Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2018.22079

111 (2018) 57–67
April

Modeling of $CaCl_2$ removal by positively charged polysulfone-based nanofiltration membrane using artificial neural network and genetic programming

Amir Dashti[a], Morteza Asghari[a,b,*], Hosna Solymani[c], Mashallah Rezakazemi[d], Ahmad Akbari[c]

[a]*Separation Processes Research Group (SPRG), Department of Engineering, University of Kashan, Kashan, Iran, email: amirdashti13681990@gmail.com*
[b]*Energy Research Institute, University of Kashan, Ghotb-e-Ravandi Ave., Kashan, Iran, Office: +98 31 5591 2427, SPRG: +98 31 5591 2823, Fax +98 31 5591 2424, email: asghari@kashanu.ac.ir*
[c]*Institute of Nanoscience and Nanotechnology, University of Kashan, Ghotb-e-Ravandi Ave., Kashan, Iran, email: hana.soleymani.s@gmail.com (H. Solymani), akbari@kashanu.ac.ir (A. Akbari)*
[d]*Faculty of Chemical and Materials Engineering, Shahrood University of Technology, Shahrood, Iran, email: rezakazemi@shahroodut.ac.ir*

Received 8 May 2017; Accepted 19 February 2018

ABSTRACT

Artificial neural network (ANN) and genetic programming (GP) models were used to predict rejection (R) and permeability coefficient of water flux (L_p) with respect to $CaCl_2$ in nanofiltration (NF) membrane process. The model inputs were concentration of the poly(ethylene imine) (PEI), p-xylene dichloride (XDC) and methyl iodide (MI), coating and crosslinking time of PEI, and pH of the solution. With this respect, ANN with 3:17:1 and 3:23:1 neurons, the lowest mean squad error (MSE) of 0.0023 and 0.000028 and the highest coefficient of determination (R^2) values of 0.9830 and 0.9990 for R and L_p, respectively, was found. In addition, the sensitivity analysis suggested that PEI coating time and pH had the significant effect on R and L_p, respectively. GP was used to make a mathematical function for prediction of R and L_p in terms of the input parameters. The GP model successfully described the R and L_p as function of input parameters. The GP results with R^2 values of more than 0.99 had an excellent preciseness.

Keywords: Artificial neural network; GP; Modeling; Nanofiltration; Membrane; Purification

1. Introduction

Many industries generate water effluents with harmful chemical compounds, and development of techniques to address treatment and removal of the pollutants is critical [1–4]. Membrane filtration process is considered a simple, versatile and efficient separation technique in the removal of suspended and/or dissolved substances from a liquid or gas phases [5–15]. Based on physical properties such as pore size range, molecular weight cut-off range and operating pressure, membranes fall into five categories of microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), forward osmosis (FO) and reverse osmosis (RO) [16]. Typically, NF membranes are fabricated by lowering pore size of a UF membrane support using chemical modification [17]. NF membranes have a composite structure with a top layer that is selective and controls solute transfer. The selectivity of the top layer in NF membranes is based on the pore size and electrostatic interactions (Donnan repulsion) [18]. NF membranes typically have a negatively charged separation layer to separate negatively charged species [19]. However, membranes with a positive charge on the surface can enhance the separation of cations in aqueous media.

*Corresponding author.

1944-3994 / 1944-3986 © 2018 Desalination Publications. All rights reserved.

For instance and biotechnological applications, positively charged membranes are a good candidate for the removal of endotoxins from solutions. Endotoxins are toxic material mostly derived from bacterial lysates.

Many efforts have been put to develop models that have a reasonably suitable explanation of an NF process and several NF models based on extended Nernst–Planck equation have been presented. With this respect, Wang et al. [20] and Bowen et al., [21,22] proposed the steric-hindrance (ES) and Donnan steric-partitioning pore model (DSPM) models, respectively. These two models demonstrated a better agreement with the experimental results as to describe the performance of NF membranes for electrolytes, particularly for 1–1 electrolytes, such as NaCl. Nevertheless, when these models are applied in NF membranes with respect to 2–1 electrolytes (such as $CaCl_2$), the predicted R tends to be lower [23], or the charge density tends to be larger [24]. Several modifications on DPSM have been performed by considering the parameters such as hindrance effect of pores to the ions [22], concentration polarization [21], and dielectric constant [25]. Aforementioned models typically are mathematically and computationally complex and need detailed knowledge of the filtration process [26]. Therefore, alternative techniques with an accurate description of an NF process using available process data and extending it to the mathematical models are considered promising.

Recent studies show that artificial neural network (ANN) model has enough simplicity and better estimations in comparison with the conventional physics-based models [27–30]. Although a considerable number of studies exist about the successful application of ANN in the successful modeling of different membrane filtration processes, only a few studies address modeling of NF. Bowen et al. [31] developed ANNs modeling in NF membranes to predict rejection (R) rate of single salts and mixtures ($MgCl_2$, NaCl, Na_2SO_4, and $MgSO_4$). Their experimental setup was a spiral wound membrane that was not simple to model with physics-based models. In a similar study by Darwish et al. [32], ANNs were employed to model the cross-flow NF of NaCl and $MgCl_2$ at seawater concentrations. They have investigated the effects of input salt concentrations and operating pressure on R rates by different NF membranes of NF 90, NF 270, and NF 30. They showed that ANN model successfully predicts the experimental R rates of NaCl and $MgCl_2$ by NF membranes. Furthermore, many researchers have successfully employed black box models for modeling of membrane filtration processes, and among them genetic programming (GP), as one of the genetic algorithm (GA) subcategory. Lee et al. [33] employed GP for modeling the fouling rate of MF membrane in a pilot-scale drinking water production system. Okhovat et al. [34] successfully applied the GP modeling in the removal of As (V), Cr (VI) and Cd (II) from wastewater as a function of transmembrane pressure (TMP) and initial pollutant concentration, using NF process. They have reported that the results gained from proposed GP models show very good agreement with the experimental results.

However, although ANN and GP models are effective and accurate models which have been used to evaluate the excessively complicated non-linear relationships, it seems that modeling of NF process using ANN and GP models is not given much attention and there is a need to address it. Hence, the primary goal of this study is to apply ANN and GP models to estimate target response parameters in NF process. With this respect, modeling of rejection (R) and permeability coefficient of water flux (L_p) with respect to the pollutant model $CaCl_2$ as a function of the solution concentration of PEI, XDC and MI, coating and crosslinking time of PEI, and pH were investigated using GP and ANN.

