic efficiency of each system.Each individual network has a hidden layer. Levenberg-marquardt training algorithm is used to train the network. By using bootstrap polymerization neural network model, the accuracy of the model is improved. The research shows that the artificial neural network can accurately simulate the energy efficiency of distillation column from the process operation data. Then, the neural network model was used to obtain the optimal distillation operation conditions, so as to maximize the energy efficiency of the distillation system, while maintaining product quality and yield. Osuolale optimized the artificial neural network model, and BANN enhanced the accuracy of model prediction.Modeling and optimization based on ANN and BANN models can help formulate energy saving operations and distillation column control.

4.2 Seawater Desalination

At present, due to limited natural resources, population growth, industrialization and climate change, the lack of fresh water has become a major problem facing the world today [44].About three quarters of the earth's surface is water, but only one percent is fresh.A separation process is needed to desalinate seawater. Membrane separation processes account for nearly 65% of the global water treatment workload, while other treatment methods account for 35%[45]. The membrane processes of seawater desalination include reverse osmosis, forward osmosis and membrane distillation.Microfiltration, ultrafiltration and nanofiltration.The establishment of mathematical model is a necessary process for the simulation and optimization of seawater desalination process. Compared with traditional models, artificial neural network has the following advantages: it can model complex nonlinear functions with high precision; Capable of multiple input multiple output (MIMO) modeling; Can handle complex and incomplete data; Less computing time; Ability to update/train models with new data.

Al-shayji et al. [46], 2002, simulated the Jeddah desalination plant in Saudi Arabia (56800m^3 per day).They built a multi-layer feedforward neural network model using BP algorithm to find the best prediction between experimental data and simulated data. The architecture consists of two hidden layers, the first containing 30 nodes and the second containing 15 nodes. They found that the model was very effective for predicting plant capacity and was able to deal with nonlinear and complex problems.They also found that such models manage complex data more effectively than statistical regression methods.

Zhao et al. [47], 2005 proposed a neural network model for accurate prediction of osmotic TDS in nanofiltration (NF)/ reverse osmosis (RO) hybrid devices.They used experimental data (from March 12, 1998 to March 16, 1999) from nine million gallon per day level 2 plants in Palm Beach, Florida. MLP neural network and NORMALIZED radial basis function (NRBF) network were used for comparative analysis. MLP is a feedforward neural network that uses a SigmoID activation function, or hyperbolic tangent.The following is a simple MLP model in which a hidden layer contains two hidden neurons. The general form of the feedforward neural network represents a linear combination of nonlinear functions of the desired target (Cp) converted to a linear combination of inputs (here feed water quality Cf, system flux (Fw) and recovery (R)).

$$g0^{-1}\big(E(C_P)\big) = w_0 + w_1 H_1 + w_2 H_2$$

$$H_1 = \tan h\big(w_{01} + w_{11} C_f + w_{21} F_w + w_{31} R\big)$$

$$H_2 = \tan h\big(w_{02} + w_{12} C_f + w_{22} F_w + w_{32} R\big)$$

Where g0^{-1}(E(Cp)) is the inverse transformation of the desired target and the output activation function, equal to the combination function (linear combination here) of the parameters of the activation function.Where, w1, W11 and w12 represent the weight, and W0, W01 and W02 represent the deviation of the data estimated by the fitting model.

The normalized Radial basis function (NRBF) network is a single hidden layer feedforward network using softmax activation function, which is suitable for the radial combination of inputs.In contrast to MLP, each basis function is the ratio of the bell-shaped Gaussian surface to the sum of the Gaussian surfaces.The NRBF model contains two hidden neurons

Data for this study came from the U.S. Environmental Protection Agency's Information Collection Rules (ICR) database on nine MGD two-stage nanofiltration plants located in Palm Beach County, Florida.Model development and validation using data from full-scale membrane plants will benefit future water communities.

Experimental data from March 12, 1998 to March 16, 1999 provided more than 8800 hours of run-time observations and 706 TDS observations for Phase 1 (354 observations) and Phase 2 (352 observations).The statistical operation data of flow rate, recovery rate, net driving force, feed flow and osmotic flow TDS are summarized in Table 1.

