Chemical Engineering Journal 162 (2010) 552–556

Contents lists available at ScienceDirect

Chemical Engineering Journal

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

Prediction of temperature elevation for seawater in multi-stage flash desalination plants using radial basis function neural network

Ali Aminian *

School of Chemical, Petroleum and Gas Engineering, Semnan University, P.O. Box 35195-363, Semnan, Iran

ARTICLE INFO

Article history:
Received 29 March 2010
Received in revised form 27 May 2010
Accepted 27 May 2010

Keywords:
Artificial neural network
MSF
Radial basis function
Temperature elevation
Boiling point temperature

ABSTRACT

In this paper, a radial basis function (RBF) neural network model was developed for estimating temperature elevation (TE) in multi-stage flash (MSF) desalination processes. The constructed artificial neural network (ANN) model use as input variables the boiling point temperature (BPT) and salinity. The developed RBF neural network was found to be precise in predicting TE from the input variables. The performance of the ANN model was analyzed by mean squared error (MSE). The developed RBF neural network was found to be highly precise in predicting TE for the new input data, which are kept unaware of the trained network showing its applicability to estimate the TE for seawater in MSF desalination plants better than the empirical correlations, thermodynamic models and MLP neural network.

1. Introduction

Desalination is the natural continuous process, which is essential for the water recycle. Desalination methods are classified into two major processes: thermal and non-thermal. Thermal distillation involves phase changes and it includes multi-stage flash (MSF), vapor-compression (VC), and multi-effect (ME). Non-thermal processes do not involve phase changes and includes reverse osmosis (RO), electro-dialysis (ED) and ion exchange (IE) [1]. RO plants can be considered as ideal processes for the seawater desalination from several viewpoints.

An ideal system requires least operating resources that are recoverable from its product if desired. Both RO and MSF plants are non-linear processes, which should operate with performance optimization under specific constraints. Although the MSF process as well as the ME process consumes a larger amount of energy than the RO process, about 18 kWh/m^3 for MSF, 15 kWh/m^3 for ME, and 5 kWh/m^3 for RO, the reliable performance of the thermal desalination processes MSF and ME made highly competitive against the RO process.

At present, MSF units with large production capacity have the largest sector in the desalination industries.

MSF plants are used for the production of potable water and process water from seawater and brackish water. Saline water is steam heated and then led into a series of stages where reduced pressure leads to immediate boiling (flash) without the need to supply additional heat [2]. A schematic diagram of a MSF desalination process is shown in Fig. 1.

In addition, MSF desalination plants especially large ones are often paired with power plants in a cogeneration configuration. Waste heat from the power plant is used to heat the seawater, providing cooling for the power plant at the same time.

This reduces the energy needed from one-half to two-thirds, which drastically alters the economics of the plant, since energy is by far the largest operating cost of MSF plants [3,4].

Modeling of MSF plants are well established in Refs. [5–12]. The steady and unsteady state models [8–12] can be used for evaluating the design characteristics of the process and study the transient behaviors, respectively. In MSF plants, the incoming seawater passes through the heating stages and is heated further in the heat recovery sections of each subsequent stage.

After passing through the last heat recovery section, and before entering the first stage where flash boiling occurs, the feed water is further heated in the brine heater using externally supplied steam. This raises the feed water to its highest temperature (boiling point temperature or top brine temperature), after which it is passed through the various stages where flashing takes place.

The seawater BPT is usually calculated by summing up the BPT of pure water at a given pressure and the TE due to salinity. It is increases the danger of corrosion and scaling in the plant. Thus, a proper knowledge of TE can lead to the better optimization and control of the system and prevent the errors in calculating the design of process equipments.

* Tel.: +98 9124319323; fax: +98 2313354089.
E-mail address: ali_aminian@ymail.com.

1385-8947/$ – see front matter
doi:10.1016/j.cej.2010.05.060

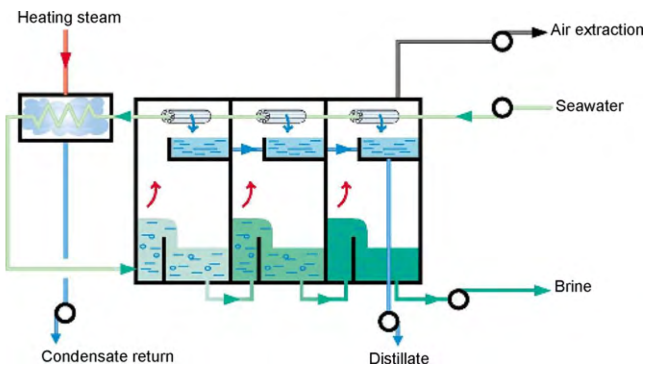

Fig. 1. An MSF desalination process.

