Desalination and Water Treatment
www.deswater.com

doi: 10/5004/dwt.2012.2748

44 (2012) 83–91
June

Prediction of permeate flux and ionic compounds rejection of sugar beet press water nanofiltration using artificial neural networks

Mostafa Shahidi Noghabi[a], Seyed Mohammad Ali Razavi[a,*], Seyed Mahmoud Mousavi[b]

[a]*Department of Food Science and Technology, Ferdowsi University of Mashhad, P.O. Box 91775-1163, Mashhad, Iran*
Tel. 98 511 8795620 Ext. 285; Fax: +98 511 8787430; emails: s.razavi@um.ac.ir, sma_razavi@yahoo.com
[b]*Chemical Engineering Department, Ferdowsi University of Mashhad, P.O. Box 91775-1111, Mashhad, Iran*

Received 23 February 2011; Accepted 31 October 2011

ABSTRACT

Artificial neural network (ANN) models were used to predict the permeate flux and rejection of ionic compounds (Na^+, K^+, Ca^{2+}, Mg^{2+}, SO_4^{2-}, Cl^-) of sugar beet press water through polyamide nanofiltration membrane. Experimental data was obtained at different transmembrane pressures (10, 15 and 20 bar), temperatures (25, 40 and 55°C) and feed concentrations (1–3 °Bx). The effect of the number of training points, the number of hidden neurons (H), type of transfer function and learning rule on the accuracy of prediction were studied. According to the results obtained for the best ANNs, 15% of the data was used to generate the model for the prediction of flux, and cross validation was performed with 40% of the total data. Independent flux predictions were also determined for the remaining 45% of the data. While for the prediction of the rejection of ionic compounds, 50%, 25% and 25% of the total data was used to learn the network, cross validation and testing ANN model, respectively. The modeling results showed that the overall agreement between ANN predictions and experimental data was excellent for both permeate flux and rejections ($r = 0.998$ and $r = 0.974$, respectively). Furthermore, sensitivity analysis indicated that temperature and Brix have the most effect on the prediction of flux and rejections (except for Ca rejection) by ANN, respectively.

Keywords: Membrane; Modeling; Flux; Rejection; Effluent; Wastewater

1. Introduction

Press water derives from the pressing station of the extracted pulp after it has passed through the extraction unit [1]. Press water essentially contains 1–3% total solids including sugars (60–80% of total solids) and salts, colloids, and suspended impurities (20–40% of total solids). This stream is very important in sugar industry, because it is produced in a very large amount (about 0.6 kg kg^{-1} of beet input). In addition, the presence of sugar and impurities in the press water affects the sugar extraction efficiency with consequent lowering of the overall productivity of the juice concentration and purity. Since press water is a very dilute stream, its direct evaporation does not represent an economic way for concentrating and recovering clean water for sugar extraction [1].

Bogliolo et al. [1] believed that the reverse osmosis (RO) treatment of press water could provide permeate, mainly consisting of clean water, to be used in the extraction unit, and a concentrate to be sent to the low grade sugar crystallization. All these may lead to important benefits in the sugar production cycle, such as a reduction in the amount of diffusion juice, a reduction in the total amount of water to be evaporated, and a higher purity of the thick juice.

*Corresponding author.

Nanofiltration (NF) membranes are relatively new classes of membranes that have properties between those of ultrafiltration (UF) and RO. It seems that NF is a suitable alternative for RO in the treatment of sugar beet press water. The advantages of NF over RO are lower energy consumption by 21%, higher flux and lower fouling [2]. The NF separation mechanisms involve both steric (Sieving) and the electrical (Donnan) effects. This combination allows NF membranes to be effective in separating mixtures of neutral and charged small organic solutes and salts [3]. Rejection of the molassogenic ions and permeate flux are very important factors to evaluate the industrial application of NF for press water treatment.

Predicting the performance of NF membrane separations is necessary for the design and optimization of the process. Artificial neural networks (ANNs) are capable of modeling highly complex and non-linear systems with large numbers of inputs and outputs [4]. ANNs have been used in several studies relevant to membrane technology, for example to predict, the evolution of membrane fouling during cross-flow microfiltration (MF) and UF of cane sugar and gum streams [5], the permeate flux and rejection for RO of ethanol and acetic acid and the UF of bleach plant effluent [6], the rate of UF of proteins and the dynamic crossflow UF rate of colloids [7,8], the evolution of flux and deposit thickness in bentonite suspension MF [9], the dynamic permeate flux, total hydraulic resistance and the milk components rejection (protein, fat, lactose, ash and total solids) as a function of UF transmembrane pressure and processing time [10], the membrane fouling during NF of ground and surface water [11], the steady-state contaminant removal during NF of ground and surface waters under conditions typical of drinking water treatment [11], the flux decline in crossflow MF of a mixture that contains phosphate and fly ash [12], cross-flow filtration of different highly concentrated salt solutions [13] and the rejection of neutral organic compounds by polyamide NF and RO membranes [14]. For a very complex system, such as a NF of real waste water, creating a mathematical model and applying a model-based control algorithm can involve hard work, whereas a ANN can be implemented in a more straightforward way [15].

