

Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models

S. Emamgholizadeh · H. Kashi · I. Marofpoor ·
E. Zalaghi

Received: 23 October 2012 / Revised: 20 May 2013 / Accepted: 8 September 2013 / Published online: 1 October 2013
© Islamic Azad University (IAU) 2013

Abstract This paper describes the application of multi-layer perceptron (MLP), radial basis network and adaptive neuro-fuzzy inference system (ANFIS) models for computing dissolved oxygen (DO), biochemical oxygen demand (BOD) and chemical oxygen demand (COD) levels in the Karoon River (Iran). Nine input water quality variables including EC, PH, Ca, Mg, Na, Turbidity, PO₄, NO₃ and NO₂, which were measured in the river water, were employed for the models. The performance of these models was assessed by the coefficient of determination R^2 , root mean square error and mean absolute error. The results showed that the computed values of DO, BOD and COD using both the artificial neural network and ANFIS models were in close agreement with their respective measured values in the river water. MLP was also better than other models in predicting water quality variables. Finally, the sensitive analysis was done to determine the relative importance and contribution of the input variables. The results showed that the phosphate was the most effective parameters on DO, BOD and COD.

Keywords ANN · ANFIS · Karoon River · Water quality

Introduction

The river water is a primary natural resource for people for different consumptions such as drinking (especially in developing countries), tourism, fish fostering and recreation; therefore, it requires at least an acceptable level of water quality. The surface water quality in a region largely depends on the nature and extent of the agricultural, industrial and other anthropogenic activities in the catchments (Singh et al. 2009). Therefore, in regard to the importance of surface water in water resource modeling, the prediction analysis of the river water quality for typical development scenarios is essential. Predicting the water quality is also the key factor in the water quality management of rivers, and it enables a manager to choose an option that satisfies a large number of identified conditions (Palani et al. 2008).

To assess river water quality, different parameters such as dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), temperature, pH and conductivity can be used. In general, the organic pollution in an aquatic system is measured and expressed in terms of declined dissolved oxygen (DO) level and the biochemical oxygen demand (BOD) (Singh et al. 2009). The DO level is a measure of the health of the aquatic system. A certain minimum level of DO in water is required for the aquatic life to survive (Basant et al. 2010). The sources of DO in a water body include re-aeration from the atmosphere, photosynthetic oxygen production and DO loading. The sinks include oxidation of carbonaceous and nitrogenous material, sediment oxygen demand and respiration by aquatic plants (Kuo et al. 2007). BOD is the amount of oxygen used by aerobic microorganism to break down the organic matters into more stable form (Cox, 2003). It measures an approximate amount of

S. Emamgholizadeh (✉) · H. Kashi
Department of Water and Soil, Faculty of Agriculture, Shahrood
University of Technology, Shahrud, Iran
e-mail: s_gholizadeh517@yahoo.com

I. Marofpoor
Department of Water, Faculty of Agriculture,
University of Kurdistan, Sanandaj, Iran

E. Zalaghi
Water and Power Authority (KWPA), Ahvaz, Khuzestan, Iran



biodegradable organic matter present in water and serves as an indicator parameter for the extent of water pollution. The BOD of any aquatic system is the foremost parameter needed for assessment of the water quality as well as development of management strategies for the protection of water resources (Singh et al. 2009). The chemical oxygen demand (COD) is used as a measure of the oxygen equivalent of the organic matter content of a sample that is susceptible to oxidation by a strong chemical oxidant. The COD is used to measure the total quantity of oxygen-consuming substances in the complete chemical breakdown of organic substances in water. It is an important parameter in measuring quality and determining what organic load is present in the water (Verma and Singh 2013).

Although parametric statistical and deterministic models have been the traditional approaches for modeling the water quality, these models require vast information on various hydrological subprocesses in order to arrive at the end results (Singh et al. 2009). Moreover, these models require precisely determined rate constants/coefficients pertaining to various hydrological, chemical, physical and biological processes, which are largely time and space specific in nature. Additionally, such models have analytical solutions, but they have boundary conditions as limitations (Basant et al. 2010). Also, since a large number of factors affect the water quality, it has a complicated nonlinear relation with the variables; therefore, traditional data processing methods are no longer good enough for solving the problem (Xiang et al. 2006; Ranković et al. 2012). In recent years, several researches have been conducted on water quality forecast models (Palani et al. 2008; Basant et al. 2010; Faruk 2010). Over the past several years, many nonlinear models such as artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) among others have been used for the prediction and forecasting of water resource variables. The literature review shows ANN models have been widely accepted as a potential useful way of modeling hydrological processes and have been applied to a range of different areas including water quality, rainfall runoff, sedimentation and rainfall forecasting. For example, Kuok et al. (2009) have used particle swarm optimization feed-forward neural network for modeling runoff. Bateni et al. (2007) have used ANNs and ANFIS models to estimate the equilibrium and time-dependent scour depth with numerous reliable database. ANNs have been used intensively in the development of a reservoir water quality simulation model (Soyupak et al. 2003; Chaves and Kojiri 2007). Palani et al. (2008) used an ANN model to predict and forecast temperature, salinity, DO and Chl-*a* in Singapore coastal. Singh et al. (2009) applied the ANN model for computing the dissolved oxygen (DO) and biochemical oxygen demand (BOD)

levels in the Gomti River (India). Faruk (2010) used a hybrid neural network and ARIMA models for water quality time series prediction. Their results showed that the hybrid model provides much better accuracy over the ARIMA and neural network models for water quality predictions by itself. Najah et al. (2009) used ANN to predict the Johor River (Malaysia) water quality parameters. Basant et al. (2010) used linear and nonlinear modeling for simultaneous prediction of the dissolved oxygen (DO) and biochemical oxygen demand (BOD) levels in the river water. They applied partial least squares (PLS) regression and feed-forward back-propagation artificial neural networks (FFBP ANNs) modeling methods to predict the DO and BOD levels by using eleven input variables which measured monthly in the river water at eight different sites over a period of ten years. Their results showed that both the linear (PLS) and nonlinear (ANN) models could predict the values of DO and BOD, but the performance of ANN was relatively better than PLS.

Although some water quality parameters like temperature, pH and DO, can be accurately measured with the exceeding simplicity, whereas on the other hand, calculation of BOD and COD is not only cumbersome but also inaccurate many times (Verma and Singh 2013). So, with regard to the importance of prediction of water quality parameters, this paper employs the application of ANN (MLB/BP and RBF) and ANFIS models to predict the DO, BOD and COD parameters of the Karoon River water simultaneously, and the results of three models are discussed and compared with those obtained through the measurements. For achievement of this purpose, the data of Karoon River with period of seventeen years (1995 to 2011) were used.