Table 1. Summary of water quality and operating condition

	Feed TDS (mg/L)	PermeateTDS (mg/L)	Flux (gsfd)	Recovery	NDF (psi)
	Stage1				
Mean	341.5	50.4	13.6	55.2%	81.3
S.D.	15.2	6.9	0.4	0.1%	1.3
Min.	301.3	42.0	12.5	54.0%	72.4
Max.	587.0	71.0	14.9	56.5%	84.8
	Stage2				
Mean	720.2	115.0	13.2	66.5%	46.1
S.D.	9.7	7.1	0.4	0.3%	1.2
Min.	690.0	104.0	12.1	65.2%	42.7
Max.	747.0	164.0	14.4	69.9%	49.0

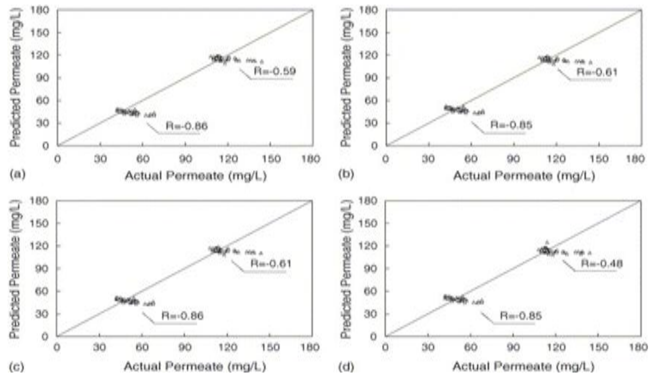

Fig. 3. Model predicted vs. actual permeate TDS concentration (mg/L) for the HSDM (a), IHSDM (b), IDM (c) and IMOP (d) models.

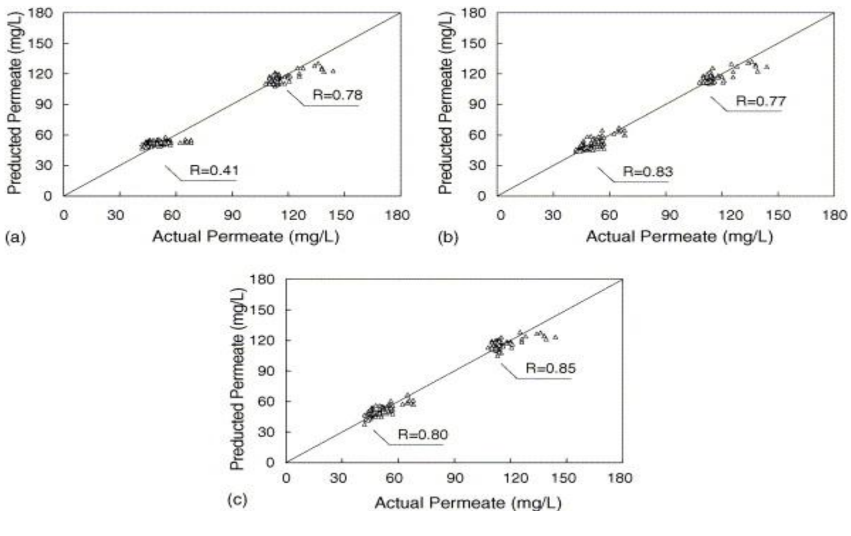

Fig. 4. Actual vs. predicted TDS using data not used for model development for the HHSDM (a), MLP (b) and NRBSEQ (c) models.

Fig. 3. shows a stage-to-stage comparison of actual and predicted osmotic flow TDS for HSDM, IHSDM, IDM, and IMOP models using independently validated data sets. The 45° line in Figure 1 is a curve representing the actual comparison predicted TDS. Predicted TDS and actual TDS are divided into stages 1 and 2. A high R2 indicates that these models can be used effectively for prediction, but a negative R at each stage indicates that the model can be improved. For low TDS observations, the model overestimates the actual TDS, while for high TDS

observations, the model underestimates the actual TDS, which is a common error of solute mass transfer prediction models in reverse osmosis or nanofiltration membranes. Of course, these models can be used, but clearly they can be improved.

The actual and predicted TDS verified by the mixed model, MLP model, and NRBSEQ model are shown in Fig. 4. The three models improved the grouping of actual and predicted TDS.

The correlation coefficients were all positive. The NRBSEQ model had the highest R^2.In addition, paired sample tests showed that there was no significant difference between the osmotic concentration predicted using any model and the actual osmotic concentration.