According to the literature, no published work has been reported on the neural network modeling of press water NF. Therefore, the aim of the present work was to develop and validate the ANN models for prediction of permeate flux and ionic compounds rejection during cross flow NF of sugar beet press water based on the experimental data, which was obtained at different transmembrane pressures, temperatures and feed concentrations.

2. Materials and methods

2.1. Membrane system

The pilot plant membrane system used in this study was equipped with a feed tank (20 l), a reciprocating pump, tubular module, two pressure gauges, a tubular heat exchanger, two control valves and a temperature sensor. The characteristics of NF membrane are summarized in Table 1. An electronic balance (±0.01 g) was used to automatically record the weight of permeate every 60 s by a computer for the flux calculating. The observed rejection (R_{obs}) of each ionic compound was calculated after the 60 min operation when fluxes were steady state according to the following equation:

$$R_{obs} = 1 - \frac{C_p}{C_f} \tag{1}$$

where C_p and C_f are the concentrations of each component (Na^+, K^+, Ca^{2+}, Mg^{2+}, SO_4^{2-}, Cl^-) in the permeate and the feed streams, respectively.

2.2. Analytical methods

Sodium and potassium ions of permeate and feed samples were measured using a Betalyser instrument (Dr. Wolfgan, Kernchen, Germany). Calcium, magnesium, sulfate and chloride were measured based on AOAC methods [16]. All measurements were carried out at least twice.

2.3. Experimental procedure

Sugar beet press water was prepared from Abkouh Sugar factory, Mashhad, Iran and then stored in several

Table 1
Characteristics of nanofiltration membrane and module used in this study

Membrane type	AFC80 (ITT PCI Membranes Ltd, UK)
Material	Polyamide
Effective area (cm^2)	24
Pore diameter (nm)	0.68
Rang of pH tolerance	1.5–10.5
Max. temperature (°C)	70
Maximum pressure (bar)	60
Apparent retention character	80% NaCl
Module	Tubular (MIC-RO 240) (ITT PCI Membranes Ltd, UK)

containers (25 l) at −20°C. For each run, one container was defrosted and the effect of varying temperature (25, 40 and 55°C), transmembrane pressure (10, 15 and 20 bar) and feed concentration (1–3 °Bx) on permeate flux, and ionic compounds rejection (Na^+, K^+, Ca^{2+}, Mg^{2+}, SO_4^{2-}, Cl^-) were studied. NF experiments were carried out in batch mode at constant pH and flow rate (5.7 and 18.07 kg min^{-1}, respectively). All experimental runs were repeated twice.

2.4. ANN modeling

In this study, fully interconnected multilayer feed-forward network was applied for modeling flux and ionic compounds rejections in NF process of press water. Multilayer perceptron (MLP) network consists of (a) an input layer with neuron(s) representing input variables, (b) an output layer with neuron(s) representing the dependent variable(s), and (c) one hidden layer containing neuron(s) to help capture the nonlinearity in the system. Figs. 1 (a) and (b) shows schematically two individual ANNs that are constructed to predict permeate flux and ions rejections, respectively. It can be seen that the four inputs including pressure, temperature, time and feed concentration were used to model the flux, while the time as input was omitted for prediction of ionic compounds rejection.

In the modeling process, there are several variables that have effects on the ANN predictability performance. Generally, these variables are the number of hidden layers (L), the number of hidden neurons (H), the type of transfer function, the type of training rule and the percentage of used data for training, validating and testing stages [4,10,17]. To find the best set of these variables and parameters, all of them must be varied and the best combination should be chosen. Principally, for an ANN to perform an acceptable prediction, a set of weights that minimizes the error between target and predicted outputs should be found. The ANN predictability can be judged by a combination of some statistical parameters such as mean-squared error (MSE), normalized mean-squared error (NMSE), mean absolute error (MAE), correlation coefficient (r), the H, the L and the number of iterations or epoch (C) [10,17]. In this study, MSE, NMSE, MAE, and r for each output were calculated by the following equations [17–19]:

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N} (O_i - T_i)^2}{N} \tag{2}$$

$$\mathrm{NMSE} = \frac{1}{\sigma^2}\frac{1}{N}\sum_{i=1}^{N} (O_i - T_i)^2 \tag{3}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} \left|O_i - T_i\right| \tag{4}$$

$$r = \sqrt{1 - \frac{\sum_{i=1}^{N}\left[O_i - T_i\right]^2}{\sum_{i=1}^{N}\left[O_i - T_m\right]^2}} \tag{5}$$

where O_i is the ith actual value, T_i is the ith predicted value, N is the number of data, σ^2 is the variance, and:

$$T_m = \frac{\sum_{i=1}^{N} O_i}{N} \tag{6}$$

The universal approximation theory suggests that a network with a single hidden layer with sufficient H can map any input to any output to any degree of accuracy [20]. Thus the ANNs used in the present work featured with a single hidden layer and bias nodes in the input and hidden layers. The bias neurons accept no input and transmit a constant output equal to one in order to preserve the universal approximation property of the network [10,17]. The optimum number of neurons in the hidden layer was determined by a trial and error procedure based on minimizing the difference between estimated ANN outputs and experimental values. In total, 1620 data for the modeling of flux and 162 data for the modeling of rejections were collected by experimental trails. First, the data order was randomized and then the data was divided into three partitions. The first partition (training data) was used to perform training the network. The second one (cross validation data) was used to evaluate the prediction quality of the network during the training stage. For the purpose of estimating the performance of the trained network on new data, a third partition, which was never seen by the ANN during the training and cross-validation steps, was used for testing.

