Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2019.23531

143 (2019) 48–58
March

A comparative study of artificial neural network models for the prediction of Cd removal efficiency of polymer inclusion membranes

Beytullah Eren[a,*], Muhammad Yaqub[b], Volkan Eyupoglu[c]

[a]*Department of Environmental Engineering, Sakarya University, Esentepe, Sakarya, Turkey, email: beren@sakarya.edu.tr*
[b]*Kumoh National Institute of Technology, Gumi, South Korea, email: yaqub92@kumoh.ac.kr*
[c]*Science Faculty, Chemistry Department, Cankiri Karatekin University, Cankiri, Turkey, email: volkan@karatekin.edu.tr*

Received 21 May 2018; Accepted 26 November 2018

ABSTRACT

In this study, three different artificial neural network (ANN) including feed forward back-propagation (FFBPNN), recurrent neural network (RNN), and generalized regression neural network (GRNN) were proposed to estimate Cd removal efficiency through polymer inclusion membranes (PIMs). A multiple linear regression (MLR) statistical technique was also applied to evaluate PIMs efficiency. The proposed ANN models and MLR results were compared regarding statistical performance criteria such as root-mean-squared error, mean absolute error and coefficient of determination (R^2). In the modeling, time, film thickness, extractant type and amount, plasticizer type and amount and polymer molecular weight were considered as inputs while Cd removal efficiency was output. Furthermore, sensitivity analysis is performed to investigate the effect of each input parameter on the output regarding magnitude. According to performance criteria of models, FFBPNN and RNN have the best prediction capability as compared with GRNN and MLR. Sensitivity analysis results demonstrated that extractant amount, plasticizer type and plasticizer amount are more influential operating parameters than time, extractant type, film thickness, and polymer molecular weight. The results of FFBPNN and RNN models are superior and reliable in the prediction of PIMs Cd removal efficiency due to the nonlinearity of data set.

Keywords: Feed forward back-propagation; Generalized regression neural network; Polymer inclusion membranes; Removal efficiency; Recurrent neural network

1. Introduction

Heavy metal pollution in water streams is mainly caused by electroplating, pigment, alloy, fertilizer and chemical industries. This is an extensively studied area of investigation [1], because heavy metals can pose serious health risks to the ecosystems that they are discharged [2]. Heavy metals are known to be accumulative within biological systems even at concentrations within regulated limits due to their potential for long-term accumulation in food chain [3–5], cadmium, zinc, copper, nickel, mercury, and chromium are often detected in industrial wastewaters, which are generated by metal plating, mining activities, smelting, battery manufacture, pesticides, and nuclear energy industries [1]. Cadmium (Cd) is one of the most toxic heavy metal and considered non-essential for living organisms [6]. The disposal of wastewater containing cadmium to the inland water has increased cadmium concentration in the food and subsequently in human bodies [1] posing a serious threat to human health [7]. The harmful effects of Cd include some acute and chronic disorders, such as "itai-itai" disease, renal damage, emphysema, hypertension, and testicular atrophy [8]. Therefore, cadmium has been added to the list of acknowledged endocrine disrupting chemicals [9].

The conventional physical and chemical treatment methods for removal of heavy metals include adsorption [10,11], bacterial biomass [1,12], electrocoagulation, electrodialysis, flotation [13], [14,15], membrane technologies such as

* Corresponding author.

nanofiltration and reverse osmosis [16–18]. However, it is uneconomical to treat low concentration (less than 5 mg/L) of Cd using chemical processes. Although reverse osmosis technique can remove Cd to regulated standards, this method has high operational and maintenance costs [8,19]. Environmental engineers are challenged to develop more efficient techniques to combat this issue. Recently, supported liquid membrane (SLM) applications have been emerging as an alternative to the conventional methods due to its high selectivity, operational simplicity, low solvent inventory, low energy consumption, zero effluent discharge, a combination of extraction and stripping in a single unit [20,21]. In this direction, liquid membranes such as polymer inclusion membranes (PIMs) have effective carrier immobilization, easy preparation, versatility, stability, excellent chemical resistance, and better mechanical properties as compared with SLM [22]. Past studies showed that PIMs can be used for metal ion extraction, separation of inorganic species, biochemical and biomedical applications [23,24].

The application of PIMs to remove Cd from aqueous solutions is a novel technique in wastewater treatment. There are various operating parameters that affect the efficiency of PIMs. To improve the performance of PIMs, estimation, optimization, and analysis of the operating parameters should be accomplished by modeling and simulation, as well as, through laboratory experiments. Modeling is a valuable approach to develop a relation between parameters and efficiency to optimize and control the process for efficient design and operation. Due to outstanding characteristics in processing the non-linear relationships among variables in complex systems with reliable and robust results, artificial neural networks (ANNs) has been successfully employed in environmental engineering [25–27]. In the literature, the ANN model has been applied for heavy metals removal through physico-chemical processes [28] from drinking water and from wastewater treatment systems [29,30]. In other studies such as biosorption efficiency of *Zea mays* for the removal of chromium from wastewater [31] and copper removal from aqueous solution by the ion-exchange process [32] was estimated by applying ANN modeling. ANN model was proposed to investigate copper removal from wastewater by adsorption on fungal biomass [33] and cadmium sorption by shelled *Moringa oleifera* seed powder from aqueous solution [34].

Moreover, ANN genetic algorithm and particle swarm optimization modeling was used for the prediction of copper removal from aqueous solutions by reduced graphene oxide-supported nanoscale zero-valent iron (nZVI/rGO) magnetic nanocomposites [35]. Same material (nZVI/rGO) was studied for optimization of rhodamine B removal from aqueous solution by using ANN-genetic algorithm [36]. Furthermore, cadmium removal from aqueous solutions using nZVI/rGO composite was studied by applying ANN modeling and genetic algorithm optimization [37]. However, the efficiency of PIMs to remove Cd has not been studied previously by applying neural network models. In addition, the removal efficiency of PIMs cannot be determined by conventional mathematical techniques alone due to the complexity of data. Therefore, ANN models were considered for the prediction of PIMs Cd removal efficiency from aqueous solutions in this study.

The objective of this study is to propose the best neural network model to predict the Cd removal efficiency from aqueous solution by using PIMs, and to optimize the process by considering the effect of various operating parameters on removal efficiency. The proposed models are trained against experimental data and then predicted results are compared with each other as well as with a statistical multiple linear regression (MLR) technique results to identify the best model. Also, sensitivity analysis has been applied to check the acceptable conformity and individual effectiveness of each operating parameter.

2. Materials and methodology

2.1. Experimental study and data preparation

In ANN model development, accurate experimental data are needed for network training, validation and testing [38]. In the present study, data were obtained from laboratory experiments regarding the selective extraction and stripping of Cd from acidic iodide solutions containing Cd by using liquid–liquid extraction technique [39]. The parameters including time, film thickness, extractant type and amount, plasticizer type and amount, and polymer molecular weight were considered as the most effective ones for the main transport of Cd ions. In this respect, key details of experimental procedures are stated briefly as follows:

The experimental design was organized to illuminate PIMs-based Cd separation from acidic aqueous solutions containing iodide as a complexing agent. The data were provided by a set of experimental study regarding the selective transport of Cd through PVDF-co-HFP-based PIMs containing symmetric imidazolium bromide salts as a carrier. For this purpose, butyl, hexyl, octyl, and decyl substituted ionic liquids were synthesized and used in the production of PVDF-co-HFP based PIMs as ion carrier. The polymer inclusion membrane was prepared by using previous methods available in the literature [40,41]. The experimental procedures are explained in as follows.

