Desalination and Water Treatment
www.deswater.com
1944-3994/1944-3986 © 2012 Desalination Publications. All rights reserved
doi: 10.1080/19443994.2012.719743
50 (2012) 317–328
December
Taylor & Francis
Taylor & Francis Group

Development of artificial neural network for prediction of salt recovery by nanofiltration from textile industry wastewaters

Beytullah Eren[a,*], Recep Ileri[b], Emrah Dogan[c], Naci Caglar[c], Ismail Koyuncu[d]

[a]*Department of Enviromental Engineering, Esentepe Campus, Sakarya University, Sakarya 54187, Turkey*
Tel. +90 505 783 6339; Fax: +90 264 295 5601; email: beren@sakarya.edu.tr
[b]*Department of Civil Engineering, Bursa Orhangazi University, Yildirim Campus, Bursa 16310, Turkey*
[c]*Department of Civil Engineering, Sakarya University, Esentepe Campus, Sakarya 54187, Turkey*
[d]*Department of Environmental Engineering, Istanbul Technical University, Maslak, Istanbul 34469, Turkey*

Received 3 November 2011; Accepted 18 June 2012

ABSTRACT

This paper presents the use of artificial neural network (ANN) to develop a model for predicting rejection rate (R_o) of single salt (NaCl) by nanofiltration based on experimental datasets. The rejection rates of NaCl were obtained when operating conditions, such as feed pressure (ΔP) and cross flow velocity (V), varied along with different physicochemical properties of feed water like salt and dye concentrations, and pH. In the modeling work, sensitivity analyses were performed to identify relative impact of each parameter and to find the best combination of input parameters in the ANN model. The optimal network architecture was developed through trial and error approach. Model predictions in each trial were compared with experimental results based on statistical evaluation such as root mean square error, mean absolute error, and coefficient of determination (R^2). Optimal network architecture was determined as one hidden layer with 25 neurons using Levenberg–Marquardt (*trainlm*) backpropagation algorithm. In this architecture, tangent sigmoid (*tansig*) in hidden layer and linear (*purelin*) in output layer was also used as transfer functions. The results showed that the developed ANN model predictions and experimental data matched well and the model can be employed successfully for the prediction of the R_o.

Keywords: Rejection rate; Nanofiltration; Neural network; Modeling; Sodium chloride

1. Introduction

Membrane filtration processes can be used effectively to separate suspended and/or dissolved substances from a liquid. They are categorized based on membrane properties, such as pore size range, molecular weight cut-off range, and operating pressure. Based on these properties, microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO) are among the membrane technologies employed in various industrial applications. In particular, they are used commonly for water reclamation, recycle, and conservation, as well as wastewater treatment and environmental pollution control [1].

Among these systems, NF membranes have pore sizes of around nanometer that can remove most natural and synthetic organic compounds as well as multivalent ions from water and wastewater. For instance, high molecular weight compounds (>1,000 g/mol), polar molecules, and polyvalent ions are captured through NF at low operating pressures. Low molecular weight compounds (200–1,000 g/mol) and monovalent ions can be filtered out under relatively high

*Corresponding author.

operating pressures [2–4]. Accordingly, a large group of substances can be removed by NF to produce high quality water flux [5]. In parallel, NF presents several advantages over other membrane technologies such as operation at relatively high temperature ranges, heat recovery, and relatively low energy consumption [5,6]. Subsequently, NF membranes are commonly used for specific industrial applications. For instance, potable water is produced by the use of NF together with RO and other processes (MF and multi-effect distillation) in an integrated treatment system. NF is also used to remove organic compounds, such as dyes and coloring agents, from industrial wastewaters [7–9].

Development of membrane processes for industrial applications has been accomplished with numerous laboratory-scale experiments that are supported with mathematical modeling of those systems. These studies identified key parameters for design, operation, and optimization of membrane processes [5,10]. For instance, Spiegler and Kedem [11] proposed that the rejection rate of an electrolyte by a membrane depends on salt permeability (Ps) of the membrane and the reflection coefficient (σ). Perry and Linder [12] suggested that the Spiegler–Kedem model considered only salt rejection; therefore, they modified this model to include filtration of salt in the presence of an organic ion. In this regard, they included the number of charges of an organic ion and its concentration in feed water into the Spiegler–Kedem model. Levenstein et al. [13] improved the Perry–Linder model to account for concentration polarization of binary salts. They showed that salt removal by NF membranes can be improved significantly by addition of a polyelectrolyte to water containing salt and organic dye. Similarly, Xu and Spencer [14] considered concentration polarization of a dye with high molecular weight. They defined two parameters to describe concentration polarization of a dye at membrane–feed interface. Dye concentrations in the bulk phase (C_D) and at the membrane–feed interface ($C_D(m)$) were the new parameters that allowed model predictions to fit experimental results accurately. Koyuncu and Topacik [2] further improved the existing models by addition of gel layer-controlled mass transport. They introduced salt and dye concentrations at membrane–feed interface into their model to predict salt rejection rate in the presence of an organic ion [2]. Overall, these models predicted accurately rejection rates of various substances by membrane systems; however, their development required in-depth understanding of the filtration process. They were developed with advanced knowledge on properties of solutes, solvents, membranes, and filtration processes. Subsequently, these models include specific parameters that vary extensively when a slight change occurs in filtration process [15–18]. Alternatively, simple models are also needed to evaluate rejection rates of substances by membrane processes.

In this regard, artificial neural networks (ANNs) are one of the simple modeling techniques to map complex relationships with large numbers of inputs and outputs. Past studies showed that ANN models offer considerable simplicity along with accurate and comparable predictions as the physics-based conventional models [19]. Thus, ANNs have been applied successfully to MF [20,21], UF [22,23], NF [5,10], and RO [24,25]. Although substantial number of studies exists on ANN applications to membranes, only a few studies are available on each case as in NF.