To select the best transfer function, the sigmoid and the hyperbolic tangent functions were tested to transfer neuron inputs to calculate flux output, while the sigmoid function, the hyperbolic tangent and linear hyperbolic tangent were evaluated to transfer neuron inputs to calculate ionic compound rejections. In addition, two learning rules (including momentum and Levenberg Marquardt) and different percentages data were examined to train, validate and test the ANNs. For validating of momentum as a learning rule, momentum value was fixed at 0.7, and learning rate was determined at level 1 on the hidden layer and 0.1 on the output layer. As a network is training, we may want to know the effect that each of the network inputs is having on the network output.

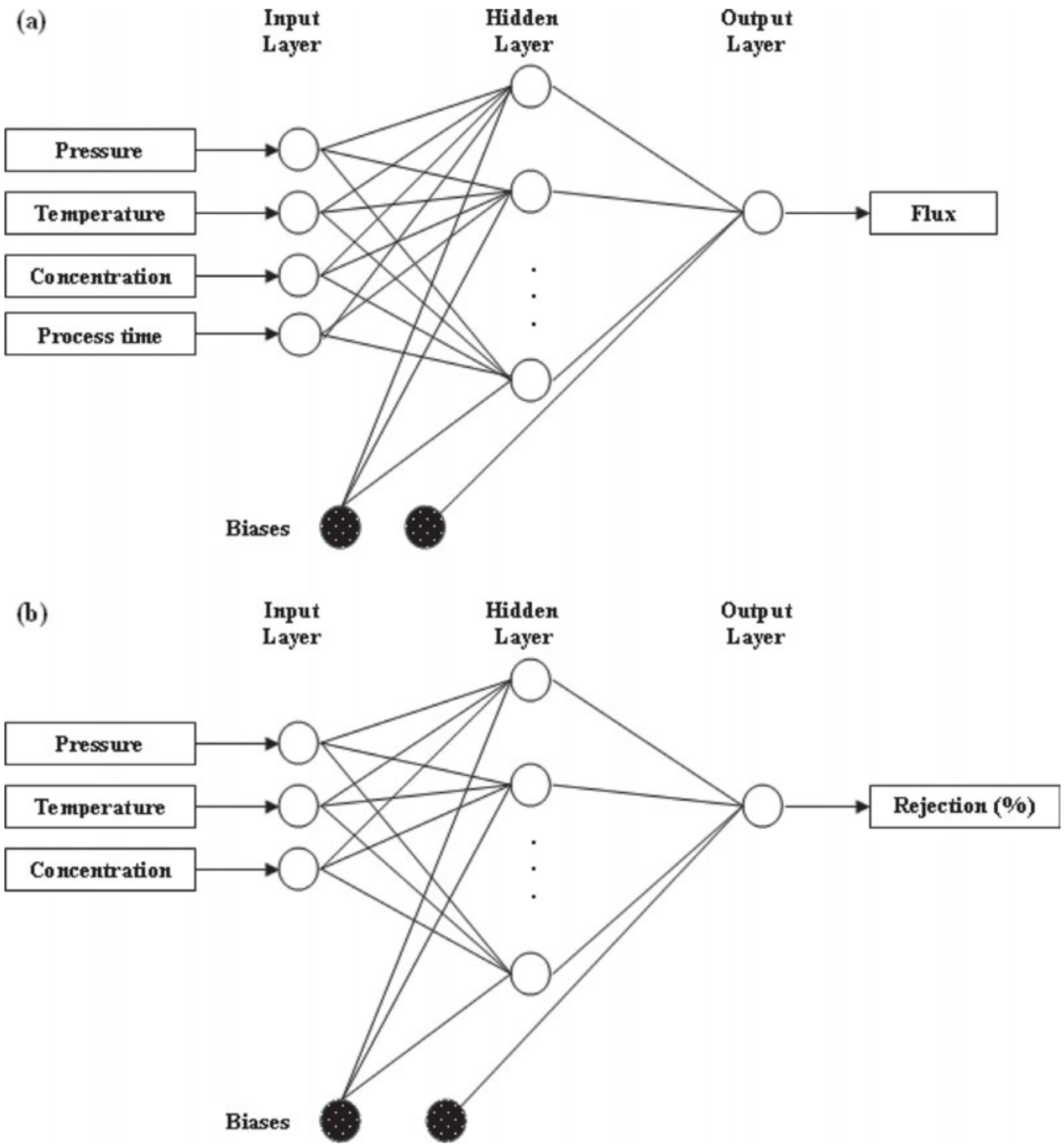

Fig. 1. Schematic of multi-layer feed forward neural network architectures with one hidden layer used for prediction of (a) permeate flux, and (b) ionic compounds rejections.

This provides feedback as to which input channels are the most significant. From there, we may decide to prune the input space by removing the insignificant channels. This will reduce the size of the network, which in turn reduces the complexity and the training times. *Sensitivity analysis* is a method for extracting the cause and effect relationship between the inputs and outputs of the network. The network learning is disabled during this operation such that the network weights are not affected. The basic idea is that the inputs to the network are shifted slightly and the corresponding change in the output is reported either as a percentage or a raw difference. In this work, the software used for the ANNs modeling was Neuro-Solutions 6 for Microsoft Windows.