2. Data attainment

A polymer dope solution made up of PSf, Polyethylene glycol (PEG), and *N*-methylpyrrolidone (NMP) (16, 14, and 70% w/w, respectively) was mixed for the preparation of the UF support membrane. Subsequently, the dope was cast onto a glass plate. Then, this solution was immersed in a water coagulant bath at environment temperature. The cast membrane was immersed in water for one day to carry out the water exchange with the solvent present in the pores. Pure PEI is so efficient for making the extra anionic colloidal charges, neutral, particularly in PH of 7 or lower. It is usually assumed that the positively charged amines in PEI raise the surface charge in the XDC membrane and MI that are employed for the crosslink and quaternization process. Initially, the PEI/PSf composite was fabricated via coating the PEI aqueous solution on the PSf surface during 1 h followed by drying for 2 h at environment temperature. Then, it was soaked in XDC/*n*-heptane solution for crosslinking. The crosslinking process involving PEI and XDC was performed at environment temperature for 5 h. Then, the prepared membrane was washed with ethanol to eliminate unwanted compounds, following the crosslinking procedure. Then, the membrane was immersed in NaOH solution to remove interfacial H^+. Eventually, for the quaternization of PEI on the upper layer, the membrane was immersed into a CH_3I solution in ethanol for 2 h. The fabrication procedure of the PEI coating of the PSf support is demonstrated in Fig. 1.

The membrane pore radius was measured on the basis of the PEG molecular weight, which was rejected at 90% by the membrane and was calculated as follows [35,36]:

$$y = -5\times10^{-8}x^2 + 5\times10^{-8}x + 0.3319 \quad (1)$$

where y is referred to pore radius (nm) and x is referred to the molecular weight of PEG (g/mol).

The size of membrane pore was predicted by the relationship involving the molecular weight cut-off (MWCO) acquired using the PEG solutions and their radius of the pores computed by Eq. (1). The MWCO of the PSf NF membranes was investigated with five various feed solutions of PEG with MWs of 1500, 2000, 3000, 4000, and 6000 Da. Based on the obtained results, for the ultimate membranes, one of them was merely crosslinked PEI using (XDC), and second one was quaternized by a MI solution, and the MWCOs were appraised to be 3670 and 3540 that yields 1.49 nm and 1.47 nm pore radius, respectively. The experimental procedure and data used in the present study are the same as in a previous [37].

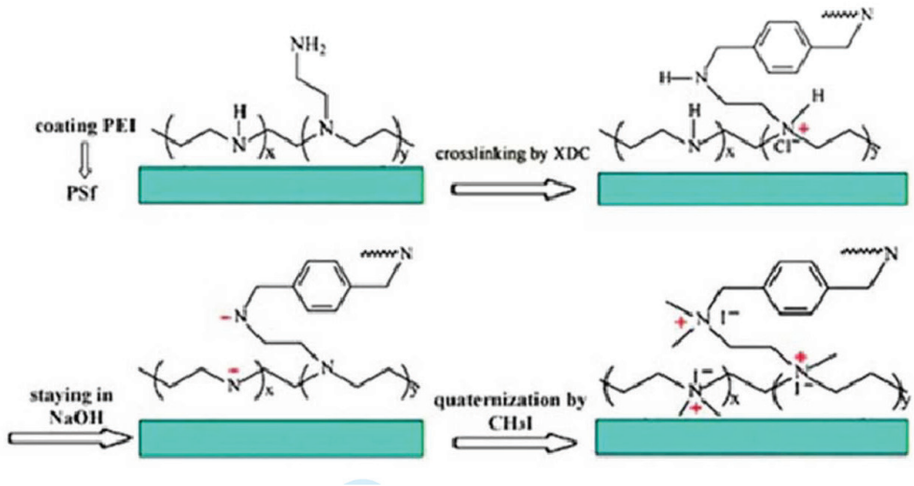

Fig. 1. Scheme of the preparation procedure of the PSf NF membrane [37]. Copyright 2015, Reproduced with permission from John Wiley and Sons.

3. Modeling theory

3.1. ANNs

Neural networks are computer algorithms originated from the way that data is processed in the biological aspect of the nervous system. ANN can be known as a neurocomputer with parallel-distributed processors [38]. Fig. 2 shows the basic ANN structure that is made up of three layers with a number of neurons in each layer. The layers contain input layer (independent variables), hidden layer, and output layer (dependent variables). In a typical network, input layer is comprised from the original experimental data (X_i) that are associated with the neurons or nodes (1, 2…, i,…, m) of the input layer. Input data are transferred to the nodes of hidden layer (1, 2…, j,…, n) and output layer (1, 2…, k,…p) by multiplying connection strength or weights (W_{ij}) between two neurons and summing using summation function. The inputs to a neuron include bias and the sum of its weighted input. The outputs of a neuron are depended on the neuron's inputs and on the transfer function of it [39]. The kind of transfer functions often employed for solving multiple regression issues are summarized in Table 1. The

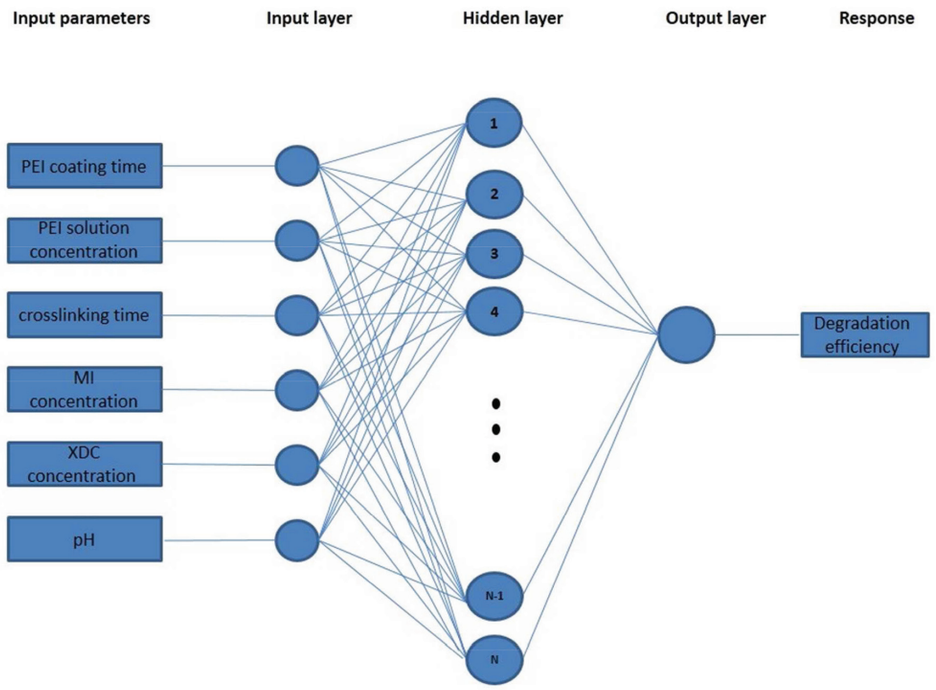

Fig. 2. Architecture of a typical ANN with an input layer, a hidden layer, and an output layer.