Materials and methods

Study area

Iran with an area of about 1,648,000 km² is located in the southwest of Asia and lies approximately between 25N and 40N in latitude and between 44E and 64E in longitude. The population of this country was estimated 75.0 million in 2011. Iran's important mountains are Alborz and Zagros ranges. Alborz and Zagros Chains stretch from northwest to northeast and from northwest to southeast, respectively. These two ranges play an important role in the non-uniform spatial and temporal distribution of precipitation in the whole country. The climate of Iran is mainly arid or semiarid, except for the northern coastal areas and some parts of western Iran. About 65 % of its territory is arid or hyperarid, and approximately 85 % has an arid, semiarid or hyperarid environment (Badripour 2006). Apart from the



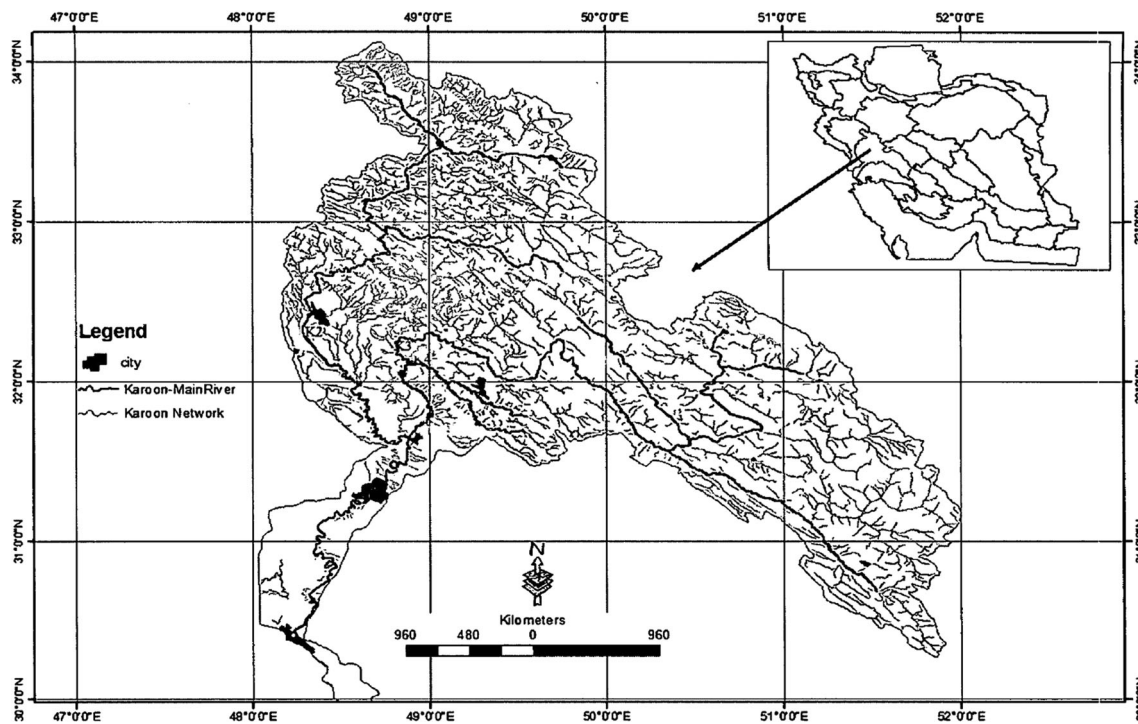


Fig. 1 Map of Karoon Basin (Iran)

coastal areas, the temperature in Iran is extremely continental with relatively large annual range about 22 to 26 °C.

The average annual rainfall of the country is about 240 mm; less than 1/3 of the average rainfall of the world (860 mm), with maximum amount in the Caspian Sea plains, Alborz and Zagros slopes with more than 1,800 and 480 mm, respectively. The annual precipitation of central and eastern plains decreases to less than 100 mm depending on their locations (Raziei et al. 2005). Since the rain in Iran is highly variable in time, space, amount and duration, therefore it is the most important limiting factor for biological and agricultural activities (Modarres and De Paulo Rodrigues da Silva 2007).

Karoon, Dez, Karkhe, Jarahi and Marooun are the main important rivers of Khuzestan province of Iran. Karoon with a basin area of 62,570 km² is the longest and most important river of Iran which collects the runoff of extensive areas and conveys to the Persian Gulf (Fig. 1). This river supplies water demands of 16 cities, several villages, thousands hectares of agricultural lands, and also with having four large dams (Karoon 1, 2, 3 and 4) has an important role in power generation. As Khuzestan Province is one of the strategic provinces of the country with a high potential for agricultural and industrial development, water pollution of the Karoon River system can significantly affect the development of this province. The increasing irregular water withdrawals for drinking, agriculture and industry, rapid population growth and urban extension, lack

of sewage networks due to the high cost of making them, unsuitable establishment of industries in the metropolitan area and non-compliance of environmental laws by large industrial units and direct wastewater discharges into Karoon River have endangered the aquatic life of this river.

The agriculture return flows by drainage have a high concentration of fertilizers, heavy metals, suspended and dissolved solids and pesticides, which violate the national effluent standard. Agricultural and agroindustrial return flows, domestic wastewater of the cities and villages and industrial effluents are the main pollution point sources of the Karoon River. Increasing water demands at the development stage including agricultural networks, fish hatchery projects and inter-basin water transfers could result in a gloomy future for water quality of the Karoon River (Karamouz et al. 2004). Approximately, the contribution of domestic, agricultural and industrial wastewater pollution is 48, 26 and 23 %, respectively. More urban sewage of Ahwaz, Gotvand, Abadan and Khorramshahr directly discharges into rivers without any treatment. Ahwaz city with 1.3 million population and discharge of three-fifths wastewater into Karun River has the highest share of pollution.

With respect to the importance of Karoon River, the study of water quality and prediction water quality parameters is very important. To achieve the goals of this research, ANN and ANFIS models were used to simulate water quality parameters such as DO, BOD and COD of the Karoon River.



Water quality data set

The water quality data set of Karoon River flowing through the southern alluvial Khuzestan plains in Iran, over a period of 17 years (1995 to 2011) was used in this study. These data are monitored regularly each month at eight different sites. The selection of an appropriate set of input variables from all possible input variables during artificial intelligence (AI) model development is important for obtaining high-quality model (Ranković et al. 2012). Many of the described methods for input variable selection are based on heuristics, expert knowledge, statistical analysis or a combination of these. However, although there is a well-justified need to consider input variable selection carefully, there is currently no consensus on how this task should be undertaken (May et al. 2008). So, in this research for prediction of the water quality parameters, i.e., dissolved oxygen (DO, mg L⁻¹), biochemical oxygen demand (BOD, mg L⁻¹) and chemical oxygen demand (COD, mg L⁻¹) based on the existing measured values of different variables and their correlative analysis, total nine factors (variables) were identified which affected the water quality parameters. These parameters are as follows: electricity conductive (EC, ds m⁻¹), water PH, calcium (Ca, mg L⁻¹), magnesium (Mg, mg L⁻¹), sodium (Na, mg L⁻¹), turbidity (NTU), phosphate (PO₄, mg L⁻¹), nitrate nitrogen (NO₃-N, mg L⁻¹), nitrite (NO₂, mg L⁻¹) (Eqs. 1, 2 and 3).

$$DO = f_1(EC, pH, Ca, Mg, Na, Turbidity, PO_4, NO_3, NO_2) \quad (1)$$

$$BOD = f_2(EC, pH, Ca, Mg, Na, Turbidity, PO_4, NO_3, NO_2) \quad (2)$$

$$COD = f_3(EC, pH, Ca, Mg, Na, Turbidity, PO_4, NO_3, NO_2) \quad (3)$$

To determine dispersion of data, two statistical measures, i.e., the coefficient of variation (CV) and standard deviation (SD), were used. The computations of these parameters for three dependent variables (DO, BOD and COD) are shown in Table 1. The results showed large variations between the samples with a high coefficient of variation (51.2 % for DO, 64.5 % for BOD and 63.7 % for COD). The reason for variation in concentration of the dependent variables is that the Karoon River during its course passes through several townships such as Ahwaz, Khorramshahr and Abadan cities and a number of wastewater drains and tributaries pour huge quantities of untreated wastewater into the Karoon River. The coefficient of variation for independent variables was also between 10.9 and 88.8 %.