3. Results and discussion

The ANN configuration has two main variables (*L* and *H*) that have a strong influence on its outcome. To find the best ANN configuration, different networks were built with different hidden neurons varying from 2 to 20. Depending on the number of neurons, different weight vectors are generated which act as starting points in the error hyper-space. The hidden layer acts as a feature detector. Therefore, the more hidden neurons, the more features can be detected, but the ANN will be more complex and as a result it will require a longer time to train. In addition, the number of local minima will increase, which requires multiple neuron runs as discussed above.

Totally, 62 runs were made to train, validate and test the data in order to find the best combination for prediction of flux (19 different neurons × 2 different transfer functions + 2 different training rules + 21 different percentages of data). However, 77 runs were made to train, validate and test the data in order to select the best combination for prediction of rejection (19 different neurons × 3 different transfer functions + 2 different training rules + 18 different percentages of data). Generally, the ANN with the lowest testing errors (MSE, NSME and MAE) and highest *r* was selected as the best. In some cases, we had to consider more factors in addition to the above parameters such as *C* of the training and cross validation steps for choosing the best combination of the mentioned factors.

The values of modeling parameters obtained for estimation of flux and ions rejection during the testing process of different ANN architectures (with 2–20 neurons in the hidden layer and different transfer functions in the hidden and output layers) are shown in Tables 2 and 3, respectively. It is worth noting that MSE, NMSE, MAE, and *r* values reported for rejections in Table 3 are means of values that are obtained for all ionic compounds rejections for example, sodium, potassium, calcium, etc.

As it is seen in Table 2, the ANN with a sigmoid function in hidden and output layers and six hidden neurons had the minimum values of MSE (0.539), NMSE (0.002), and MAE (0.541), and maximum *r* value (0.998) for prediction of the flux. Although the ANN with a tanh function and 10 neurons in the hidden layer showed very similar results, the first one was chosen as the best ANN model to predict the permeate flux in press water NF because less neurons were used.

As shown in Table 3, the ANN with a linear hyperbolic tangent function in hidden and output layers and 16 hidden neurons had the lowest values of MSE (0.193), NMSE (0.086) and MAE (0.304) for prediction of the ions rejection. Although, this ANN did not show the highest *r* value among the ANNs obtained, the difference between *r* values of this architecture and two other proper ANNs, for example, ANN with a hyperbolic tangent or sigmoid functions in hidden and output layers and two hidden neurons, was not considerable. More details about the ANN parameters of three selected ANNs for predictions of the rejection are given in Table 4. It can be found that all parameters obtained for the ANN with linear hyperbolic tangent function in hidden and output layers and 16 hidden neurons, were the lowest in comparison with

Table 2
Different architectures of ANN with different neurons in the hidden layer and transfer functions in the hidden and output layers used for prediction of permeate flux in sugar beet press water nanofiltration

Hidden layer neurons	Tanh function				Sigmoid function			
	MAE	NMSE	MAE	*r*	MSE	NMSE	MAE	*r*
2	197.03	0.532	1.99	0.684	5.011	0.027	1.688	0.986
3	195.99	0.517	1.73	0.69	5.123	0.025	1.707	0.987
4	3.803	0.022	1.467	0.988	2.6	0.014	1.299	0.994
5	2.447	0.013	1.205	0.993	1.458	0.007	0.914	0.996
6	1.374	0.007	0.885	0.996	**0.539**	**0.002**	**0.541**	**0.998**
7	1.797	0.009	1.052	0.996	2.917	0.016	1.286	0.991
8	5.163	0.028	1.468	0.987	3.092	0.017	1.309	0.991
9	3.049	0.014	1.271	0.992	0.701	0.003	0.66	0.998
10	**0.552**	**0.002**	**0.509**	**0.998**	3.161	0.018	1.347	0.992
11	2.901	0.016	1.252	0.992	192.51	0.497	1.126	0.709
12	4.55	0.026	1.456	0.987	0.821	0.004	0.674	0.998
13	3.927	0.022	1.32	0.988	3.076	0.017	1.407	0.991
14	4.972	0.027	1.575	0.986	0.861	0.005	0.725	0.998
15	4.218	0.023	1.344	0.988	3.264	0.019	1.423	0.99
16	0.462	0.002	0.51	0.998	3.904	0.021	1.542	0.989
17	6.171	0.035	1.616	0.983	0.709	0.003	0.639	0.998
18	0.386	0.002	0.496	0.998	3.126	0.017	1.422	0.991
19	3.909	0.02	1.376	0.989	3.493	0.018	1.427	0.99
20	4.468	0.024	1.453	0.989	1.418	0.008	0.948	0.997

Table 3
Different architectures of ANN with different neurons in the hidden layer and transfer functions in the hidden and output layers used for prediction of ionic compounds rejections in sugar beet press water nanofiltration