Assuming that TDS in the feed flow of hybrid, MLP and NRBFEQ models is 1000mg/L, the prediction of flux and recovery by Cp is shown in Fig. 5. The Cp prediction plane of the mixed model is very smooth, and there is no significant Cp change for varying flux and recovery. However, the mixed model has the lowest prediction accuracy for Cp.The prediction of Cp by MLP model is almost like a step function. This trend can be clearly seen in Fig.5(b).Similar changes can be seen in the later part of the Cp plane predicted in Fig.5(b). Compared with THE MLP model, the Cp changes predicted by NRBFEQ model were obvious but more stable. NRBFEQ predicted an increase in Cps recovery from 50mg/L at 80% to 130mg/L at 90%.Model development using actual data from an extended range of independent variables is likely to result in changes in the neural network model. The neural network model established in this study is based on SAS Enterprise Miner software.

The results show that the hybrid model and the artificial neural network model are more accurate than any diffusion-based model in predicting infiltration TDS, and will not exceed or fall below the infiltration TDS at low and high infiltration TDS. Although artificial neural network models are not based on mechanical principles, these models do predict significant changes in penetration TDS under certain flux and recovery combinations.A large number of hybrid and neural network models can be developed and prediction of membrane performance can be significantly improved.

Therefore, Salami et al. [48] proposed a BP neural network model in 2016 to simulate 8 kinds of SWRO membranes for seawater desalination.The established model is used to simulate and estimate the output TDS corresponding to the input flow rate, temperature and recovery rate. The data set was generated by ROSA software. It involves a BP algorithm, feedforward structure, and two hidden layers with variable nodes.The model results are in good agreement with the measured data. It is concluded that BP artificial neural network can be used in the simulation calculation of seawater desalination process, and the results obtained are more accurate.

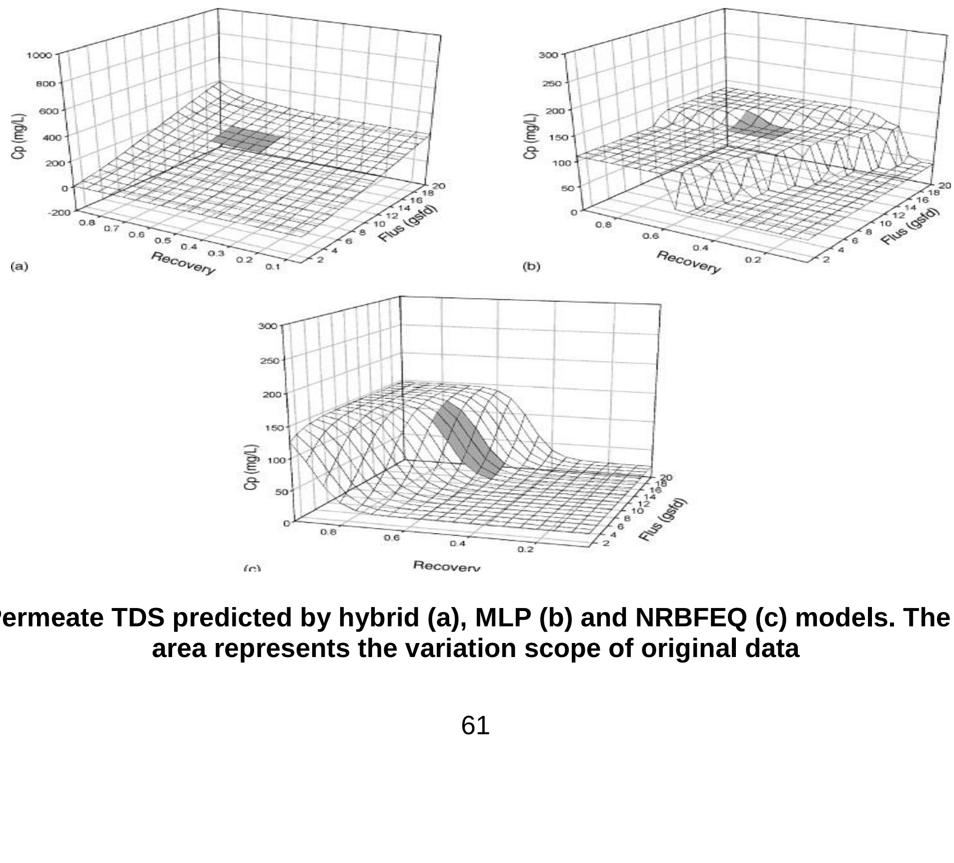

Fig. 5. Permeate TDS predicted by hybrid (a), MLP (b) and NRBFEQ (c) models. The shaded area represents the variation scope of original data

4.3 Application of Artificial Neural Network in Wastewater Treatment

Phenolic compounds, especially the highly toxic chlorophenol, are one of the common pollutants found in the wastewater of various chemical industries. Even at very low concentrations, they can have a significant impact on human health [49]. Therefore, several treatment methods have been developed to reduce phenolic compounds in industrial wastewater. Among them, oxidation, adsorption, distillation and membrane separation are the most effective methods to remove phenolic compounds in wastewater [50]. In particular, reverse osmosis (RO) process has been widely used to treat phenolic compounds in wastewater and achieved great success [51]. Artificial neural networks are widely used in modeling complex processes due to their high accuracy [52]. Artificial neural network can predict several complex processes including reverse osmosis according to its nonlinear characteristics.