In this study, ANNs and ANFIS were identified to predict the water quality (DO, BOD and COD) of the Karoon

Table 1 Basic statistics of the measured water quality variables in Karoon River, Iran

Variable	Unit	Min	Max	Average	SD	CV (%)
EC	ds m ⁻¹	1.7	9.26	4.55	1.71	83.2
pH	–	5.1	8.71	7.21	0.79	41.7
Ca	mg L ⁻¹	1	75	10.02	7.88	78.6
Mg	mg L ⁻¹	2.1	60	11.41	10.13	80.9
Na	mg L ⁻¹	1.42	40	18.23	8.13	44.6
Turbidity	NTU	1	25	7.81	6.32	88.8
PO ₄	mg L ⁻¹	0.13	3.21	1.14	0.89	78.6
NO ₃	mg L ⁻¹	0.34	2.7	1.01	0.42	10.9
NO ₂	mg L ⁻¹	0.08	2.1	0.46	0.38	37.5
COD	mg L ⁻¹	1.06	34.2	13.12	8.36	63.7
BOD	mg L ⁻¹	3.7	40.6	16.91	10.90	64.5
DO	mg L ⁻¹	3.1	29.4	12.28	6.28	51.2

SD standard deviation, CV coefficient of variation

River (Iran). For these models identification, the complete river water quality data set of 17 years (200 samples × 11 variables) was divided into two subsets. The training and testing data subsets comprised of 160 (80 %) and 40 (20 %) samples, respectively. Finally, the model input (independent variables), and the training and testing data sets had dimensions of 160 samples × 11 variables and 40 samples × 11 variables, respectively. The output variables (DO, BOD and COD) corresponding to the input variables belonged to the same water sample which were measured in same time and space.

Artificial neural network (ANN)

Artificial neural network (ANN) was first introduced in 1943 (McCulloch and Pitts 1943), and it is a massively parallel-distributed information processing system that has certain performance characteristics resembling biological neural networks of the human brain (Haykin 1999). ANN works very well when the volume and number of variables or diversity of the data are very large; the relationships between variables are vaguely understood, or the relationships are difficult to describe adequately with conventional approaches. Thus, ANN provides an analytical alternative to conventional techniques, which are often limited by strict assumptions of normality, linearity, variable independence, etc. For the training of the ANN model, considerable amounts of data set are needed. According to the learning algorithm, the ANN models can be classified to back-propagation, feed-forward back-propagation, feed-forward cascade correlation, radial basis function and conjugate gradients. (ASCE 2000a, b). These models were used by many researchers for engineering problems (Yilmaz and Kaynar 2011; Wang et al. 2010; Bandyopadhyay



and Chattopadhyay 2007), water quality study (Singh et al. 2009 and Chau 2006) and hydrological and hydraulic modeling (ASCE 2000a, b). In the present study, two kinds of ANN, i.e., the multi-layer perceptron (MLP) with back-propagation algorithm and Radial basis neural networks (RBNN), were used.

Multi-layer perceptron (MLP/BP)

MLP has generally three layers: input, output and intermediate or hidden layers. Back-propagation was proposed by Rumelhart et al. (1986), and it is the most popular algorithm for training of a MLP network (Wasserman 1989; Fausett 1994; Haykin 1999). The back-propagation algorithm involves two steps. The first step is a forward pass, in which the effect of the input is passed forward through the network to reach the output layer. After the error is computed, a second step starts backward through the network. The errors at the output layer are propagated back toward the input layer with the weights being modified according to Eq. (4).

$$\Delta w_{ij}(n) = -\varepsilon \times \frac{\partial E}{\partial w_{ij}} + \alpha \times \Delta w_{ij}(n-1) \quad (4)$$

where $\Delta w_{ij}(n)$ and $\Delta w_{ij}(n-1)$ are, respectively, weight increments between node i and j during the n th and $(n-1)$ th pass, or epoch, E is Error; ε and α are called learning rate coefficient and momentum factor, and control the algorithm's rate of learning. To optimize the rate at which a network learns, these factors must be set and/or adjusted properly during the training process. The valid range for both ε and α is between 0 and 1 (Qnet 2000 Manual 1999). In the typical configuration for MLP, a set of data ($X_1, X_2, \dots, X_n$) is first fed directly into the MLP through the input layer, and subsequently, the multi-layer perceptron produces an expected result y in the output layer. The number of hidden layers exhibits the complexity of the MLP, because a greater number of hidden layers increase the number of connections in the ANN. The number of nodes in each layer is evaluated by trial and error. The MLP is trained with a training set of input and known output data.

Radial basis network (RBF)

Radial basis neural networks (RBNN) were first introduced into the neural network literature in late 1980s by Broomhead and Lowe (1988). The RBF network is similar in topology to the MLP network (Fernando and Jayawardena 1998). RBF networks have the advantage of non-suffering from local minima in the same way as multi-layer perceptrons (Haykin 1999; Orr 1996). RBF networks are also good at modeling nonlinear data and can be trained in one stage rather than using an iterative process as in MLP and also

learn the given application quickly (Venkatesan and Anitha 2006). Also, an important property of RBF neural networks is that a high-dimensional space nonlinear problem can be easily broken down through a set of combination of radial basis functions, besides they are the beneficiary of the ability to be quickly trained (Chang and Chen 2003). The RBF network can be considered as a three-layer network, in which the hidden layer performs a fixed nonlinear transformation with no adjustable parameters (Leonardis and Bischof 1998). This layer consists of a number of nodes and a parameter vector called a "center," which can be considered the weight vector of the hidden layer. The standard Euclidean distance is used to measure how far an input vector is from the center. For each node, the Euclidean distance between the center and the input vector of the network input is computed and transformed by a nonlinear function that determines the output of the nodes in the hidden layer. The output layer then combines these results in a linear fashion. The output y of an RBF network is computed by the Eq. (5) (ASCE 2000a):

$$y = f(u) = \sum_{i=1}^n w_i R_i(x) + w_0 \quad (5)$$

where w_i = connection weight between the hidden neuron and output neuron; w_0 = bias; x = input vector and R is radial basis function.

