



# Application of artificial neural networks to predict total dissolved solids in the river Zayanderud, Iran

Asadollahfardi Gholamreza<sup>†</sup>, Meshkat-Dini Afshin, Homayoun Aria Shiva, Roohani Nasrin

Department of Civil Engineering, Kharazmi University, Tehran 15719-14911, Iran

## ABSTRACT

An Artificial Neural Network including a Radial Basis Function (RBF) and a Time Delay Neural Network (TDNN) was used to predict total dissolved solid (TDS) in the river Zayanderud. Water quality parameters in the river for ten years, 2001-2010, were prepared from data monitored by the Isfahan Regional Water Authority. A factor analysis was applied to select the inputs of water quality parameters, which obtained total hardness, bicarbonate, chloride and calcium. Input data to the neural networks were pH, Na<sup>+</sup>, Mg<sup>2+</sup>, Carbonate (CO<sub>3</sub><sup>2-</sup>), HCO<sub>3</sub><sup>-1</sup>, Cl<sup>-</sup>, Ca<sup>2+</sup> and Total hardness. For learning process 5-fold cross validation were applied. In the best situation, the TDNN contained 2 hidden layers of 15 neurons in each of the layers and the RBF had one hidden layer with 100 neurons. The Mean Squared Error and the Mean Bias Error for the TDNN during the training process were 0.0006 and 0.0603 and for the RBF neural network the mentioned errors were 0.0001 and 0.0006, respectively. In the RBF, the coefficient of determination (R<sup>2</sup>) and the index of agreement (IA) between the observed data and predicted data were 0.997 and 0.999, respectively. In the TDNN, the R<sup>2</sup> and the IA between the actual and predicted data were 0.957 and 0.985, respectively. The results of sensitivity illustrated that Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup> parameters had the highest effect on the TDS prediction.

**Keywords:** Neural network, RBF, TDNN, TDS, Zayanderud

## 1. Introduction

The river Zayanderud is the life of Isfahan province. Therefore, protecting the water quality of the river for drinking, agriculture and industry is vital. The first step in the proper and sustainable management of water resources is to analyze water quality, changes in the time and place and to identify the main sources and types of water pollutants. The main sources of pollution in river Zayanderud are agricultural land uses, domestic and industrial wastewater. Agricultural lands provide major cations, anions, nitrogen and phosphorus to the River. Domestic sewage and industry add pollutants such as phosphorus and heavy metals to the pollution in the river [1].

Salinity in the surface waters is a significant issue of concern in various agricultural purposes and domestic consumption. The amount of total dissolved solid (TDS) is as an indicator of salinity and determining the changes that occurred over time are very important to the planning and management of water usage. The River's salinity is a problem of great importance and sensitivity in all Iranian Rivers and can be caused by several factors, for

example, minerals in river water and catchment soils that contain both suspended and soluble particles [2].

Interpretation of water quality is a very important part of water quality management. Several methods are available to analyze water quality data. Maier and Dandy [3] used the Artificial Neural Networks (ANNs) to predict salinity in the Murray River in South Australia. Zhang et al. [4] applied the ANN to predict water quality in the North Saskatchewan River. Huang and Foo [5] used the ANN for assessing variations in salinity in the Apalachicola River in Florida. Misaghi and Mohammadi [6] studied the water quality of the river Zayanderud, Iran, using an ANN. They employed a Generalized Regression Neural Network (GRNN) to predict the Biochemical oxygen demand (BOD) and the Dissolved oxygen (DO) of the river. Kanani et al. [7] used a Multilayer Perceptron (MLP) and an Input Delay Neural Network (IDNN) to predict TDS in the Ajichay River. Their results indicated a good performance and acceptable accuracy to predict salinity in the river. Asadollahfardi et al. [2] applied a MLP and a recurrent neural network (RNN) to predict TDS in the Talkheh Rud River, Iran. They reported that the results of the RNN had a good agreement to the field monitoring. Nemati et al. [8] studied the salinity of



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2016 Korean Society of Environmental Engineers

Received August 11, 2015 Accepted June 8, 2016

<sup>†</sup> Corresponding author

Email: asadollahfardi@yahoo.com

Tel: +98-9121192424 Fax: +98-2188830891

**Table 1.** Summary of Statistical Data Used in This Study

Parameters	pH	Ca <sup>2+</sup> (mg/L)	Na <sup>+</sup> (mg/L)	Cl <sup>-</sup> (mg/L)	SO <sub>4</sub> <sup>2-</sup> (mg/L)	HCO <sub>3</sub> <sup>-1</sup> (mg/L)	Mg <sup>2+</sup> (mg/L)	K <sup>+</sup> (mg/L)	TH (mg/L)	TDS (mg/L)
Number (mon)	120	120	120	120	120	120	120	120	120	120
Average	7.82	3.40	2.20	1.83	2.14	2.98	1.37	0.04	2.35	432.98
Minimum	6.90	1.70	0.10	0.20	0.10	1.10	0.10	0.00	1.40	232.05
Maximum	8.80	7.80	8.30	9.60	9.51	5.30	5.50	0.50	5.20	1220.70
Standard deviation	0.41	0.96	1.78	1.48	1.53	0.67	0.74	0.08	0.65	187.51
Mode	7.80	2.9	1	0.9	1.4	2.9	1.0	0.01	2.1	320.45