Salgado-reyna et al.[53] studied “a four-layer feedforward network with back propagation algorithm and established a BP artificial neural network model, which was composed of two hidden layers with 4 and 3 neurons respectively. The algorithm is used to analyze the observation data of osmotic flow rate in the pilot ro membrane treatable wastewater treatment plant under continuous flow and different operating conditions”. It is proved that there is a good agreement between model predictions and available data. Salam et al. [54] successfully achieved “the development of artificial neural network models for different types of reverse osmosis membranes under certain operating conditions. In addition to the influence of experimental data, the influence of hidden layers and the number of neurons should also be considered in model prediction. The comparison between the experimental value and the predicted value of the model shows that the error is very small”. The study shows that the reliability and accuracy of the BP artificial neural network model can be used in wastewater treatment.

Mohammad et al. [55] innovatively combined artificial neural network and genetic algorithm, and developed the first comprehensive neural network model that used artificial neural network and genetic algorithm to predict the pilot-scale reverse osmosis process's chlorophanol removal amount in wastewater. The ANN model has two structures (4-2-2-1) and (4-8-8-1), which are used to predict the removal of chlorophenol from wastewater, and to study the effect of multiple neurons in the hidden layer on the matching of measured values and ANN predicted values. The data set includes input parameters such as inlet flow rate (Fb), inlet pressure (Pb), inlet temperature (Tb), inlet concentration (Cb) and chlorophenol rejection response (Rej). The input data were randomly divided into training set, verification set and test set (48 records in training set, 11 records in verification set and 11 records in test set), and the training set, verification set and test set were 68.57%, 15.71% and 15.71%, respectively.

The neural network model was developed based on genetic and back propagation algorithms to improve the generalization performance of the neural network and to generate a network model to predict the performance indicators of chlorophenol removal by reverse osmosis process.As mentioned earlier, this study utilized actual observational data on chlorophenol removal from wastewater.The neural network model adopts the multi-layer structure of two-layer structure.The steps, design and characteristics of the neural network model proposed in Alyuda simulation can be summarized as follows.

1. In the analysis and division steps, the input data were randomly divided into training set, verification set and test set (68.57% in training set, 11 in verification set and 15.71% in test set, respectively 48 records).
2. Pre-processing step to scale data to the range [-1,1].
3. Genetic algorithm is used as the input feature selection method to improve the generalization performance of neural network.Genetic methods are superior in determining mutually needed inputs, but time-consuming in selecting the best configuration. This could mean a flaw in the artificial neural network.However, it is effective to select suitable networks by identifying input data that contribute significantly to the performance of neural networks.
4. In the design stage, the number of hidden layers, the number of activation functions and the number of error functions are selected as multilayer, hyperbolic tangent and sum of squares respectively.
5. The number of hidden layer neurons was selected according to the number of inputs

in the equation proposed by Sanjay et al. j = n/2 and 2*n n and j are the number of inputs and neurons respectively.Thus, two structures were established (4-2-2-1) and (4-8-8-1).

6. In the training algorithm step, training parameters such as (weight matrix is calculated by fast back propagation), the number of iterations =500, and the random range (W)=±3.

Generally speaking, the good characteristic of genetic algorithm in input feature selection determines the speed of training and verification and the number of iterations. The main task of genetic algorithm is to identify the input data which contributes little to the generalization process in training stage. Therefore, genetic algorithm has the advantage of determining mutually needed inputs with small relative errors and fewer iterations, but it takes time to select the best configuration. On the other hand, the improvement in the absolute error of the structure (4-8-8-1) at the verification stage can be attributed to the influence of the number of neurons in the hidden layer. It is worth mentioning that the evaluation of neural network performance does not depend particularly on validation sets. However, the combination of training sets and validation sets will provide a useful tool to measure the performance of neural network structures.