The membrane glue was made by dissolution of PVDF-co-HFP, purchased from Sigma-Aldrich (Sleaze, Germany) in different molecular weights, plasticizers "2-nitrophenyl octyl ether (ONPOE), 2-nitrophenyl pentyl ether (2-NPPE), bis (2-ethylhexyl) adipate (B2EHA), tris (2-ethylhexyl) phosphate (TEHP)" are also purchased from Sigma-Aldrich (Sleaze, Germany), and the synthesized room temperature ionic liquids in the literature in our previous paper [42] in 20 mL of acetone. The polymer solution was vigorously agitated by a magnetic stirrer at ambient temperature (25°C ± 0.5°C) for an hour, and after that, the glue was sonicated for 30 min to obtain a homogeneous polymer solution. The PVDF-co-HFP membrane solution was poured into a glass petri dish having 15-cm internal diameter and placed on the well-adjusted flat table by water gauge to mold membrane film as more smooth and uniform thickness. After polymer solution was evaporated overnight, the remaining transparent membrane film was peeled from the glass surface in cold water. The membrane was washed with deionized water (Milli-Q ultrapure water) and stored in a desiccator to provide their dryness. The membrane films waited in the 0.1 mol HCl solution overnight. The membrane was placed into a semi-cell of the diffusion type transport unit and then clamped with the other semi-cell. First, the diffusion type

transport cell was controlled with deionized water toward any liquid leakage, and then 100 mL feed and stripping solutions were filled into each semi-cells mutually. After that, the samples were taken from each compartment of the set-up for quantitative elemental analysis by ICP-MS Agilent 7700 (Santa Clara, United States). All of the quantitative elemental analysis was triplicated in the standard deviation range of ±0.001 – 0.003.

The Cd transport conditions through the membrane were optimized with changing PIMs properties (time, membrane thickness, extractant type and amount, plasticizer type and amount along with polymer molecular weight) vs. constant aqueous phase (feed and stripping phase) properties, investigated in our previous study [43]. The experimental setup is shown in Fig. 1.

The experimental data for this study were obtained under different operating conditions such as time (range 0–8 h), film thickness of membrane (range 25.17–151.24 μm), extractant type range 1 to 4 (1-butyl, 2-decyl, 3-hexyl, and 4-octyl) and extractant amount (ranges 0–0.367 [w/w]), plasticizer type range 1–4 (1-TEHP, 2-NPPE, 3-B2EHA, and 4-ONPOE), plasticizer amount (ranges 0–0.338 [w/w]) and polymer molecular weight range of 1–4 (1-low Mw; 2-Mw of 43,000; 3-Mw of 62,000; and Mw of 80,000) were used as inputs and removal efficiency of Cd (range 0.13–1.00 [C/C_0]) was used as output variable. Dataset is preprocessed before giving to the input layer. Some experimental data of operating parameters such as extractant type, plasticizer type, and polymer molecular weight were converted from characters to numerical data. Statistical information related to the data of each experiment is summarized in Table 1. In this work, three different ANN models including feed forward back-propagation (FFBPNN), recurrent neural network (RNN), and generalized regression neural network were developed to predict Cd removal efficiency of PIMs. These approaches are briefly described as follows.

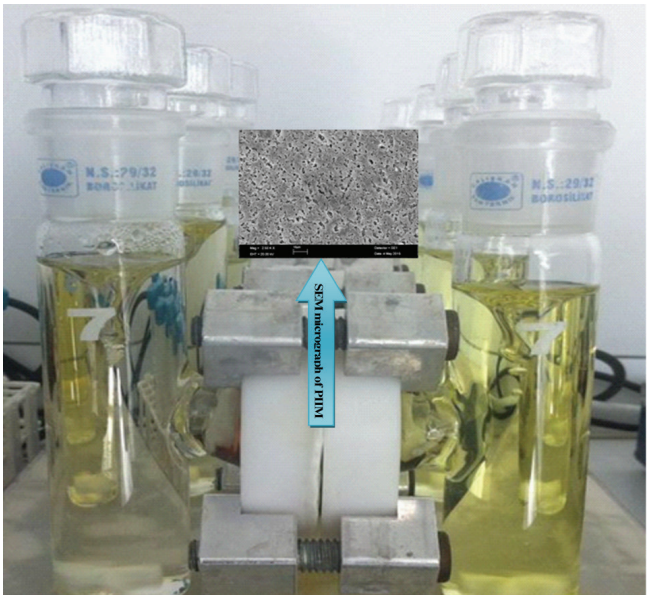

Fig. 1. Experimental setup for Cd transport through PIMs with SEM micrograph (2,500×).

2.2. Feed forward back-propagation neural network

FFBPNN is a supervised learning procedure [44], and recently it has been practiced for the approximation of nonlinear data sets. Back-propagation of errors is a frequently applied technique in ANN training to calculate the gradient of predicted and measured values on all weights in the developed network. The topology of FFBPNN comprises an input, an output, and one or more hidden layers as shown in Fig. 2. During the training phase, an input prototype is specified as the input layer for the computation of output. Then the predicted output of the network is compared with the desired output pattern [45]. The endeavor of the back-propagation learning rule is to describe a technique of adjusting the weights of networks to minimize output error on the weights and thresholds [46]. Eventually, the network will give the output that matches the desired output pattern given any input pattern in the training set. The input I_k and output O_k to the k_{th} neuron are determined by the following equations:

$$I_k = \sum_{i=1}^{n} \mu_{i,k} O_i \tag{1}$$

$$O_k = f\left(I_k + \theta_k\right) \tag{2}$$

where $\mu_{i,k}$ is the weight to f the connection from the i_{th} neuron in the previous layer to the k_{th} neuron, $f(I_k + \theta_k)$ represents the activation function of the neurons, O_k is the output of neuron, and the θ_k is the biases input to the neuron [46]. In order to improve performance, we adopted the bipolar sigmoid activation function, which is defined as follows:

$$f\left(x\right) = \frac{2}{1+e^{-x}} - 1 \tag{3}$$

FFBPNN method used to compute the network parameters by an iterative Levenberg–Marquardt algorithm: from an initial set of parameters. The backpropagation learning technique is considered the most efficient technique for obtaining good results [47].

Table 1
Data statistics of model variables (n = 460)

Variables	Data statistics			
	x_{min}	x_{max}	x_{mean}	Σ
Input layer				
Time	0.00	8.00	4.000	2.832
Extractant type	1.00	4.00	2.500	1.119
Extractant rate	0.00	0.367	0.261	0.065
Film thickness (μm)	25.17	151.24	49.000	21.159
Plasticizer type	1.00	4.00	3.739	0.736
Plasticizer rate	0.00	0.338	0.236	0.058
Polymer molecular weight	1.00	5.00	3.00	0.660
Output layer				
Removal efficiency	0.215	1	0.845	0.144

x_{min}, x_{max}, x_{mean}: minimum, maximum and mean values.
σ: standard deviation.

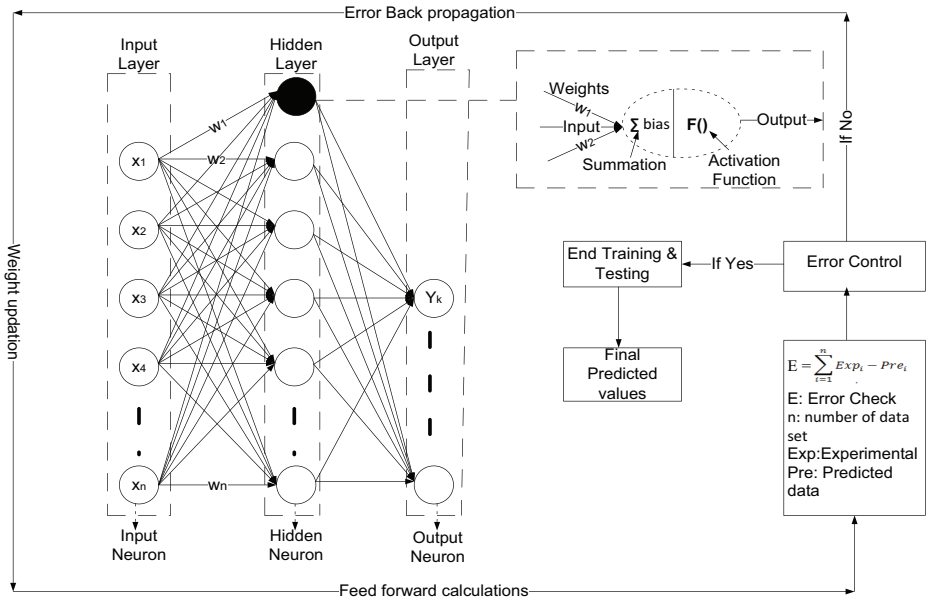

Fig. 2. A feed forward back-propagation neural network schematic diagram.