No. of neurons	Linear tanh function				Tanh function				Sigmoid function			
	MSE	NMSE	MAE	r	MSE	NMSE	MAE	r	MSE	NMSE	MAE	r
2	0.787	0.203	0.64	0.956	**0.355**	**0.105**	**0.409**	**0.977**	**0.401**	**0.113**	**0.404**	**0.975**
3	1.02	0.248	0.688	0.918	1.02	0.21	0.642	0.918	0.97	0.183	0.588	0.944
4	1.205	0.482	0.754	0.818	0.524	0.322	0.531	0.847	2.878	1.205	0.855	0.437
5	2.465	2.0608	1.159	0.521	0.709	0.379	0.615	0.832	1.012	0.6	0.689	0.778
6	0.69	0.164	0.619	0.946	0.419	0.13	0.485	0.958	0.982	0.25	0.668	0.932
7	1.067	0.595	0.747	0.804	2.474	0.932	1.037	0.68	0.286	0.196	0.36	0.932
8	1.07	0.744	0.698	0.907	1.604	0.871	0.748	0.883	0.958	0.679	0.721	0.892
9	1.683	0.484	0.829	0.935	0.589	0.197	0.487	0.929	1.08	0.304	0.669	0.959
10	1.159	0.362	0.664	0.883	0.461	0.205	0.471	0.899	0.679	0.199	0.583	0.917
11	0.762	0.63	0.642	0.825	0.671	0.46	0.554	0.853	0.403	0.424	0.472	0.823
12	0.888	0.745	0.641	0.841	0.756	0.5	0.616	0.856	1.227	1.081	0.81	0.753
13	1.869	0.99	1.022	0.688	0.306	0.171	0.413	0.943	0.492	0.293	0.524	0.909
14	1.433	0.534	0.831	0.8	0.559	0.591	0.548	0.718	0.929	0.595	0.652	0.781
15	1.878	0.388	0.945	0.895	1.123	0.196	0.709	0.953	1.231	0.208	0.798	0.955
16	**0.193**	**0.086**	**0.304**	**0.974**	0.616	0.159	0.506	0.964	0.375	0.135	0.444	0.952
17	1.786	0.522	0.937	0.799	1.246	0.307	0.698	0.951	1.325	0.3	0.702	0.947
18	1.359	0.349	0.77	0.883	0.295	0.123	0.381	0.961	0.992	0.23	0.677	0.929
19	0.739	0.331	0.626	0.866	0.414	0.181	0.464	0.923	1.521	0.585	0.85	0.762
20	1.522	0.557	0.831	0.85	1.913	0.76	0.936	0.811	2.007	0.68	0.936	0.859

Table 4
More information about three selected architectures of ANN selected to model the rejection that have the minimum MSE with different transfer function and number of neurons

No. of neurons	Transfer function	Training epoch	Validating epoch	Minimum MSE of training	Final MSE of training	Minimum MSE of validation	Final MSE of validation
16	Linear Tanh	**42**	**17**	**0.0008**	**0.0008**	**0.0260**	**0.0360**
2	Tanh	125	25	0.0240	0.0240	0.0433	0.0473
2	Sigmoid	107	17	0.0063	0.0063	0.0158	0.0309

other networks. Therefore, this one was chosen as the best ANN to model the ionic compounds rejection.

Comparing two learning rules used for selected ANNs, it was found that the Levenberg Marquardt rule demonstrates better results than the momentum rule in modeling the flux and rejection (Table 5).

In ANNs modeling, the less data used to train, the better ANN results will be for industrial applications, since the time and the cost of experiments would be reduced. In fact, when the percentage of training/validating data is less and the percentage of testing data is more, the ANN is better if the power of network to predict the output based on the new input will be higher. Therefore, different percentages of data (5–60% of total data) was used in this study for training step and MSE, NMSE, MAE and r values of each run was determined (Table 6). It can be seen that 15% of the whole data was enough to obtain the best ANN (lowest MSE, NMSE, MAE and highest r) for flux modeling. In the next step, 15% of total data was used for training step of each run and the percentage of used data for cross validation was changed from 5% to 45%. As it is shown in Table 7, when 40% of total data was applied for cross validating the networks, the minimum MSE (0.677), NMSE (0.003), MAE (0.633) and maximum r (0.998) values were achieved for modeling the permeate flux.

Table 5
Comparison of two learning rules used for selected ANN architectures to predict the permeate flux and ionic compounds rejection in sugar beet press water nanofiltration

Parameter	No. of neurons	Levenberg Marquardt				Momentum			
		MSE	NMSE	MAE	r	MSE	NMSE	MAE	r
Flux	6	**0.539**	**0.002**	**0.541**	0.998	13.312	0.072	2.853	0.970
Rejection	16	**0.193**	**0.086**	**0.304**	0.974	0.495	0.125	0.462	0.962

Table 6
Comparison of different percentages of data used for training of selected ANN architectures to model the permeate flux

Training data (%)	Validation data (%)	Testing data (%)	MSE	NMSE	MAE	r
5	47.5	47.5	2.019	0.01	1.011	0.994
10	45	45	0.956	0.005	0.7	0.997
15	**42.5**	**42.5**	**0.57**	**0.003**	**0.572**	**0.998**
20	40	40	1.243	0.007	0.868	0.997
25	37.5	37.5	0.756	0.004	0.586	0.997
30	35	35	1.014	0.006	0.712	0.997
35	32.5	32.5	0.437	0.002	0.504	0.998
40	30	30	4.161	0.023	1.534	0.988
45	27.5	27.5	1.868	0.01	1.121	0.997
50	25	25	3.728	0.02	1.242	0.99
55	22.5	22.5	1.713	0.009	0.938	0.995
60	20	20	2.411	0.012	0.913	0.993