For NF, Bowen et al. [10] developed the first application of ANNs to predict rejection rates of single salts (NaCl, Na_2SO_4, $MgCl_2$, and $MgSO_4$) and their mixtures by NF membranes. Their experimental system was a spiral wound membrane that was difficult to model with conventional physics-based models. Thus, the ANNs served as an alternative to predict experimental rejection rates of single salts and of the salt mixtures by this system. In a similar study by Darwish et al. [5], ANNs are applied to cross-flow NF of two salts (NaCl and $MgCl_2$) at typical seawater concentrations. This study elucidated the effects of input salt concentrations and different operating pressures on rejection rates by various NF membranes. They demonstrated that ANN model predicts successfully the experimental rejection rates of NaCl and $MgCl_2$ by three different NF membranes (NF 90, NF 270, and NF 30). The success of above studies supports potential application of NF membranes to actual wastewaters. However, industrial applications of NF membranes have remained as a challenge to this day. Thus, this study presents the first application of NF membranes to actual wastewater from textile industry. Because the experimental part of this work has been published separately [26], we present here development of our ANN model for assessment of salt recovery with NF from textile industry wastewater.

In textile industry, large quantities of salt are used to enhance dying of fabrics by promoting "salting out" of dyestuff precipitates. The extra salt also removes organic and inorganic contaminants in water as salt precipitates [27]. Although unit cost of salt (NaCl) is relatively low, its recovery from wastewater generates significant amounts of annual savings as well as reducing environmental pollution [28]. Therefore, NF has been proposed as a feasible process to recover salt from textile industry wastewater. For this purpose, Koyuncu [26] conducted experimental studies to determine salt recovery by NF membranes from

textile industry wastewater. In particular, they assessed NaCl rejection rates under various feed phase concentrations of salt (C_{sb}) and dye (C_{db}), feed pressure (ΔP), cross-flow velocity (V), and pH. In this study, we used their experimental work as our basis and we developed an ANN model to predict salt rejection rates (R_o) by NF membranes from textile industry wastewaters. Through rigorous analyses, we identified the best network structure, the best training algorithm, the optimum number of neurons for each layer of the network, and the optimum transfer functions for neurons of each layer. We tested our model with experimental results of Koyuncu [26] indicating that our ANN model can predict accurately salt rejection rates of NF membranes from actual wastewaters. We used our model to evaluate relative impact of each experimental variable. This analysis indicated that initial dye (C_{db}) and salt concentrations (C_{sb}) are relatively more important than feed pressure (ΔP), cross-flow velocity (V), and pH. Collectively, our results from this study offer key information for design, operation, and optimization of industrial level NF processes. Our study supports and indicates that NF can serve as a viable option to recover salt from industrial wastewaters.

2. Material and methods

2.1. Artificial neural networks

The ANNs are computational modeling tools with a flexible structure that is able to capture and simulate complex input/output relationships. The ANN modeling is usually suitable for any specific problem that is not compatible or difficult to model by conventional statistical and mathematical methods [29].

In general, the development of a neural network model consists of the following steps: data collection; analysis and preprocessing of the data; creation and configuration of the network, training, and validation of the network; and finally, simulations and predictions with validated network [30].

The critical step is the creation and configuration of the network. Fig. 1 shows the basic ANN structure that is composed of three layers with number of neurons in each layer. The layers include input layer (independent variables), hidden layer, and output layer (dependent variables). To create the best ANN structure, number of hidden layers, neurons in each hidden layer, training algorithm, and transfer functions must be selected carefully [30]. These tasks can be accomplished with commercially available neural network software packages such as the neural

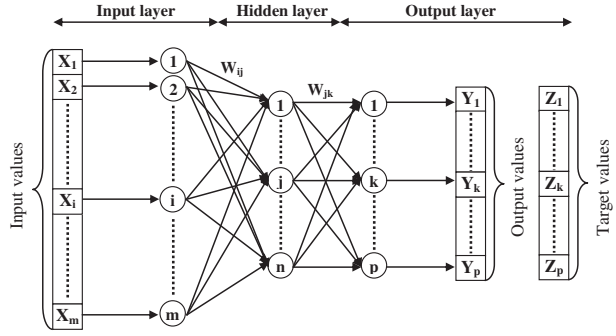

Fig. 1. Architecture of a typical ANN with an input layer, a hidden layer, and an output layer.

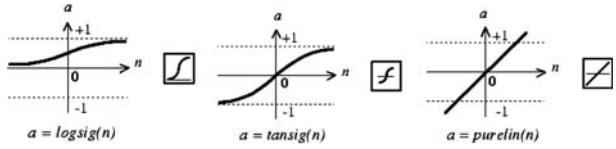

Fig. 2. Graphical representation for activation (transfer) functions (MATLAB).

network toolbox of Matlab v. 7.0 (the Math Works Inc., Natick, MA, USA). We used this software package to construct our ANN model.

In a typical network, input layer is composed of the original experimental data (X_i) that are associated with the neurons or nodes ($1, 2 \ldots, i, \ldots m$) of input layer. Input data are transferred to the nodes of hidden layer ($1, 2 \ldots, j, \ldots n$) and output layer ($1, 2 \ldots, k, \ldots p$) by multiplying connection strength or weights (W_{ij}) between two neurons and summing using summation function. The input data of each layer are processed or converted to outputs by using an activation function that is a nonlinear mathematical function known as a transfer function. The most widely used transfer functions are the tangent sigmoid (*tansig*), the logarithmic sigmoid (*logsig*), and the linear (*purelin*) transfer functions. The *logsig*, *tansig*, and *purelin* transfer functions are mathematically described by Eqs. (1)–(3) and illustrated in Fig. 2, respectively. Among them, *tansig* function offers slightly better predictions than the others [30].

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{1}$$

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

$$f(x) = x \tag{3}$$

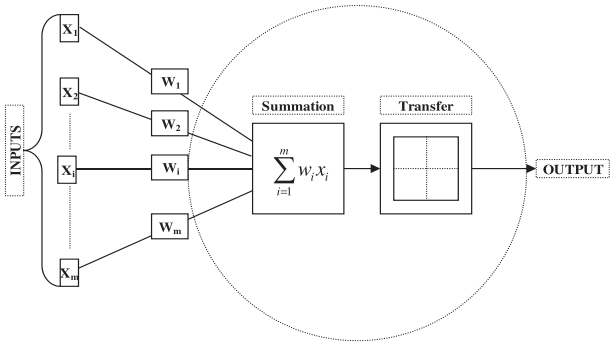

Fig. 3. The representation of the information flow in a neuron.