2.3. Recurrent neural network

The uniqueness of this RNN is that its structure is more analogous to the original biological neural network concept of massively parallel processing. In Fig. 3 basic architecture has shown a network of neuron-like units, in which each unit is connected to every other unit. At any specified time step, each non-input unit evaluates its current activation as a nonlinear function of the weighted sum of the activations of all concerned units. In each sequence, the generated error is the sum of the deviations of all activations calculated by the network from the related target signals. The total error of a training set in many sequences is the sum of the errors of all individual sequences. The gradient descent is used to vary each weight about its derivative concerning the error, to minimize the total error and provided the non-linear activation functions are differentiable [48].

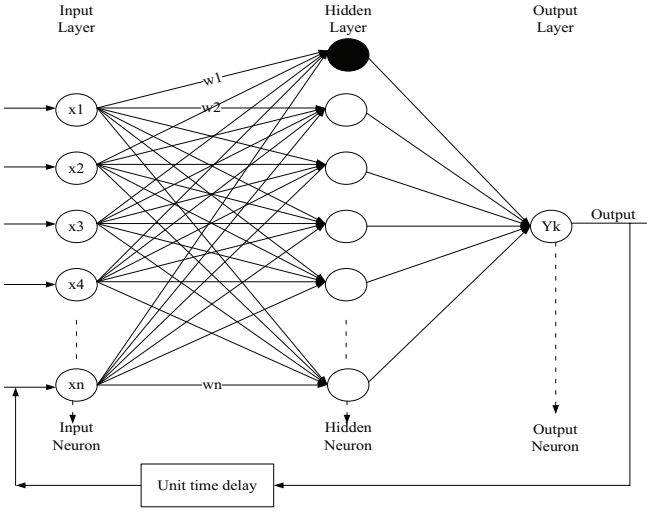

Fig. 3. An RNN schematic diagram.

2.4. Generalized regression neural network

The term GRNN for Nadaraya–Watson kernel regression [49,50], also reinvented in the neural network literature [51] and there is no need for iterative training procedure. It estimates any arbitrary function between input parameters and output, by drawing a function that predicts directly from training data. Furthermore, it is consistent; as the size of the training set increases, the estimation error approaches zero, it is related to the radial basis network and is based on kernel regression statistical technique. GRNN basic architecture is shown in Fig. 4. The model has four layers, the first layer is the input layer, the second is pattern layer, the third layer is the summation layer and final layer is the output layer. The number of input units in the first layer is equal to independent variables x_i and connected to the pattern layer (hidden layer). Each pattern layer unit is attached to the two neurons in the summation layer, known as S and D summation neuron. The S summation neuron computes the sum of the weighted outputs of the pattern layer while the D summation

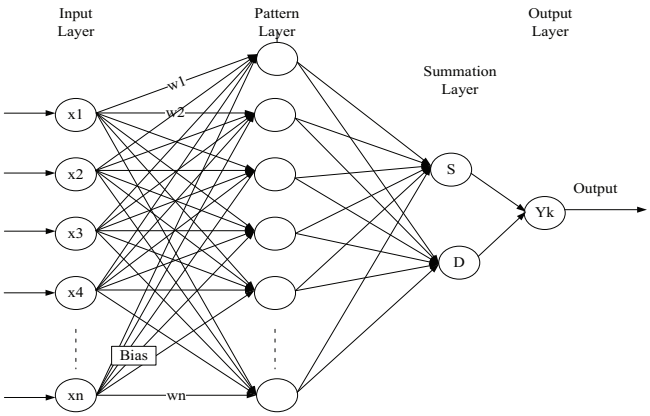

Fig. 4. Schematic diagram of GRNN.

neuron calculates the un-weighted outputs of the pattern neurons [52]. The output layer merely divides the output of each S summation neuron by that of each D summation neuron, yielding the predicted value expressed as:

$$y_i = \frac{\sum_{i=1}^{n} y_i \exp\left[-D\left(x, x_i\right)\right]}{\sum_{i=1}^{n} \exp\left[-D\left(x, x_i\right)\right]} \tag{4}$$

where n is the number of independent input variables, y_i is the target output value corresponding to the ith input pattern and the Gaussian D function is defined as:

$$D\left(x, x_i\right) = \sum_{j=1}^{p} \left(\frac{\left(x_{j,} - x_{ij}\right)}{\sigma_j} \right)^2 \tag{5}$$

where p is the total number of training patterns, and σ_j is referred to the smoothing parameter (width or spread), whose optimal value is often determined experimentally as presented in the literature [53].

2.5. Multiple linear regression

MLR is a mathematical equation that expresses the relationship between a dependent variable and some independent variables. This technique is based on least squares: the model is fit such that the sum of squares of differences between observed and predicted values are minimized. MLR can be formulated by using general equation (Eq. (6)) as follows:

$$Y = \beta_o + \beta_1 X_1 + \ldots + \beta_n X_n + \varepsilon \tag{6}$$

In this equation, Y denotes dependent the variable, X denotes independent variables, β denotes predicted parameters, and ε is the error term.

MLR calculations are made by using Eq. (7) where Y represents dependent variables (removal efficiency), and X_i represents independent variables (respectively, time, extractant type, extractant amount, film thickness, plasticizer type, plasticizer amount, and polymer molecular weight).

$$\begin{aligned} Y = {} & 1.02053 - 0.00397X_1 + 0.00291X_2 - 0.0759X_3 \\ & + 0.000258X_4 - 0.00285X_5 - 0.05166X_6 - 0.00846X_7 \end{aligned} \tag{7}$$

2.6. Proposed model architectures

Neural networks have brought about breakthroughs in processing images, video, speech, and audio, whereas recurrent nets have shown light on sequential data such as text and speech [54]. Since our dataset is in sequential format, recurrent neural networks are applied along with other models such as FFBPNN, GRNN which were selected by a literature review [47,55–62]. The purpose of using multiple models is to evaluate the model efficiency and propose the best model. In our work, we compared the model efficiencies by training the dataset and identified the best model for Cd removal efficiency. MATLAB programming was used for the configuration, training, and optimization of ANN models. Experimental data were randomly divided into three groups for training (70%), validating (15%), and testing (15%). The dataset contains 460 rows; each row comprises seven features and an output label. All the 460 rows were used in MLR, since, it is a linear regression technique; therefore, a linear relationship is assumed between the dependent variable and the independent variables. In case of ANN, 322 data rows were selected for the ANN training process, 69 data rows for the validation process and the remaining 69 rows for testing the model.