Table 7
Comparison of different percentages of data used for cross-validation and testing of selected ANN architectures to model the permeate flux

Training data (%)	Validation data (%)	Testing data (%)	MSE	NMSE	MAE	r
15	5	80	61.233	0.249	0.896	0.866
15	10	75	64.838	0.251	0.862	0.865
15	15	70	1.013	0.005	0.786	0.997
15	20	65	0.887	0.004	0.677	0.997
15	25	60	1.606	0.008	0.967	0.996
15	30	55	88.35	0.316	0.946	0.827
15	35	50	1.272	0.007	0.844	0.997
15	**40**	**45**	**0.677**	**0.003**	**0.633**	**0.998**
15	45	40	1.187	0.006	0.853	0.997

Similar to the procedure used for flux, different percentages of data (15–60% of total data) were applied to train the ANN selected for prediction of ions rejection and the obtained MSE, NMSE, MAE and r values of each run were presented in Table 8. It can be found that 50%

Table 8
Comparison of different percentages of data used for training of selected ANN architectures to model the ionic compounds rejection

Training data (%)	Validation data (%)	Testing data (%)	MSE	NMSE	MAE	r
15	42.5	42.5	5.068	1.208	1.482	0.673
20	40	40	3.951	1.526	1.414	0.586
25	37.5	37.5	1.895	0.547	0.855	0.908
30	35	35	1.312	0.412	0.785	0.865
35	32.5	32.5	2.294	0.605	0.897	0.826
40	30	30	1.487	0.267	0.718	0.946
45	27.5	27.5	1.164	0.196	0.631	0.926
50	**25**	**25**	**0.193**	**0.086**	**0.304**	**0.974**
55	22.5	22.5	0.517	0.204	0.51	0.938
60	20	20	0.214	0.148	0.354	0.965

of the whole data was enough to obtain the best ANN based on the lowest MSE, NMSE, MAE and highest r values. In the next step, 50% of total data was used for training step of each run and different percentages of data (5–40%) were applied for cross validation step. As it is shown in Table 9, when 25% of the total data was used for cross validating, the resulted ANN represented the lowest MSE (0.193), NMSE (0.086), MAE (0.304) and the highest r (0.974) values.

As it is shown in Fig. 2, symbols were used to identify experimental data points of flux which were obtained at TMP 20 bar, temperature 40°C and different feed concentration (1–3 °Bx), while the lines show data points which were predicted by the chosen ANN in the testing phase. It can be seen that the ANN successfully predicts the dynamic behavior of flux both for validating and testing data. Similar results were obtained for other operating conditions (figures not shown).

The prediction efficiency of the chosen ANN model (3/16/1) for rejection of Na^+, K^+, Ca^{2+}, Mg^{2+}, SO_4^{-2} and Cl^- is presented in Figs. 3 (a)–(f), in which the predicted rejection values are plotted against their experimentally measured values. The calculated r values for estimation of Na^+, K^+, Ca^{2+}, Mg^{2+}, SO_4^{-2} and Cl^- were obtained 0.99,

Table 9
Comparison of different percentages of data used for cross-validation and testing of selected ANN architectures to model the ionic compounds rejection

Training data (%)	Validation data (%)	Testing data (%)	MSE	NMSE	MAE	r
50	5	45	3.51	0.91	1.379	0.583
50	10	40	0.847	0.187	0.603	0.932
50	15	35	0.668	0.211	0.578	0.926
50	20	30	0.228	0.12	0.348	0.946
50	**25**	**25**	**0.193**	**0.086**	**0.304**	**0.974**
50	30	20	0.552	0.163	0.479	0.942
50	35	15	0.535	0.118	0.556	0.97
50	40	10	0.926	0.352	0.568	0.931

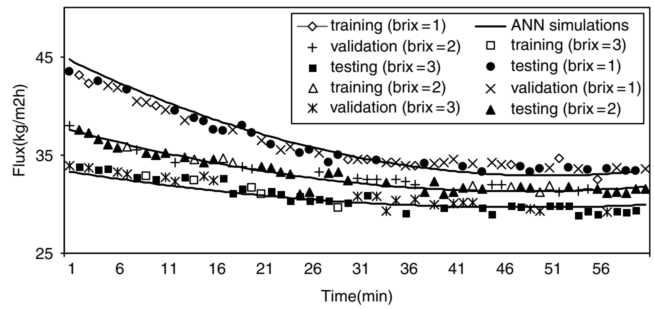

Fig. 2. Dynamic prediction of permeate flux during nanofiltration of sugar beet press water in different concentrations. Lines show the ANN prediction and symbols show the experimental data that was used for training (■), validating (×) and testing (○). ANN used: 4/6/1. Training points/validation points/testing points: 243/648/729 (operating conditions: TMP = 20 bar and T = 40°C).