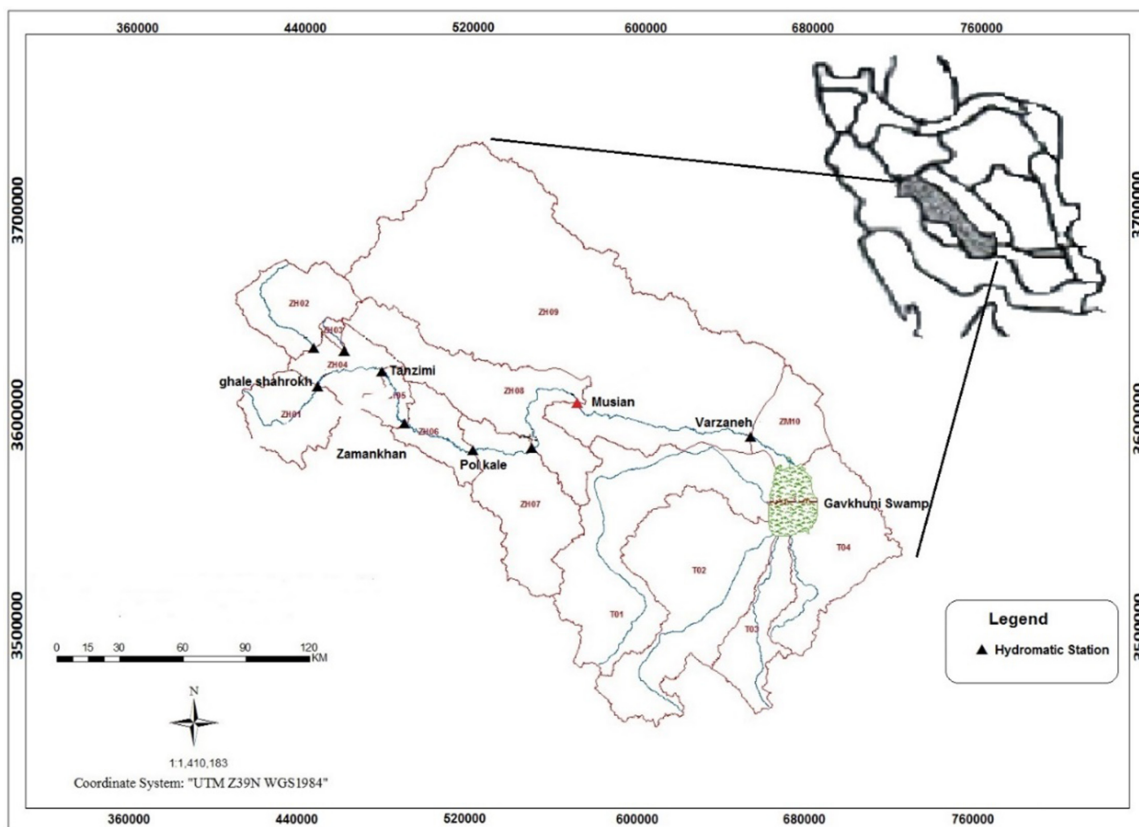
the Siminrud River in Iran using an ANN. They concluded that magnesium had a high impact in predicting salinity.

The objective of our study was to predict TDS in the river Zayanderud, Iran, using a Radial Basis Function (RBF) and a Time Delay Neural Network (TDNN). We also applied a sensitivity analysis to find the effect of each parameter in predicting of TDS.

### 1.1. Study Area

The river Zayanderud basin includes the southwestern region of Iran located between 31°30'N and 33°32'N and 49°30'E and 49°52'E. The area includes four cities including Shahrekord, Frieden, Lenjan and Isfahan and covers part of the Chahar Mahal Bakhtiari province (1.7% of the total area). Fig. 1 indicates the study area [1]. Average annual rainfall varies from 1,600 mm in the Zard Kuh Mountains to less than 40 mm in the eastern regions of Isfahan [9].

Generally, the amount of rainfall in the catchment area of the river Zayanderud decreases from West to East. The mean annual air temperature in the northwestern highlands reaches 3.5°C and in the eastern parts of the central region temperatures can reach 21.5°C. The relative humidity in January is at its highest and at its lowest in July. From a geological perspective, the studied area consists of three main geological zones: Zagros, Sanandaj-Sirjan and Central Iran. Virtually each of these zones has affected the area according to their specific characteristics. Jurassic metamorphic and sedimentary rock and new Quaternary alluvium are the most abundant constituents of river rocks in the area. Two aspects of fine-grained rock (sedimentary and metamorphic) have great importance. First, the geochemistry of the rock is partly effective in increasing of natural concentrations of minerals in the river Zayanderud, and second erodibility has



**Fig. 1.** Location of the Mosian Station study area on the river Zayanderud - Iran.

an important role in increasing fine-grained particles in the river. In general, fine-grained particles cause an uptake of potentially toxic elements in sediments in the river bed due to a high adsorption capacity. The Iranian Central region, including the river Zayanderud basin from the western highlands to the eastern regions consists of areas with mild, cool and dry summers, areas with cold winters and very hot and dry summers.