Table 1
Transfer functions of artificial neurons used for solving multiple regression problems

Transfer function	Notation	Transfer function equation	Output range
Linear	purelin	$f(\mathrm{A_i}) = \mathrm{A_i}$	[−∞,+∞]
Log-sigmoid	logsig	$f(\mathrm{A_i}) = \frac{1}{1+exo(-\mathrm{A_i})}$	[0,1]
Hyperbolic tangent sigmoid	tensing	$f(\mathrm{A_i}) = \frac{1-exo(-\mathrm{A_i})}{1+exo(-\mathrm{A_i})}$	[–1,+1]

most crucial stage for building ANN model is the training of the network. In the training process, the weights and biases of a feed-forward neural network are modified methodically in order to decrease the remainder error between network outputs (predictions) and targets (experimental data) [40,41]. There are many different training algorithms. The most prevalent training algorithms for feed-forward neural networks are the back-propagation (BP) method [42]. Training of ANN using BP algorithm is definitely an iterative optimization process used for performance function minimization by modifying the network weights and biases correctly. The most applied performance function is the mean-squared-error (MSE) and the coefficient of determination (R^2). In the case of a single output neuron, MSE and R^2 might be written as [43–45]:

$$MSE = \frac{1}{n}\sum_{q=1}^{n}\left(Y_q^{exp} - Y_q^{pred}\right)^2 \tag{2}$$

$$R^2 = 1 - \frac{\sum_{q=1}^{n}\left(Y_q^{exp} - Y_q^{pred}\right)^2}{\sum_{q=1}^{n}\left(Y_q^{exp} - \bar{Y}_q\right)^2} \tag{3}$$

where Y_q^{exp} is the experimental response (target), Y_q^{pred} is the predicted response by ANN (network output), n is the number of experimental data points and q is the iteration index (positive integer number) and $\bar{Y}_q$ is the average of the observed experimental data acquired by following equation:

$$\bar{Y}_q = \sum_{q=1}^{n} Y_q^{exp} \tag{4}$$

There are several modifications of BP algorithm. A lot of them utilize the gradient descent approach for iterative updating of weights and biases before the convergence is satisfied. When the ANN was trained the optimal weights and biases are stored and the neural network model can be utilized for simulation and optimization [40].

3.2. GP

GP is a type of genetic-evolutionary algorithms that was developed by Koza. He introduced it in his book "Genetic Programming" in 1992 [46]. The GP has been utilized in fields such as control, robotics, games, and symbolic regression. GP is based on rules of biological evolution [47]. GP is an advanced method for supplying nonlinear input–output empirical models in engineering applications [48]. It offers good solutions for several problems and is a promoted development of GA. The output of the GP is a computer program, whilst the output of the GA is a value. GP is significantly more powerful than GA and is a device learning approach for optimizing a particular aspect of the system based on a fitness criterion supplied by the user. GP frequency changes a population of computer programs right into a novel generation of the population. In each generation of the algorithm, closely favorable individuals are chosen as "parents" for the next generation and create a new reproduction source. A fresh generation of solutions grows mutation and reproduction that using one of three genetic operations include crossover. After many generations, a program would develop providing solutions for the issue [49]. Functions connecting nodes of inputs and constants generate a preliminary model population, whose convolution is determined by the user. The setting of terminals (the independent variables of the problem and random constants) and original functions for every single branch of the program, measuring the fitness of individuals in the population, validating the parameters to manage the program being run and the technique estimating the goodness of fit are the main four fundamental steps necessary in GP. Population size, the maximum number of generations and the probability of crossover and mutation are determining parameters in a GP [50]. Every program in the process of GP is expressed as a tree. For example, Fig. 3 illustrates the representation of the function $\exp(X^2) + \cos(X + 3)$. This study uses GP to discover a mathematical function of six input variables. The terminal set includes the independent variables, PEI solution concentration, PEI coating time, XDC concentration, MI concentration, pH and crosslinking time. Thus the terminal is set as $(x_1, x_2, x_3, x_4, x_5, x_6)$, the function is set as (+,−,×,/ and exp) and the fitness function is the difference between the developed output and the target output. Desired functions are those with lower fitness functions values [51]. A software program (Matlab GPTIPS toolbox) was employed to utilize GP on a computer.

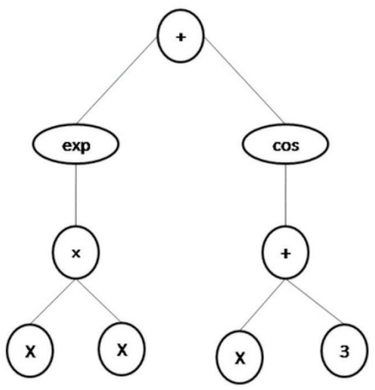

Fig. 3. Tree representation of $\exp(X^2) + \cos(X + 3)$.

4. Results and discussion

4.1. Modeling results and evaluation of ANN prediction

The experimental data used to build the ANN model for NF process are summarized in Table 2. In order to predict the values of flux utilizing the ANN model, 70% of the data were employed for training purpose. The remainders were used for testing and validation data, equally. Inputs of the neural network were six variables, i.e., the PEI solution concentration, PEI coating time, crosslinking time, XDC concentration, MI concentration, and pH. L_p and R were considered as a response (output or target). In order to prevent from overfitting, equally input and output result were normalized. The input factors were normalized so that they differ in the range of [0–1] based on the subsequent relationship [40]:

$$x_j = \frac{\left(z_j - z_j^{min}\right)}{\left(z_j^{max} - z_j^{min}\right)} \tag{5}$$

where x_j refers to the normalized input variable or parameter, while z_j, z_j^{min} and z_j^{max} are the actual, maximum and minimum values of the input variable. So as to construct the ANN model our network was developed applying MATLAB computer software. The normalized values of the inputs and output were applied to feed and train the ANN. In this study, the LM back-propagation algorithm was used [52]. All neurons

Table 2
Experimental data used in ANN modelling [37]