Several investigations are performed to model the TE predictions of various source of seawater based on the empirical correlations [9,14,15]. El-Dessouky and Ettouney [13] developed an empirical correlation for calculating TE as a function of BPT and the salinity in weight percent of seawater. In addition, a detailed model incorporating neural networks for physical properties estimation describes the MSF desalination process. In previous work [16], for estimating TE in MSF plants, for each source of experimental data an MLP neural network model was constructed.

The choice of the input variables is the key to insure complete description of the systems, whereas the quality and the number of the training observations (experimental data) have a critical impact on the reliability and the performance of the neural network. The most important variables that affect the TE estimations are the BPT and salinity of seawater [16]. Therefore, experimental data are required for accurate estimating of TE for given BPT (degree of Celsius) and salinity (weight percent).

Experimental data are usually not available in the wide range of operating conditions and therefore should be predicted using accurate models. Using experimental data, the TE for seawater was evaluated from such models.

ANN is another type of modeling procedure. ANNs have superiority as compared with other conventional modeling techniques. The advantage of ANN is that it does not need any knowledge about the process. ANN, however, is capable of modeling highly complex and non-linear systems with large numbers of inputs and outputs. ANNs have been widely used in many fields such as process modeling, control, optimization and prediction [17–19].

In this paper, an RBF neural network model as an adequate powerful tool was developed for predicting TE over a wide range of operating conditions, which are based on the available experimental data.

2. Neural network modeling

2.1. RBF neural network background

The objective of this work is to explore the use of a RBF neural network for the prediction of TE in MSF desalination plants. A RBF consists of an input layer, hidden layer and output layer with the activation function of the hidden units being radial basis functions (Fig. 2).

Normally, an RBF consists of one hidden layer, and a linear output layer. One of the most common kinds of radial basis function is the Gaussian bell-shaped distribution. The response of the hidden layer unit is dependent on the distance an input is from the centre represented by the radial basis function (Euclidean Distance) [20]. Each radial function has two parameters: a centre and a width. The width of the basis function determines the spread of the function and how quickly the activation of the hidden node decreases with the input being an increased distance from the centre [21]. The output layer neurons are weighted linear combination of the RBF in the hidden layer. An RBF network can be modeled by the following equations:

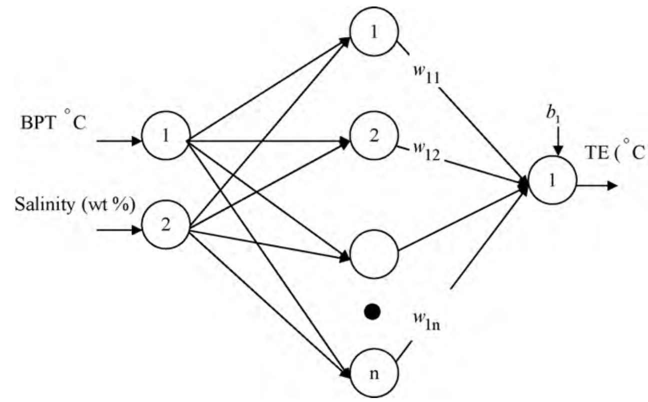

Fig. 2. The structure of RBF network.

$$y_j(x) = \sum_{i=1}^{n} w_{ji}\psi_i(x) + b_j \tag{1}$$

where $y_j(x)$ is the output at the jth node in the output layer, n is the number of hidden nodes, w_{ji} is the weight factor from the ith hidden node to the jth output node, $\psi_i(x)$ is the radial basis activation function of the hidden layer and b_j is the bias parameter of the jth output node. Some of the common types of RBF are linear function, Duchon radial cubic, radial quadratic plus cubic and Gaussian activation function. The last function has the form:

$$\psi_i(x) = \exp\left(\frac{-||X - u_i||^2}{2\sigma_i^2}\right) \tag{2}$$

where X is the input vector, u_i is the center vector of ith hidden node and σ is the width of the basis function. There are two distinct types of Gaussian RBF architectures. The first type uses the exponential activation function, so the activation of the unit is a Gaussian bump as a function of the inputs. The second type of Gaussian RBF architecture uses the softmax activation function, so the activations of all the hidden units are normalized to sum to one. This type of network is often called a “normalized RBF” or NRBF network. An NRBF network with unequal widths and equal heights can be written in the following form:

$$\psi_i(x)(\text{softmax}) = \frac{\exp(h_i)}{\sum_{i=1}^{n}\exp(h_i)} \tag{3}$$

$$h_i = \left(-\sum_{l=1}^{2}\frac{(X_l - u_{il})^2}{2\sigma_i^2}\right) \tag{4}$$

Again, X is the input vector (Salinity, BPT), u_{il} is the center of the ith hidden node (i = 1,. . .,12) that is associated with the lth (l = 1,2) input vector, σ_i is a common width of the ith hidden node in the layer and softmax (h_i) is the output vector of the ith hidden node. The radial basis activation function used in this study is the softmax activation function [22]. The NRBF neural network developed during this study consists of an input layer, a hidden layer and an output layer, which include 2, 12 and 1 node, respectively. At first, the input data is used to determine the centers and the widths of the basis functions for each hidden node. The second step includes the procedure, which is used to find the output layer weights that minimize the quadratic error between the predicted values and the target values. Mean square error (the average sum of squares error)

Table 1
Center and width of the hidden layer nodes and the weights of the hidden-output layer.

Hidden layer parameters												
Node (i)	1	2	3	4	5	6	7	8	9	10	11	12
Salinity (center)	0.08	0.36	0.78	0.86	0.25	0.48	0.06	0.37	0.88	0.66	0.29	0.08
BPT (center)	0.17	0.13	0.13	0.46	0.45	0.42	0.51	0.70	0.91	0.79	0.96	0.80
Width	0.93	1.13	1.09	1.09	0.58	0.80	0.68	0.93	0.96	1.08	0.78	0.86
Hidden – output layers weights												
w_{ij}	2.5	6.8	−10	8.8	1.2	−3.6	−1.3	4.4	27.1	−24	−1.6	−3.8

is defined as

$$\mathrm{MSE} = \frac{1}{N}\sum_{k=1}^{N}((\mathrm{TE})_k^{\mathrm{exp}} - (\mathrm{TE})_k^{\mathrm{cal}})^2 \qquad (5)$$

The NRBF neural network modeling and measuring the network performance was implemented under the MATLAB environment.

2.2. The parameters of the NRBF network

When designing an NRBF network, the most critical task is certainly the determination of the parameters of the hidden layer. There are a number of centers selection techniques, such as K-means clustering, linear regression, C-means clustering and Kohonen algorithm. In each of these techniques, the parameters of the basis functions are determined through the unsupervised or supervised training algorithms [23]. Means clustering looks for convenient clusters of data and places an RBF center in the middle of each cluster. Genetic algorithm or unsupervised neural network can be used to find optimal center vectors. Whenever the Gaussian function has been selected as a basis function the width has to be adjusted to control the amount of overlapping. The width can be set to a constant value, calculated using the gradient descent method or using a heuristic algorithm. After the centers and widths of the hidden layer determined, the next stage involves finding the final layer weights that minimize the error between the network's output and the target values using supervised learning algorithm. In the present work, fuzzy C-means clustering was used to determine the parameters of the basis function. Fuzzy C-means is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. This technique starts to work with an initial guess for the cluster centers. The initial guess for these cluster centers is most likely incorrect. By iteratively updating the cluster centers and the membership grades for each data point, fuzzy C-means iteratively moves the cluster centers to the right location within a data set. This iteration is based on minimizing a cost function that represents the distance from any given data point to a cluster center weighted by that data point's membership grade [24,25]. In mathematical expression it can be written in the following form:

$$J = \sum_{k=1}^{N}\sum_{i=1}^{n}\mu_{ik}^2(\mathrm{x}_k - u_i)^2, \quad \sum_{i=1}^{n}\mu_{ik} = 1 \quad \text{which } 1 \le k \le N \qquad (6)$$

Which is must be minimized through a non-linear optimization technique that is contains the new values of the membership and the center at each iteration. μ_{ik} is the degree of membership of X_k in the cluster ith. The widths and the centers of the hidden layer nodes associated with each input vector and the weights between the hidden and output layers are shown in Table 1.