## 2. Materials and Methods

We used monthly water quality data from the river Zayanderud from 2001 to 2010, about 120 data for each variable. The data were monitored by the Isfahan Regional Water Authority and the Department of Environment (Iran). Table 1 presents the summary of statistical data in the Mosian monitoring station.

Selection of suitable input parameters for an artificial neural network is a very important step. One of the techniques to identify the relationship between water quality parameters is a factor analysis method.

### 2.1. Factor Analysis

A factor analysis is a technique for reducing the input parameters to the artificial neural network and finding a more effective method to predict TDS parameter. Factor analysis variables (water quality parameter) are located in the factors, so that the first factor decreases the next variance factor. The variables that are located in the first factor are the most effective ones. To perform a factor analysis, we applied SPSS version 21 (2014). Before performing a factor analysis, we must ensure that we apply appropriate variables of the factor analysis. For this purpose, we used a Kaiser-Meyer-Olkin (KMO) index, Eq. (1) [10].

$$KMO = \frac{\sum \sum r_{ij}^2}{\sum \sum r_{ij}^2 + \sum \sum a_{ij}^2} \quad (1)$$

Where  $r_{ij}$  is the correlation coefficient of indicator  $x_i$  and indicator of  $x_j$ ;  $a_{ij}$  is the offset correlation coefficient of index  $x_i$  and indicator  $x_j$ .  $KMO$  values of close to one indicate that the correlation between pairs of variables can be explained by other variables. Therefore, justifying the application of variable factor analysis is provable. The following steps should be carried out for factor analysis.

- 1 - Creating a matrix of correlations between the water quality parameters which is a square matrix of correlation coefficients.
- 2 - Determining  $KMO$  to demonstrate the suitability of factor analysis.
- 3 - Factors should be partially rotated around the origin, to obtain a new position.
- 4 - Finally, the number of factors equal of the correlation matrix that is considered to be greater than one [11-15].

### 2.2. ANN Modeling

An ANN is a computing technique to help the learning process

and tries to map the input space (input layer) and a favorable environment (output layer) by using processors called neurons by identifying inherent relationships between data [16]. A hidden layer receives data from the input layer, processes it and sends it to the output layer. Each network receives training through examples. Network learning is carried out when the connection weight between the layers change in a way that the difference between the predicted and the measured values are within acceptable limits. Achievement to this condition fulfills the learning process. This expresses the weight of memory and network knowledge. Trained neural networks can be used to predict outputs corresponding to actual new data [17]. Due to the structure of ANNs, features such as high-speed processing, template learning ability by a template pattern method, generalization of ability after learning, flexibility against unexpected failures and lack of significant disruption on the part of the connection are due to network weight distribution [18].

### 2.3. Architecture of the Network

We used two hidden layers of TDNNs with a sigmoid tangent transfer function and linear output. The number of neural in the hidden layer can be determined by the output component. The only lasting selection in network architecture would be the number of hidden layers which we selected according to a minimum error between one, two and three hidden layer neural network. To determine the number of neural in the hidden layer, the method of trial and error was applied to arrive at best network. Another point is the simplicity of the network. Between two alternatives, the one which has less neurons in the hidden layers was selected for TDNNs to avoid network complication. For adequacy of model during training, Mean Bias Error (MBE) and Mean Squared Error (MSE) were used. The MBE indicates the adequacy of the model. If MBE equal to zero; our model is adequate and if MBE less than zero, the model is underestimated. The MBE greater than zero shows overestimating of the network. MBE is a tool to prevent overestimating of neural network. In the training process, the weights and biases were adjusted using momentum methods to minimize the network performance, and the performance was evaluated with MSE between the network outputs and the target outputs. If the calculated MSE found small enough and stable at the end of each learning epoch, by adjusting the learning rate, epochs, number of hidden layers and neurons, the parameter set is determined and a post-process is carried out.

In learning parameters of a prediction function and testing, a model would just repeat the labels of the samples that it has just seen would have a perfect score but would fail to predict anything useful on yet-unseen data. This situation is called overfitting. To avoid it, k-fold cross validation is common practice. In this method, the original sample is randomly split into k almost equal sized subsamples. A model is trained using k-1 of the folds as training data; the resulting model is validated on the remaining part of the data as testing data for calculating the accuracy of the model. The cross-validation process is then repeated k times, with each of the k subsamples used exactly once as the validation data. The k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. This approach can be

computationally expensive, but does not waste too much data, which is a major advantage in problem such as inverse inference where the number of samples is small.

As usual, the true error is estimated as the average error rate on test examples, Eq. (2).

$$Error = \frac{1}{N} \sum_{i=1}^N E_i \quad (2)$$

Where N is the number of subsamples.