PEI concentration	PEI coating time	XDC concentration	MI concentration	pH	Crosslinking time	R	L_p
25	60	5	16	7	300	94	3.78
25	60	7	0	7	300	80.50	2
25	150	5	0	7	300	83.20	3.70
25	60	5	0	7	240	81.10	5.29
25	60	3	0	7	300	77.70	4.20
25	60	5	8	3.30	300	94.10	4.42
15	120	5	0	7	300	82.90	1.50
25	60	5	2	7	300	86	4.70
25	60	5	8	10	300	84.80	4.25
25	60	5	0	7	300	86	5.10
25	60	5	8	7	300	93.60	4.25
10	120	5	0	7	300	73	0.90
25	60	4	0	7	300	84	4.40
25	60	5	0	7	180	79.10	5.76
25	0	5	0	7	300	10	54
30	120	5	0	7	300	86	3.90
20	120	5	0	7	300	84	2.70
25	120	5	0	7	300	87	4.60
25	60	6	0	7	300	84.30	4.16
25	60	5	0	7	30	72	9.23
5	120	5	0	7	300	0	0.57
25	90	5	0	7	300	83	4.70
25	60	5	0	7	300	84.60	4.64
25	60	1	0	7	300	72.30	3.80
25	60	5	0	7	420	87.87	4.53
25	60	5	0	7	300	87.80	5.10
25	30	5	0	7	300	61	9.96
25	120	5	0	7	300	84	4.10
25	60	5	8	7	300	93.40	4.88
25	60	5	4	7	300	93	4.57
25	60	5	0	7	90	73.73	8.61

Table 3
Weight and bias values of the ANN model for *R* prediction

Neuron	Hidden layer							Output layer	
	Weights						Bias	Weights	Bias
	PEI solution concentration	PEI coating time	XDC concentration	MI concentration	pH	Crosslinking time			
1	–0.1883	–0.7687	–1.0031	–1.4509	1.3625	–0.5919	2.0036	0.0563	–0.2867
2	–0.8998	0.1724	–0.3103	–0.8751	0.2837	–1.1435	2.4677	–1.0447	
3	–0.3949	0.0133	–1.6019	–0.5459	–1.2074	1.0780	1.4807	0.0250	
4	–0.1008	0.6195	0.7541	0.9450	0.7638	–1.2534	1.5267	–0.54549	
5	0.7767	–0.6152	1.1201	–0.7795	–0.5753	1.2302	–0.7831	–0.0673	
6	–0.8591	0.0227	–0.6379	–1.7136	0.4856	1.1182	0.6379	0.2501	
7	0.3311	–0.4661	0.8851	–0.3641	1.0888	–1.4411	–0.6846	–0.1770	
8	0.2188	–1.3205	1.0873	–0.1704	1.1855	–1.0378	–0.6484	0.4412	
9	–0.7825	0.6169	1.1890	0.5749	–0.8627	–1.1081	–0.1148	0.0792	
10	0.7441	0.5873	0.4609	–0.1565	–0.4150	1.3834	0.8707	–0.6279	
11	–0.0050	0.4704	–1.1257	–1.1209	1.1766	0.8650	0.5366	–0.0516	
12	1.0402	0.1961	–1.1916	1.4055	0.5905	–0.4307	0.6738	0.1839	
13	–2.1073	–2.2833	–0.4541	–1.1898	–0.9968	–0.8322	–1.2996	–1.8008	
14	0.4291	0.6690	1.1903	1.1104	–1.3383	–0.5866	1.4785	0.4733	
15	–1.1801	0.8752	–0.5186	–0.9062	–1.5330	–0.2441	–1.5334	0.4193	
16	–0.6449	–0.8711	1.5704	0.5421	0.5601	1.3501	–1.7701	–0.5231	
17	2.0186	–0.6909	–0.6639	–0.4886	–0.9051	0.0205	2.8436	0.8839	

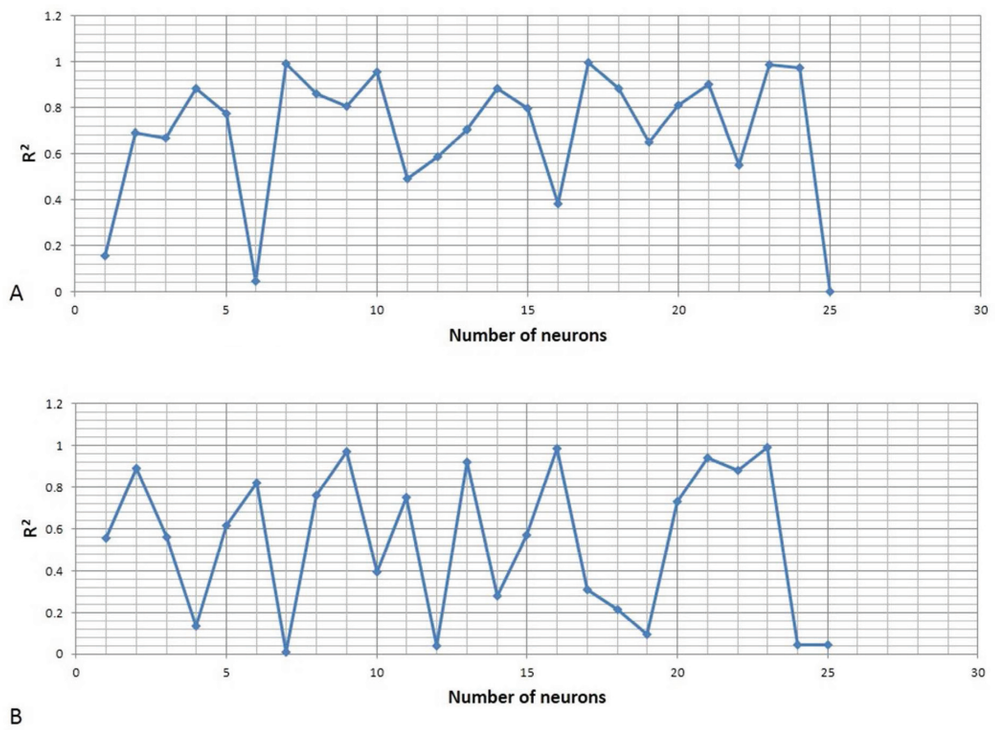

Fig. 4. Variation of R^2 with number of neurons in models for (a) *R*, (b) L_p.

of the hidden layer have the tansig and single neuron from the output layer has the linear (purelin) transfer function. Among different transfer function in the hidden layer, tansig function offers slightly better predictions than others [30]. As a way to optimize the ANN structure, the computations began using one neuron in the hidden layer. Figs. 4a and 4b and show the variation of R^2 with the number of hidden neurons in the hidden layer for R and L_p, respectively. It is obvious from the figure that the maximum R^2 for a structure of R and L_p are with 17 and 23 neurons in the hidden layer.