2.3. Training result

In order to predict the TE in MSF desalination plants, the required input/target data are taken from two sources [14,15] that are used for the training, validating and testing the NRBF network. The BPT and salinity of seawater have been selected as input variables, while TE (°C) selected as output variable. Experimental data sets were chosen with BPT range from 20 to 150 °C and salinity range from 1.5 to 13 wt%. The total number of data was 380, 266 sets of training samples were used to train network and 114 sets of validating and testing samples were used to show the generalization capability of the trained NRBF neural network. Some of these data are shown in Table 2.

Before training the network with input/target data, the input and target vectors are required to be normalized. Thus, neural network predict the output with mean and standard deviation of zero and one, respectively. One of the problems that occur during neural network training is called over fitting. The error on the training set is driven to a very small value, but when new data is presented to the network the error is large. Therefore, Experimental data should be divided into training, validation and test set. Training, validating and testing data sets were obtained using specified indices and partitioned to 70%, 15% and 15% of experimental data, respectively.

Each data set corresponds to one TE for seawater. The developed NRBF neural network is trained using hybrid learning algorithm to obtain the center and width of each hidden layer nodes and the least squares method to obtain the output weights.

Fig. 3 shows the error between the experimental data and the neural network predicted temperature elevation. From Fig. 3, it can be observed that the deviation of neural network predictions

Table 2
A group of data is used for training.

Salinity (%)	BPT	TE
1.50	70.0	0.183
1.50	80.0	0.196
1.50	90.0	0.209
1.50	100	0.222
3.00	20.0	0.247
3.00	30.0	0.274
3.00	40.0	0.300
5.00	120	0.898
5.00	130	0.945
5.00	140	0.991
5.00	150	1.038
6.00	100	0.989
6.00	110	1.046
6.00	120	1.103
8.00	20.0	0.759
8.00	30.0	0.836
8.00	40.0	0.914
9.50	70.0	1.408
9.50	80.0	1.503
9.50	90.0	1.598
9.50	100	1.693
10.50	130	2.232
10.50	140	2.339
10.50	150	2.446
12.00	100	2.252
12.00	110	2.377
12.00	120	2.502
13.00	120	2.765
13.00	130	2.902
13.00	140	3.041

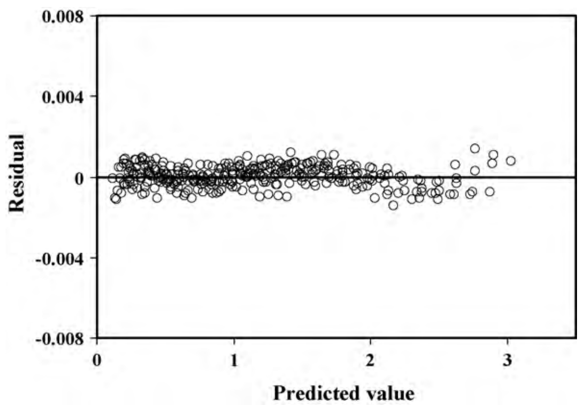

Fig. 3. Plot of error between experimental data and the neural network results.

with the experimental data was very low. The MSE of the optimal RBF network architecture for both training and testing data was obtained near to zero. Note that, the time taken to optimize the parameters of the hidden layer and train the network using the training data set is about 30 s.

3. Simulation results

It should be mentioned that the empirical correlations such as the El-Dessouky and Ettouney correlation [13], could not provide an adequate predictions of temperature elevation for the wide range of operating conditions. However, the proposed NRBF neural network model as an adequate powerful tool can be used to predict the TE in the range of available experimental data as well as for some seawater compositions, which are not represented in the training data set. Table 3 represents the comparison between the results of this work and El-Dessouky and Ettouney correlation with experimental values obtained using BPT at 60 °C, which are indicated in Refs. [14,15] (these data are excluded and not used in the training phase to show the predictability of NRBF neural network). The time taken to execute the trained network for estimating the outputs for the new inputs data was less than 1 s.

Tanvir and Mujtaba [16] used the multi-layer perceptron (MLP) neural network based correlations (model) to improve the predictions of temperature elevation for seawater in MSF desalination plants.