## 2.4. The TDNN Network

A time delay neural network is a multilayer neural network that is able to deal with the dynamic nature of sample data and to hold input signals. A TDNN consists of three layers whose weights are coupled with time delay cells. Each cell's transfer function has a TDNN Sigmoid tangent function and a D (N + 1) weighted input. A TDNN is a dynamic network output, which depends on the network's previous inputs and outputs in addition to its current inputs. Since networks have dynamic memories, they can be used for learning time-varying, sequential patterns. A TDNN operator receives an input signal and keeps it for a time step and in the next time step, the input signal emerges as an output result. By connecting an N series to a TDNN operation, a Tapped Delay Line (TDL) will be obtained. The output is a vector with N + 1 component. The N + 1 components include the inputs in the current time step and the previous N time steps [19]. The present study evaluated the modeling of a new TDNN network with a training function trainlm to predict TDS variation trends. We applied a new TDNN function network, a sigmoid tangent transfer function, a training trainlm function and two hidden layers to predict TDS using 10 years of data.

## 2.5. RBF Neural Network

The Gaussian RBF neural network is a non-normalized form of a Gaussian distribution nonlinear function and has good features for enhanced learning. Gaussian neural networks that used for complex mapping can also learn, identify, synchronize and to control, nonlinear dynamic systems [16].

An RBF network is naturally derived from an interpolation problem. The RBF has a non-linear input layer and a Gaussian hidden layer. Fig. 2 depicts a view of an RBF neural network.

According to Fig. 2, an RBF neural network input layer is directly connected to the hidden layer. The output of the j-th hidden layer is obtained from Eq. (3) [20].

$$h_j = \frac{\varphi(X - c_j)}{\delta_j} \quad (3)$$

Where  $h_j$  is the output of the j-th neuron;  $\varphi$  is a nonlinear function of RBF; X is an input vector;  $c_j$  is a neuron center and  $\delta_j$  is the neuron's central span. Nonlinear function is due to  $\Phi$  functions. Neurons have a linear function in the output layer and the output of  $y_k$  in neuron  $k$  in the output layer is obtained from Eq. (4) [20]:

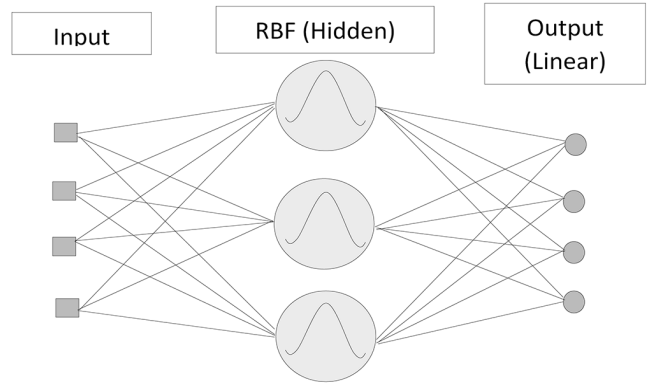


Fig. 2. Schematic view of the RBF network.

$$y_k = \sum W_{kj} \cdot h_j \quad (4)$$

Where  $W_{kj}$  is a synaptic weight connecting of j-th of hidden layer and neuron  $k$  of output layer and  $m$  is the number of hidden layer neurons.

We used a newrb function network and a radbas transfer function. The number of neurons in the network can be added sequentially and continue until MSE approaches the set target.

## 2.6. Model Efficiency

To compare different prediction results, forecast errors of different periods need to be considered to be regarded as benchmark comparisons. Among these criteria, we used the root mean squared error (RMSE) (Eq. (5)). An MBE is used to calculate the adequacy of the model. (Eq. (5)) [21].

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (A_t - F_t)^2} \quad (5)$$

$$MBE = \frac{1}{N} \sum_{t=1}^N \left| \frac{F_t - A_t}{A_t} \right| \quad (6)$$

Where  $N$  is the number of data;  $A_t$  is the actual data and  $F_t$  is the predicted data.

Also the coefficient of determination ( $R^2$ ) and the Index of Agreement (IA) indicate the reliability of the model [22].  $R^2$  and IA can be illustrated as follows:

$$R^2 = \left( \frac{\sum_{i=1}^i (A_t - \bar{A})(F_t - \bar{F})}{(\sum_{i=1}^i Z(A_t - \bar{A})^2) \sum_{i=1}^i ((F_t - \bar{F})^2)^{0.5}} \right)^2 \quad (7)$$

$$IA = d = 1.0 - \frac{\sum_{i=1}^N (A_t - F_t)^2}{\sum_{i=1}^N (|F_t - \bar{A}| + |A_t - \bar{A}|)^2} \quad (8)$$

Where  $\bar{A}$  and  $\bar{F}$  are the means of the actual data and the predicted data, respectively.

**Table 2.** The Correlation Coefficient Matrix between the Water Quality Parameters in the Mosian Station

	EC	pH	CO <sub>3</sub> <sup>-2</sup>	HCO <sub>3</sub> <sup>-1</sup>	Cl <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>	Ca <sup>2+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>	K <sup>+</sup>	TH
EC	1.000	-.190	-.008	.469	.932	.867	.817	.616	.960	.176	.950
pH	-.190	1.000	.548	-.282	-.100	-.201	-.284	-.019	-.190	.048	-.234
CO <sub>3</sub>	-.008	.548	1.000	-.333	.018	.036	-.059	.082	-.021	-.117	.001
HCO <sub>3</sub>	.469	-.282	-.333	1.000	.384	.180	.521	.218	.421	-.001	.511
Cl	.932	-.100	.018	.384	1.000	.703	.735	.571	.906	.184	.866
SO <sub>4</sub>	.867	-.201	.036	.180	.703	1.000	.674	.603	.867	.168	.836
Ca	.817	-.284	-.059	.521	.735	.674	1.000	.164	.760	.036	.842
Mg	.616	-.019	.082	.218	.571	.603	.164	1.000	.532	.160	.667
Na	.960	-.190	-.021	.421	.906	.867	.760	.532	1.000	.182	.866
K	.176	.048	-.117	-.001	.184	.168	.036	.160	.182	1.000	.109
TH	.950	-.234	.001	.511	.866	.836	.842	.667	.866	.109	1.000