Fig. 6 shows the scatter diagram of network output and target (experimental data) of 4-2-2-1 and 4-8-8-1 structures in the test phase. All the points are near the straight line, which means that the network has learned the input-output mapping with good accuracy. The scatter diagram of structure 4-8-8-1 is closer to the experimental data than that of structure 4-2-2-1. This indicated that the mean error of 4-8-8-1 was 0.042(R^2 = 0.999).The mean error of structure 4-8-8-1 in the test stage is the smallest, 0.67(R^2 = 0.989). Therefore, it can be considered that 4-8-8-1 structure has the best R^2 value, correlation coefficient and mean error in four stages, and is nominated as the best model for predicting chlorophenol rejection.

Figs. 7 and 8 generate the predicted values of chlorophenol rejection for 4-2-2-1 and 4-8-8-1 structures and experimental data (targets) in each stage of the training set and artificial neural network, respectively. The pattern of the ANN structure shows a high degree of consistency, where the lines produced are similar to the experimental data. In addition, significant differences in error values and sum of squares between 4-2-2-1 and 4-8-8-1 structures can confirm the accuracy of 4-8-8-1 structures.

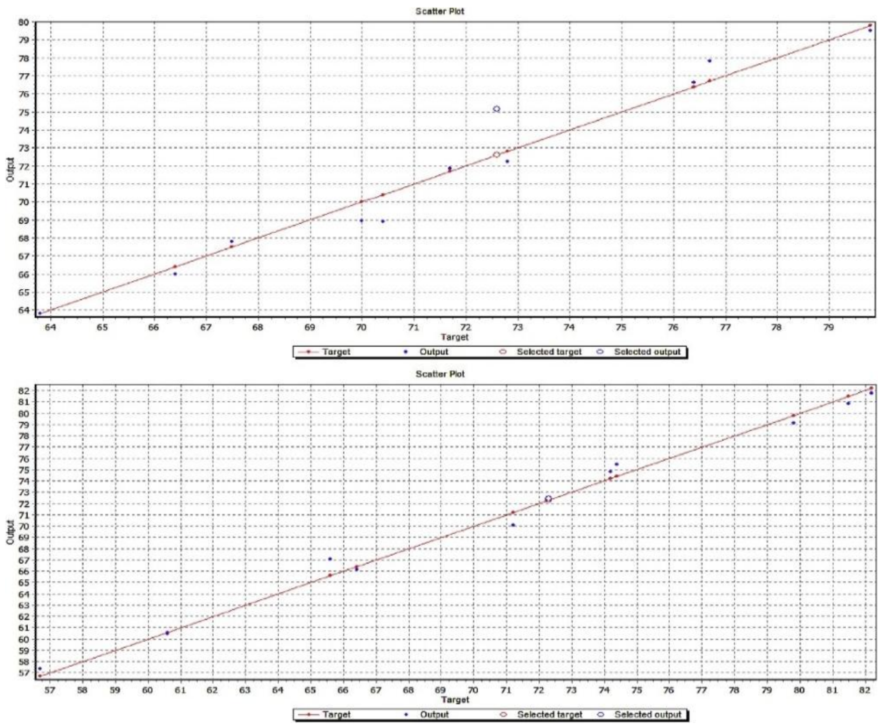

Fig. 6. Scatter plot of 4-2-2-1 and 4-8-8-1 ANN structures in testing stage.

Table 2. Performance of ANN structures

Structure	Performance Indicators	Training stage	Validation stage	Testing stage	All stages
	R^2	0.996	0.968	0.955	0.989
4-2-2-1	Coefficient of correlation	0.998	0.991	0.978	0.994
	Average error	0.3077	0.830	0.737	0.457
	R^2	0.999	0.946	0.989	0.990
4-8-8-1	Coefficient of correlation	0.999	0.975	0.995	0.996
	Average error	0.042	1.047	0.670	0.299
	R^2	0.999	0.950	0.990	0.989
4-15-15-1	Coefficient of correlation	0.997	0.979	0.997	0.994
	Average error	0.051	1.038	0.692	0.311