The number of layers and the number of nodes in each layer is used to determine the topology of the ANN model. The number of neurons (N) in the hidden layer is determined according to the minimum error prediction and considered a basic parameter for ANN structure. There is no standard method for finding the ideal number of hidden nodes in ANN modeling to determine the best activation function. A trial-and-error method can lead to inelegant ANN designs, but it is a time-consuming technique [63]. To determine the optimum number of neurons in the hidden layer, different topologies were examined, in which the number of nodes was varied from 3 to 45. Each topology was repeated three times, and RMSE was used as the error function [64] and found that 15 neurons in each hidden layer are the best topology due to minimum RMSE for training and cross-validation. By using trial and error technique, it was found that (7-15-1) is the best topology in this study as represented in Fig. 5. In the proposed models, the input layer consists of 7 neurons, the hidden layer contains 15 neurons, and the final layer is an output layer with 1 neuron. In our proposed model different parameters are used, tuned and tested to attain the best performance. These parameters include optimizers, activation functions, learning rate, and momentum. Table 2 shows the trained algorithms with different best-fit parameters where scaled conjugate gradient back-propagation performed better during testing and selected for proposed ANN models. The optimum structure of FFBPNN model is presented in Fig. 6. Similarly, Fig. 7 shows the proposed RNN model, and GRNN model architecture is shown in Fig. 8. Furthermore, the MLR technique was also applied to check the non-linearity of the dataset. SPSS statistics was used for MLR.

2.7. Model performance evaluation

In this study, three performance measures have been used to validate models and their predictions. The root

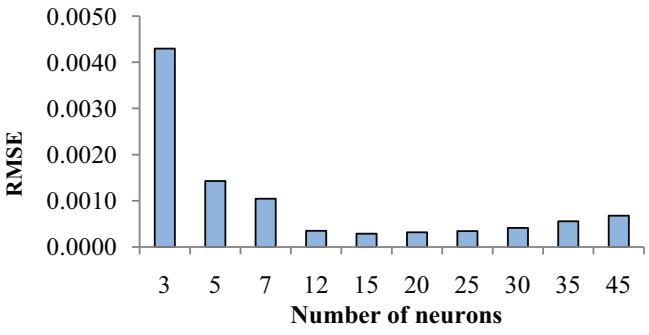

Fig. 5. RMSE vs. number of neurons.

Table 2
Comparison of trained algorithms with best fit parameters using 15 neurons in the hidden layer

Algorithms	Function	RMSE		
		Train	Validation	Test
BFGS quasi-Newton back-propagation	trainbfg	0.0432	0.0720	0.0554
Powell–Beale conjugate gradient back-propagation	traincgb	0.0442	0.0607	0.0464
Fletcher–Reeves conjugate gradient back-propagation	traincgf	0.0416	0.0712	0.0664
Polak–Ribiére conjugate gradient back-propagation	traincgp	0.0450	0.0540	0.0391
Gradient descent back-propagation	traingd	0.1373	0.0804	0.1682
Gradient descent with momentum back-propagation	traingdm	0.1558	0.0903	0.1731
Gradient descent with adaptive learning rate back-propagation	traingda	0.0769	0.0350	0.1145
Levenberg–Marquardt back-propagation	trainlm	0.0398	0.0842	0.0926
Resilient back-propagation	trainrp	0.0488	0.0464	0.0906
Scaled conjugate gradient back-propagation	trainscg	0.0445	0.0478	0.0353
One-step secant back-propagation	trainoss	0.0459	0.0593	0.0494

RMSE: root mean square error.

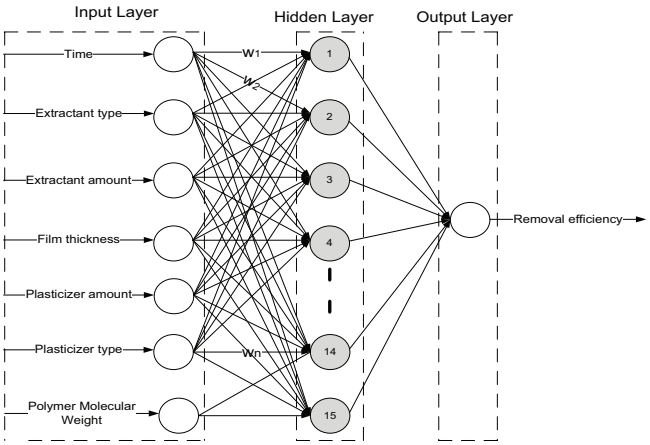

Fig. 6. Optimum structure of FFBPNN model.

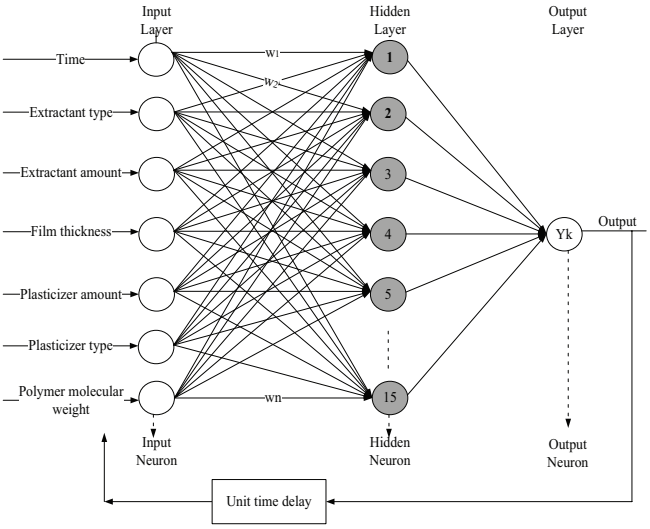

Fig. 7. Optimum structure of RNN model.

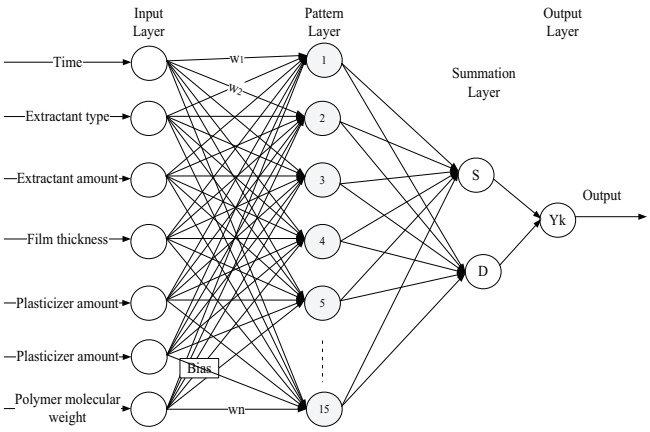

Fig. 8. Optimum structure of GRNN model.

means square error (RMSE), the coefficient of determination (R^2) and the mean absolute error (MAE) are the common performance measures for the evaluation of model performance. These are described as follows:

- The RMSE represents the error between model predictions and target values. It ranges from 0 to 1, where lower RMSE values are preferable and can be computed using Eq. (8) [65].

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{\mathrm{sim}}^{i} - Y_{\mathrm{obs}}^{i}\right)^2}{n}} \tag{8}$$

In this equation n represents the number of target values, X_{sim}^{i} and Y_{obs}^{i} are model predictions and their corresponding target values, respectively.

- The coefficient of determination (R^2) is estimated through Eq. (9). 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%) and how

much this value closes to 1 means the greater correlation and the stronger relationship between predictions and target values [65].

$$R^2 = \left[\frac{n\sum_{i=1}^{n} y_{\text{obs},i} y_{\text{sim},i} - \left(\sum_{i=1}^{n} y_{\text{obs},i}\right)\left(\sum_{i=1}^{n} y_{\text{sim},i}\right)}{\sqrt{\left[n\sum_{i=1}^{n} y^2_{\text{obs},i} - \left(\sum_{i=1}^{n} y_{\text{obs},i}\right)^2\right]\times\left[n\sum_{i=1}^{n} y^2_{\text{sim},i} - \left(\sum_{i=1}^{n} y_{\text{sim},i}\right)^2\right]}}\right]^2 \quad (9)$$

The RMSE and R^2 values provide information on general error ranges between model predictions and target values. Also, we used MAE to estimate the distribution of errors between model predictions and target values.

- The MAE can be computed with Eq. (10) and its values can range from 0 to 1. Similar to RMSE, lower values of MAE indicate a good correlation between model predictions and experimental data [65].

