Desalination and Water Treatment
www.deswater.com

39 (2012) 45–51
February

doi: 10/5004/dwt.2012.2558

Taylor & Francis
Taylor & Francis Group

Studies on prediction of separation percent in electrodialysis process using neural networks

Jing Guolin*, Du Wenting, Xing Lijie

College of Chemistry and Chemical Engineering, Northeast Petroleum University, Daqing 163318, PR China
Tel. +86-459-650-3502; Fax: +86-459-650-3502; e-mail: jingguolin@yahoo.cn

Received 29 December 2010; Accepted 18 August 2011

ABSTRACT

In the electrodialysis process there is a nonlinear relationship between a variety of influencing factors and separation percent (SP), and the relationship is hard to predict. This paper discusses a predictable method via back propagation (BP) neural networks and introduces BP neural networks to forecast separation percent in the electrodialysis process. Moreover, the paper aims to predict the nonlinear relationships between separation percent and its four influencing factors (voltage, concentration, temperature, flow rate). Back propagation neural networks is based on multilayer feedforward neural networks, and achieves nonlinear mappings from inputs to outputs, it is called BP neural networks, hence this method is suitable to predict the nonlinear relationship of separation percent and its influencing factors in the electrodialysis process. We obtained predictable values of separation percent using BP neural networks. Separation percent from experiments compared with its predictable values, and the correlation coefficient was more than 0.99, MSE and MSRE were less than 0.2. Prediction accuracy was high in the range of permissible error, and acquired a good fitting. Therefore, it is verified that BP neural networks is a nice prediction performance and reference value in the electrodialysis process.

Keywords: Electrodialysis; BP neural networks; Separation percent; Predictable value; Concentration; Desalinated water

1. Introduction

Desalinated water is providing an increasing portion of the total fresh water supply in a growing number of countries [1]. Electrodialysis (ED) is an electrochemical process for separation of ions across charged membranes from one solution to another under the influence of an electrical potential difference used as a driving force. This process has been widely used for production of drinking and process water from brackish water and seawater, treatment of industrial effluents, recovery of useful materials from effluents and salt production [2–4].

The prediction of electrodialysis process is an important issue in the realm of membrane science, and there are many influencing factors for electrodialysis, such as voltage, concentration, temperature, flow rate, membrane permeability and so on. And voltage, concentration, temperature and flow rate are the most remarkable factors. Separation percent is considered as a function of concentration, temperature, flow rate and voltage, but an exact formula cannot expressed, namely, there is a nonlinear relationship between them. Thus, it is so hard to predict separation percent.

Artificial neural network (ANN) utilizes interconnected mathematical nodes or neurons to form a network that can model complex functional relationships [5].

*Corresponding author.

Its development started in the 1940s to help cognitive scientists to understand the complexity of the nervous system. It has been evolved steadily and was adopted in many areas of science. Basically, ANNs are numerical structures inspired by the learning process in the human brain. They are constructed and used as alternative mathematical tools to solve a diversity of problems in the fields of system identification, forecasting, pattern recognition, classification, process control and many others [6]. Artificial neural networks have powerful, nonlinear, mapping ability, and the modeling process is easier and more direct than for empirical models [7]. Artificial neural networks can treat with nonlinear relationships, so in the recent years, it has used widely in the realm of different membrane processes, such as reverse osmosis, nanofiltration, ultrafiltration, microfiltration, membrane separation, gas separation, membrane bioreactor and fuel cell.

2. BP neural networks

The argument of Minsky and Papert made many people lose confidence on neural networks, but many scholars still had insisted on the academic research. Rumelhart, McClelland and their colleagues had insight into the importance of information processing of neural networks, and a PDP team was found in 1982, they studied how to deal with parallel-processing information. In 1985, they developed learning algorithm of BP networks and achieved multi-layer networks which Minsky had assumed.

There are many types of neural networks, mainly two kinds: one is feedback model which is represented by Hopfield networks model, it is useful to memory and solve nonlinear optimization; the other is feedforward model which bases on multi-layer perceptron, and uses mainly for classification, pattern recognition, self-organization and memory. BP neural networks is typical of this kind [8].