### 3. Results and Discussion

The first step is to identify the parameters which contribute to the prediction of TDS. We used factor analysis to reach this objective. The first stage in factor analysis is to construct the correlation parameter matrix. If the correlation coefficients are less than 0.3, using factor analysis is questionable [13, 23]. However, according to Table 2, many of the correlation coefficient between the water quality parameters in the Mosian Station are larger than 0.3.

According to Table 3, the data are suitable for factor analysis because the KMO index of all stations is greater than 0.5 [11]. Bartlett's test results indicate that when the *p* value less than 0.05, a null hypothesis is confirmed, and a significant correlation exists between the variables.

Table 4 describes the eigenvalues for the factor analysis of hydro chemical data for the Mosian Station. To determine the number of factors, we selected the eigenvalues which were bigger than 1 [10].

Fig. 3 indicates the screen plot, in which the horizontal axis determines the factor number and the vertical axis presents the eigenvalues. As presented in Fig. 3, the eigenvalues are in descending tradition and a sudden drop between eigenvalue 1 and eigenvalue 2 confirms the existence of at least two of the main factors. In general, factors with a steep slope are most helpful in analysis and factors with a low slope have less impact on the analysis.

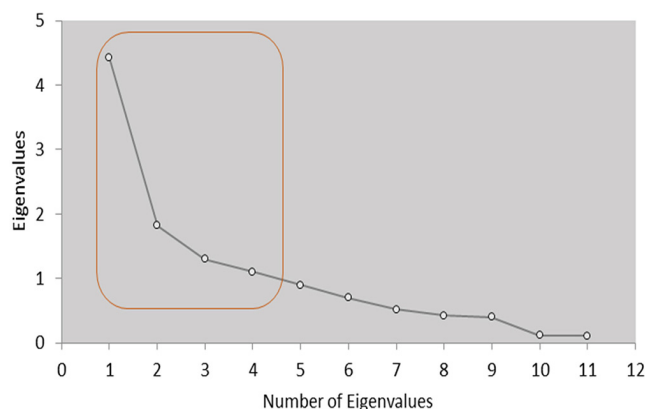
The first three factors include 80.63% of the total variance (Table 4). Table 5 illustrates the results of rotated factor loading using a Varimax method. We used the result of factor analysis to select the proper input to the ANN. As indicated in Table 5, the first factor is total hardness (TH), bicarbonate (HCO<sub>3</sub><sup>-1</sup>), chloride (Cl<sup>-</sup>) and calcium (Ca<sup>2+</sup>), which are the most important parameters in water quality of the river Zayanderud. We selected the mentioned parameters as input parameters to the ANN.

**Table 3.** Coefficient of the KMO and Bartlett Test Results

Test statistic	The Mosian Station
KMO	0.817
Significance	0.00

**Table 4.** Individual Eigenvalues and the Cumulative Variance of Water Quality Observations in the Mosian Station

Factors	Eigenvalue	Variance %	Cumulative variance %
1	6.871	57.255	57.255
2	1.732	14.437	71.692
3	1.073	8.945	80.637
4	.783	6.523	87.161
5	.721	6.006	93.167
6	.377	3.142	96.309
7	.244	2.034	98.343
8	.168	1.402	99.745
9	.021	.175	99.920
10	.008	.066	99.987
11	.002	.013	100.000
12	-1.004E-013	-1.033E-013	100.000



**Fig. 3.** Screen plot of eigenvalues of the Mosian Station.

#### 3.1. The ANN Results

Input data for training the ANN are based on the results of factor analysis. We selected pH, Na<sup>+</sup>, Mg<sup>2+</sup>, carbonate parameters and

the results of factor analysis, which were  $\text{HCO}_3^{-1}$ ,  $\text{Cl}^{-}$ ,  $\text{Ca}^{2+}$  and TH as input data to the ANN. In the sensitivity analysis the accuracy of the factor analysis will be examined. To avoid overfitting, in this study, we used 6-fold cross validation to compute the true error estimation. We applied six sets of data, in which the testing data were changed. The results of each five subsamples validating for both TDNN and RBF networks are listed in the Table 6.

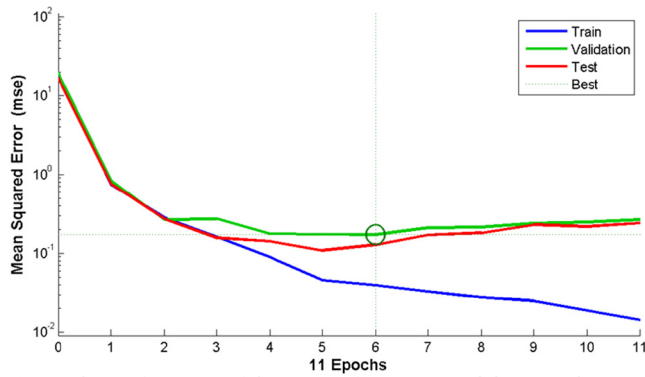


Fig. 4. The performance of the TDNN for training, validation and testing.