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|X^i_{\text{sim}} - Y^i_{\text{obs}}\right| \quad (10)$$

2.8. Sensitivity analysis

Sensitivity analysis computes the uncertainty in all complex models that is beneficial to recognize the impact of critical input parameters to output [66]. To identify primary input–output association, the technique of considering one parameter at a time was employed manually in the present study. It provides information about the significance of each parameter among the seven operating parameters to predict PIMs efficiency. The sensitivity was defined as the R^2 value, which indicates the performance of the network and the more important parameter shows higher R^2 values stating that the network is affected to a greater extent [67].

3. Results and discussion

3.1. Modeling results

The predicted results of PIMs removal efficiency of different models are evaluated by RMSE, MAE, and R^2 values. RMSE and R^2 values provide information on general error range between model predictions and measured results to assess the performance of developed ANN model while MAE value estimates the distribution of errors between model predictions and target values [68]. Similar to RMSE, lower values of MAE indicate a good correlation between model predictions and experimental data [69].

Results of the proposed ANN models and MLR are shown in Table 3. FFBPNN is the best option with maximum R^2, and minimum RMSE and MAE followed by RNN model while GRNN was the least ANN model. The correlation coefficient R^2 for MLR results showed the lowest value of 0.787 as compared with ANN models having values of 0.988, 0.981, and 0.861 for FFBPNN, RNN, and GRNN models, respectively. MLR results showed that it is not an appropriate technique to predict PIMs removal efficiency because no linear relationship was found due to non-linearity of data sets.

The comparison of the experimental results of three ANN models for the testing data is graphically presented in Fig. 9. It showed that ANN model results are well distributed around $X = Y$ line in a narrow area. Experimental results also show that FFBPNN model is very appropriate for the prediction of Cd removal from aqueous solution by using PIMs with seven input variables (time, film thickness, extractant type and amount, plasticizer type and the amount and polymer molecular weight). From this comparative study, it is shown that overall ANN prediction values are closer to the experimentally measured values. The reliability of the proposed FFBPNN model is shown in Fig. 10. It shows that FFBPNN is the best fit model since the model predicted values are closer to the experimental values.

3.2. Sensitivity analysis

Table 3 presents the results of the determination coefficient for each developed model for all parameters affecting Cd removal efficiency of PIMs. The sensitivity analysis was performed using one by one approach that considers only one variable at a time to determine the percentage of contribution posed by the variable that would affect the R^2 values.

FFBPNN proposed model shown best match between predicted and experimental results and considered for sensitivity analysis. In sensitivity analysis, one by one technique was applied, and the results are presented in Fig. 11 for all operating parameters. It was observed that three of them including extractant amount, plasticizer type, and plasticizer amount are significant parameters as indicated by higher R^2 values shown in Figs. 11(c), (e) and (f). In contrast, time, extractant type, film thickness, and polymer molecular weight do not have much influence on removal efficiency as presented in Figs. 11(a), (b), (d) and (g). The extractant amount, plasticizer type, and plasticizer amount are more significant due to their effects on transport phenomenon depending on the Fick's second law. Also, these parameters affect the membrane pore structure, and extraction capacity. On the other hand, extractant type, film thickness, and polymer molecular weight are also insignificant because of their

Table 3
Evaluation criteria used in prediction of PIMs Cd removal efficiency by the ANN and MLR models

Model	Data	Performance criteria		
		RMSE	MAE	R^2
FFBPNN	Training	0.01370	0.00926	0.98856
	Validating	0.01165	0.00804	0.98625
	Testing	0.01419	0.00575	0.98866
RNN	Training	0.02537	0.01641	0.97257
	Validating	0.01365	0.01086	0.98181
	Testing	0.01842	0.01034	0.98137
GRNN	Training	0.09215	0.05483	0.70102
	Validating	0.04808	0.02946	0.75759
	Testing	0.05950	0.03501	0.86148
MLR	Testing	0.07774	0.05213	0.78772

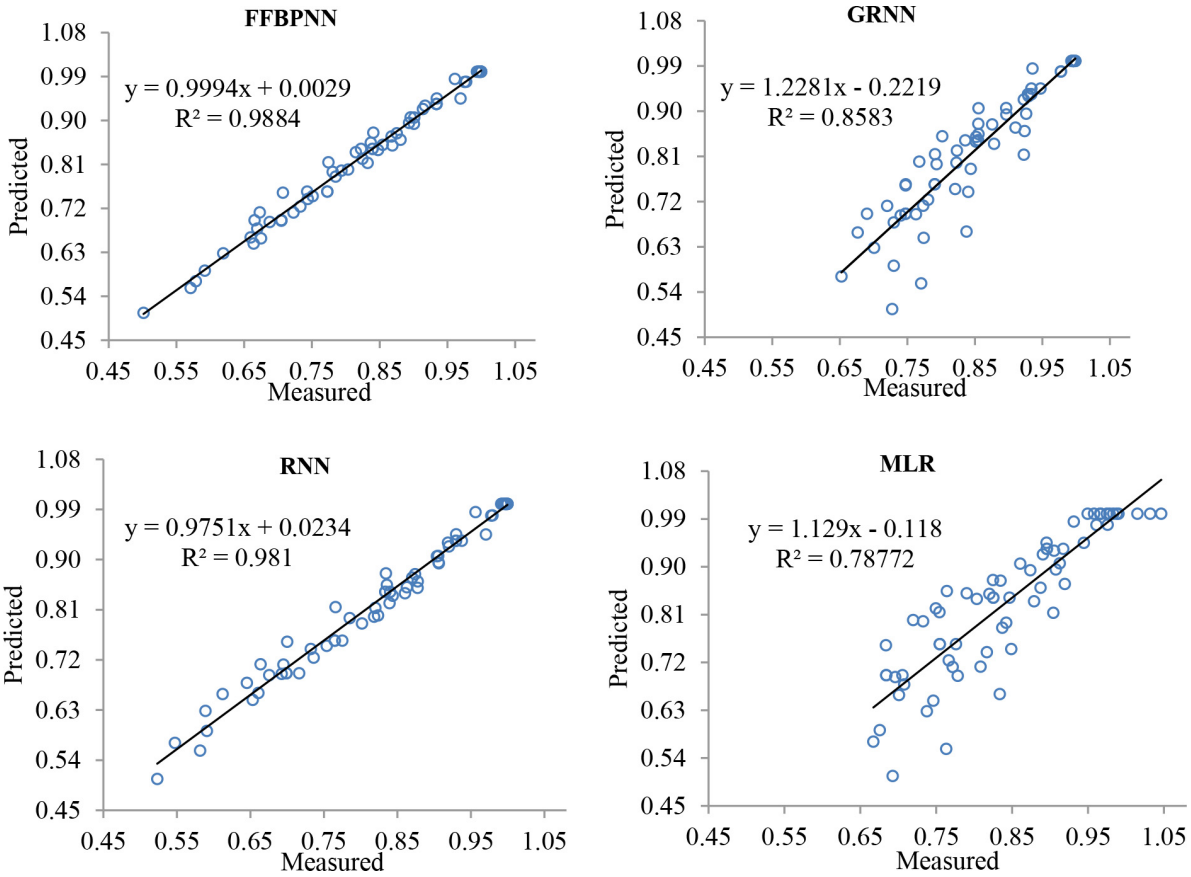

Fig. 9. Comparison of the measured, all ANN and MLR predicted results of testing data.