Back propagation neural networks is a multi-layer feedforward neural networks, transfer functions of neurons are S-type functions, outputs are continuous between 0 and 1, and BP neural networks can achieve any nonlinear mapping from inputs to outputs. Owing to adjustment of weights by learning algorithm of BP, therefore, it is called BP networks.

At present, in the practical application of artificial neural networks, the vast majority of neural networks models employ BP networks. It is also the core of the feedforward networks, and embodies the essence of artificial neural networks [9].

The BP networks are extensively employed in the back-analysis because of its simplicity and power to extract useful information from patterns. It allows specification of multiple input criteria and generation of multiple output recommendations without pre-assumptions regarding the form of functions related to input and output variables. The BP model eliminates the limitations of the traditional regression methods, and accurately establishes the mapping between the input and output variables. It can approximate an arbitrary nonlinear function with better precision [10].

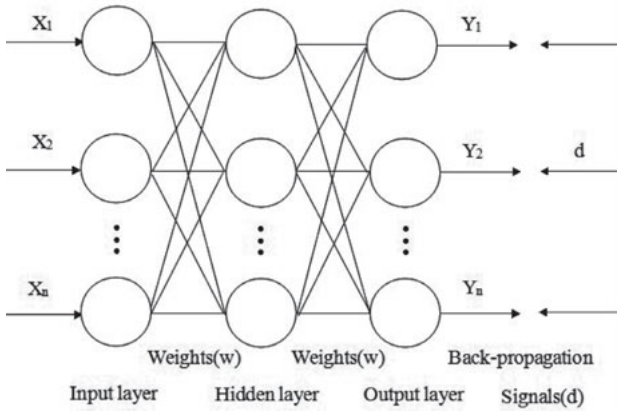

Fig. 1. The structure of BP neural networks.

BP networks is a typical feedforward networks, including input layer, hidden layer and output layer. All connections are between an upper layer and a lower one, but neurons of every layer are no connection. Weights of each layer can be regulated by learning rules, and the basic processing unit of networks (excluding input units) is used to express the nonlinear relationships of inputs and outputs. Learning processing of BP networks contains output calculation and backward error propagation. When input modes are given, they transport from input-layer unit to hidden-layer unit, and output-layer unit generates output modes, this process is named as feedforward propagation. The errors are between output responses and the desired output modes. If the errors are beyond the given scope, it will transfer through the opposite direction. When the errors transport layer-by-layer inversely, and modify weights of each layer. For a given set of samples, outputs calculate repeatedly in the processing of BP networks, until fulfilling the demand of training, then training is completed (Fig. 1).

3. Experimental

3.1. Experimental principle

Electrodialysis technology has been widely used in different realms for its high efficiency and low energy consumption. Under the direct electric field, electrodialysis utilizes the permeability of ion-exchange membrane to separate the electrolyte from the solution, with the potential difference, in order to realize the purposes of dilution,

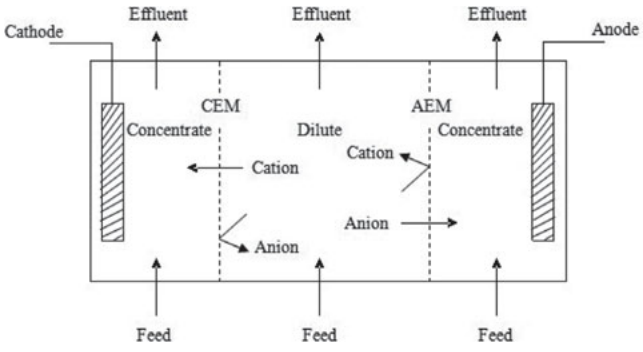

Fig. 2. The principle of electrodialysis process.