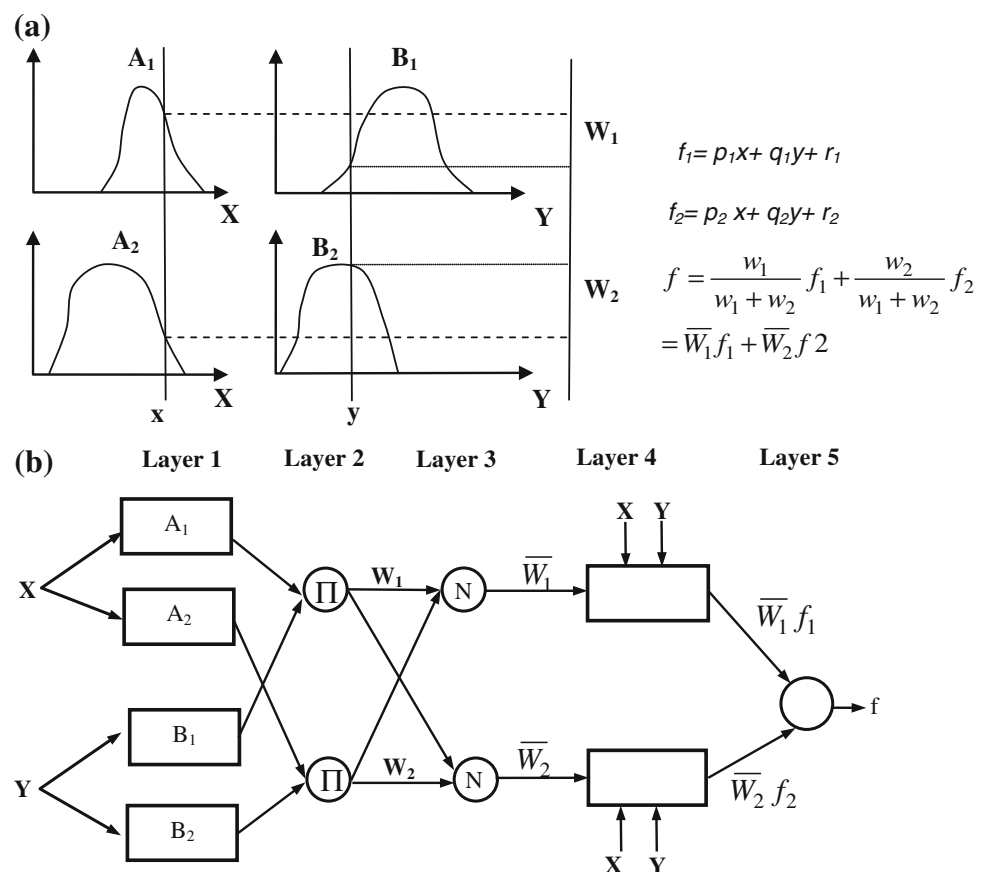
Adaptive neuro-fuzzy inference systems (ANFIS)

The Adaptive neuro-fuzzy inference system (ANFIS) first was introduced by Jang (1993). The neuro-fuzzy approach combines ANN and fuzzy logic. It effectively integrates the learning capability of neural networks into a fuzzy inference system (FIS). It can be used to approximate any real continuous function on a compact set to any degree of accuracy (Jang et al. 1997). Depending on the types of inference operations upon if-then rules, most FIS can be classified into three types: Tsukamoto's system, Mamdani's system and Sugeno's system (Kisi 2007). In this study, the first-order Sugeno fuzzy model is used because it has been used widely in engineering problems. The ANFIS model is able to use two different optimization methods (hybrid and back-propagation) to tune member function (MF) and generate fuzzy rules. The hybrid method is a combination of least squares estimation combined with back-propagation method (MATLAB User Manual 2008). In first-order Sugeno's system, if FIS has two inputs x and y and one output f , a typical rule set with two fuzzy IF/THEN rules can be expressed as:

$$\begin{aligned} \text{Rule 1 : If } x \text{ is } A_1 \text{ and } y \text{ is } B_1; \text{ then } f_1 &= p_1x + q_1y + r_1 \\ \text{Rule 2 : If } x \text{ is } A_2 \text{ and } y \text{ is } B_2; \text{ then } f_2 &= p_2x + q_2y + r_2 \end{aligned} \quad (6)$$



Fig. 2 **a** Sugeno's fuzzy if-then rule and fuzzy reasoning mechanism; **b** equivalent ANFIS



where x and y are input variables. A_i and B_i are the linguistic labels (low, medium, high, etc.) characterized by convenient membership functions, f_i are the outputs within the fuzzy region specified by the fuzzy rule; p_i , q_i and r_i are the parameters of the output function ($i = 1$ or 2).

Figure 2a shows the first-order Takagi–Sugeno fuzzy model with a two input and one output system. Also, the equivalent ANFIS architecture of the first order is shown in Fig. 2b.

As can be seen from Fig. 2a, ANFIS consists of five layers, and the basic functions of each layer are the input, fuzzification, rule inference, normalization and defuzzification. The relationship between the input and output of each layer is summarized as follows:

Layer 1 (Input nodes) Every node I in this layer is an adaptive node, and each node in this layer generates membership grades of crisp inputs variable which belong to each of convenient fuzzy sets by using the membership functions. Parameters in this layer are called premise parameters. The output of each node in this layer O_i^1 is defined by:

$$\begin{aligned} O_i^1 &= \mu_{A_i}(x) \quad \text{for } i = 1, 2 \text{ or} \\ O_i^1 &= \mu_{B_{i-2}}(y) \quad \text{for } i = 3, 4 \end{aligned} \quad (7)$$

where x and y are the crisp input to the i th node, and A_i and B_{i-2} are the linguistic labels characterized by MFs μ_{A_i} and

μ_{B_i} , respectively. Another major consideration that should be taken into account while designing ANFIS is the selection of proper MF. Several MFs were reported in the literature such as Gaussian, generalized bell-shaped, trapezoidal and triangular. Selection among the abovementioned MFs is generally based on the trial and error. Assuming a generalized bell-shaped as the MF, the node output in first layer O_i^1 can then be calculated by:

$$\begin{aligned} O_i^1 &= \mu_{A_i} = \frac{1}{1 + ((x - c_i)/a_i)^{2b_i}} \\ O_i^1 &= \mu_{B_{i-2}} = \frac{1}{1 + ((y - c_i)/a_i)^{2b_i}} \end{aligned} \quad (8)$$

where $\{a_i, b_i$ and $c_i\}$ are the parameters in generalized bell-shaped membership function that changes the shape of membership function with minimum and maximum equal to 0 and 1, respectively.

Layer 2 (Rule nodes) Every node in this layer is a fixed node labeled π , and the output is the product of all the incoming signals. Each node output that represents the firing strength of a rule is computed as.

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1 \text{ and } 2 \quad (9)$$

Layer 3 (Average nodes) In this layer, the nodes calculate the ratio of the i th rule's firing strength to the



sum of all rules' firing strengths; the firing strength in this layer is normalized and taken as $\bar{w}_i$:

$$O_i^3 = \bar{w}_i = \frac{w_i}{\sum w_i}, \quad i = 1 \text{ and } 2 \quad (10)$$

Layer 4 (Consequent nodes) This layer's nodes are adaptive with node functions,

$$O_i^4 = \bar{w}_i \cdot f_i = w_i(p_i x + q_i y + r_i), \quad i = 1 \text{ and } 2 \quad (11)$$

where $\bar{w}_i$ is the i th node's output from the previous layer and f_i is a linear function of input variables.

Layer 5 (Output nodes) This layer's single fixed node computes the final output as the summation of all incoming signals

$$O_i^5 = \frac{\sum \bar{w}_i \cdot f_i}{\sum \bar{w}_i} \quad (12)$$

Modeling performance criteria

The performance of ANN and ANFIS configurations was assessed based on three error measures, namely correlation coefficient, R^2 , which presents the degree of association between predicted and measured values; root mean square error, RMSE, which is preferred in many iterative prediction and optimization schemes; and mean absolute error, MAE, which is a parameter commonly understood in engineering applications. Expressions for these measures are given as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (13)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |O_i - P_i| \quad (14)$$

$$R^2 = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2}} \quad (15)$$

where N is the number of data, O observed values, P predicted values and the bar denotes the mean of the variable.