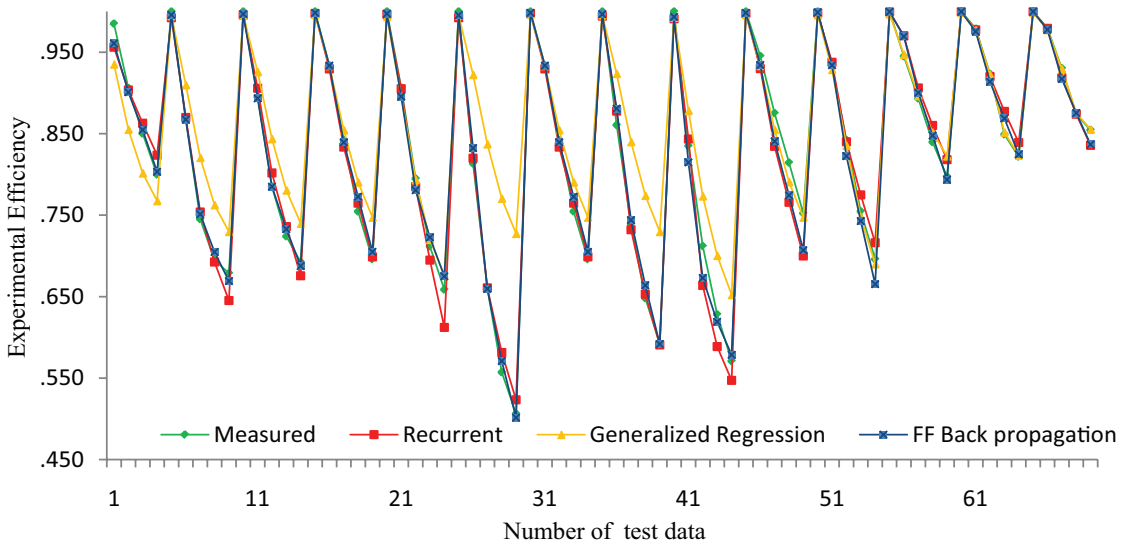

Fig. 10. Comparison of experimental and predicted results by ANN models and MLR.

less effectiveness in the membrane structure compared with the other membrane operating parameters.

Graphical representation showed the results of SA depicting the influence of operating parameters on removal efficiency and comparison between predicted and experimental values. The regression coefficient R^2 for time, extractant type, film thickness, and polymer molecular weight was observed as 0.4208, 0.2984, 0.538, and 0.0076, respectively, as shown in Figs. 11(a), (b), (d) and (g). The effect of extractant amount, plasticizer type, and plasticizer amount are major operating parameters showing R^2 values as 0.8834, 0.8574, and 0.8941 presented in Figs. 11(c), (e) and (f).

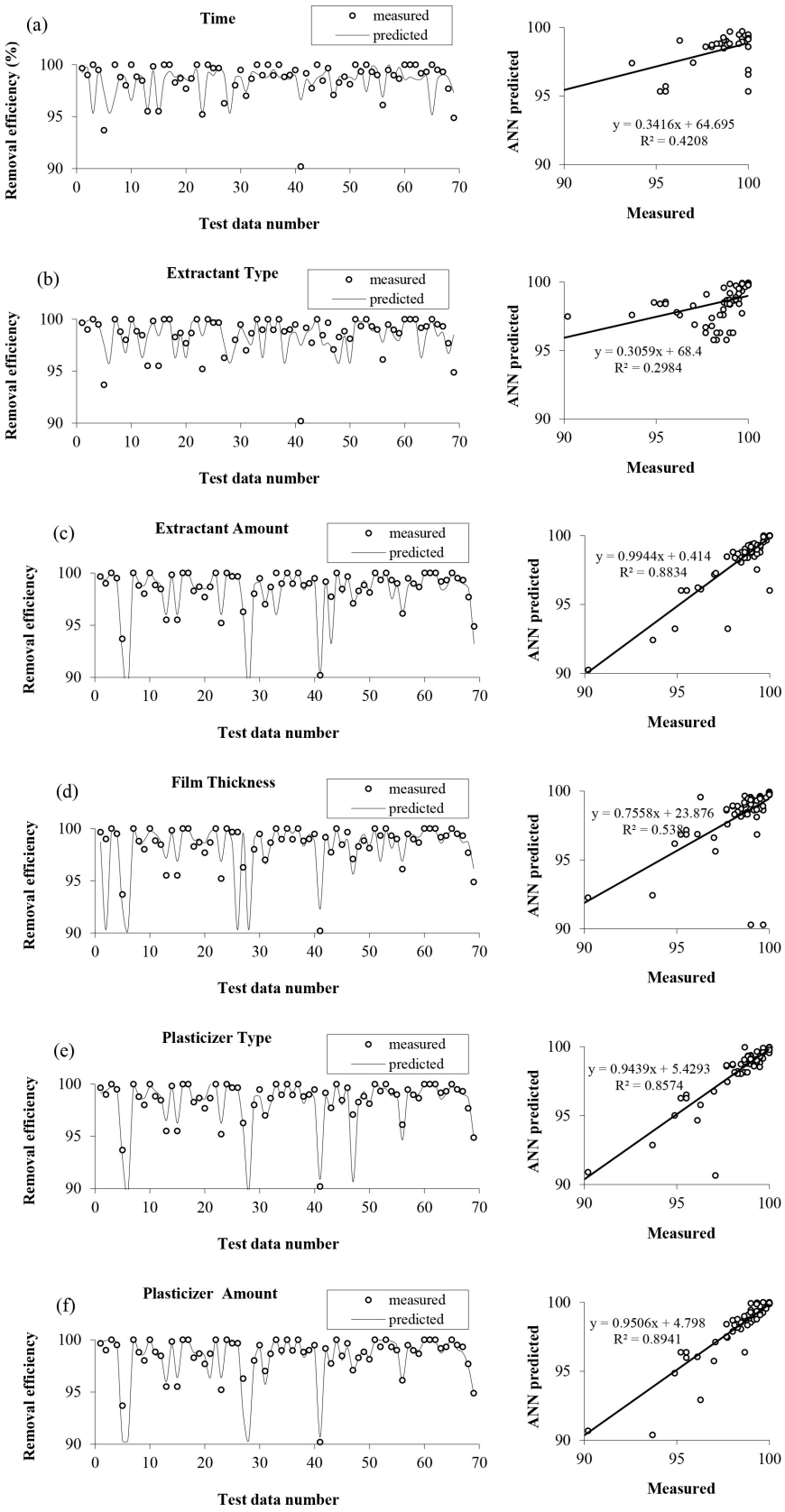

Fig. 11. Sensitivity analysis of FFBPNN model results for each input parameter (a) time, (b) extractant type, (c) extractant amount, (d) film thickness, (e) plasticizer type, (f) plasticizer amount, and (g) polymer molecular weight.

4. Conclusion

In this paper, we studied ANN models and MLR technique for the prediction of Cd removal efficiency of the PIMs and sensitivity analysis to find major affecting parameters. Three different ANN models were proposed to predict Cd removal efficiency of PIMs. The developed FFBPNN and RNN models predicted Cd removal efficiency with the highest R^2 and the lowest RMSE and MAE values rather than GRNN and MLR model. FFBPNN model results are in perfect match with an R^2 value of 0.988 that provides a meaningful supplement for the conventional and complicated mathematical patterns in the prediction of removal efficiency. Moreover, a sensitivity analysis was performed using one by one approach. Sensitivity analysis showed that extractant amount, plasticizer type, and plasticizer amount are major operating parameters while extractant type and polymer molecular weight were the least impacting operating parameters. Also, time and film thickness presented R^2 value approximately 0.50, so they could be considered for future studies to investigate their impact on PIMs removal efficiency of other heavy metals.

In future work, the least important parameters can be excluded for further research work while other parameters may be included at laboratory scale level. Furthermore, we will explore feasible, stable, and more economical models at industrial scale with a focus on other heavy metals removal by using optimum and significant inputs to maximize the efficiency.

Acknowledgments

The authors are thankful for the financial support provided by the Scientific and Technological Research Council of Turkey (TÜBİTAK), grant no: TBAG-112T806.