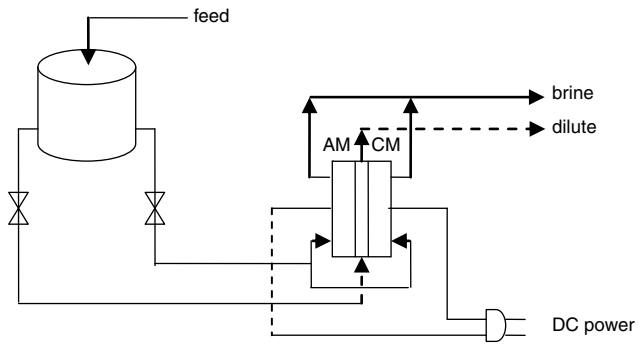

Fig. 3. Electrodialysis cell.

concentration or purification of the solution (Fig. 2). Moreover, cation exchange membrane penetrates only cation ions, and it is the same with anion exchange membrane.

3.2. Experimental design

This experiment used self-designed electrodialysis cell which was made of PMMA. Electrodialysis cell composed of two clamping devices (A, B) and a partition. The side of Device A had two inlets and two outlets, and the side of Device B had an inlet and an outlet. Feed entered ED cell from three inlets which were under two clamping devices, and discharged from three outlets which were the upper end of clamping devices. Feed was divided into two parts through the electrodialysis cell: one part was dilute water and needed to measure flow rates and electrical conductivities of its outlet; the other was brine which recycled at outlet and needed to supply solution which was the same concentration as feed, then added to feed for reuse (Fig. 3).

Experiments were carried out under limiting current density. Four factors were studied: feed concentration (500 ppm, 1000 ppm and 1500 ppm), flow rate of dilute compartment (0.05 ml/s, 0.5 ml/s and 1 ml/s), reaction temperature (288.15 K, 308.15 K and 323.15 K), applied voltage (2 V, 5 V and 8 V).

4. BP neural networks model predict separation percent of NaCl solution

4.1. Establishment of the samples

In the experiments, 81 sets of data were obtained. Based on these data, BP algorithm predicted separation percent of NaCl solution. Obviously, four inputs were concentration *C*, temperature *T*, voltage *V* and flow rate *Q*, and one output was separation percent (*SP*) of NaCl solution. Samples were divided into two categories, respectively, training data and test samples. Doing this was to prevent over-fitting phenomenon and improve forecast accuracy. In all the experiments, 18 sets of data were considered as testing data (Table 1), and the remaining 63 sets were training samples (Table 2).

4.2. Preprocessing of data

In the learning process, Pretreatment of data is of importance, this is due to training and reflecting relationships between various factors [11]. Different input

Table 1
Test samples

T(K)	*C*(ppm)	*E*(V)	*Q*(ml/s)	*SP*(%)	*T*(K)	*C*(ppm)	*E*(V)	*Q*(ml/s)	*SP*(%)
288.15	500	5	1	7.32	308.15	1500	8	0.5	18.47
288.15	500	8	1	10.98	323.15	500	2	0.5	18.24
288.15	1000	2	0.5	4	323.15	500	5	0.5	24.53
288.15	1000	5	0.5	8	323.15	1000	8	0.5	25.6
288.15	1000	8	0.5	8.67	323.15	1000	2	1	8.65
308.15	500	2	0.05	41.32	323.15	1000	5	1	15.22
308.15	500	5	0.05	57.85	323.15	1000	8	1	15.57
308.15	500	5	1	12.46	323.15	1500	2	0.05	52.87
308.15	500	8	1	13.22	323.15	1500	5	0.05	73.1