Fig. 3 shows an example of the information flow for a single neuron in the network. Here, each input coming from previous layer is multiplied by a weight (W) and summed using summation function. Then, this single value is passed through a transfer function to produce the output value of a neuron [30].

In order to assess model predictions, the output values ($Y_1, \ldots Y_k, \ldots Y_p$) are compared with the target values ($Z_1, \ldots Z_k, \ldots Z_p$) such as experimental results. The differences between model predictions and target values are evaluated against the modeling performance criteria that are established within the ANN algorithm. Reprocessing of the output values is required if the modeling performance criteria are not met [30].

2.2. Modeling performance criteria

The root mean square error (RMSE), the coefficient of determination-R^2 value, statistical T value, and the mean absolute error (MAE) are the common performance criteria for the evaluation of an ANN model performance.

The RMSE represents the error between model predictions and target values. It can be computed with Eq. (4) with a range from 0 to ∞. Lower values of RMSE are preferable as there is no absolute criterion for a "good" value [30]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(y_{\mathrm{model},i} - y_{\mathrm{obs},i}\right)^2}{n}} \tag{4}$$

where n is the number of target values, $y_{\mathrm{model},i}$ and $y_{\mathrm{obs},i}$ are model predictions and their corresponding target values, respectively.

The coefficient of determination (R^2) is estimated through Eq. (5). This value shows the percentage of variability between experimental data and model predictions. R^2 values range between 0 and 1 (i.e. 0–100%). An R^2 value $\approx$1 means the greater correlation and the stronger relationship between predictions and actual data [30]:

$$R^2 = \left[\frac{n\sum_{\mathrm{i}=1}^{n} y_{\mathrm{obs},i} y_{\mathrm{model},i} - \left(\sum_{\mathrm{i}=1}^{n} y_{\mathrm{obs},i}\right)\left(\sum_{\mathrm{i}=1}^{n} y_{\mathrm{model},i}\right)}{\sqrt{\left[n\sum_{\mathrm{i}=1}^{n} y_{\mathrm{obs},i}^2 - \left(\sum_{\mathrm{i}=1}^{n} y_{\mathrm{obs},i}\right)^2\right] \times \left[n\sum_{\mathrm{i}=1}^{n} y_{\mathrm{model},i}^2 - \left(\sum_{\mathrm{i}=1}^{n} y_{\mathrm{model},i}\right)^2\right]}}\right]^2 \tag{5}$$

The RMSE and R^2 values provide information on general error ranges between model predictions and target values.

In parallel, T value measures the scattering around the line (1:1) that can be obtained by a plot of target values vs. model predictions. Computed with Eq. (6), T values close to 1 indicate that a good fitting is achieved between model predictions and experimental data [31]:

$$T = 1 - \frac{\sum_{i=1}^{n} \left(y_{\mathrm{model},i} - y_{\mathrm{obs},i}\right)^2}{\sum_{i=1}^{n} \left(y_{\mathrm{model},i} - \bar{y}\right)^2} \tag{6}$$

where, $\bar{y}$ represents the arithmetic average of n samples.

The above common criteria are used to assess performance of our ANN model. In addition, we also used MAE to estimate the distribution of errors between model predictions and target values. The MAE can be computed with Eq. (7) and its values can range from 0 to ∞. Similar to RMSE, lower values of

MAE indicate a good correlation between model predictions and experimental data [30]:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_{\mathrm{obs},i} - y_{\mathrm{model},i}| \quad (7)$$

The above criteria are used commonly for validation of models and their predictions. However, we must note that the quality of experimental data is the fundamental requirement for any type of modeling work. Otherwise, the results of statistical tests and model predictions will be inaccurate.

2.3. Experimental procedure and data

The experimental part of this work was conducted by Koyuncu [26]. Details of the experimental system and data collection methods are presented in previously published studies [2,18,28,32]. Briefly, the laboratory-scale experiments were conducted to filter typical textile wastewater with NF membranes under different operational conditions. The experimental variables were concentrations of salt (C_{sb}) and dye (C_{db}) in feed water, cross-flow velocity (V), feed pressure (ΔP), and pH of feed water. These parameters were varied in the following ranges: 1–80 mg/L for C_{sb}, 0.1–50 mg/L for C_{db}, 0.11–1.11 m/s for V, 8–24 bars for ΔP, and 4–10 for pH. The rejection rates (R_o) of single salt (NaCl) were measured as a result of the changes in input parameters and operational conditions.

Because the quality of experimental results is important, we conducted a preliminary assessment of experimental data through statistical parameters including minimum (x_{min}), maximum (x_{max}), and mean (x_{mean}) values, standard deviation (σ) of experimental data-sets, the coefficient of variation (C_v), and the coefficient of skewness (C_{sx}). Among them, the standard deviation, σ, shows the deviation of data from their mean value. The coefficient of variation, C_v, the ratio between standard deviation and mean value, represents dispersion of data points around their mean value. This parameter is used generally to compare the degree of variation from one data series to another. The coefficient of skewness, C_{sx}, indicates the degree of symmetry in variable distribution. It shows the degree that data are skewed or asymmetric with respect to their mean value. The value of C_{sx} will equal to zero for data of normal distribution, and it will have positive or negative values if the data distribution is asymmetric.

3. Results and discussion

We present our results in the following three sections: Section 3.1 presents the preliminary assessment of experimental data-sets, Section 3.2 presents the development of ANN model, and Section 3.3 presents the application of the model to the experimental system. Section 3.1 presents the results of our statistical assessment of experimental data-sets. Section 3.2 presents construction of our ANN model in several steps including, selection of input variables through sensitivity analyses, selection of the best training algorithm, and determination of optimal number of neurons in the hidden layer. The training and validation of the ANN model are also presented here. Section 3.3 presents our model predictions for the experimental work and detailed discussion of these results.