Thus, in this case, the ideal structure of the ANN model includes six inputs (i.e., variables), one hidden layer with 17 and 23 neurons (for R and L_p, respectively) and one output layer with an individual neuron. Hence, the best structure for R and L_p prediction was 6:17:1 and 6:23:1, respectively. The network was evaluated to possess converged once the test set error is lowest. Tables 3 and 4 show the weights and biases of the optimum architecture.

The statistical data for training, validation and test data for both ANN networks are demonstrated in Table 5. The values of R^2 for training, validation, test and the overall data set (0.9868, 0.8731, 0.9944, and 0.983, respectively) and (0.9999, 0.9633, 0.9955, and 0.9995, respectively) for R and L_p predictions, respectively represent an efficient network prediction using ANN in comparison with experimental results. It can be concluded that there is a good accordance between experimental data and ANN results.

4.2. Importance of operating factors from ANN

Determination of the relative importance of input variables is a technique for systematically changing input variables in a model to determine the effects of such changes on the output of the model and showing how the input variables can be quantitatively apportioned. In order to evaluate the relative importance of various operating variables on output variables, a neural net weight matrix [53] was used. Fig. 5 shows the importance of various operating variables on output variables. PEI coating time with a relative

Table 4
Weight and bias values of the ANN model for L_p prediction

Neuron	Hidden layer							Output layer	
	Weights						Bias	Weights	Bias
	PEI solution concentration	PEI coating time	XDC concentration	MI concentration	pH	Crosslinking time			
1	0.8347	–2.6862	0.5813	0.9790	–0.6316	0.3709	–2.7174	1.6015	–0.3487
2	1.1123	–0.5744	0.0962	0.9359	1.7327	0.3155	–2.1438	0.002	
3	–1.7161	0.0904	–1.0678	–1.0461	0.7743	0.7373	1.8385	0.6289	
4	1.1838	–1.4083	0.0717	–0.0354	–1.1540	–1.4231	–1.9535	0.8974	
5	–0.7200	0.3146	0.2035	1.3694	1.4702	–0.8500	1.4711	0.1651	
6	–0.7348	0.5959	1.2714	–1.1290	–1.4081	–0.0930	1.1579	0.5651	
7	0.8269	–1.5858	0.9384	1.2344	0.7268	–0.6090	–1.1200	0.8505	
8	–0.3496	–1.1384	1.6105	1.0702	–0.1827	–0.9895	0.4939	–0.2745	
9	–0.4037	0.9654	1.5462	–1.0957	–0.4364	0.1297	0.9030	–0.3352	
10	1.3929	–1.0274	–0.7063	1.1887	0.2135	–0.7841	–0.4049	0.0507	
11	1.0500	–0.1114	–1.0412	–0.9708	1.2582	1.1230	–0.4183	0.9891	
12	–0.1234	–1.8424	–0.6712	0.8237	–1.2713	–1.3781	–0.5079	0.8186	
13	0.6454	0.3072	–1.4935	–1.3873	0.2095	–0.9028	0.4055	0.1989	
14	–1.0651	–0.5820	–0.8142	–1.3666	–0.8571	–0.8826	–0.2682	–0.1290	
15	0.2574	–1.6798	0.8073	–0.0521	–1.1051	0.3983	0.4795	–0.0467	
16	–0.9032	–1.0118	–0.4599	0.7479	0.3243	–1.7887	–1.1334	–0.5624	
17	–1.1460	–0.8036	0.8996	–1.1135	–1.1103	–0.5217	–1.3394	0.8380	
18	–0.3702	0.5584	0.9614	–0.3775	1.5357	–1.3580	–1.6396	–0.8280	
19	–1.1558	–2.122	–0.0431	–0.9021	–0.0060	–0.1688	–1.4766	–0.7662	
20	–0.9954	1.1070	–0.1044	1.2861	0.5777	–1.1425	–1.7030	–1.0300	
21	0.9409	0.9186	–0.6677	–0.8880	–1.2409	1.0282	1.9571	0.5797	
22	0.7726	–0.4637	–1.5687	0.2585	1.2992	0.5192	2.3144	–0.7787	
23	1.4034	–0.5032	0.3912	1.0349	–0.9130	1.1980	2.3973	–0.5720	

Table 5
Comparison of performance of optimum ANN model with different algorithms

	R		L_p	
	R^2	MSE	R^2	MSE
Train	0.9868	0.0011	0.9998	2.09E-06
Validation	0.8731	0.0003	0.9279	9.25E-05
Test	0.9944	0.0096	0.9910	7.21E-05
All	0.9830	0.0023	0.9990	2.80E-05

importance of 19.74% was the most influential parameter on the R and pH with a relative importance of 73.5% was the most influential parameter on the L_p. As observed, the contribution of PEI coating time and pH in the membrane structure has a significant effect on the membrane performance. In other words, this means that the smaller changes in PEI coating time and pH make the bigger changes in R and L_p values, respectively.

4.3. GP results

248 data was utilized in this model. The population size or number of individuals creating a population in each generation, number of generations to run for including generation, probability of GP tree mutation (%), probability of GP tree crossover (%), maximum depth of trees were set at 1000, 200, 10%, 85% and 8, respectively. Crossover and mutation were selected in this program. The best tree depth number was 8 with the lowest runtime, complexity and error values.