Table 2
Training data

T(K)	*C*(ppm)	*E*(V)	*Q*(ml/s)	*SP*(%)	*T*(K)	*C*(ppm)	*E*(V)	*Q*(ml/s)	*SP*(%)
288.15	500	5	1	7.32	308.15	1000	8	0.5	17.18
288.15	500	8	1	10.98	308.15	1000	2	1	5.73
288.15	1000	2	0.5	4	308.15	1000	5	1	10.53
288.15	1000	5	0.5	8	308.15	1000	8	1	11.89
288.15	1000	8	0.5	8.67	308.15	1500	2	0.05	23.57
288.15	500	2	0.05	41.32	308.15	1500	5	0.05	69.1
288.15	500	5	0.05	57.85	308.15	1500	8	0.05	88.54
288.15	500	5	1	12.46	308.15	1500	2	0.5	7.64
288.15	500	8	1	13.22	308.15	1500	5	0.5	15.6
288.15	500	2	0.05	37.04	308.15	1500	2	1	3.18
288.15	500	5	0.05	59.76	308.15	1500	5	1	9.87
288.15	500	8	0.05	62.2	308.15	1500	8	1	12.42
288.15	500	2	0.5	3.66	323.15	500	2	0.05	61
288.15	500	5	0.5	12.2	323.15	500	5	0.05	71.7
288.15	500	8	0.5	17.07	323.15	500	8	0.05	77.98
288.15	500	2	1	3.66	323.15	500	8	0.5	27.04
288.15	1000	2	0.05	27.3	323.15	500	2	1	12.03
288.15	1000	5	0.05	55.3	323.15	500	5	1	14.56
288.15	1000	8	0.05	64	323.15	500	8	1	21.54
288.15	1000	2	1	2	323.15	1000	2	0.05	58.82
288.15	1000	5	1	5.3	323.15	1000	5	0.05	77.85
288.15	1500	8	1	6.28	323.15	1000	8	0.05	81.66
308.15	500	8	0.05	67.77	323.15	1000	2	0.5	12.11
308.15	500	2	0.5	9.92	323.15	1000	5	0.5	21.45
308.15	500	5	0.5	17.35	323.15	1500	8	0.05	87.56
308.15	500	8	0.5	21.48	323.15	1500	2	0.5	17.32
308.15	500	2	1	7.44	323.15	1500	5	0.5	23.67
308.15	1000	2	0.05	30.4	323.15	1500	8	0.5	25.45
308.15	1000	5	0.05	67.84	323.15	1500	2	1	10.83
308.15	1000	8	0.05	69.74	323.15	1500	5	1	16.74
308.15	1000	2	0.5	7.93	323.15	1500	8	1	25.8
308.15	1000	5	0.5	15.86					

variables exist magnitude difference, and this difference makes larger input and output nodes remain in the range of the largest gradient of transfer functions [12].

Inputs normalize to [0,1]. When the data are [0,0.1] and [0.9,1], transfer functions change slowly, so we keep inputs between 0.1 and 0.9, specific methods are as follows:

```
% p stands for normalization of inputs
% t stands for normalization of outputs
for i = 1:4
p(i,:) = 0.1+0.8*(P(i,:)-min(P(i,:)))/(max(P(i,:))-min(P(i,:)));
end
```

```
for i = 1
t(i,:) = 0.1+0.8*(T(i,:)-min(T(i,:)))/(max(T(i,:))-min(T(i,:)));
end
```

4.3. Design of BP networks

Owing to four dimensional input vectors, input layer has four neurons. Likewise, output layer has one neuron. And the neurons of hidden layer are calculated [13]:

$$z = \sqrt{x + y} + a \tag{1}$$

where x is the neurons of input layer; y is the neurons of output layer; and a is the constants [1,10].

In the paper we chose eight neurons of hidden layer, so the structure of net was $4 \times 8 \times 1$. Transfer functions are linked to the neurons of two layers. According to general design, the transfer function of input layer was S-type function, the transfer function of output layer was linear function. We chose "tansig" transfer function of input layer, for its output range was between –1 and 1. This paper used default training function, namely, "trainlm". "trainlm" makes use of Levenberg-Marquardt algorithm to deal with medium-sized BP neural networks well. It can reduce calculation during training process, but need a larger amount of memory spaces. The training parameters of "trainlm": maximum time of training was 2000, minimum mean square error was 0.00001, minimum step of training was 50, minimum gradient (min_grad) was 1e-20, other parameters were default.

4.4. BP networks training

The experimental results were processed by Matlab 7.8.0 (R2009a). Each initialization was random. When training was terminated, the errors were not the same, likewise, weights and thresholds were slightly different, so the results of each training were different. During several trainings, a team was acquired (namely, the best team of training).

4.5. BP networks simulation

Simulation function (sim) was to simulate networks. After training, additional data tested (test samples are shown in Table 1), predictable values of separation percent were treated by anti-normalization, compared predictable values of separation percent with its experimental values. If the errors of the two values were relatively low, the preformance met with practical applications. Simulation was expressed as follows:

Y = sim(net, A).