References

[1] S.J. Kulkarni, J.P. Kaware, Adsorption for cadmium removal from effluent - a review, Int. J. Sci. Eng. Technol. Res., 2 (2013) 1840–1844.
[2] U.M. Jibesh Datta, A comparative study of chromium and cadmium removal from their common aqueous solution by batch operation using tea factory waste as an adsorbent, Int. J. Eng. Res. Appl., 4 (2014) 98–105.
[3] S.-A. Ong, C.-E. Seng, P. Lim, Kinetics of adsorption of Cu (II) and Cd (II) from aqueous solution on rice husk and modified rice husk, Electron. J. Environ. Agric. Food Chem., 6 (2007) 1764–1774.
[4] A.J.M. Barros, S. Prasad, V.D. Leite, A.G. Souza, The process of biosorption of heavy metals in bioreactors loaded with sanitary sewage sludge, Braz. J. Chem. Eng., 23 (2006) 153–162.
[5] E.A. Silva, L.G.L. Vaz, M.T. Veit, M.R. Fagundes-Klen, E.S. Cossich, C.R.G. Tavares, L. Cardozo-Filho, R. Guirardello, Biosorption of chromium(III) and copper(II) ions onto marine alga *Sargassum* sp. in a fixed-bed column, Adsorpt. Sci. Technol., 28 (2010) 449–464.
[6] Z.R. Holan, B. Volesky, Biosorption of lead and nickel by biomass of marine algae, Biotechnol. Bioeng., 43 (1994) 1001–1009.
[7] APHA, Water Environment Federation and American Water Works Association, Standard Methods for the Examination of Water and Wastewater Part 4000 Inorganic Nonmetallic Constituents Standard Methods for the Examination of Water and Wastewater, 1999, 733 pp.
[8] R. Ali, Study on Removal of Cadmium from Water Environment by Adsorption on Gac, Bac and biofilter, Diffuse Pollution Conference, Vol. 1, 2003, pp. 35–39.
[9] M.C. Henson, P.J. Chedrese, Endocrine disruption by cadmium, a common environmental toxicant with paradoxical effects on reproduction., Exp. Biol. Med. (Maywood), 229 (2004) 383–392.
[10] H.K. Boparai, M. Joseph, D.M. O'Carroll, Kinetics and thermodynamics of cadmium ion removal by adsorption onto nano zerovalent iron particles, J. Hazard. Mater., 186 (2011) 458–465.
[11] U. Kumar, Agricultural products and by-products as a low cost adsorbent for heavy metal removal from water and wastewater: a review, Sci. Res. Essay, 1 (2006) 33–37.
[12] P. Kaewsarn, Q. Yu, Cadmium(II) removal from aqueous solutions by pre-treated biomass of marine alga *Padina* sp., Environ. Pollut., 112 (2001) 209–213.
[13] G.H. Azarian, A.R. Mesdaghinia, F. Vaezi, R. Nabizadeh, D. Nematollahi, Algae removal by electro-coagulation process application for treatment of the effluent from an industrial wastewater treatment plant, Iran. J. Public Health, 36 (2007) 57–64.
[14] Y.S. Dzyazko, V.N. Belyakov, Purification of a diluted nickel solution containing nickel by a process combining ion exchange and electrodialysis, Desalination, 162 (2004) 179–189.
[15] V.C. Srivastava, I.D. Mall, I.M. Mishra, Removal of cadmium(II) and zinc(II) metal ions from binary aqueous solution by rice husk ash, Colloids Surf., A, 312 (2008) 172–184.
[16] M. Mohsen-Nia, P. Montazeri, H. Modarress, Removal of Cu^{2+} and Ni^{2+} from wastewater with a chelating agent and reverse osmosis processes, Desalination, 217 (2007) 276–281.
[17] G.T. Ballet, L. Gzara, A. Hafiane, M. Dhahbi, Transport coefficients and cadmium salt rejection in nanofiltration membrane, Desalination, 167 (2004) 369–376.
[18] M. Ulewicz, W. Walkowiak, Separation of zinc and cadmium ions from sulfate solutions by ion flotation and transport through liquid membranes, Fizykochem. Probl. Miner., 37 (2003) 77–86.
[19] T.J. Butter, L.M. Evison, I.C. Hancock, F.S. Holland, K.A. Matis, A. Philipson, A.I. Sheikh, A.I. Zouboulis, The removal and recovery of cadmium from dilute aqueous solutions by biosorption and electrolysis at laboratory scale, Water Res., 32 (1998) 400–406.
[20] B. Swain, K. Sarangi, R. Prasad Das, Separation of cadmium and zinc by supported liquid membrane using TOPS-99 as mobile carrier, Sep. Sci. Technol., 39 (2005) 2171–2188.
[21] B. Swain, K. Sarangi, R.P. Das, Effect of different anions on separation of cadmium and zinc by supported liquid membrane using TOPS-99 as mobile carrier, J. Membr. Sci., 277 (2006) 240–248.
[22] A.Ö. Saf, S. Alpaydin, A. Coskun, M. Ersoz, Selective transport and removal of Cr(VI) through polymer inclusion membrane containing 5-(4-phenoxyphenyl)-6H-1,3,4-thiadiazin-2-amine as a carrier, J. Membr. Sci., 377 (2011) 241–248.
[23] R.F. Karimi, A study of the heavy metal extraction process using emulsion liquid membranes, Master's thesis, no. Chalmers University of Technology, Göteborg, Sweden, 2012, pp. 1–61.
[24] V.S. Kislk, Liquid Membrane Principles and Application in Chemical Separation and Waste Water Treatment, First Edition, Burlington, Elsevier, 2010.
[25] Y.-S. Park, T.-S. Chon, I.-S. Kwak, S. Lek, Hierarchical community classification and assessment of aquatic ecosystems using artificial neural networks, Sci. Total Environ., 327 (2004) 105–122.
[26] L. Belanche, J.J. Valdes, J. Comas, I.R. Roda, M. Poch, Prediction of the bulking phenomenon in wastewater treatment plants, Artif. Intell. Eng., 14 (2000) 307–317.
[27] G.R. Shetty, S. Chellam, Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks, J. Membr. Sci., 217 (2003) 69–86.
[28] E.S. Elmolla, M. Chaudhuri, M.M. Eltoukhy, The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process, J. Hazard. Mater., 179 (2010) 127–134.
[29] M. Cote, B. Grandjean, P. Lessard, J. Thibault, Dynamic modelling of the activated sludge process: improving prediction using neural networks, Water Res., 29 (1995) 995–1004.
[30] C. Gontarski, P. Rodrigues, Simulation of an industrial wastewater treatment plant using artificial neural networks, Comput. Chem. Eng., 24 (2000) 1719–1723.
[31] A. Kardam, K.R. Raj, J.K. Arora, S. Srivastava, ANN modeling on predictions of biosorption efficiency of zea mays for the removal of Cr (III) and Cr (VI) from waste water, Int. J. Math. Trends Technol., (2011) 23–29.