3.1. Preliminary assessment of experimental data-sets

Table 1 shows the experimental variables and their ranges used in experiments. The range values represent typical concentrations found in textile industry wastewater and typical operating conditions of NF process. According to minimum, maximum, and mean values in Table 1, salt concentration (C_{sb}) of 1–80 mg/L, dye concentrations of 0.1–50 mg/L, feed pressure (ΔP) of 8–24 bar, flow velocity of 0.11–1.11 m/sn, and pH of 4–10 were used as the input parameters of experimental system.

The minimum, maximum, and mean values of salt (C_{sb}) and dye concentrations (C_{db}) in Table 1 present a broad range, because they are the independent

Table 1
The statistical parameters of each data-set

Parameters		x_{min}	x_{max}	x_{mean}	σ	$C_v(\sigma/x_{mean})$	C_{sx}	Correlation with R_o
Input	C_{sb} (mg/l)	1.000	80.00	28.50	30.74	1.080	0.890	−0.858
	C_{db} (mg/l)	0.100	50.00	7.150	13.17	1.840	2.610	−0.303
	ΔP (bar)	8.000	24.00	15.65	5.750	0.367	0.080	0.208
	V (m/s)	0.110	1.110	0.679	0.260	0.382	−0.740	0.015
	pH	4.000	10.00	6.985	0.760	0.109	−0.257	−0.028
Output	R_o	0.033	0.837	0.431	0.251	0.582	−0.003	1.000

Notes: C_{sb}: Salt concentration; C_{db}: Dye concentration; ΔP: Feed pressure; V: Cross-flow velocity; and R_o: Rejection rate.

variables of the experimental system. Although operational pressure, flow velocity, and pH also are independent variables, their typical value ranges are well-known for NF processes; and therefore, the variations in these parameters are within relatively narrow ranges. This difference is indicated by standard deviations (σ) of salt (C_{sb}) and dye (C_{db}) concentrations. The σ values are 30.74 and 13.17 for C_{sb} and C_{db}, respectively, which are greater than the values for other three experimental parameters. Similarly, the coefficient of variation (C_v) is 1.84 for dye concentrations, which shows that this is the most varied parameter in these experiments.

The coefficient of correlation values in Table 1 indicates how each experimental parameter affects rejection rate (R_o) by NF. While the positive values indicate a direct relationship between an experimental parameter and R_o, the negative values indicate an inverse relationship between R_o and the corresponding experimental parameter. For instance, the positive coefficient of correlation between ΔP, V, and R_o indicates that the higher the feed pressure and cross-flow velocity is the higher the rejection rate will be. Subsequently, negative correlation values between C_{sb}, pH, C_{db}, and R_o indicate that lower values of these parameters will result in higher rejection rates in this experimental setup. The relatively large correlation values between C_{sb} and R_o (−0.858) and between C_{db} and R_o (−0.303) point out that these parameters affect rejection rate much more than the other parameters.

3.2. Development of ANN model

3.2.1. Preprocessing of experimental data

In the experimental part of this work, Koyuncu et al. [26,28] varied feed water concentrations of salt (C_{sb}) and dye (C_{db}), cross-flow velocity (V), feed pressure (ΔP), and pH of feed water. The critical focus of their experiments was salt recovery as it was important for textile industry; and therefore, they measured salt concentrations in feedwater and effluent of NF system. By using these values, they estimated salt rejection rates as percentage for NF membranes operated under various experimental conditions. A total of 218 experimental data-sets were available from their experimental work for this study. The neural networks require that all training data must be preprocessed before training. Therefore, the raw data sets must be normalized before they are used for training of a network. Normalization is a widely used preprocessing method that scales the data-sets into an acceptable range. Thus, experimental data were normalized into unitless numbers within the range of 0–1 using the following equation:

$$x_i = \frac{(x - x_{min})}{(x_{max} - x_{min})} \tag{8}$$

where x_i is the normalized value of a certain parameter, x is the measured value for this parameter, and x_{min} and x_{max} are the minimum and maximum values in the data-set for this parameter, respectively.

The normalized experimental data are grouped into three segments as the training set (142 data), validation set (55 data), and test set (21 data). Our network is trained with the training data, while the performance of the network is tested with validation data-set. And finally, the test data-set is used for the performance assessment of our ANN model predictions.

3.2.2. Selection of input parameters for the ANN

The basic structure of an ANN model starts with development of input layer with experimental variables that have any degree of effect on the outcome (Fig. 3). The relative impact of each input parameter can be identified with a single-input ANN model that has one hidden layer with 25 neurons and a sensitivity analysis, which require sufficient numbers of experimental data for these tests. Consequently, the ineffective variables of the process can be excluded from the input layer which leads to a more compact neural network.

In this work, the rejection rate is a function of the input parameters—concentrations of salt and dye in feed water, cross-flow velocity, pH, and feed pressure—which can be represented simply as $R_o = f(C_{sb}, C_{db}, V, \text{pH}, \text{and } \Delta P)$. In order to identify how each input variable affects the process, we build separate ANN models with a single input parameter. Fig. 4 presents the model predictions of these individual models. The top two panels on the left side of Fig. 4 show model predictions of rejection rate by using only salt (C_{sb}) and dye (C_{db}) concentrations. The modeling results for these cases match the experimental data (R_o) relatively well. Below these panels in the same figure are the model predictions of rejection rate with cross-flow velocity, pH, and feed pressure. These predictions present significant discrepancy with the experimental results indicating that these parameters are less effective than salt and dye concentrations.