As indicated in table 6, the second set had the best performance. The true RMSE acquires from the average of the each subsamples error. The true RMSE were 0.843 and 0.516 for TDNN and RBF network, respectively, which proves the accuracy of the RBF is better than TDNN model. The results of the two RBF and TDNN models are as follows:

### 3.2. The Time Delay Neural Network (TDNN) Results

Fig. 4 presents the MSE between the actual and the simulated data during the modeling process in the TDNN method for training, validation and testing. Changing the rate of the error after epoch 6 is negligible. Therefore, we stopped the training process at epoch 11. The TDNN contains 2 hidden layers and 15 neurons in each layer. For TDNNs, we reached a minimum error when 2 hidden layers and 15 neurons in each layer were applied.

As indicated in Fig. 5, the coefficient of determination ( $R^2$ ) and the IA between the predicted TDS and the observed data were 0.957 and 0.986 which means the accuracy of the model in predicting TDS parameters was acceptable.

Table 5. Rotated Factors Loading for Water Quality Observations in the Mosian Station Using a Vartimax Method

TH	$\text{K}^+$	$\text{Na}^+$	$\text{Mg}^{2+}$	$\text{Ca}^{2+}$	$\text{SO}_4^{2-}$	$\text{Cl}^{-}$	$\text{HCO}_3^{-1}$	$\text{CO}_3^{-2}$	pH	TDS	Factors	Station
0.82	-	-	-	0.83	-	0.68	0.65	-	-	0.78	1	Mosian
-	0.63	0.75	-	-	0.85	-	-	-	-	-	2	
-	-	-	0.96	-	-	-	-	0.84	0.78	-	3	

Table 6. The Result of 5-fold Cross Validation for TDNN and RBF Networks

Subsamples	TDNN		RBF	
	$R^2$	RMSE	$R^2$	RMSE
NO.1	0.932	0.801	0.985	0.612
NO.2	0.957	0.823	0.997	0.431
NO.3	0.928	0.854	0.981	0.689
NO.4	0.952	0.917	0.976	0.456
NO.5	0.942	0.836	0.989	0.458
NO.6	0.944	0.829	0.987	0.451
True RMSE	0.843		0.516	

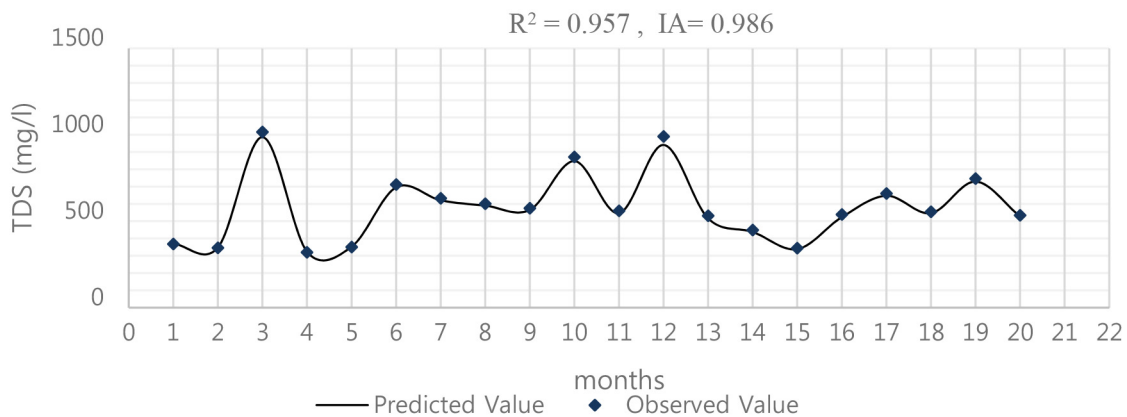


Fig. 5. Comparisons of the observed and the predicted results using the TDNN neural network.

### 3.3. The RBF Neural Network Results

As illustrated in Fig. 6, the amount of errors decreased sharply at first and then gradually declined until the amount of error approach zero during the training process of the network. The amount of errors after Epoch 90 was approximately the same for, validation and testing. As indicated in Fig. 6, using one hidden layer with 100 neurons, we reached an MSE equal to 0.0001, after which training was stopped.

Fig. 7. presents the actual and the predicted data of the TDS. As illustrated in Fig. 7, the  $R^2$  between the predicted data and the observed data for TDS in the Mosian station was 0.997 and the IA was 0.999.