The primary purpose of data transformation is to modify the distribution of input variables so that they can match outputs better. The performance of a neural network is often improved through data transformations (Shi 2000). Therefore, in this research for development of ANN and ANFIS models, the raw data of both the independent and dependent variables were normalized to an interval by transformation by means of Eq. (16):

$$X_N = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (16)$$

where X_N is the normalized value of X , $X_{\max}$ and $X_{\min}$ represents the maximum and minimum value of each variable of original data, respectively; Eq. (16) makes data within the range of [0–1].

Results and discussion

MLP/BP

The learning of ANN model depends on the number of nodes in the hidden layer and transfers functions. Selection of an appropriate number of nodes in the hidden layer is a very important factor since a larger number of these nodes may result in over-fitting, while a smaller number of nodes may not capture the information adequately. The number of nodes in the hidden layer(s) can be determined by trial and error (Bateni et al. 2007; Eberhart and Dobbins 1990). Fletcher and Goss (1993) suggested that the appropriate number of nodes in a hidden layer ranges from $(2n^{1/2} + m)$ to (2^{n+1}) , where n is the number of input nodes and m is the number of output nodes. Therefore, in this study for the computation of DO, BOD and COD parameters, different MLP/BP models were constructed and tested in order to determine the optimum number of nodes in the hidden layer and transfer functions. The network was trained using the training data set, and then, it was validated with the testing data set. The optimal structure was determined by varying the number of hidden layers from 1 to 8 and picking the MLP/BP structure which led to the best results. The training of the ANN model was stopped when the number of iterations exceeded a prescribed maximum of 10,000. The learning rate of 0.01 was also used. Besides the number of hidden nodes, the transfer function between nodes also affects the performance of MLP/BP model. In order to investigate their effect, the model was run with different transfer functions including sigmoid ($f(x) = 1/(1 + \exp(-x))$), Gaussian ($f(x) = e^{-x \cdot x}$), hyperbolic tangent ($f(x) = \tanh(x)$) and hyperbolic secant ($f(x) = \text{Sech}(x)$). The optimal network size was selected from the one which resulted in minimum mean square error (MSE) and root mean square error (RMSE) and high regression coefficient (R^2) in training and testing data sets.

Comparison of results with the aforementioned transfer functions and different hidden nodes indicted for DO the MLP/BP model was composed of hyperbolic tangent transfer function, and three hidden layers gave the best results. Also for BOD and COD, MLP/BP model composed of hyperbolic tangent transfer function and three hidden layers and hyperbolic tangent transfer function and four hidden layers, respectively, gave the best results. The coefficient of determination (R^2), RMSE and MAE as computed for the best of three models (DO, BOD and COD) is presented in Table 2. Figure 3 shows the plots between measured and model computed values of DO, BOD and COD in training and testing data stages. The results showed that MLP/BP model had very small MAE and RMSE during training and testing stage. In the training stage, RMSE for DO, BOD and COD was 3.37, 3.99 and 3.27 mg L⁻¹, respectively. Also, in the testing stage, RMSE for DO, BOD and COD was 3.15, 2.57 and

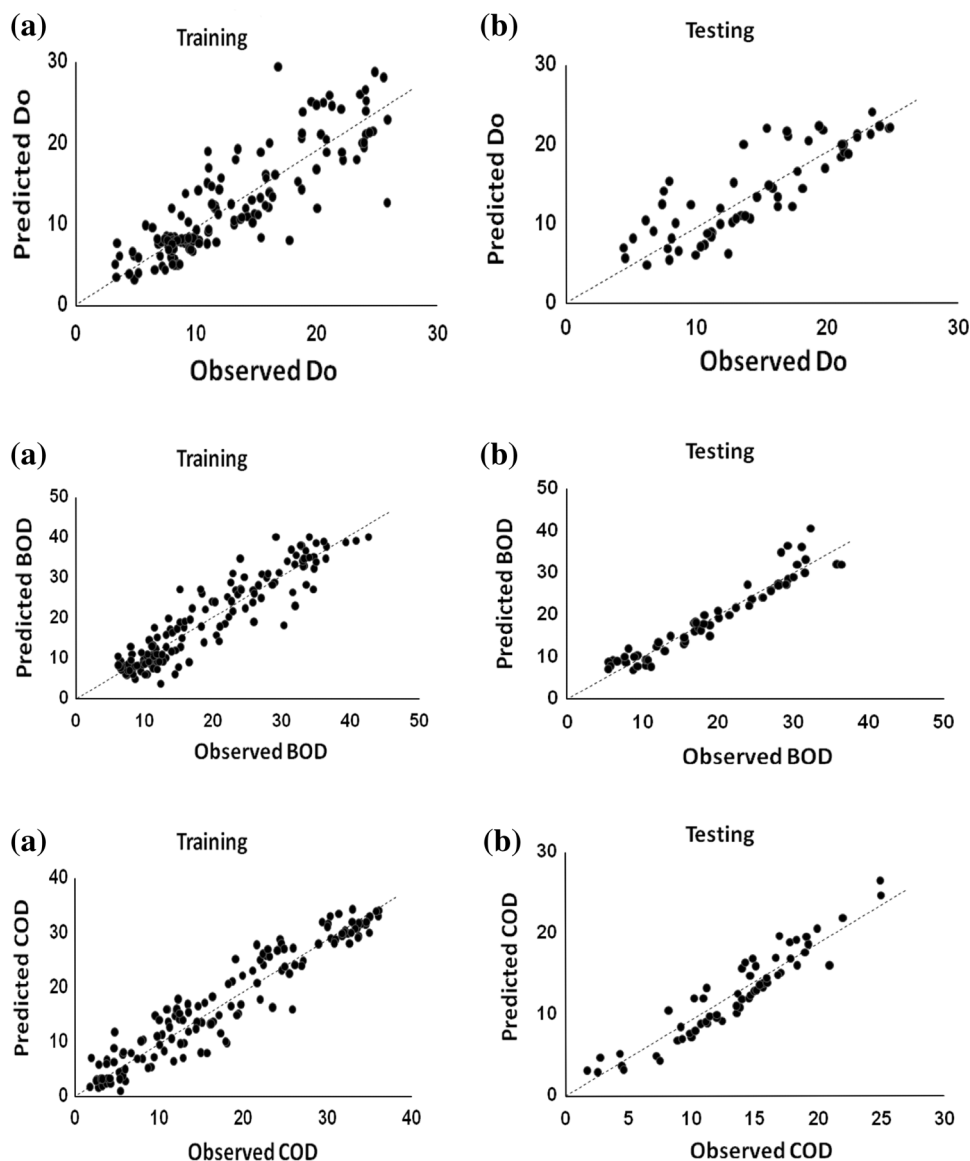


1.99 mg L⁻¹, respectively. Also, the results of MLP/BP model for BOD, DO and COD showed consistently good correlation throughout the training (>0.86) and testing (>0.85). Overall, with respect to the statistical parameters, the results showed that the MLP/BP model can predict the BOD, DO and COD parameters.