In addition to ANN modeling, the sensitivity analysis was carried out by evaluating the differences in the network output. In this analysis, one variable at

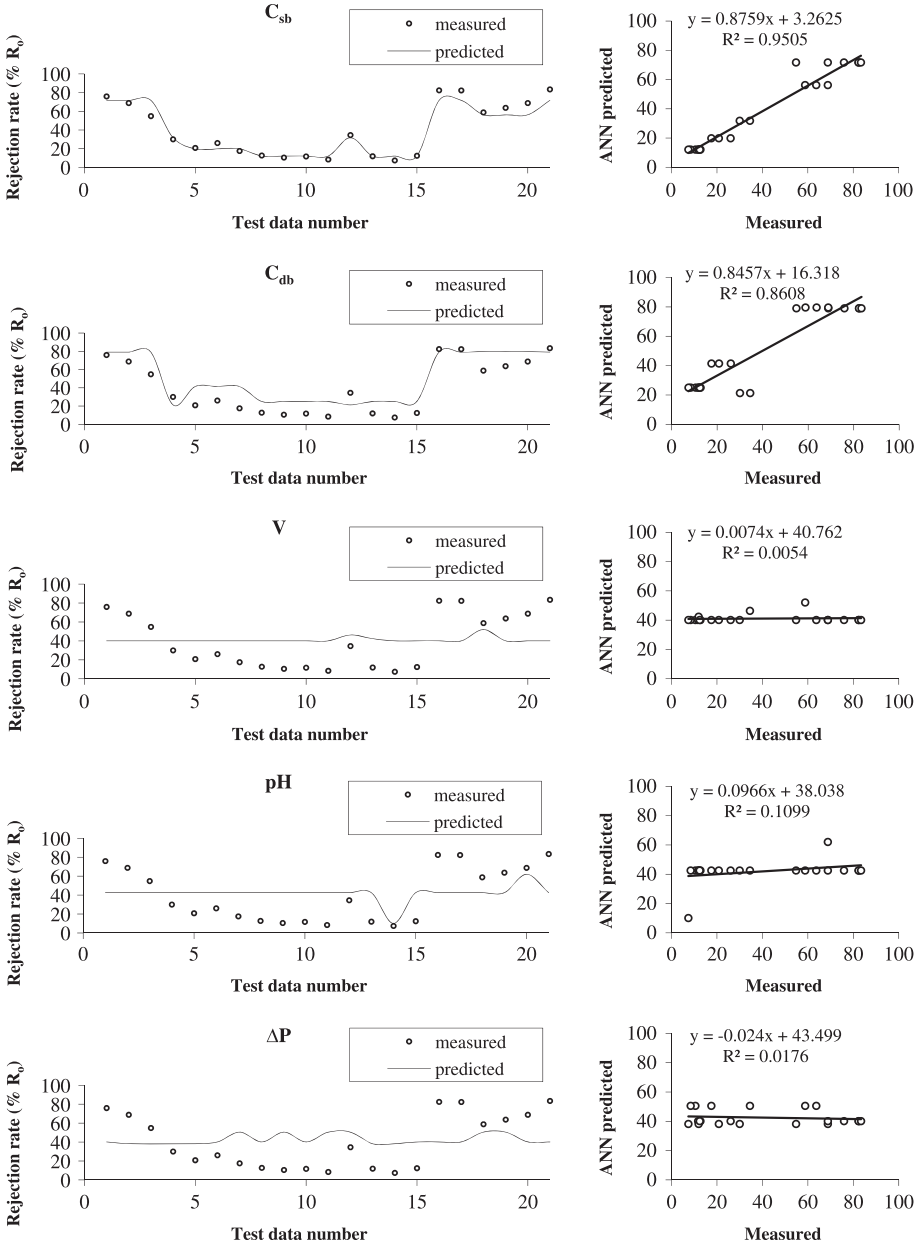

Fig. 4. Comparison of ANN results and observed rejection rate (% R_o) depending on each input parameter.

the time is set as input parameter to determine its impact on modeling results. Thus, it indicates which parameter has key importance in the model and in the experimental system. The sensitivity analyses are presented on the right side of Fig. 4 with plots of the correlation between model predictions and experimental data. The coefficients of determination (R^2) between model predictions and experimental data are shown in each panel in Fig. 4. These values indicate that the most effective input parameter is the salt concentration (C_{sb}) with R^2 value of 0.95, while the others in descending order of effectiveness are C_{db} ($R^2 = 0.86$), pH ($R^2 = 0.11$), ΔP ($R^2 = 0.02$), and V ($R^2 = 0.01$).

Because the salt concentration was the most effective parameter, we identified relative impacts of other input parameters while having salt concentration as the common parameter in all possible combinations. We started with salt concentration and added a parameter to form new combinations of input parameters as shown in the top row of Table 2. In all of these analyses, we compared our model predictions with the experimental data and we estimated statistical

Table 2
Performance evaluation of the effective parameters for sensitivity analysis

Performance	C_{sb}	$C_{sb} + C_{db}$	$C_{sb} + C_{db}$ + pH	$C_{sb} + C_{db}$ + pH + ΔP	$C_{sb} + C_{db}$ + pH + $\Delta P + V$
MAE	0.041	0.035	0.033	0.031	0.026*
RMSE	0.064	0.052	0.050	0.041	0.035*
R^2	0.950	0.961	0.963	0.972	0.990*

*Indicates best results.

Table 3
Comparison of different training algorithms with 15 neurons in the hidden layer

Training algorithms	Function	IN	RMSE	R^2
BFGS quasi-Newton backpropagation	Trainbfg	54	0.058	0.983
Powell–Beale conjugate gradient backpropagation	Traincgb	82	0.068	0.976
Fletcher–Reeves conjugate gradient backpropagation	Traincgf	105	0.080	0.967
Polak–Ribiére conjugate gradient backpropagation	Traincgp	50	0.055	0.984
Gradient descent backpropagation	Traingd	159	0.362	0.294
Gradient descent with momentum backpropagation	Traingdm	85	0.148	0.884
Gradient descent with adaptive learning rate backpropagation	Traingda	93	0.122	0.922
Gradient descent with momentum and adaptive learning rate backp	Traingdx	139	0.320	0.537
LM backpropagation	Trainlm	25	0.031	0.995
Resilient backpropagation	Trainrp	54	0.032	0.994
Scaled conjugate gradient backpropagation	Trainscg	36	0.042	0.991
One-step secant backpropagation	Trainoss	43	0.055	0.984

Note: IN: iteration number; RMSE: root mean square error; and R^2: coefficient of determination.

parameter values for each combination case. As we increased the number of input variables in the ANN model, the model predictions and the statistical parameters improved significantly. Table 2 indicates that the ANN model with five inputs (C_{sb}, C_{db}, ΔP, V, and pH) gives the best results, with R^2 value of 0.99. Therefore, all of the experimental parameters are used as the input parameters in our ANN model.