Table 7 indicates the MSE, the MBE and the RMSE for the TDNN and RBF. As described in the Table 7, the amount of errors in the RBF is lower than in the TDNN. Using the RBF neural network is more acceptable than using a TDNN to predict the TDS of the river. If we compared the  $R^2$  of TDNNs with RBF

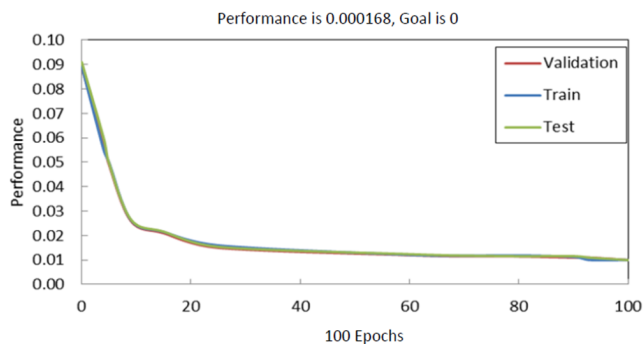


Fig. 6. The performance of training, validation and testing for the RBF neural network.

Table 7. The Amounts of Different Errors in the TDNN and the RBF

RMSE	MBE	MSE	Type of network
0.843	0.0603	0.0006	TDNN
0.516	0.0006	0.0001	RBF

neural network, RBF predicting of TDS indicates more accurate than TDNNs (Fig. 5, Fig. 7).

The prediction of TDS concentrations in the river may be beneficial for water quality management to make proper decisions in using river water for irrigation.

A similar research for predicting of TDS in the Ajichay River, Iran, was carried out by Kanani et al. (2008) using MLP and TDNN methods in one station. The  $R^2$  between the observed and the predicted TDS concentrations was 0.859 and 0.949, respectively. Their input data to the model was only the amount of flow. However, in our study, eight parameters including pH,  $Na^+$ ,  $Mg^{2+}$ , Carbonat( $CO_3^{2-}$ ),  $HCO_3^{-1}$ ,  $Cl^{-}$ ,  $Ca^{2+}$  and TH were used as input parameters to the ANN. We also carried out a sensitivity analysis to determine the roles of each input parameter in predicting of TDS in the river. Asadollahfardi et al. [2] predicted the TDS of the Talkherud River, Iran, using an MLP and ELMAN methods. They studied two stations and the  $R^2$  between the observed and the predicted TDS concentrations was 0.964 and 0.96, respectively. Their input data was rate of flow. The  $R^2$  in our study is larger than their study and the input parameter for their work was only the amount of flow in the river. Nemati et al. [8] applied an ANN to predict TDS of the Siminehrud River, Iran. Its  $R^2$  was 0.841. The  $R^2$  of our study was 0.999 and for selecting of suitable input parameters, we applied factor analysis.

### 3.4. Sensitivity Analysis

Sensitivity analysis is a method to assess the importance of each input parameter on the concentration of the output parameter. We declined and increased each of the input parameters 20%, while the other input parameter data were kept unchanged. After that, the impact of each parameter in prediction of the TDS concentrations was identified.

Fig. 8 presents the results of sensitivity analysis, using TDNN network. As indicated in Fig. 8,  $Ca^{2+}$  and  $SO_4^{2-}$  had the greatest impact on the TDS prediction. After that,  $Cl^{-}$ ,  $HCO_3^{-1}$  and TH were effective, respectively. Except the  $SO_4^{2-}$ , the results are the same as factor analysis results for selection of input parameters.

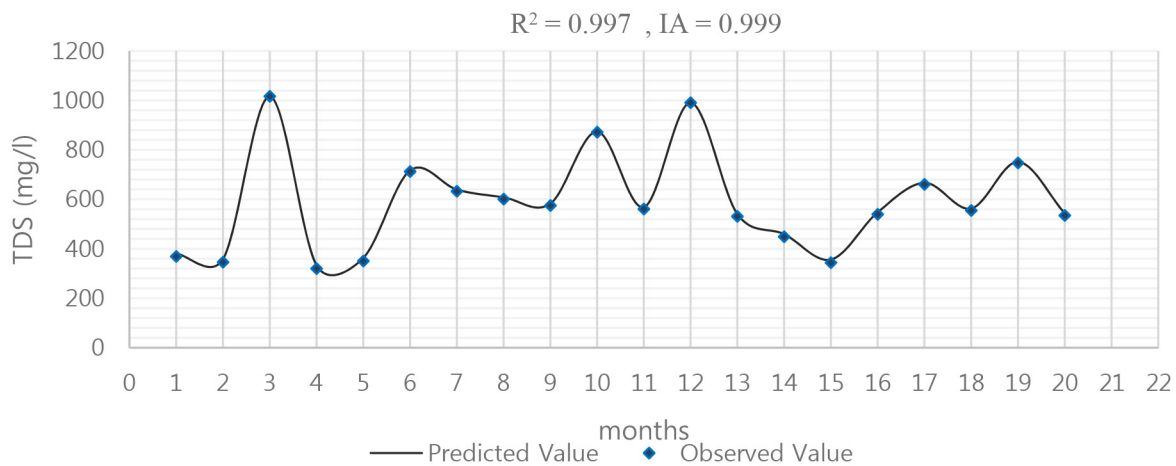


Fig. 7. Comparisons of actual and simulated data using the RBF neural network.