The response of the NRBF neural network to the prediction data indicates that the NRBF neural network model provides a maximum error of 0.00095 compared to the MLP neural network based correlations [16] and the empirical correlations [9,13–15]. The results are summarized in Table 4.

The results obtained by the NRBF neural network model predicting TE as a function of BPT and salinity for the whole set of experimental data which are studied in this work is shown in Fig. 4. From Fig. 4, it is evident that the interpolative capability of the NRBF neural network is quite believable.

Again, experimental data [14,15] at different salinity and BPT that are indicated in Fig. 4 are not used in the training phase. As mentioned before, the results show that the trained NRBF neural network can predict TE accurately from a set of new source of seawater. Because of the high operating cost of MSF desalination process, it is necessary to determine the optimized operating conditions for these processes. Therefore, in development of predictive model for desalination plants, application of neural network is essential due to non-linearity and complexity of interactions between operating variables.

However, the energy optimization in MSF desalination plants requires the modeling of MSF processes either from process models, thermodynamic analysis, empirical correlations or from artificial intelligence, leading to low-cost production of fresh water.

Table 3
Comparison of predicted TE (°C) and experimental values at BPT = 60 °C.

Salinity (wt%)	This work	Empirical correlation[a]	Experimental
1.5	0.170070	0.136261	0.170
2.5	0.290554	0.315323	0.291
3.0	0.353429	0.389233	0.354
4.0	0.484162	0.548672	0.484
5.0	0.621246	0.724247	0.621
6.5	0.838281	1.019555	0.838
7.0	0.913643	1.126897	0.913
8.0	1.068942	1.355518	1.069
9.5	1.313646	1.734747	1.314
10.5	1.484867	2.012891	1.485
11.5	1.662615	2.312197	1.663
12.5	1.846662	2.633438	1.847
13.0	1.940867	2.802525	1.941

[a] El-Dessouky and Ettouney correlation: $TE = Ax + Bx^2 + Cx^3$. where $A = 0.083 + 0.00018 \times T + 4 \times 10^{-6} \times T^2$; $B = -0.00076 + 0.00009 \times T + 5 \times 10^{-7} \times T^2$; $C = 0.00015 - 3 \times 10^{-6} \times T^{-3} \times 10^{-8} \times T^2$; T = BPT in °C and x = salinity in wt%.

Table 4
Comparison of maximum error between different methods for data reported in [14,15].

This work	MLP neural network based correlations [16]	El-Dessouky and Ettouney correlation
0.00095	0.024	3.50

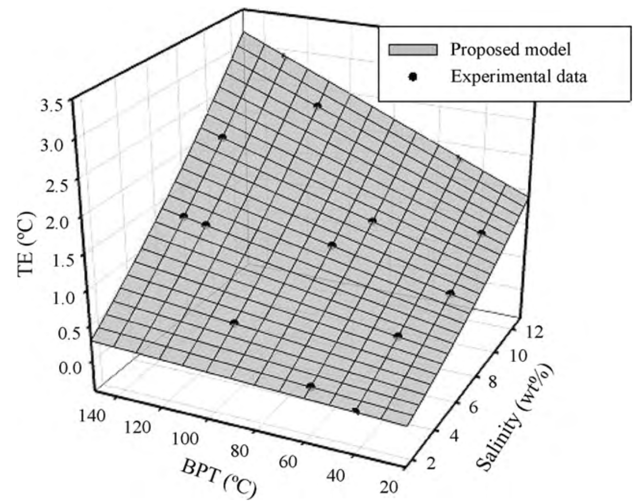

Fig. 4. Comparison between experimental data and predicted with the RBF network.

4. Conclusions

In order to proper operation of MSF desalination plants, an RBF neural network has been used to predict the temperature elevation during the MSF process. In an MSF process, top brine temperature is one of the important parameters, which could be obtained at a proper TE estimation. Furthermore, the danger of corrosion and the energy consumption reduced and the design of process equipments (e.g., heat transfer area, the size of flash chamber) will calculate in such a way to minimize the total cost of MSF unit.

Experimental data sets were chosen with BPT range from 20 to 150 °C and salinity range from 1.5 to 13 wt%. The developed neural network can be used in an MSF unit for pattern recognition based on C-mean clustering of input data set. Because of the wide range of operating conditions studied, the proposed neural network model can be used, through an unsupervised pattern classification, for accurate prediction of TE as well as for the input variables, which are kept unaware of the trained neural network.