3.2.3. Selection of the best training algorithm

As input parameters in our ANN model is established, several backpropagation training algorithms were tested to determine the best training algorithm for our model. The choice of training algorithm and transfer function is of crucial importance for robust performance of the ANN model. In general, linear functions are frequently used for input and output layers, while nonlinear transfer functions are preferred for hidden layers [33]. Therefore, we maintained the nonlinear tangent sigmoid function (*tansig*) as the transfer function in hidden layer with 15 neurons and a linear transfer function (*purelin*) for the output layer of our ANN model. Table 3 presents statistical evaluation of the performance of ANN models for each selected training algorithm. The Levenberg–Marquardt (LM) algorithm is a standard technique for nonlinear least square problems. LM algorithm is the most widely used optimization algorithm for a variety of ANN problems. It is mostly preferred due to its speed and stability in training of ANN model. Table 3 shows that the LM backpropagation algorithm (*trainlm*) resulted in the smallest RMSE, 0.031, and the highest R^2 value, 0.995. Therefore, this algorithm was selected as the best training algorithm. *Trainrp* and *trainscg* functions followed this *trainlm* function with a RMSE of 0.032 and 0.042, respectively. The other functions such as *traincgp*, *trainbfg*, and *traingda* resulted in higher levels of RMSE than the *trainlm* function.

3.2.4. Determination of the number of neurons in hidden layer

The optimal ANN architecture can be identified by selecting the optimum number of neurons in the hidden layer. To do this, two neurons are used in hidden layer as the initial guess and then, the number of neurons is increased two at a time from 2 to 30. *Trainlm*

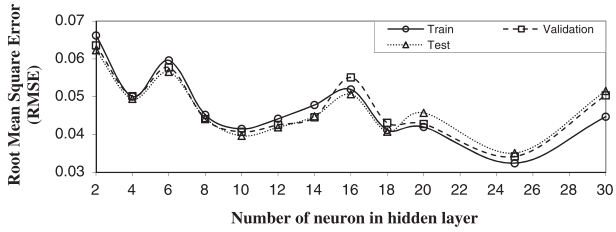

Fig. 5. Effect of the number of the neurons in the hidden layer on the performance of the network.

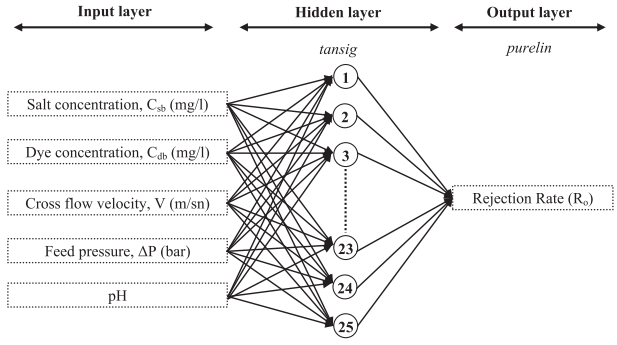

Fig. 6. Optimal ANN architecture developed in this study.

function was used as the best training algorithm for all trials. The minimum value of RMSE between the measured and the predicted values of R_o is used as the criteria for selecting number of neurons. Fig. 5 shows the RMSE values vs. the number of neurons in the hidden layer. Fig. 5 indicates that three local minimum RMSE values are observed at neuron numbers of 4, 10, 18, and 25. However, the neural network architecture with 25 hidden neurons reached the minimum RMSE in training, validation, and test of the ANN model. Thus, 25 neurons were chosen as the optimum neuron number for the hidden layer.

Fig. 6 shows the overall configuration of our ANN model that is consisted of five neurons in the input layer, 25 neurons in the hidden layer operating with tangent sigmoid (*tansig*) transfer function, and one neuron in the output layer operating with linear transfer function (*purelin*).

3.3. ANN modeling results of NF process

Fig. 7 shows the results of this modeling exercise which illustrates how model predictions compare with experimental data during training, validation, and testing phases of the modeling work. The panels on

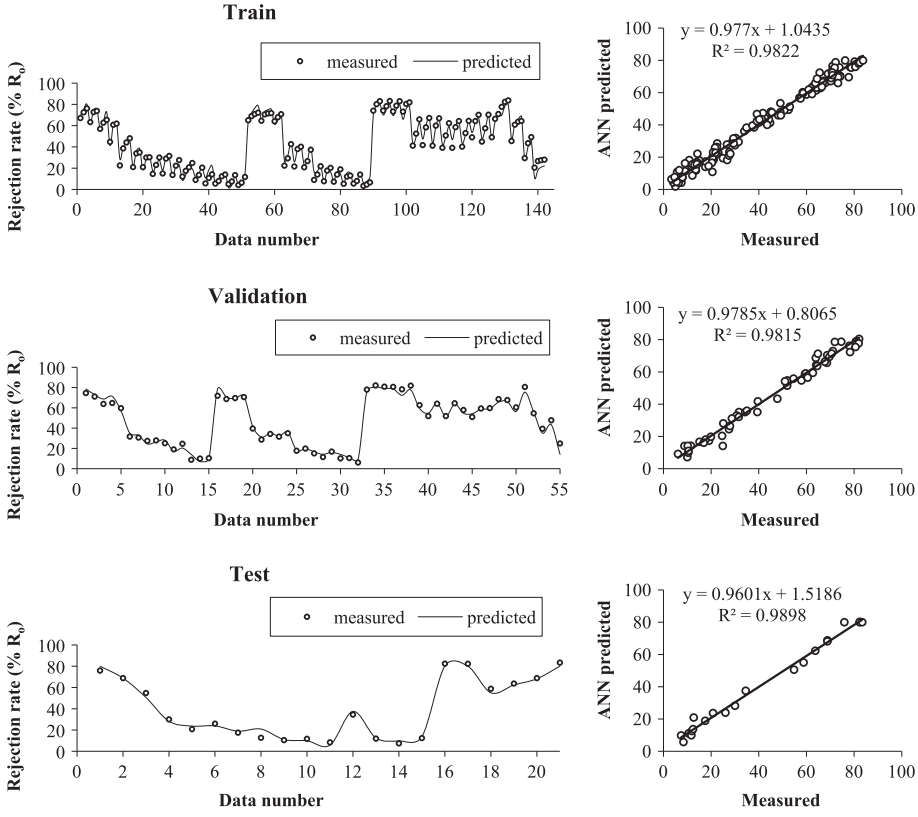

Fig. 7. The ANN predicted vs. measured values for train, validation, and test sets.