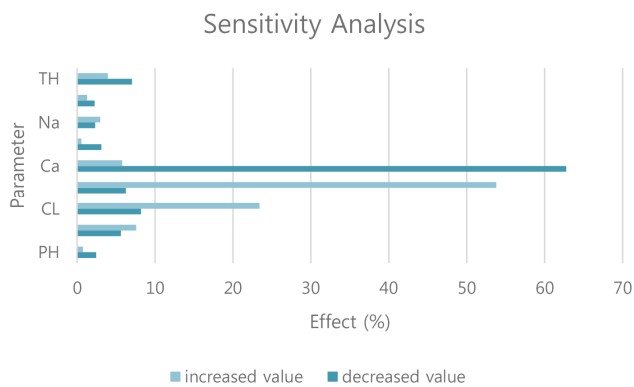


Fig. 8. Sensitivity analysis using TDNN network.

## 4. Conclusions

We summarized the results and discussions of using the RBF and the TDNN to predict TDS of the river Zayandehrud in the Mosian monitoring station as follows:

- For TDS prediction in TDNN,  $R^2$  and IA between the predicted data and the observed data were 0.957 and 0.986, respectively, which mean that our two neural network results are acceptable.
- The  $R^2$  and the IA between the predicted and the observed data for predicting TDS in the RBF was 0.997 and 0.999. The TDNN contained 2 hidden layers with 15 neurons in each layer and the RBF with one hidden layer containing 100 neurons.
- The MSE, RMSE and MBE for the TDNN were 0.0006, 0.0603 and 0.843, respectively. For the RBF neural network the mentioned errors were 0.0001, 0.43 and 0.516, respectively.
- The result of the RBF is more accurate than the TDNN in the prediction of TDS in Zayanderud River.
- The results of sensitivity analysis indicated that  $Ca^{2+}$  and  $SO_4^{2-}$  had the highest effect on the TDS prediction. According to its results, all the parameters from factor analysis had an important role in changes of TDS. The  $SO_4^{2-}$  was not mentioned in the results of factor analysis.

## Acknowledgements

I wish to acknowledge Mr. Ernest Rammel's assistance in editing our manuscript.

## References

- Hossaini Abari H. Zayandehrood from source to swamp. 2nd ed. Esfahan: Golha press; 2000.
- Asadollahfardi G, Taklify A, Ghanbari A. Application of artificial neural network to predict TDS in Talkheh Rud River. *J. Irrig. Drain. Eng.* 2012;138:363-370.
- Maier HR, Dandy GC. The use of artificial neural networks for the prediction of water quality parameters. *J. Water Resour. Res.* 1996;32:1013-1022.
- Zhang Q, Stanley SJ. Forecasting raw-water quality parameters in the North Saskatchewan River by neural network modeling. *Water Res.* 1997;1354:72-79.
- Huang W, Foo S. Neural network modeling of salinity variation in the Apalachicola River. *Water Res.* 2000;36:356-362.
- Misaghi F, Mohammadi K. Estimating water quality changes in the Zayandeh Rud River using artificial neural network model. 2000; written for presentation at CSAE/SCGR.
- Kanani SH, Asadollahfardi G, Ghanbari A. Application of artificial neural network to predict total dissolved solid in Achechay River basin. *J. World Appl. Sci.* 2008;4:646-654.
- Nemati S, Naghipour L, Fazeli Fard MH. Artificial neural network modeling of total dissolved solid in the Simineh River, Iran. *J. Civil Eng. Urban.* 2014;4:8-14.
- Moienian M. The natural landscape of the Zayandehrood River. Esfahan: Jahad Daneshgahi; 1999.
- Kaiser HF. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* 1960;20:141-151.
- Sarmad Z, Bazargan A, Hejazi A. Research methods in the behavioral sciences. 2nd ed. Tehran: Nshragah Institute; 1997.
- Lapin L. Probability and statistics for modern engineering, translation: Tamuri M, Rezaeian M. 1st ed. Tehran: Univ. of Technology; 2007.
- Johnson RA, Wichem DW. Applied multivariate statistical analysis, translation: Nirumand H. Mashhad: Ferdowsi Univ.; 2007.
- Meshkani M. Time series analysis, forecasting and control. Tehran: Univ. of Shahid Beheshti; 1992.
- Sobhanifar Y, Kharazian M. Factor analysis, structural equation modeling and multilevel. 1st ed. Tehran: Univ. of Emam Sadegh; 2012.
- Kia M. Neural network in MATLAB. 2nd ed. Tehran: Kian Univ. Press; 2012.
- Mahdinejad M. The prediction of air pollution of Tehran based on artificial neural network. Master Thesis. Tehran: Univ. of Kharazmi; 2013.
- Taherion M. Artificial neural network and its application in environmental engineering. First conference on environmental engineering. Tehran: Tehran Univ.; 2006.
- Hornik KM, Stinchcombe M, White H. Multilayer feed forward networks are universal approximator. *Neural Networks* 1989;2: 359-366.
- Menhaj M. Computational intelligence, fundamentals of artificial neural networks. Vol. 1. Farsi: Amirkabir Univ. Publisher; 1998.
- Kennedy JB, Neville AD. Basic statistical methods for engineers and scientists. 2nd ed. New York: Harper and Ro.; 1964.
- Krause P, Boyle DP, Base F. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* 2005;5:89-97.
- Tabachnick BG, Fidell LS. Using multivariate statistic. 6th ed. Boston: Pearson/Allyn and Bacon; 2012.