PEI solution concentration (x_1), PEI coating time (x_2), XDC concentration (x_3), MI concentration (x_4), pH (x_5) and crosslinking time (x_6) were independent variables to acquire a model for prediction of R and L_p by using GP. Figs. 6a and 6b illustrates the prediction of R and L_p using GP and comparison is performed between experimental data and results of the model. As observed in Fig. 6, there is great agreement between model and experimental results and the model demonstrating an excellent prediction of the system behaviors. The best-so-far GP model for prediction of R and L_p which were obtained after satisfying termination criterion are:

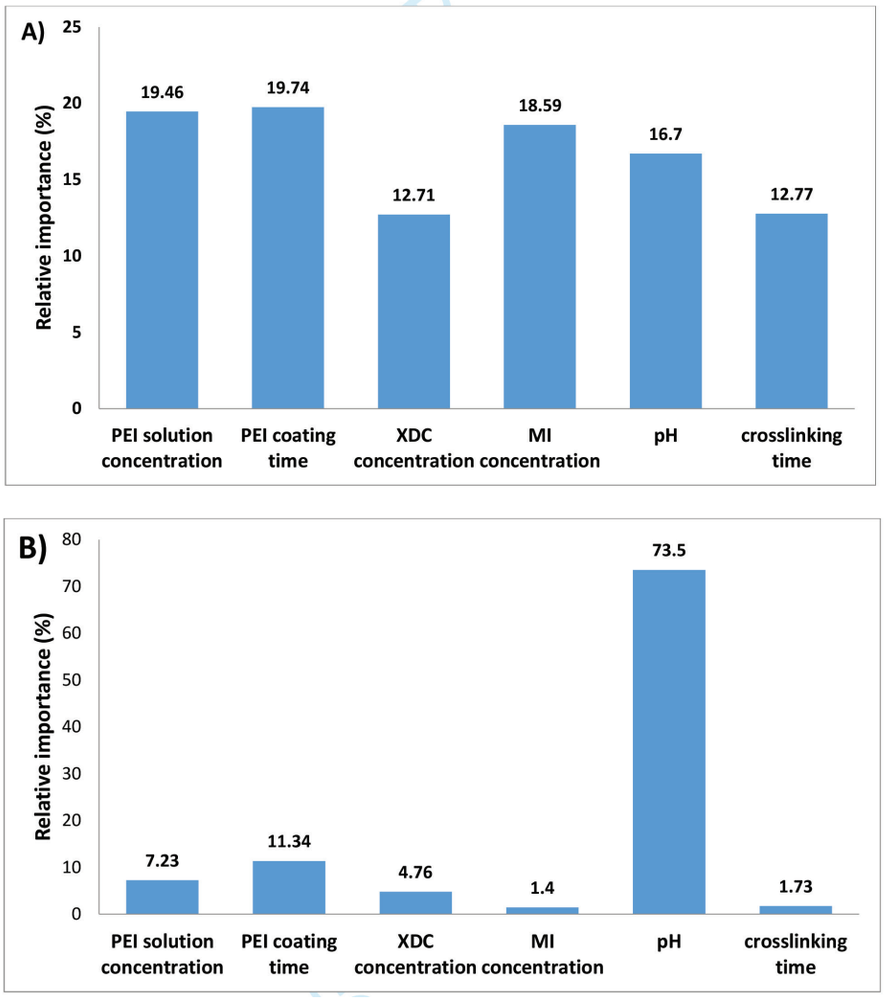

Fig. 5. The importance of various operating variables on output variables: (a) R, (b) L_p.

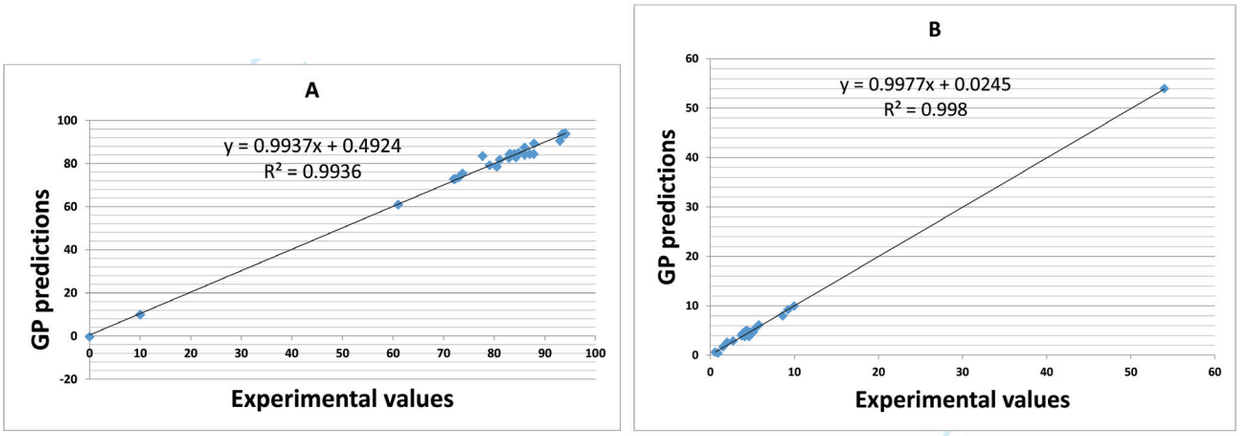

Fig. 6. Experimental results and GP model prediction of (a) R, (b) L_p.

Table 6
R^2 and MSE for permeate flux and rejection prediction by GP

	R	L_p
R^2	0.9936	0.9980
MSE	2.6900	0.1617

$$R = (2.15 \times (x_1 - 9.444) \times (3 \times x_1 + x_6 - (x_6 + \exp(x_3))/(x_3 + 2 \times x_4) - 38.27))/(x_{12}) - (2.799 \times (x_6 + \exp(x_{22}/x_6) + \exp(x_{22}/x_6) \times \exp(9.495 \times x_4/x_3) + ((x_1 + x_6)/((x_1 - x_2) \times (x_1 - x_2 + x_5)))))/(\exp(x_{22}/x_6) \times \exp(x_4 \times (x_4 + 9.576)/x_5) + 9.576) + 76.74 \quad (6)$$

$$L_p = 53.88 - (0.01132 \times (x_2 + (x_3 - 3.985) \times (x_3 - x_5)) \times (x_3 - x_2 - 3 \times x_1 + 2 \times x_4 + x_6 + x_3 \times (x_3 - 4.067) + (2 \times x_1)/(x_3) - 3.985))/(\exp((x_1 - 3.985)/x_6) \times (x_1 - x_5 + 0.1001 \times x_6 - x_3 \times (x_3 - 3.86) + (x_6/x_{12}))) - 126.6 \times \exp(((2 \times \exp(x_1 - x_3 + x_5) - 2 \times x_1 + \exp(x_2) + (x_1 - 3.985)/(x_3 \times x_5) - 2))/(\exp(x_3 \times (x_3 - 3.985) - \exp(x_5)) - \exp(x_1) - \exp(x_2) + (x_1 - x_5)/(x_1)) \quad (7)$$

The execution time of running the evolutionary algorithm for this case was approximately 2 min on a Sony PC (Core i7, RAM 4GB, Windows Seven). The R-square (coefficient of fitting illustration) value (0.993 and 0.998) for R and L_p indicates the GP model results are fitted to the experimental data very well.