Table 2 Performance parameters of MLP model for computation of the BOD, DO and COD in Karoon River water (Iran)

ANN-structure	Variable	Stage	RMSE	MAE	R ²
MLP	DO	Training	3.37	2.6	0.86
		Testing	3.15	2.7	0.85
	BOD	Training	3.99	3.09	0.93
		Testing	2.57	2.05	0.96
	COD	Training	3.27	2.75	0.95
		Testing	1.99	1.77	0.94

Fig. 3 Scatter plots of observed and predicted DO (*top panel*), BOD (*middle panel*) and COD (*bottom panel*) using MLP/BP model: **a** training and **b** testing



RBF

In addition to MLP/BP, another type of ANN namely RBF was developed using the same input variables to predict DO, BOD and COD parameters. In the RBF model, the center selection process found an appropriate tolerance value of 0.005 and the radial basis spread of 40. To evaluate the performance of the RBF model, the predicted DO, BOD and COD were shown versus the observed ones in Fig. 4 for both training and testing data sets. Also R², RMSE and MAE as computed for the training and testing data are presented in Table 3. As the results of Fig. 4 and Table 3 showed, RBF had performed relatively well in predicting the DO concentrations in the training stage (MAE = 3.52 mg L⁻¹, RMSE = 4.38 mg L⁻¹ and R² = 0.75), but its performance was not good in the testing stage (MAE = 3.65 mg L⁻¹, RMSE = 4.29 mg L⁻¹ and



Fig. 4 Scatter plots of observed and predicted DO (*top panel*), BOD (*middle panel*) and COD (*bottom panel*) using ANFIS model: **a** training and **b** testing

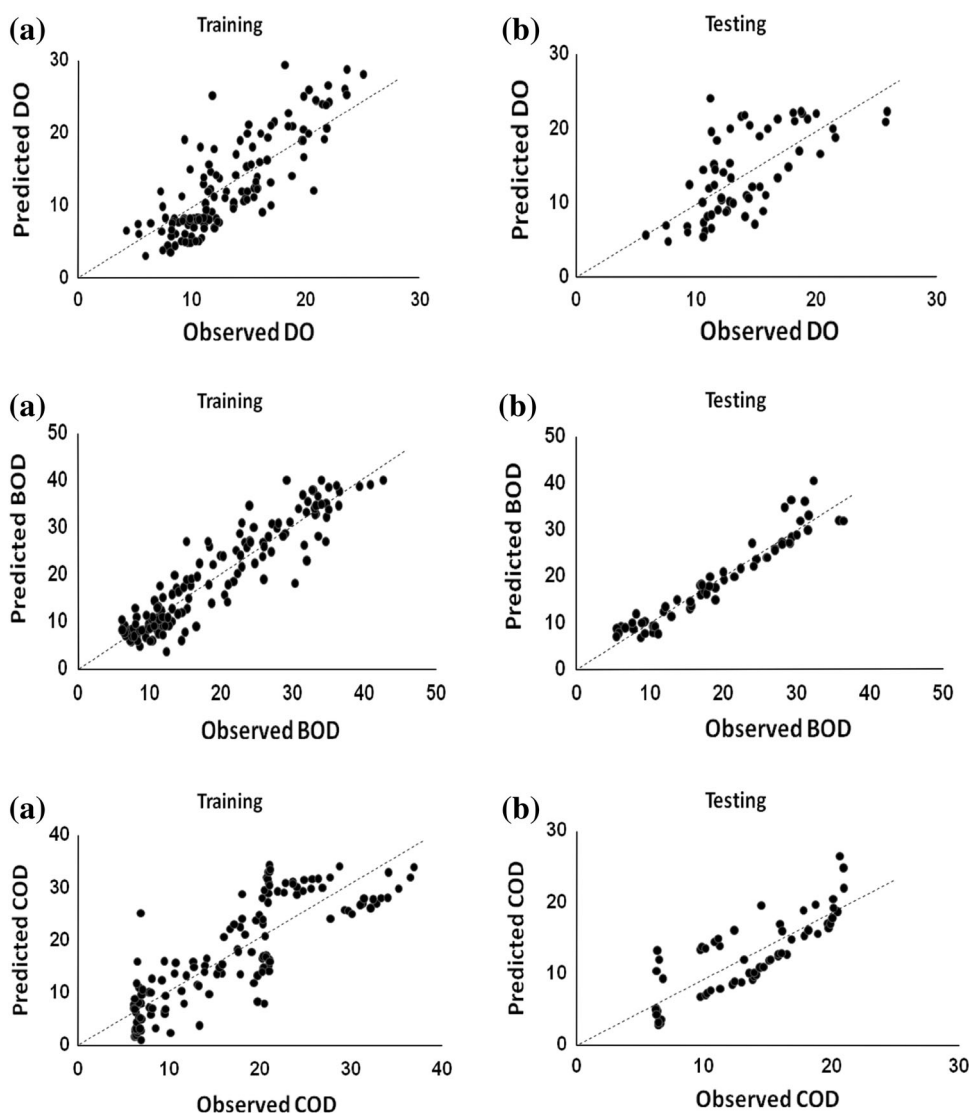


Table 3 Performance parameters of the RBF model for computation of the BOD, DO and COD in Karoon River water (Iran)

ANN-structure	Variable	Stage	MAE	RMSE	R^2
RBF	DO	Training	3.52	4.38	0.75
		Testing	3.65	4.29	0.66
	BOD	Training	4.58	5.75	0.85
		Testing	4.43	4.86	0.85
	COD	Training	4.62	5.57	0.84
		Testing	3.05	3.31	0.81

Table 4 Performance parameters of the RBF model for computation of the BOD, DO and COD in Karoon River water (Iran)

ANN-structure	Variable	Stage	MAE	RMSE	R^2
ANFIS	DO	Training	3.14	3.75	0.83
		Testing	3.65	4.07	0.71
	BOD	Training	4.11	5.7	0.89
		Testing	3.52	4.72	0.89
	COD	Training	3.75	4.21	0.92
		Testing	2.56	2.92	0.85

$R^2 = 0.66$). Also the model could well predict the BOD and COD concentrations with RMSE of 5.57 and 5.57 mg L^{-1} in the training stage, and 4.86 and 3.31 mg L^{-1} in the testing stage, respectively.

Anfis

For the ANFIS model, the Gaussian, triangular, trapezoidal, sigmoid and bell-shaped input membership functions were



used. Because the ANFIS only operates on Sugeno-type system, therefore, two types of constant and linear functions were used for output membership function. Moreover, the performances of the ANFIS model with two hybrid and back-propagation learning algorithms were investigated. It was found that, the best results were achieved when the input membership function was gaussian, the learning algorithm was hybrid and the output membership function was constant. To evaluate the performance of the ANFIS model, the predicted DO, BOD and COD were shown versus the observed ones in Fig. 5 and Table 4 for both training and testing stages. As the results showed, ANFIS performed well in predicting the DO parameter in the training stage ($\text{RMSE} = 3.75 \text{ mg L}^{-1}$ and $R^2 = 0.83$), but its performance was not good in the testing stage ($\text{RMSE} = 4.07 \text{ mg L}^{-1}$ and $R^2 = 0.71$). As shown in Fig. 5, ANFIS performed well in predicting the BOD both in training ($\text{RMSE} = 5.7 \text{ mg L}^{-1}$ and $R^2 = 0.89$) and testing stages ($\text{RMSE} = 4.72 \text{ mg L}^{-1}$ and $R^2 = 0.89$). Also, ANFIS performed well in predicting the COD both in training

($\text{RMSE} = 4.21 \text{ mg L}^{-1}$ and $R^2 = 0.92$) and testing stages ($\text{RMSE} = 2.92 \text{ mg L}^{-1}$ and $R^2 = 0.85$). Overall, the comparison of the results of ANFIS and RBF models with ANN revealed that the ANN performed better than ANFIS and RBF models in both training and testing.