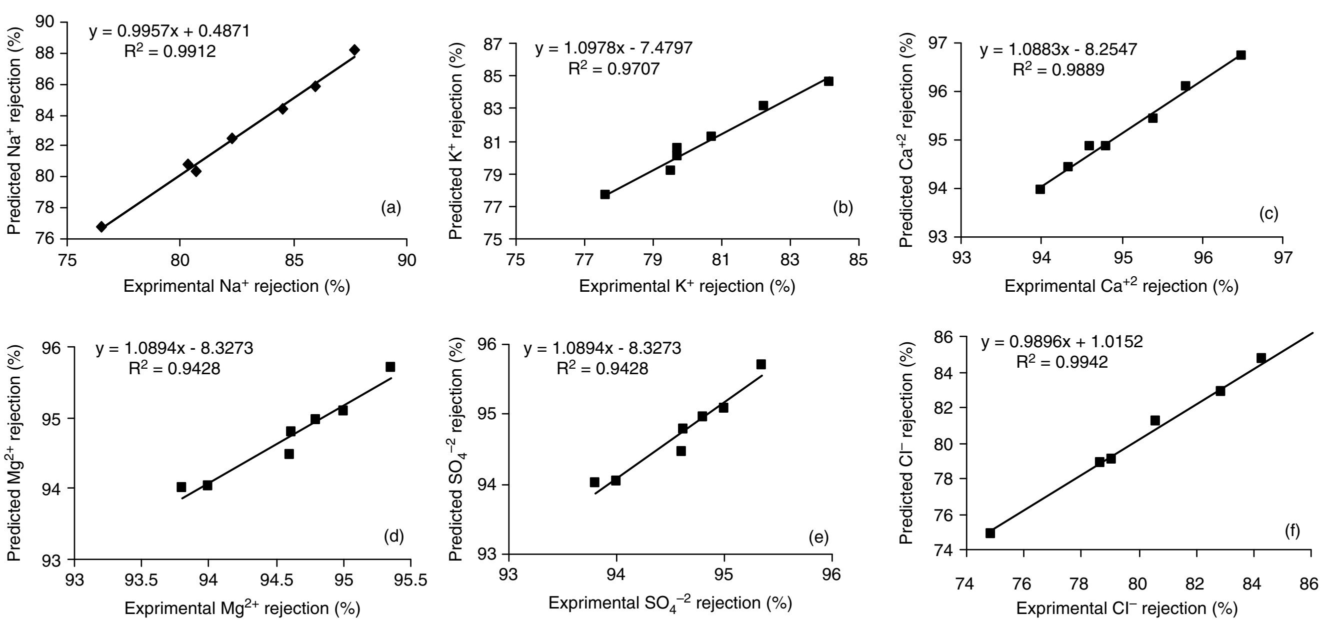

Fig. 3. Experimental and predicted ionic compounds rejections of sugar beet press water nanofiltration. ANN used: 3/16/1. Training points/validation points/testing points: 81/40/41.

0.97, 0.98, 0.94, 0.94 and 0.99, respectively, which are acceptable and revealed excellent agreement between predicted and experimental values.

Finally, sensitivity analysis was tested in order to study the sensitivity of neural network models toward different inputs. This testing process provides a measure of the relative importance among the inputs of the neural model and illustrates how the model output varies in response to variation of an input. As shown in Fig. 4, the temperature was the most effective factor in predicting the permeate flux by the selected ANN. Whereas, feed concentration (Brix) was the main sensitive factor for prediction of ionic compounds rejection (except for Ca rejection) by the chosen ANN (Fig. 5).

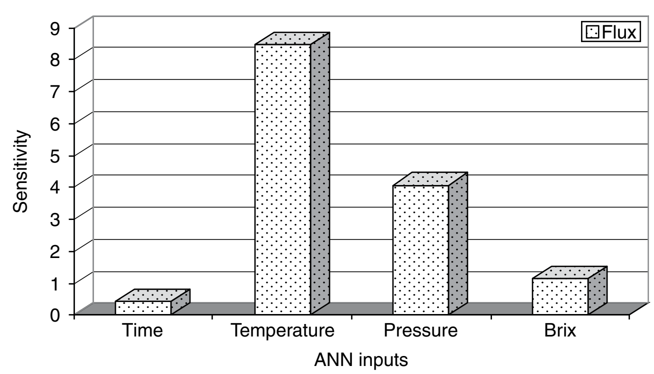

Fig. 4. Sensitivity of the best chosen ANN (4/6/1) toward the inputs for flux prediction.

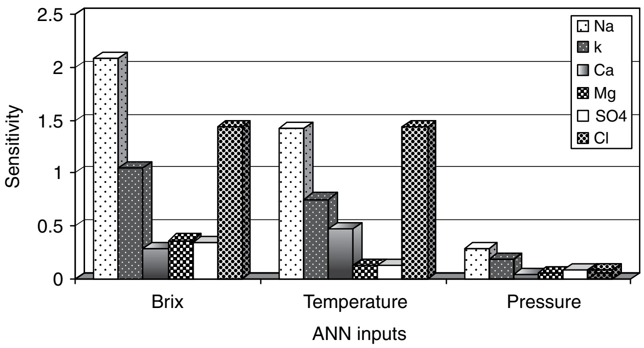

Fig. 5. Sensitivity of the best selected ANN (3/16/1) toward the inputs for prediction of ionic compounds rejection.