5. Conclusions

PEI solution concentration, PEI coating time, XDC concentration, crosslinking time, MI concentration and pH closely affected R and L_p. NF process data was successfully described with ANN because of the highest determination of coefficient values between network prediction and corresponding experimental data. Results of this model showed that PEI coating time was the most influential parameter on the R and pH was the most influential parameter on the L_p. GP model successfully described the NF process as a function of PEI solution concentration, PEI coating time, XDC concentration, crosslinking time, MI concentration and pH in a single equation. GP gave a unique model to calculate R and L_p at all studied conditions to describe permeability and R data. Artificial intelligence indicated that NF process could be precisely modeled and well described by GP and ANN. These artificial intelligence methods can be used for any kind of NF process without any dependence on the feed or membrane structure.

Acknowledgment

The authors are grateful to Energy Research Institute at University of Kashan for supporting this work. We would also like to appreciate Mr. Vahid Jabbari in The University of Texas at El Paso for checking the English of the paper. Also, we would like to acknowledge Mr. Hadi Halakoei for his help in the preparation of the work.

References

[1] M. Rezakazemi, A. Khajeh, M. Mesbah, Membrane filtration of wastewater from gas and oil production, Environ. Chem. Lett., (2017) 1–22.
[2] A. Azimi, A. Azari, M. Rezakazemi, M. Ansarpour, Removal of heavy metals from industrial wastewaters: a review, ChemBioEng. Rev., 4 (2017) 37–59.
[3] M. Rezakazemi, A. Ghafarinazari, S. Shirazian, A. Khoshsima, Numerical modeling and optimization of wastewater treatment using porous polymeric membranes, Polym. Eng. Sci., 53 (2013) 1272–1278.
[4] S. Shirazian, M. Rezakazemi, A. Marjani, S. Moradi, Hydrodynamics and mass transfer simulation of wastewater treatment in membrane reactors, Desalination, 286 (2012) 290–295.
[5] T. Mohammadi, M. Maghami, M. Rezakazemi, High loaded synthetic hazardous wastewater treatment using lab-scale submerged ceramic membrane bioreactor, Periodica Polytech. Chem. Eng., (2017) 1–6.
[6] M. Rezakazemi, S. Shirazian, S.N. Ashrafizadeh, Simulation of ammonia removal from industrial wastewater streams by