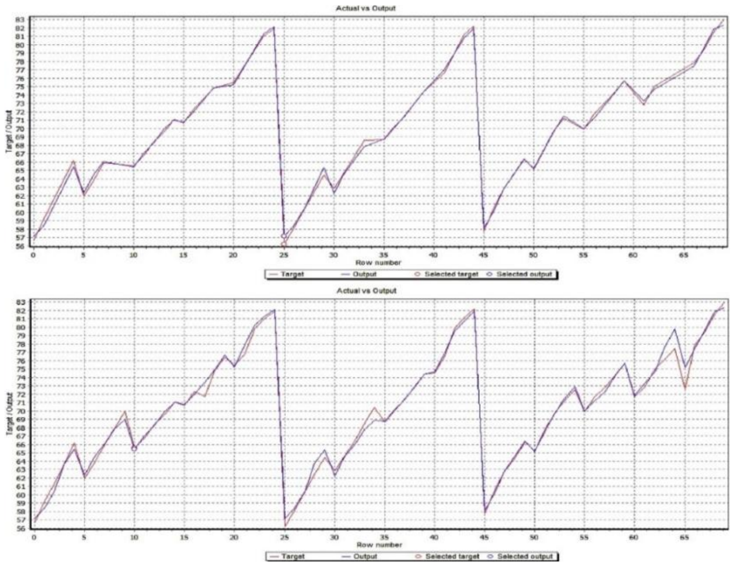

Fig. 7. Experimental and predicted of 4-2-2-1 ANN structure for training and all stage

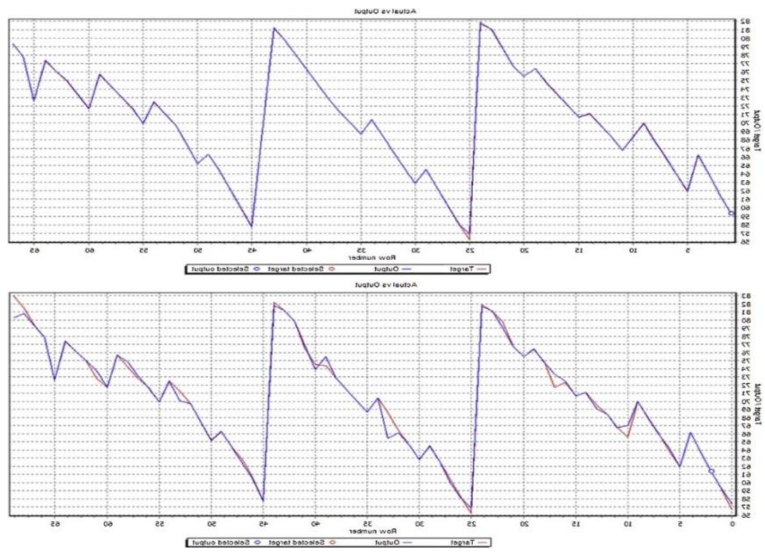

Fig. 8. Experimental and predicted of 4-8-8-1 ANN structure for Training and All stages

The research results show that the accuracy of artificial neural network combined with genetic algorithm is satisfactory, which is consistent with experimental data.It also shows that the prediction accuracy depends on the number of hidden layer neurons. Combined with artificial neural network genetic algorithm, a high-precision model can be established to predict the performance of complex and nonlinear systems. The proposed comprehensive model can optimize the operating variables and remove chlorophenol to the maximum extent to minimize energy consumption.

5. CONCLUSION

After years of research, artificial neural network has been fully developed in pattern recognition, signal processing, knowledge engineering, expert system, optimization combination, robot control and other fields. The application of artificial neural network in chemical separation process is becoming more and more extensive, which improves the accuracy and speed of data simulation processing. Of course, artificial neural network still has some shortcomings and has a great space for development. However, with the continuous progress of neural network theory itself and related theories and technologies, the application of neural network will be more in-depth. Artificial neural network will become an indispensable force in the development of chemical separation field in the future.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Liu LH, Zhang SZ. Separation Engineering [M], Chemical Industry Press, Beijing; 2021.
2. Qi Y, Zheng Z. Neural Network Algorithm and Its Application in Supercritical Extraction Process, Asian Journal of Chemical Sciences. 2021;9(1):19-28. 5
3. Wang H, Mo R. Review of Neural Network Algorithm and Its Application in Reactive Distillation. Asian Journal of Chemical Sciences. 2021;9(3):20-29. 6
4. Zhao N, Lu J. Review of Neural Network Algorithm and its Application in Temperature Control of Distillation Tower, Journal of Engineering Research and Reports. 2021;20(4):50-61.
5. Li C, Wang C. Application of artificial neural network in distillation system: A critical review of recent progress. Asian Journal of Research in Computer Science. 2021; 11(1):8-16.
6. Zheng Z, Qi Y. Study on the Simulation Control of Neural Network Algorithm in Thermally Coupled Distillation. Asian Journal of Research in Computer Science. 2021;10(3): 53-64.
7. Shi F, Gao J, Huang X. An affine invariant approach for dense wide baseline image matching. International Journal of Distributed Sensor Networks (IJDSN). 2016;12(12).
8. Qin C, Schlemper J, Caballero J, et al. Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction, in IEEE Transactions on Medical Imaging. 2019;38(1):280-290.
9. Ivanov AI, Kulagin VP, Kuznetsov YM, Chulkova GM, Ivannikov AD. High-dimensional neural-network artificial intelligence capable of quick learning to recognize a new smell, and gradually expanding the database, 2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC). 2016;332-335.
10. Gao J, Shi F. A Rotation and Scale Invariant Approach for Dense Wide Baseline Matching. Intelligent Computing Theory - 10th International Conference, ICIC (1) 2014: 345-356.
11. Hsu YY, Lu FC. A combined artificial neural network-fuzzy dynamic programming approach to reactive power/voltage control in a distribution substation[J]. IEEE Transactions on Power Systems. 1998; 13(4):1265-1271.
12. Tan TZ, Quek C, Ng GS. Ovarian cancer diagnosis with complementary learning fuzzy neural network [J]. Artificial Intelligence in Medicine. 2008;43(3):207-222.
13. Gao J, Xu Y, Barreiro-Gomez J, Ndong M, Smyrnakis M, Tembine H. (September 5th 2018) Distributionally Robust Optimization. In Jan Valdman, Optimization Algorithms, IntechOpen.
DOI: 10.5772/intechopen.76686. ISBN: 978-1-78923-677-4.
14. Lu J, Zhao N. Application of Neural Network Algorithm in Propylene Distillation, Journal of Engineering Research and Reports. 2021;20(12):53-63.
15. Sun J, Tang Q. Review of Artificial Neural Network and Its Application Research in