The proposed method was compared with the MLP neural network and the empirical correlation, the results showed that the proposed method was superior for prediction.

After selecting the several network architectures, it was found that a network with one hidden layer of twelve neurons is the optimum network architecture. The MSE is closely near to zero with the proposed RBF network. The constructed RBF neural network was found to be precise in predicting TE for the new input data. The results demonstrate that this new developed RBF neural network is successful and can be implemented within an MSF process model because of its computational efficiency and accuracy.

References

[1] W.T. Hanbury, T. Hodgkeiss, R. Morris, Desalination Technology, Porthan Limited, Glasgow, UK, 1993.
[2] M.A. Darwish, A.A. El-Hadik, The multieffect boiling desalting system and its comparison with the multistage flash system, Desalination 60 (1986) 251–265.
[3] E. Drioli, C. Alessandra, C. Efrem, Integrated membrane operations for seawater desalination, Desalination 147 (2002) 77–81.
[4] J.H.S.C. Low Tay, S. Jeyaseelan, Vacuum desalination for water purification using waste heat, Desalination 106 (1996) 131–135.
[5] M.A. Darwish, Thermal analysis of multi-stage flash desalination systems, Desalination 85 (1991) 59–79.
[6] A.M. Omar, Simulation of M.S.F. desalination plants, Desalination 45 (1983) 65–76.
[7] H.T. El-Dessouky, H.I. Shaban, H. Al-Ramadan, Steady state analysis of multi-stage flash desalination, Desalination 103 (1995) 271–277.
[8] N.M. Abdel-Jabbar, H.M. Qiblawey, F.S. Mjalli, H. Ettouney, Simulation of large capacity MSF brine circulation plans, Desalination 204 (2007) 501–514.
[9] A.M. Helal, M.S. Medani, M.A. Soliman, J.R. Flower, Tridiagonal matrix model for multi-stage flash desalination plants, Comp. Chem. Eng. 10 (1986) 327–342.
[10] M. Rosso, A. Beltramini, M. Mazzotti, M. Morbidelli, Modeling multistage flash desalination plants, Desalination 108 (1997) 365–374.
[11] A. Husain, A. Hassan, D.M.K. Al-Gobaisi, A. Al-Radif, A. Woldai, C. Sommariva, Modeling, simulation, optimization and control of multistage flashing (MSF) desalination plants. Part I. Modeling and simulation, Desalination 92 (1993) 21–41.
[12] M.F Falcetta, E. Sciubba, Transient simulation of a real multi-stage flashing desalination process, Desalination 122 (1999) 263–269.
[13] H.T. El-Dessouky, H.M. Ettouney, Fundamentals of Salt Water Desalination, Elsevier Science, Amsterdam, 2002.
[14] L.A. Bromley, D. Singh, P. Ray, S. Sridhar, S.M. Read, Thermodynamic properties of sea salt solutions, AIChE J. 20 (2) (1974) 326.
[15] B.M. Fabuss, Properties of seawater Principles of Desalination, vol. 2, 2nd ed., Academic Press, New York, 1980.
[16] M.S. Tanvir, I.M. Mujtaba, Neural network based correlations for estimating temperature elevation for seawater in MSF desalination process, Desalination 195 (2006) 251–272.
[17] D. Barba, G. Linzzo, G. Tagliferri, Proceedings of the 4th International Symposium on Fresh Water from the sea, vol. 1, 1973, pp. 153–157.
[18] M.A. Mandil, E.E. Ghafour Abdel, F. Maidani, Chem. Eng. Sci. 25 (1970) 611–615.
[19] E.D. Howe, Fundamentals of Desalination, Dekker, New York, 1974.
[20] Statsoft, Neural Networks, Statsoft Electronic Textbook, 2003.
[21] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press Inc., New York, 1995.
[22] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Upper Saddle River, NJ, 1999.
[23] A.B. Bulsari, Neural networks for chemical engineers Computer-aided Chemical Engineering, 6, Elsevier, Amsterdam, 1995.
[24] J.C. Bezdek, Pattern Recognition with Fuzzy Logic Objective Function Algorithm, Plenum Press, New York, 1981.
[25] J.C. Bezdek, Fuzzy mathematics in pattern classification, PhD Thesis, Cornell University, Ithaca, New York (1973).