Table 4
The structure and performance of the final selected ANN model with optimum values of model parameters

Optimum structure	Transfer function	Mean value						
		Training		Validation		Testing		
5-25-1	*Tansig* for hidden layer and *purelin* for output layer	RMSE	MAE	RMSE	MAE	RMSE	MAE	T
		0.032	0.024	0.034	0.025	0.035	0.026	0.993

Notes: RMSE: root mean square error; MAE: mean absolute error; and T: scattering value.

the left side show that the model predictions are in close agreement with the experimental data in all cases. The panels on the right side show the correlation between model predictions and measured values. The coefficients of determination ($R^2 > 0.98$) in these panels indicate strong correlations for all three stages. These results verify that our ANN model structure is correct and it can predict the salt rejection rate of NF process under various conditions.

To evaluate the performance of our model, the RMSE, MAE, and T values are estimated between model predictions and experimental values. Table 4 presents the results of this statistical analysis. The RMSE values were 0.032 for training, 0.034 for validation, and 0.035 for testing phases of the model. The MAE values for the same stages were 0.024, 0.025, and 0.026, respectively. Similarly, T value of 0.993 reflects high quality of fitting between our model predictions and the experimental results.

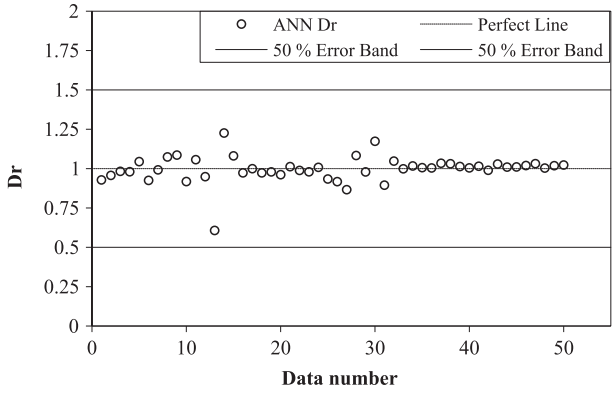

Fig. 8. Distribution of the ANN model discrepancy ratios.

Table 5
Accuracy of model for R_o prediction

Presented model	Discrepancy ratio		
	Number of test data	Mean	Standard Deviation
ANN model	55	1.014	0.126

As part of the performance evaluation, discrepancy ratios between model predictions and experimental results are estimated by using the equation $D_r = R_{o_{estimated}}/R_{o_{measured}}$, where $R_{o_{estimated}}$ is the estimated or predicted value of rejection rate and $R_{o_{measured}}$ is the measured one. As the predicted and experimental values get closer to each other, then this ratio approaches to 1. Fig. 8 shows that the discrepancy ratios were mostly around one in our work, with a few exceptions of 25% discrepancy and with one datum around 40% of its target value.

The mean value of discrepancy ratios and the standard deviation of these values are presented in Table 5. Here, the mean value of discrepancy ratio is estimated with $\bar{D}_r = \Sigma D_{ri}/N$, while its standard deviation is estimated with the equation $\sigma = (\Sigma(D_{ri} - \bar{D}_r)^2/N-1)^{1/2}$. These results show that the mean value of discrepancy ratios is around 1 and the standard deviation is 12%.

4. Conclusions

This modeling exercise is the first application of ANNs to predict salt rejection rate (R_o) from actual wastewaters by NF process. Our modeling work identified several key findings about this process as follows:

- The input concentrations of salt and dye affect directly the removal of salt by NF.
- Our sensitivity analysis showed that input parameters such as feed pressure, cross-flow velocity, and pH have relatively minimal effect on performance of NF.
- Although salt concentration has the highest effect on model predictions, the model with five input variables (C_{sb}, C_{db}, pH, ΔP, and V) has the best prediction performance among all combinations of input parameters. Therefore, we developed the input layer of our ANN model with five neurons corresponding to five input parameters.

- We note that experimental data must be normalized to unitless values in order to reduce and eliminate data redundancy.
- The best algorithm for training of ANN has been identified as the LM backpropagation algorithm (*trainlm*) in this work.
- We estimated that 25 neurons in the hidden layer served as an optimal number of neurons for this modeling exercise.
- The transfer functions for neurons were nonlinear tangent sigmoid (*tansig*) for the hidden layer and linear transfer function (*pureline*) for the output layer of our ANN model.

In addition to these key highlights, ANN modeling of NF processes can provide valuable information on membrane design, operation, maintenance, and process optimization. A good predictive model will provide information on membrane characteristics, process performance as well as optimization of the process. In parallel, ANN modeling of NF processes will result in lower number of experiments and subsequently saving time and money during the development of a NF process. In practice, the ANN approach has the advantage of only requiring simple and readily available inputs and a minimum understanding of the complex phenomena controlling NF process.

Our findings support that ANN models can serve as useful tools to identify effective parameters for NF as well as to predict the rejection rate of a key substance by this system. Our results show that the outcome of NF process can be predicted with the ANN model without in-depth knowledge of the modeled system. However, ANN models require sufficient number of data for training, validation, and testing stages of the modeling work. In this study, we have just focused on prediction of rejection rate in NF membranes. Future studies should focus on how properties of membranes and pollutants can be evaluated together to optimize removal performance. In this process, ANN models can serve to obtain rapid results for an evaluation of membrane performance for the removal of a specific substance.

Acknowledgments

The first author gratefully acknowledges a graduate studies scholarship (National Scholarship Programme for MSc and PhD Students) from The Scientific and Technological Research Council of Turkey (TUBITAK). We thank Dr. Fatih Karadagli for his valuable suggestions and his remarkable support to improve our manuscript.