Distillation, Journal of Engineering Research and Reports, 2021;21(3):44-54.

16. MCCULLOCH WS， PITTS W.A logical calculus of the ideas immanent in nervous activity[J].Bulletin of Mathmatical Biophysic. 1943;5:115-133.
17. Wang H, Chen Y. Research on Application of Artificial Neural Network in Fault Diagnosis of Chemical Process, Asian Journal of Chemical Sciences. 2021;10(4): 90-97.
18. Rumelhart D, Hinton G, Williams R. Learning representations by back-propagating errors [J]. Nature. 1986;(6088): 533-536.
19. Birajdar GK, Mankar VH. Subsampling-Based Blind Image Forgery Detection Using Support Vector Machine and Artificial Neural Network Classifiers [J]. Arabian Journal for Science & Engineering. 2018; 43(2):555-568.
20. HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities [J]. Proceedings of the National Academy of Science. 1982; 79(8):2254-2558.
21. King T D, El-Hawary ME, El-Hawary F. Optimal environmental dispatching of electric power systems via an improved Hopfield neural network model [J]. IEEE Transactions on Power Systems. 2002; 10(3):1559-1565.
22. Han H G, Qiao J F, Chen Q L. Model predictive control of dissolved oxygen concentration based on a self-organizing RBF neural network [J]. Control Engineering Practice. 2012. 20(4):465-476.
23. Gt A. Research on Building Tax Predictive Model Based on RBF Neural Network [J]. Computer Engineering. 2002;28(5):22-1444.
24. Gao J, Chakraborty D, Tembine H and Olaleye O. Nonparallel Emotional Speech Conversion. INTERSPEECH 2019, Graz, Austria, September; 2019.
25. Chen Z, Qi Z, Wang B, et al. Learning with label proportions based on nonparallel support vector machines [J]. Knowledge-Based Systems. 2017;119(MAR.):126-141.
26. Grossberg S. Adaptive Resonance Theory: how a brain learns to consciously attend, learn, and recognize a changing world [J]. Neural Networks. 2013;37(none):1-47.
27. JIANG B, ZHANG L. Study of Algorithms of Clustering Based on Multi-dimensional Self-organizing Feature Mapping [J]. Computer Science. 2008;35(6):181-182.
28. [28] Liang Y N. Development and application of artificial neural network [J]. Silicon Valley. 2014; 7(12):2.
29. Jin X J, Jia W W. Overview of artificial neural network research [J] Forestry science and technology information. 2008;40(1):2.
30. Wang G G, Guo T, Yu Y, et al. Multilayer Perceptron Generation Adversarial network Based on Semi supervised Learning [J]. Minicomputer system. 2019;40(11):7.
31. Gao, J. Game-theoretic approaches for generative modeling [D]. New York University, Tandon School of Engineering Pro Quest Dissertations Publishing. 2020; 27672221.
32. Li M Q. Analysis of artificial neural network and its application model [J] Science and technology communication. 2019;11(8):3.
33. Yang Y, Cheng He, Chen F F. Rolling Bearing Based on Wavelet Packet and BP Neural Network [J]. Coal mine electromechanical, 2008(01):50-52.
34. Zhao J, Zhang ZY, Xie ZY, et al .Water Quality Modeling for Water works in the Minjiang River estuary Based on BP Neural Network Model [J]. Environmental Science and technology, 2020;(S01):6.
35. Wang X P, Sun J Y, Jin X. Prediction of water quality index of Yutang River Based on BP neural network [J]. Journal of Zhejiang University: Engineering Edition. 2007;41(2):361-364.
36. Chu W H. Application of BP artificial neural network in stock forecasting [J] Journal of Harbin University of Commerce: Natural Science Edition. 2019;35 (2): 4.
37. Tan JY, Chen XF, Lei YG, et al. Application of adaptive frequency shift variable scale stochastic resonance in fault diagnosis [J] Journal of Xi’an Jiaotong University, 2009, 43(7): 5.
38. Zhang FM, Tang SW. Application of RBF artificial neural network in enterprise credit evaluation [J] Statistics and decision making. 2010;(11): 3.
39. Mo JW, Wu HL. Research on housing construction cost prediction based on RBF artificial neural network [J] Project cost management. 2018;(6): 9.
40. Cai YL, Wu GX, Huang J. Application of RBF artificial neural network in sitting