5. Results and discussion

Using MATLAB, we obtained final predictable values. And compared experimental values with predictable values, the data are listed in Table 3 and Table 4:

Training data and test samples trained to acquire the best results, finally, we obtained predictable values of training data and test samples, compared experimental values of separation percent with its predictable values, X axis represented the numbers of test samples or training data, Y axis was output values (Fig. 4).

Table 3
Comparison of experimental values and predictable values of test samples

Experimental values	Predictable values	Experimental values	Predictable values
7.32	7.2772	18.24	18.5926
10.98	10.7279	24.53	24.0769
4	4.324	25.6	25.9908
8	8.287	8.65	8.7559
8.67	8.9077	15.22	15.3516
41.32	41.5814	15.57	15.7658
57.85	57.4849	52.87	53.1205
12.46	12.8475	73.1	72.8075
13.22	13.7618	18.24	18.5926
18.47	18.0401	–	–

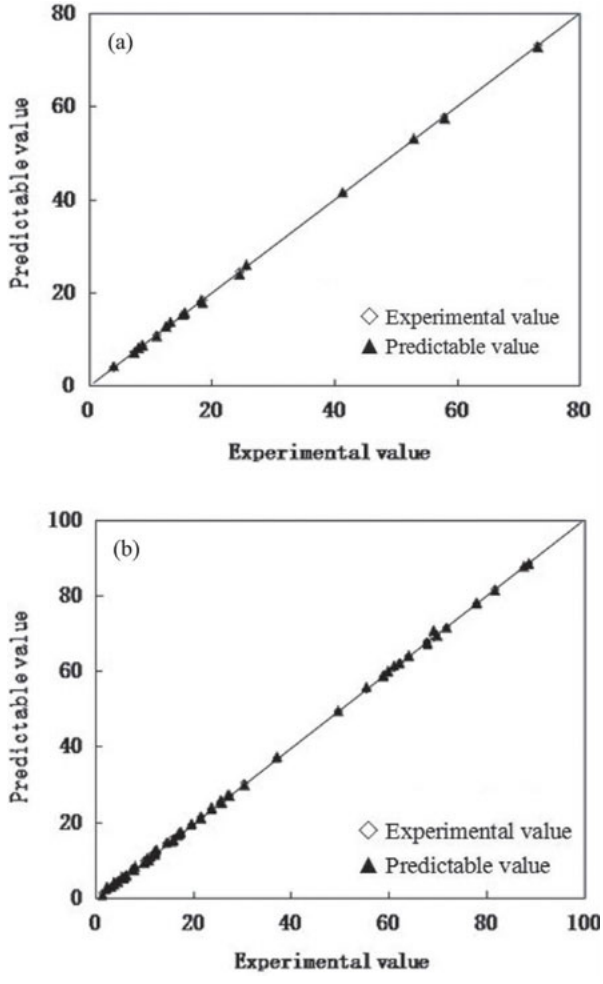

Fig. 4. Fitting of experimental values and predictable values. (a) Fitting of experimental values and predictable values of test samples. (b) Fitting of experimental values and predictable values of training data.

Table 4
Comparison of experimental values and predictable values of training data

Experimental values	Predictable values	Experimental values	Predictable values	Experimental values	Predictable values
37.04	37.4277	6.28	6.0883	9.87	9.3332
59.76	60.0509	67.77	67.8288	12.42	12.8889
62.2	62.1455	9.92	9.6023	61	61.5352
3.66	3.7768	17.35	17.4744	71.7	71.5966
12.2	12.4256	21.48	21.5103	77.98	77.9643
17.07	16.9505	7.44	7.806	27.04	27.3677
3.66	4.14	30.4	30.0105	12.03	11.7832
27.3	27.1435	67.84	67.2323	14.56	14.7492
55.3	55.8219	69.74	69.5436	21.54	21.1143
64	64.185	7.93	7.4376	58.82	59.0518
2	2.4729	15.86	15.4873	77.85	78.1495
5.3	5.671	17.18	17.5463	81.66	81.4415
8	8.4549	5.73	5.2469	12.11	11.8691
19.56	19.7281	10.53	10.4877	21.45	21.4707
49.55	49.5247	11.89	12.2654	87.56	87.7923
58.74	58.8225	23.57	23.6153	17.32	17.5223
2.24	2.9588	69.1	70.8494	23.67	24.1559
7.62	7.8081	88.54	88.5281	25.45	25.8196
11.21	10.9765	7.64	7.753	10.83	10.2092
1.35	1.0281	15.6	15.1456	16.74	16.7155
4.48	4.4539	3.18	3.0102	25.8	25.2615