Sensitivity analysis

To assign the relative significance of each of the independent parameters (input variables), on DO, BOD and COD (outputs), a sensitivity analysis was applied. As the MLP was better than ANFIS and RBF models, the analysis was conducted with the MLP model in the absence of every parameter in the data set. Table 5 gives the sensitivity analysis results for the parameters in Eqs. (1), (2) and (3), respectively. Although, the network does not necessarily represented physical meaning through the weights, it suggested that all the input variables had direct relevance with the dependent variables in water and played a significant role in determining the outputs, i.e., DO, BOD and

Fig. 5 Scatter plots of observed and predicted DO (*top panel*), BOD (*middle panel*) and COD (*bottom panel*) using ANFIS model: **a** training and **b** testing

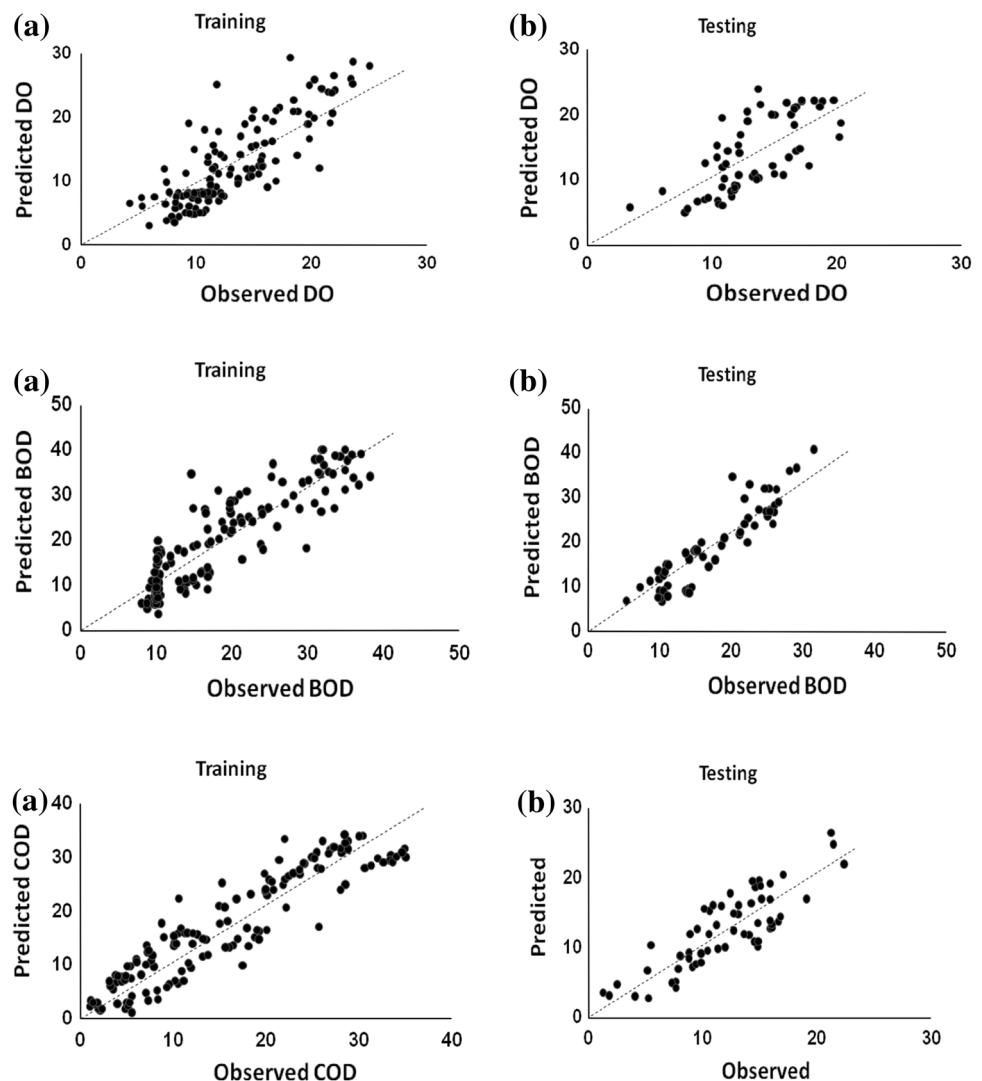


Table 5 Sensitivity analysis of the governing variables on DO, BOD and COD

Method	DO (mg L ⁻¹)			BOD (mg L ⁻¹)			COD (mg L ⁻¹)		
	MAE	RMS	R ²	MAE	RMS	R ²	MAE	RMS	R ²
The best ANN	2.70	3.40	0.85	3.00	3.75	0.94	2.55	2.9	0.95
ANN no EC	3.23	3.82	0.80	3.36	4.23	0.92	3.29	3.84	0.89
ANN no PH	3.05	3.77	0.80	3.31	4.19	0.92	3.09	3.69	0.89
ANN no NO ₂	3.90	4.40	0.74	6.90	9.02	0.71	7.51	8.16	0.74
ANN no NO ₃	3.50	4.17	0.75	5.05	7.13	0.80	4.25	4.84	0.86
ANN no Turbidity	3.60	4.20	0.76	4.11	4.46	0.90	3.66	4.1	0.87
ANN no Na	3.30	3.95	0.78	3.72	4.36	0.91	3.55	4	0.88
ANN no PO ₄	4.20	5.32	0.51	8.07	11.35	0.60	11.29	13.21	0.68
ANN no Mg	2.85	3.43	0.85	3.11	4.10	0.92	2.27	2.83	0.93
ANN no Ca	3.03	3.61	0.83	3.21	4.16	0.92	2.96	3.37	0.92

COD levels simultaneously (Basant et al. 2010). However, as the results of Table 5 showed, all the parameters had relatively a high contribution ($R^2 > 0.51$) to the networks, but phosphate (PO₄), nitrite (NO₂), nitrate (NO₃), turbidity and salinity (Na) have a greater impact on the model results. In other word, PO₄ was the most effective parameter on DO, BOD and COD. This finding is consistent with that of Talib and Amat (2012) and Singh et al. (2009). The second and third factors to contribute to the DO, BOD and COD values detected were NO₂ and NO₃ parameters. Increased levels of nutrients such as phosphate and nitrogen compounds (PO₄, NO₂, NO₃) in a water source can lead to increased rates of primary production by phytoplankton and organic plants, and therefore, the amount of organic matter increased in the water source. Levels of phosphates and nitrate that are intolerable to local organisms have been known to deplete dissolved oxygen levels by causing algae blooms. The COD value tends to be higher when the phosphate concentration is high. As phosphate takes important part in oxidization and energy-release process, there is a direct link between increase in phosphate concentrations and microorganisms, thus, higher growth of microorganisms (Talib and Amat 2012). The turbidity is an important parameter in the water quality assessment because the degree of turbidity of stream water is often taken to be an approximate measure of the intensity of the pollution. Indeed, this parameter can be used to evaluate the effects of pollution by waste waters and even to follow the course of self-purification of streams (Klein et al. 1962). The sensitivity analysis shows that turbidity has effect on DO, BOD and COD values. Turbidity can impact the level of dissolved oxygen in water in multiple ways. Increased light absorbency by turbid water can lead to increased water temperatures and decreased oxygen levels. The contents of the particles causing turbidity can also lead to a decrease in dissolved oxygen levels (Zhang and Li 2010). The various substances making up turbidity causing particles can have a high biological oxygen demand (BOD), which is a measure of the amount of oxygen

used to decompose a substance. Increased turbidity is therefore associated with increased BOD and decreased oxygen levels (Singh et al. 2009). Salinity also has an impact on the model results. From chemical aspects, when an ionic salt like NaCl is added to water, the ions from the salt introduced will attract the water molecules in an effort to “solvate” the ions. This has the tendency to decrease the weak affinity of non-polar oxygen molecules to water and drive the dissolved oxygen out of the polar water. In general, the solubility of a gas in a solvent is affected significantly by the presence of other solutes in the solution (Comer 1996).