[32] J. Kabuba, A.F. Mulaba-bafubiandi, The Use of Neural Network for Modeling of Copper Removal from Aqueous Solution by the Ion-Exchange Process, International Conference on Mining, Mineral Processing and Metallurgical Engineering (ICMMME'2013), Johannesburg, SouthAfrica, April 15–16, 2013, pp. 131–135.
[33] H.M. Madhloom, Modeling of Copper removal from simulated wastewater by adsorption on to fungal biomass using artificial neural network, Glob. J. Adv. Pure Appl. Sci., 5 (2015) 35–44.
[34] A. Kardam, Artificial neural network modeling for sorption of cadmium from aqueous system by shelled *Moringa oleifera* seed powder as an agricultural waste, J. Water Resour. Prot., 2 (2010) 339–344.
[35] M. Fan, J. Hu, R. Cao, K. Xiong, X. Wei, Modeling and prediction of copper removal from aqueous solutions by nZVI/rGO magnetic nanocomposites using ANN-GA and ANN-PSO, Sci. Rep., 7 (2017) 18040.
[36] X. Shi, W. Ruan, J. Hu, M. Fan, R. Cao, X. Wei, Optimizing the removal of rhodamine B in aqueous solutions by reduced graphene oxide-supported nanoscale zerovalent iron (nZVI/rGO) using an artificial neural network-genetic algorithm (ANN-GA), Nanomaterials, 7 (2017) 134.
[37] M. Fan, T. Li, J. Hu, R. Cao, X. Wei, X. Shi, W. Ruan, Artificial neural network modeling and genetic algorithm optimization for cadmium removal from aqueous solutions by reduced graphene oxide-supported nanoscale zero-valent iron (nZVI/rGO) composites, Materials (Basel), 10 (2017) 544.
[38] D. Bastani, M.E. Hamzehie, F. Davardoost, S. Mazinani, A. Poorbashiri, Prediction of CO_2 loading capacity of chemical absorbents using a multi-layer perceptron neural network, Fluid Phase Equilib., 354 (2013) 6–11.
[39] V. Eyupoglu, Ağır Metallerin Seçici Ekstraksiyonu için İmidazolyum Tuzları İçeren Polimer İçerikli Membranların Üretimi Karakterizasyonu ve Taşınım Verimlerinin Yapay Sinir Ağları ile Modellenmesi, Grant No: TBAG-112T806, Sci. Technol. Res. Counc. Turkey (TÜBİTAK), (2015).
[40] M. Zawadzki, L. Niedzicki, W. Wieczorek, U. Domańska, Estimation of extraction properties of new imidazolide anion based ionic liquids on the basis of activity coefficient at infinite dilution measurements, Sep. Purif. Technol., 118 (2013) 242–254.
[41] B. Wassink, D. Dreisinger, J. Howard, Solvent extraction separation of zinc and cadmium from nickel and cobalt using Aliquat 336, a strong base anion exchanger, in the chloride and thiocyanate forms, Hydrometallurgy, 57 (2000) 235–252.
[42] H.I. Turgut, V. Eyupoglu, R.A. Kumbasar, I. Sisman, Alkyl chain length dependent Cr(VI) transport by polymer inclusion membrane using room temperature ionic liquids as carrier and PVDF-co-HFP as polymer matrix, Sep. Purif. Technol., 175 (2017) 406–417.
[43] V. Eyupoglu, E. Polat, Evaluation of Cd(II) transport with imidazolium bromides bearing butyl and isobutyl groups as extractants from acidic iodide solutions by liquid-liquid solvent extraction, Fluid Phase Equilib., 394 (2015) 46–60.
[44] Y. Zhang, D. Guo, Z. Li, Common nature of learning between back-propagation and hopfield-type neural networks for generalized matrix inversion with simplified models, IEEE Trans. Neural Networks Learn. Syst., 24 (2013) 579–592.
[45] S. Liu, L. Xu, D. Li, Multi-scale prediction of water temperature using empirical mode decomposition with back-propagation neural networks, Comput. Electr. Eng., 49 (2016) 1–8.
[46] N. An, W. Zhao, J. Wang, D. Shang, E. Zhao, Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting, Energy, 49 (2013) 279–288.
[47] Y. Srinivas, A.S. Raj, D.H. Oliver, D. Muthuraj, N. Chandrasekar, A robust behavior of Feed Forward Back propagation algorithm of artificial neural networks in the application of vertical electrical sounding data inversion, Geosci. Front., 3 (2012) 729–736.
[48] Z.C. Lipton, A Critical Review of Recurrent Neural Networks for Sequence Learning, arXiv Prepr., 4 (2015) 1–35.
[49] E.A. Nadaraya, On Estimating Regression, Theory of Probability and Its Applications, 9 (1964) 141–142.
[50] G.S. Watson, Smooth regression analysis, Indian J. Stat., 26 (1964) 359–372.
[51] H. Schløler, U. Hartmann, Mapping neural network derived from the Parzen window estimator, Neural Networks, 5 (1992) 903–909.
[52] M.H. Ali, M. Park, I.K. Yu, T. Murata, J. Tamura, B. Wu, Enhancement of transient stability by fuzzy logic-controlled SMES considering communication delay, Int. J. Electr. Power Energy Syst., 31 (2009) 402–408.
[53] Z. Mingyu, D. Coggeshall, E. Ghaneie, T. Pope, M. Rivera, M. Georgiopoulos et al., Gap-based estimation: Choosing the smoothing parameters for probabilistic and general regression neural networks, Neural Comput., 19 (2007) 2840–2864.
[54] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature, 521 (2015) 436.
[55] A.J. Al-Mahasneh, S.G. Anavatti, M.A. Garratt, Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems, arXiv Prepr., (2018).
[56] A.L. Betker, T. Szturm, Z. Moussavi, Application of Feedforward Backpropagation Neural Network to Center of Mass Estimation for Use in a Clinical Environment, in Engineering in Medicine and Biology Society, Proc. 25th Annual International Conference of the IEEE, Vol. 3, 2003, pp. 2714–2717.
[57] G.J. Bowden, G.C. Dandy, H.R. Maier, Input determination for neural network models in water resources applications. Part 1—background and methodology, J. Hydrol., 301 (2005) 75–92.
[58] G.J. Bowden, H.R. Maier, G.C. Dandy, Input determination for neural network models in water resources applications. Part 2. Case study: forecasting salinity in a river, J. Hydrol., 301 (2005) 93–107.
[59] F. Benvenuto, A. Marani, Neural networks for environmental problems: data quality control and air pollution nowcasting, Glob. NEST Int. J., 2 (2000) 281–292.
[60] M. Firat, M. Gungor, Generalized regression neural networks and feed forward neural networks for prediction of scour depth around bridge piers, Adv. Eng. Softw., 40 (2009) 731–737.
[61] A.A. Konate, H. Pan, N. Khan, J.H. Yang, Generalized regression and feed-forward back propagation neural networks in modelling porosity from geophysical well logs, J. Pet. Explor. Prod. Technol., 5 (2015) 157–166.
[62] S. Asante-Okyere, Q. Xu, R.A. Mensah, C. Jin, Y.Y. Ziggah, Generalized regression and feed forward back propagation neural networks in modelling flammability characteristics of polymethyl methacrylate (PMMA), Thermochim. Acta, 667 (2018) 79–92.
[63] H. Jaeger, Artificial Neural Networks, Talk, 17 (2008) 281–289.
[64] M.R. Dehghani, H. Modarress, A. Bakhshi, Modeling and prediction of activity coefficient ratio of electrolytes in aqueous electrolyte solution containing amino acids using artificial neural network, Fluid Phase Equilib., 244 (2006) 153–159.
[65] 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.
[66] R. Haghbakhsh, H. Hayer, M. Saidi, S. Keshtkari, F. Esmaeilzadeh, Density estimation of pure carbon dioxide at supercritical region and estimation solubility of solid compounds in supercritical carbon dioxide: correlation approach based on sensitivity analysis, Fluid Phase Equilib., 342 (2013) 31–41.
[67] R.H. Myers, S.L. Myers, Probability & Statistics for Engineers & Scientists, Pearson Prentice Hall, Vol. 6, 2007.
[68] A. Giacomino, O. Abollino, M. Malandrino, E. Mentasti, The role of chemometrics in single and sequential extraction assays: a review. Part II. Cluster analysis, multiple linear regression, mixture resolution, experimental design and other techniques, Anal. Chim. Acta., 688 (2011) 122–139.
[69] I. Yildirim, S. Ozsahin, K.C. Akyuz, Prediction of the financial return of the paper sector with artificial neural networks, BioResources, 6 (2011) 4076–4091.