In statistics, correlation coefficient (R) describes linear correlation between two variables, the correlation coefficient calculates as follows:

$$R = \frac{N\sum_{i=1}^{N}(SP_{\mathrm{cal},i} \times SP_{\mathrm{exp},i}) - \sum_{i=1}^{N} SP_{\mathrm{cal},i}\sum_{i=1}^{N} SP_{\mathrm{exp},i}}{\sqrt{N\sum_{i=1}^{N}(SP_{\mathrm{cal},i})^2 - (\sum_{i=1}^{N} SP_{\mathrm{cal},i})^2}\sqrt{N\sum_{i=1}^{N}(SP_{\mathrm{exp},i})^2 - (\sum_{i=1}^{N} SP_{\mathrm{exp},i})^2}} \quad (2)$$

where "cal" is the predictable values of separation percent; and "exp"-stands for experimental values of separation percent.

Generally $|R|$ is more than 0.8, it is treated as linear correlation between the two variables.

R^2 can be positive values, ranging from $R^2 = +1.0$ for a perfect correlation (positive or negative) down to $R^2 = 0.0$ for a complete absence of correlation. The advantage of R^2 is that it can indicate the method of coefficient strength, namely, the closest to the line of best fit [14].

Using MATLAB toolbox, we got the correlation diagram of experimental values and predictable values (Fig. 5).

Different groups of training data were examined and with respect to the mean squared error (MSE) of testing samples, the proper model was developed. MSE is calculated as follows:

$$\mathrm{MSE} = \frac{\sum_N (SP_{\mathrm{cal}} - SP_{\mathrm{exp}})^2}{N} \quad (3)$$

where subscripts "cal" and "exp" denote predictable values and experimental values of SP, respectively. N is the numbers of testing samples and training data.

Another measure standard of fit is MSRE, the formula is as follows:

$$\mathrm{MSRE} = \frac{1}{N}\sum_N \left(\frac{SP_{\mathrm{cal}} - SP_{\mathrm{exp}}}{SP_{\mathrm{exp}}}\right)^2 \quad (4)$$

Put experimental values and predictable values of training data and test samples into formulas, and the results are as follows:

According to Table 5, it is found that there is an accep agreement (R>0.9 and R^2>0.81) between experimental values and predictable values. The lower the values of MSE and MRSE, the higher the model fitting. So experimental values and predictable values performed fitting effect greatly.

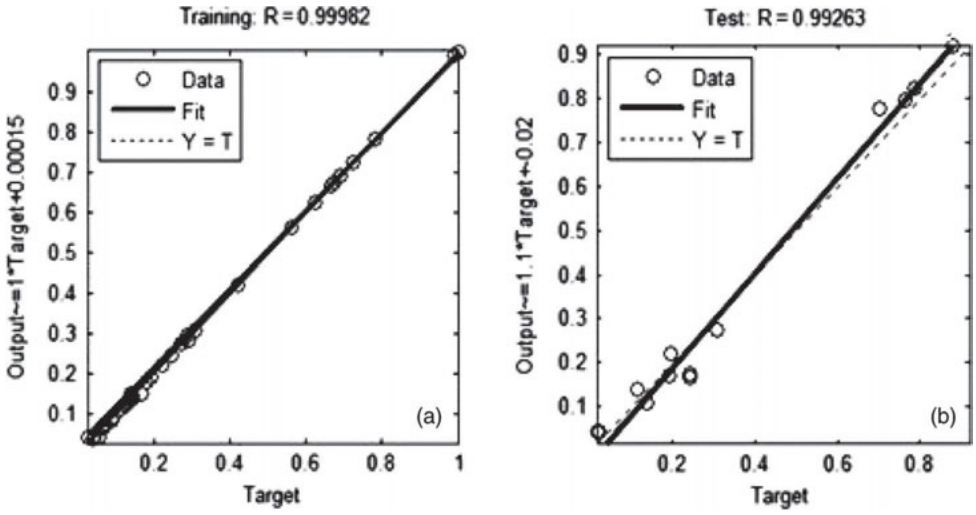

Fig. 5. The correlation diagram of experimental values and predictable values. (a) the correlation diagram of experimental values and predictable values of training data (b) the correlation diagram of experimental values and predictable values of test samples.