means of a hollow-fiber membrane contactor, Desalination, 285 (2012) 383–392.
[7] M. Rezakazemi, M. Sadrzadeh, T. Mohammadi, In: R. Wilson, A.K.S.S.C. George, Transport Properties of Polymeric Membranes, Elsevier, Amsterdam, (2018) 243–263.
[8] M. Rezakazemi, K. Shahidi, T. Mohammadi, Synthetic PDMS composite membranes for pervaporation dehydration of ethanol, Desal. Water Treat., 54 (2014) 1–8.
[9] B. Baheri, M. Shahverdi, M. Rezakazemi, E. Motaee, T. Mohammadi, Performance of PVA/NaA mixed matrix membrane for removal of water from ethylene glycol solutions by pervaporation, Chem. Eng. Commun., 202 (2014) 316–321.
[10] M. Shahverdi, B. Baheri, M. Rezakazemi, E. Motaee, T. Mohammadi, Pervaporation study of ethylene glycol dehydration through synthesized (PVA-4A)/polypropylene mixed matrix composite membranes, Polym. Eng. Sci., 53 (2013) 1487-1493.
[11] M. Rezakazemi, M. Iravaninia, S. Shirazian, T. Mohammadi, Transient computational fluid dynamics modeling of pervaporation separation of aromatic/aliphatic hydrocarbon mixtures using polymer composite membrane, Polym. Eng. Sci., 53 (2013) 1494–1501.
[12] M. Rezakazemi, M. Shahverdi, S. Shirazian, T. Mohammadi, A. Pak, CFD simulation of water removal from water/ethylene glycol mixtures by pervaporation, Chem. Eng. J., 168 (2011) 60–67.
[13] A. Dashti, M. Asghari, Recent progresses in ceramic hollow-fiber membranes, Chem. Bio. Eng. Rev., 2 (2015) 54–70.
[14] M. Rezakazemi, A. Dashti, M. Asghari, S. Shirazian, H2-selective mixed matrix membranes modeling using ANFIS, PSO-ANFIS, GA-ANFIS, Intl. J. Hydrogen Energy, 42 (2017) 15211–15225.
[15] V. Zargar, M. Asghari, A. Dashti, A review on chitin and chitosan polymers: structure, chemistry, solubility, derivatives, and applications, Chem. Bio. Eng. Rev., 2 (2015) 204–226.
[16] R. Singh, Production of high-purity water by membrane processes, Desal. Water Treat., 3 (2009) 99–110.
[17] B. Van der Bruggen, M. Mänttäri, M. Nyström, Drawbacks of applying nanofiltration and how to avoid them: A review, Sep. Purif. Technol., 63 (2008) 251–263.
[18] J. Schaep, C. Vandecasteele, Evaluating the charge of nanofiltration membranes, J. Membr, Sci., 188 (2001) 129–136.
[19] Q. Zhang, H. Wang, S. Zhang, L. Dai, Positively charged nanofiltration membrane based on cardo poly (arylene ether sulfone) with pendant tertiary amine groups, J. Membr. Sci., 375 (2011) 191–197.
[20] X.L. Wang, T. Tsuru, M. Togoh, S.I. Nakao, S. Kimura, Transport of organic electrolytes with electrostatic and steric-hindrance effects through nanofiltration membranes, J. Chem. Eng. Japan., 28 (1995) 372–380.
[21] W.R. Bowen, A.W. Mohammad, N. Hilal, Characterisation of nanofiltration membranes for predictive purposes – use of salts, uncharged solutes and atomic force microscopy, J. Membr. Sci., 126 (1997) 91–105.
[22] W.R. Bowen, H. Mukhtar, Characterisation and prediction of separation performance of nanofiltration membranes, J. Membr. Sci., 112 (1996) 263–274.
[23] A. Szymczyk, Y. Lanteri, P. Fievet, Modelling the transport of asymmetric electrolytes through nanofiltration membranes, Desalination, 245 (2009) 396–407.
[24] C. Labbez, P. Fievet, F. Thomas, A. Szymczyk, A. Vidonne, A. Foissy, P. Pagetti, Evaluation of the "DSPM" model on a titania membrane: measurements of charged and uncharged solute retention, electrokinetic charge, pore size, and water permeability, J. Colloid. Interface Sci., 262 (2003) 200–211.
[25] S. Bandini, D. Vezzani, Nanofiltration modeling: the role of dielectric exclusion in membrane characterization, Chem. Eng. Sci., 58 (2003) 3303–3326.
[26] H. Al-Zoubi, N. Hilal, N.A. Darwish, A.W. Mohammad, Rejection and modelling of sulphate and potassium salts by nanofiltration membranes: neural network and Spiegler–Kedem model, Desalination, 206 (2007) 42–60.
[27] M. Rostamizadeh, M. Rezakazemi, K. Shahidi, T. Mohammadi, Gas permeation through H_2-selective mixed matrix membranes: Experimental and neural network modeling, Int. J. Hydrogen Energy, 38 (2013) 1128–1135.
[28] M. Rezakazemi, T. Mohammadi, Gas sorption in H_2-selective mixed matrix membranes: Experimental and neural network modeling, Int. J. Hydrogen Energy, 38 (2013) 14035–14041.
[29] M. Rezakazemi, S. Razavi, T. Mohammadi, A.G. Nazari, Simulation and determination of optimum conditions of pervaporative dehydration of isopropanol process using synthesized PVA–APTEOS/TEOS nanocomposite membranes by means of expert systems, J. Membr. Sci., 379 (2011) 224–232.
[30] N. Azizi, M. Rezakazemi, M.M. Zarei, An intelligent approach to predict gas compressibility factor using neural network model, Neural Comput. Applications, (2017) 1–10.
[31] W. Richard Bowen, M.G. Jones, H.N.S. Yousef, Prediction of the rate of crossflow membrane ultrafiltration of colloids: A neural network approach, Chem. Eng. Sci., 53 (1998) 3793–3802.
[32] N.A. Darwish, N. Hilal, H. Al-Zoubi, A.W. Mohammad, Neural networks simulation of the filtration of sodium chloride and magnesium chloride solutions using nanofiltration membranes, Chem. Eng. Res. Des., 85 (2007) 417–430.
[33] T.M. Lee, H. Oh, Y.K. Choung, S. Oh, M. Jeon, J.H. Kim, S.H. Nam, S. Lee, Prediction of membrane fouling in the pilot-scale microfiltration system using genetic programming, Desalination, 247 (2009) 285–294.
[34] A. Okhovat, S.M. Mousavi, Modeling of arsenic, chromium and cadmium removal by nanofiltration process using genetic programming, Appl. Soft Comput., 12 (2012) 793–799.
[35] R. Du, J. Zhao, Properties of poly (N, N-dimethylaminoethyl methacrylate)/polysulfone positively charged composite nanofiltration membrane, J. Membr. Sci., 239 (2004) 183–188.
[36] Y.C. Chiang, Y.Z. Hsub, R.C. Ruaan, C.J. Chuang, K.L. Tung, Nanofiltration membranes synthesized from hyperbranched polyethyleneimine, J. Membr. Sci., 326 (2009) 19–26.
[37] A. Akbari, H. Solymani, S.M.M. Rostami, Preparation and characterization of a novel positively charged nanofiltration membrane based on polysulfone, J. Applied. Polym. Sci., 132 (2015).
[38] S. Bila, Y. Harkouss, M. Ibrahim, J. Rousset, E. N'Goya, D. Baillargeat, S. Verdeyme, M. Aubourg, P. Guillon, An accurate wavelet neural-network-based model for electromagnetic optimization of microwave circuits, Intl. J. RF Microwave Computer-Aided Eng., 9 (1999) 297–306.
[39] M. Kimura, R. Nakano, Dynamical systems produced by recurrent neural networks, Syst. Comput. Japan, 31 (2000) 77–86.
[40] M. Khayet, C. Cojocaru, M. Essalhi, Artificial neural network modeling and response surface methodology of desalination by reverse osmosis, J. Membr. Sci., 368 (2011) 202–214.
[41] C. Cojocaru, M. Macoveanu, I. Cretescu, Peat-based sorbents for the removal of oil spills from water surface: Application of artificial neural network modeling, Colloids Surfaces A: Physicochem. Eng. Asp., 384 (2011) 675–684.
[42] M. Khayet, C. Cojocaru, M. Essalhi, Artificial neural network modeling and response surface methodology of desalination by reverse osmosis, J. Membr. Sci., 368 (2011) 202–214.
[43] E. Soroush, S. Shahsavari, M. Mesbah, M. Rezakazemi, Z. Zhang, A robust predictive tool for estimating CO2 solubility in potassium based amino acid salt solutions, Chin. J. Chem. Eng., (2017).
[44] R. Foroutan, H. Esmaeili, M. Abbasi, M. Rezakazemi, M. Mesbah, Adsorption behavior of Cu(II) and Co(II) using chemically modified marine algae, Environ. Technol., (2017) 1–9.
[45] M. Mesbah, E. Soroush, M. Rezakazemi, Development of a least squares support vector machine model for prediction of natural gas hydrate formation temperature, Chin. J. Chem. Eng., 25 (2017) 1238–1248.
[46] J.R. Koza, Genetic Programming, The MIT Press, Cambridge MA, USA (1992).
[47] B. Grosman, D.R. Lewin, Automated nonlinear model predictive control using genetic programming, Comput. Chem. Eng., 26 (2002) 631–640.
[48] M. Rezakazemi, A. Dashti, M. Asghari, S. Shirazian, H_2-selective mixed matrix membranes modeling using ANFIS,

PSO-ANFIS, GA-ANFIS, Int. J. Hydrogen Energy, 42 (2017) 15211–15225.
[49] X.H. Wang, Y.G. Li, Y.D. Hu, Y.L. Wang, Synthesis of heat-integrated complex distillation systems via Genetic Programming, Comput. Chem. Eng., 32 (2008) 1908–1917.
[50] W. Yuan, A. Odjo, N.E. Sammons, J. Caballero, M.R. Eden, Process structure optimization using a hybrid disjunctive-genetic programming approach, Comput. Aided Chem. Eng., 27 (2009) 669–674.
[51] A. Das, M. Abdel-Aty, A genetic programming approach to explore the crash severity on multi-lane roads, Accid. Analysis & Prevention, 42 (2010) 548–557.
[52] H. Demuth, M. Beale, Neural network toolbox for use with MATLAB, Mathworks, Massachusetts, (1993).
[53] A. Khataee, M. Kasiri, Artificial neural networks modeling of contaminated water treatment processes by homogeneous and heterogeneous nanocatalysis, J. Mol. Catal. A: Chem., 331 (2010) 86–100.