4. Conclusions

Totally, the following conclusions can be drawn from this investigation:

1. The multilayer feed forward neural network based on four inputs (pressure, temperature, time and Brix), sigmoid function as the transfer function, six neurons in the single hidden layer and Levenberg Marquardt method as the learning rule was found to be the best ANN for predicting permeate flux of sugar beet press water NF, which showed the lowest MSE (0.677) and the highest r (0.998) values. These results could be obtained by 15%, 40% and 45% of total data for training, cross validation and testing, respectively.
2. The multilayer feed forward neural network based on three inputs (pressure, temperature and Brix), linear hyperbolic tangent function as the transfer function, 16 neurons in the single hidden layer and Levenberg Marquardt as the learning rule was found to be the best ANN for predicting ionic compounds rejection of sugar beet press water NF, which showed minimum MSE (0.173) and maximum r (0.974) values. These results could be obtained using 50%, 25% and 25% of total data for training, cross validation and testing, respectively.
3. There was an excellent agreement between the experimental and the predicted values of permeate flux and ionic compounds rejection, then the efficiency of the selected ANN models were completely acceptable for prediction of both permeate flux and ionic compounds rejection in different conditions (pressure, temperature and concentration).
4. Temperature and Brix were determined as the most sensitive inputs of the best chosen ANNs for the prediction of the flux and ionic compounds rejection (except for Ca rejection), respectively.

Acknowledgments

The authors are greatly thankful to the Iranian Nanotechnology Initiative Council for their financial support and to Mr. Salehi, Miss Zaerzadeh and Miss Mohamadzadeh for their assistance during the experimental works.

References

[1] M. Bogliolo, A. Bottino, G. Capannelli, M. Depetro, A. Servida, G. Pezzi and G. Vallini, Clean water recycle in sugar extraction process: performance analysis of revers osmosis in the treatment of sugar beet press water, Desalination, 108 (1996) 261–271.
[2] J. Warczok, M. Ferrando, F. Lopez and C. Güell, Concentration of apple and pear juices by nanofiltration at low pressures, J. Food Eng., 63 (2004) 63–70.
[3] L.P. Raman, M. Cheryan and N. Rajagopalan, Consider nanofiltration for membrane separations, Chem. Eng. Prog., 90 (1994) 68–74.
[4] W. Richard Bowen, G.J. Meirion, S. Julian, S. Welfoot and N.S.Y. Haitham, Predicting salt rejections at nanofiltration membranes using artificial neural networks, Desalination, 129 (2000) 147–162.
[5] M. Dornier, M. Decloux, G. Trystram and A. Lebert, Dynamic modelling of crossflow microfiltration using neural networks, J. Membr. Sci., 98 (1995) 263–273.
[6] H. Niemi, A. Bulsari and S. Palosaari, Simulation of membrane separation by neural networks, J. Membr. Sci., 102 (1995) 185–191.
[7] W.R. Bowen, M.G. Jones and H.N.S. Yousef, Dynamic ultrafiltration of proteins—a neural network approach, J. Membr. Sci., 146 (1998) 225–235.
[8] W.R. Bowen, M.G. Jones and 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.
[9] M. Hamachi, M. Cabassud, A. Davin and M.M. Peuchot, Dynamic modelling of crossflow microfiltration of bentonite suspension using recurrent neural network, Chem. Eng. Process., 38 (1999) 203–210.
[10] S.M.A. Razavi, S.M. Mousavi and S.A. Mortazavi, Dynamic prediction of milk ultrafiltration performance: a neural network approach, Chem. Eng. Sci., 58 (2003) 4185–4195.
[11] R.S. Grishma and C. Shankararaman, Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks, J. Membr. Sci., 217 (2003) 69–86.
[12] A. Coskun, D. Ibrahim and Y. Ergun, Modeling of flux decline in crossflow microfiltration using neural networks: the case of phosphate removal, J. Membr. Sci., 248 (2005) 53–62.
[13] H. Al-Zoubi, N. Hilal, N.A. Darvish and 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.
[14] V. Yangali-Quintanilla, A. Verliefde, T.U. Kime, A. Sadmani, M. Kennedy and G. Amy, Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes, J. Membr. Sci., 342 (2009) 251–262.
[15] M. Bravi and A. Chianese, Neouro-fuzzy control of a continuous cooled MSMPR crystallizer, Chem. Eng. Technol., 26 (2003) 3.
[16] AOAC, Official methods of analysis of AOAC International vol 1, 18th ed. AOAC International, Gaithersburg, MD, Official Methods (920.199, 920.200, 925.54 and 973.51) (2005).
[17] M.Z. Abdullah, J. Mohamad-Saleh, A.S. Fathinul-Syahir and B.M.N. Mohd-Azami, Discrimination and classification of fresh cut tar fruits (Averrhoa carambola L.) using automated machine vision system, J. Food Eng., 76 (2006) 506–523.
[18] A. Mohebbi, M. Taheri and A. Soltani, A neural network for predicting saturated liquid density using genetic algorithm for pure and mixed refrigerants, Int. J. Refrig., 31 (2008) 1317–1327.
[19] K. Movagharnejad and M. Nikzad, Modeling of tomato drying using artificial neural network, Comput. Electron. Agric., 59 (2007) 78–85.
[20] S.S. Tambe, B.D. Kulkami and P.B. Deshpande, Elements of artificial neural networks with selected applications in chemical engineering, and chemical and biological sciences, Simulation & Advanced Controls Ltd., Louisville, KY, (1996).