Table 5
Statistical criteria for evaluation of the fuzzy model

Evaluation standards	Training data	Test samples
MSE	0.15604	0.10205
R	0.99926	0.92695
R^2	0.99852	0.85924
MSRE	0.00439	0.00767

6. Conclusions

This paper discussed mainly about the structure and algorithm of BP neural networks. Neural networks predicts and constructs unknown objects, so it can apply to multivariable, nonlinear and random factors in the electrodialysis process, and get a higher precision. BP neural networks determines the optimal nodes numbers of hidden layer and a better structure of networks through iterative experiments. BP neural networks has ability of learning, feasibility and effectiveness, reduces subjective factors significantly, uses exhaustion to determine the optimal conditions of the process, improves fitting, and simplifies calculation process.

The results were shown that BP neural networks simulated the trend of separation percent well in the networks model, approached changes of training sets, judged test samples accurately. In the range of permissible errors, the correlation coefficient was above 0.9, MSE and MSRE were below 0.2. Owing to little probability event predictable values could not be in accord with actual results, but the general trend would not have too much deviation. Neural networks meets with practical applications and provides with an effective way to predict electrodialysis process. Meanwhile, this theory may spread to other similar fields.

References

[1] Liat Birnhack, Noga Fridman and Ori Lahav, Potential applications of quarry dolomite for post treatment of desalinated water, Desalin. Water Treat., 1 (2009) 58–67.
[2] W.S. Winston Ho and K.K. Sirkar, Membrane Handbook, Champan & Hall, 1992.
[3] T. Mohammadi and A. Kaviani, Water shortage and seawater desalination by electrodialysis, Desalination, 158 (2003) 267–270.
[4] T. Mohammadi, A. Moheb, M. Sadrzadeh and A. Razmi, Separation of copper ions by electrodialysis using Taguchi experimental design, Desalination, 169 (2004) 21–31.
[5] W. Sha and K.L. Edwards, The use of artificial neural networks in materials science based research, Mater. Des., 28 (2007) 1747–1752.
[6] F.S. Mjalli, S. Al-Asheh and H.E. Alfadala, Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, J. Environ. Manage., 83 (2007) 329–338.
[7] Y.P. Liu, M.G. Wu and J.X. Qian, Predicting coal ash fusion temperature based on its chemical composition using ACO-BP neural network, Thermoch. Acta, 454 (2007) 64–68.
[8] Q. Wu, Y.W. Liu and Z.L. Yu, et al. Based on BP neural networks statistics of accidents of safety in production analysis, China Saf. Sci. J., 19 (2009) 47–52.
[9] C.H. Dong, Matlab neural networks and application, Beijing: National Defence Industry Press, 2005.
[10] B.H.M. Sadeghi, A BP-neural network predictor model for plastic injection molding process, J. Mater. Proces. Technol., 103 (2000) 411–416.
[11] L. Ma and J.W. He, Based on grey correlation analysis and BP neural networks model prediction of daily reference crop evapotranspiration, Water Res. Water Eng., 20 (2009) 34–37.
[12] Q. Yang and L.S. Zhang, Based on BP neural networks separation model of imidacloprid pesticide wastewater nanofiltration, Saf. Env. Eng., 17 (2010) 22–30.
[13] C.R. Yuan, Artificial neural networks and its application, Tsinghua University Press, China, 1999.
[14] Mohtada Sadrzadeh, Ali Ghadimi and Toraj Mohammadi, Coupling a mathematical and a fuzzy logic-based model for prediction of zinc ions separation from wastewater using electrodialysis, J. Chem. Eng., 151 (2009) 262–274.