References

[1] R. Singh, Production of high-purity water by membrane processes, Desal. Water Treat. 3 (2009) 99–110.
[2] I. Koyuncu, D. Topacik, Effect of organic ion on the separation of salts by nanofiltration membranes, J. Membrane Sci. 195 (2002) 247–263.
[3] P. Aptel, Membrane pressure driven processes in water and wastewater treatment, in Proceedings of the NATO advanced study institute on membrane processes in separation and purification, Curia, Portugal, 1983, pp. 263–281.
[4] M. Mulder, The use of membrane processes in environmental problems, in Proceedings of the NATO advanced study institute on membrane processes in separation and purification, Curia, Portugal, 1983, pp. 229–262.
[5] 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.
[6] C.R. Wu, S.H. Zhang, D.L. Yang, X.G. Jian, Preparation, characterization and application of a novel thermal stable composite nanofiltration membrane, J. Membrane Sci. 326 (2009) 429–434.
[7] I. Koyuncu, D. Topacik, Effects of operating conditions on the salt rejection of nanofiltration membranes in reactive dye/salt mixtures, Sep. Purif. Technol. 33 (2003) 283–294.
[8] A.F. Ismail, W.J. Lau, Influence of feed conditions on the rejection of salt and dye in aqueous solution by different characteristics of hollow fiber nanofiltration membranes, Desal. Water Treat. 6 (2009) 281–288.
[9] M. Turek, M. Chorazewska, Nanofiltration process for seawater desalination-salt production integrated system, Desal. Water Treat. 7 (2009) 178–181.
[10] W.R. Bowen, M.G. Jones, J.S. Welfoot, H.N.S. Yousef, Predicting salt rejections at nanofiltration membranes using artificial neural networks, Desalination 129 (2000) 147–162.
[11] K.S. Spiegler, O. Kedem, Thermodynamics of hyperfiltration (reverse osmosis): Criteria for efficient membranes, Desalination, 1 (1966) 311–326.
[12] M. Perry, C. Linder, Intermediate reverse osmosis ultrafiltration (RO UF) membranes for concentration and desalting of low molecular weight organic solutes, Desalination 71 (1989) 233–245.
[13] R. Levenstein, D. Hasson, R. Semiat, Utilization of the Donnan effect for improving electrolyte separation with nanofiltration membranes, J. Membrane Sci. 116 (1996) 77–92.
[14] X.T. Xu, H.G. Spencer, Dye-salt separations by nanofiltration using weak acid polyelectrolyte membranes, Desalination 114 (1997) 129–137.
[15] G. Vakili-Nezhaad, Z. Akbari, Modification of the extended Spiegler-Kedem model for simulation of multiple solute systems in nanofiltration process, Desal. Water Treat. 27 (2011) 189–196.
[16] Z.V.R. Murthy, L.B. Chaudhari, Separation of binary heavy metals from aqueous solutions by nanofiltration and characterization of the membrane using Spiegler-Kedem model, Chem. Eng. J. 150 (2009) 181–187.
[17] 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.
[18] I. Koyuncu, Influence of dyes, salts and auxiliary chemicals on the nanofiltration of reactive dye baths: Experimental observations and model verification, Desalination 154 (2003) 79–88.
[19] T.M. Hwang, Y. Choi, S.H. Nam, S. Lee, H. Oh, K. Hyun, Y. K. Choung, Prediction of membrane fouling rate by neural network modeling, Desal. Water Treat. 15 (2010) 134–140.
[20] E. Piron, E. Latrille, F. Rene, Application of artificial neural networks for crossflow microfiltration modelling: Black-box and semi-physical approaches, Comput. Chem. Eng. 21 (1997) 1021–1030.

[21] C. Aydiner, I. Demir, E. Yildiz, Modeling of flux decline in crossflow microfiltration using neural networks: The case of phosphate removal, J. Membrane Sci. 248 (2005) 53–62.
[22] W.R. 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.
[23] N. Delgrange, C. Cabassud, M. Cabassud, L. Durand-Bourlier, J.M. Laine, Modelling of ultrafiltration fouling by neural network, Desalination 118 (1998) 213–227.
[24] D. Libotean, J. Giralt, F. Giralt, R. Rallo, T. Wolfe, Y. Cohen, Neural network approach for modeling the performance of reverse osmosis membrane desalting, J. Membrane Sci. 326 (2009) 408–419.
[25] Q.F. Liu, S.H. Kim, S. Lee, Prediction of microfiltration membrane fouling using artificial neural network models, Sep. Purif. Technol. 70 (2009) 96–102.
[26] I. Koyuncu, Effect of organic ion on the separation of salts by nanofiltration, PhD dissertation, Istanbul Technical University, 2001.
[27] D.S. Kostick, Salt, U.S.G.S. 2008 Minerals Yearbook (Advance Release), 2010, 63.1–63.5.
[28] I. Koyuncu, D. Topacik, E. Yuksel, Reuse of reactive dyehouse wastewater by nanofiltration: Process water quality and economical implications, Sep. Purif. Technol. 36 (2004) 77–85.
[29] G. Singh, J. Kandasamy, H.K. Shon, J. Cho, Measuring treatment effectiveness of urban wetland using hybrid water quality—artificial neural network (ANN) model, Desalin. Water Treat. 32 (2011) 284–290.
[30] E. Dogan, A. Ates, E.C. Yilmaz, B. Eren, Application of artificial neural networks to estimate wastewater treatment plant inlet biochemical oxygen demand, Environ. Prog. 27 (2008) 439–446.
[31] N.S. Kaveh, S.N. Ashrafizadeh, F. Mohammadi, Development of an artificial neural network model for prediction of cell voltage and current efficiency in a chlor-alkali membrane cell, Chem. Eng. Res. Des. 86 (2008) 461–472.
[32] I. Koyuncu, D. Topacik, Effect of cross flow velocity, feed concentration, and pressure on the salt rejection of nanofiltration membranes in reactive dye having two sodium salts and NaCl mixtures: Model application, J. Environ. Sci. Health A 39 (2004) 1055–1068.
[33] W. Duch, N. Jankowski, Survey of neural transfer functions, Neural Comput. Surv. 2 (1999) 163–212.