posture evaluation [J] Journal of Beijing University of information technology: Natural Science Edition. 2015;30 (6):5.

41. Araújo AB, Brito RP, Vasconcelos LS. Exergetic analysis of distillation processes-A case study [J]. Energy. 2007;32(7):1185-1193.
42. Samtana EI, Zemp RJ. Thermodynamic Analysis of a Crude-Oil Fractionating Process [J]; 2005.
43. Osuolale F, Jie Z. Energy Efficiency Optimisation for Distillation Column Using Artificial Neural Network Models [J]. Energy. 2016;106:562-578.
44. Kim YM , Yun SL , Lee YG , et al. Development of a package model for process simulation and cost estimation of seawater reverse osmosis desalination plant[J]. Desalination. 2009;247(1-3):326-335.
45. Kress N. Desalination Technologies [J]. Marine Impacts of Seawater Desalination; 2019.
46. Al-Shayji KA, Liu YA. Predictive Modeling of Large-Scale Commercial Water Desalination Plants: Data-Based Neural Network and Model-Based Process Simulation [J]. Industrial & Engineering Chemistry Research. 2002;41(25):6460-6474.
47. Zhao Y, Taylor J, Chellam S. Predicting RO/NF water quality by modified solution diffusion model and artificial neural networks [J]. Journal of Membrane Science. 2005;263(1-2):38-46.
48. Salami E S, Ehetshami M, Karimi-Jashni A, et al. A mathematical method and artificial neural network modeling to simulate osmosis membrane's performance [J]. Modeling Earth Systems & Environment, 2016;2(4):207.
49. Gami AA. Phenol and its toxicity [J]. Journal of Environmental Microbiology and Toxicology, 2014;2(1).
50. B, Awad E. Mohammed A , et al. Optimal design and operation of an industrial three phase reactor for the oxidation of phenol [J]. Computers & Chemical Engineering. 2016;94:257-271.
51. Al-Obaidi MA, Kara-ZaTri C, Mujtaba I M. Simulation and 67 ptimization of a two-stage/two-pass reverse osmosis system for improved removal of chlorophenol from wastewater [J]. Journal of Water Process Engineering. 2018; 22:131-137.
52. Basturk E, Alver A. Modeling azo dye removal by sono-fenton processes using response surface methodology and artificial neural network approaches [J]. Journal of Environmental Management, 2019;248:109300.1-109300.9.
53. Salgado-Reyna, Soto-Regalado, Gomez-Gonzalez, et al. Artificial neural networks for modeling the reverse osmosis unit in a wastewater pilot treatment plant [J]. Desalin Water Treat. 2015;53 (5):1177-1187.
54. Salami ES, Ehetshami M , Karimi-Jashni A , et al. A mathematical method and artificial neural network modeling to simulate osmosis membrane's performance [J]. Modeling Earth Systems & Environment, 2016;2(4):207.
55. Mohammad AT, Al-Obaidi MA, Hammed EM, et al. Modelling the chlorophenol removal from wastewater via reverse osmosis process using a multilayer artificial neural network with genetic algorithm [J]. Journal of Water Process Engineering. 33:100993.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/89365