Conclusion

In this study, the applications of two types of ANNs, namely MLP/BP and RBF along with ANFIS, were used to estimate the DO, BOD and COD parameters. The measured data of Karoon River (Iran) over a period of 17 years were collected for networks training and testing. The results showed that a single-layer MLP/BP with 12 neurons, a single-layer RBF with 25 neurons and an ANFIS model with 3 membership functions were considered as the best and the most executive models. Also, the comparison of RMSE, MAE indices and coefficient of correlations (R^2) for predicting DO, BOD and COD showed that the performance of the MLP/BP model was higher than those of ANN-RBF and ANFIS models. The R^2 values for a good fit of the DO, BOD and COD models to the data set were obtained 0.86, 0.93 and 0.95 for the training stage and 0.85, 0.96 and 0.94 for testing, respectively (Table 4). Based on the results of this research, the ANN can be used as an effective tool to compute and predict the river water quality parameters.

Acknowledgments The authors would like to thank Khozestan Water and Power Authority (KWPA) of Iran for the provided data.



References

- ASCE Task Committee (2000a) Artificial neural networks in hydrology. I: preliminary concepts. *J Hydrol Eng* 5(2):115–123. doi:[10.1061/\(ASCE\)1084-0699\(2000\)5:2\(115\)](https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(115))
- ASCE Task Committee (2000b) Artificial neural networks in hydrology. II: hydrologic applications. *J Hydrol Eng* 5(2):124–137. doi:[10.1061/\(ASCE\)1084-0699\(2000\)5:2\(124\)](https://doi.org/10.1061/(ASCE)1084-0699(2000)5:2(124))
- Badripour H (2006) Country pasture/forage resource profiles. FAO, Rome
- Bandyopadhyay G, Chattopadhyay S (2007) Single hidden layer artificial neural network models versus multiple linear regression model in forecasting the time series of total ozone. *Int J Environ Sci Technol* 4(1):141–149
- Basant N, Gupta S, Malik A, Singh KP (2010) Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water—A case study. *Chemometr Intell Lab Syst* 104:172–180
- Bateni SM, Borghei SM, Jeng DS (2007) Neural network and neuro-fuzzy assessments for scour depth around bridge piers. *Eng Appl Artif Intell* 20:401–414
- Broomhead D, Lowe D (1988) Multivariable functional interpolation and adaptive networks. *Complex Syst* 2:321–355
- Chang FJ, Chen YC (2003) Estuary water-stage forecasting by using radial basis function neural network. *J Hydrol* 270:158–166
- Chau KW (2006) A review on integration of artificial intelligence into water quality modeling. *Mar Pollut Bull* 52:726–733
- Chaves P, Kojiri T (2007) Conceptual fuzzy neural network model for water quality simulation. *Hydrol Process* 21:634–646
- Comer PJ (1996) Wetland trends in Michigan since 1800: a preliminary assessment. United States. Environmental Protection Agency, Michigan. Land and Water Management Division. Michigan Natural Features Inventory, Lansing
- Cox BA (2003) A review of dissolve oxygen modeling techniques for lowland rivers. *The Sci Total Environ* 314–316:303–334
- Eberhart RC, Dobbins RW (1990) Neural network PS tools: a practical guide. Academic press, San Diego
- Faruk DO (2010) A hybrid neural network and ARIMA model for water quality time series prediction. *Eng Appl Artif Intell* 23:586–594
- Fausett L (1994) Fundamentals of neural networks: architectures, algorithms, and applications. Prentice-Hall, Englewood Cliffs, NJ
- Fernando AK, Jayawardena AW (1998) Runoff forecasting using RBF networks with OLS algorithm. *J Hydrol Eng ASCE* 3:203–209
- Fletcher D, Goss E (1993) Forecasting with neural networks: an application using bankruptcy data. *Inf Manag* 24:159–167
- Haykin S (1999) Neural network a comprehensive foundation. Prentice-Hall, New Jersey
- Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23:665–685
- Jang JSR, Sun CT, Mizutani E (1997) Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice Hall, Upper Saddle River, NJ
- Karamouz M, Mahjouri N, Kerachian R (2004) River water quality zoning: a case study of Karoon and Dez River system. *Iran J Environ Health Sci Eng* 1:16–27
- Kisi O (2007) Stream flow forecasting using different artificial neural network algorithms. *J Hydrol Eng ASCE* 12:532–553
- Klein L, Erichsen Jones JR, Hawkes HA, Downing AL (1962) River pollution, causes and effects II. Butterworths and Co (Publishers) Ltd., London, p 456
- Kuo JT, Hsieh MH, Lung WS, Shed N (2007) Using artificial neural network for reservoir eutrophication prediction. *Ecol Model* 200:171–177
- Kuok KK, Harun S, Shamsuddin SM (2009) Particle swarm optimization feedforward neural network for modeling runoff. *Int J Environ Sci Technol* 7(1):67–78
- Leonardis A, Bischof H (1998) An efficient MDL-based construction of RBF networks. *Neural Netw* 11:963–973
- Matlab User Manual (2008) Fuzzy Logic Toolbox software with Matlab
- May RJ, Maier HR, Dandy GC, Fernando TMKG (2008) Non-linear variable selection for artificial neural networks using partial mutual information. *Environ Model Softw* 23(10–11):1312–1326
- McCulloch WS, Pitts W (1943) A logical calculus of the ideas imminent in nervous activity. *Bull Math Biophys* 5:115–133
- Modarres R, De Paulo Rodrigues da Silva V (2007) Rainfall trends in arid and semi-arid regions of Iran. *J Arid Environ* 70:344–355
- Najah A, Elshafie A, Karim OA, Jaffar O (2009) Prediction of Johor River water quality parameters using artificial neural networks. *Eur J Sci Res* 28:422–435
- Orr JL (1996) Introduction to radial basis function networks. Technical report, Institute of Adaptive and Neural Computation. Edinburgh University. <http://www.anc.ed.ac.uk/mjo/papers/intro.ps>
- Palani S, Liong SY, Tkachik P (2008) An ANN application for water quality forecasting. *Mar Pollut Bull* 56:1586–1597
- Qnet 2000 (1999) Qnet 2000 neural network modelling for windows 95/98/NT, QnetToll User's Guide and Datapro User's Guide, Vesta Services, Inc., USA
- Ranković V, Radulović J, Radojević I, Ostojić A, Čomić L (2012) Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system. *J Hydroinform* 14(1):167–179
- Raziei T, Daneshkar Arasteh P, Saghfian B (2005) Annual rainfall trend in arid and semi-arid regions of Iran. ICID 21st European Regional Conference 2005. 15–19 May 2005
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representation by error back propagation. In: Rumelhart DE, McClelland JL (eds) Parallel distributed processing. MIT Press, Cambridge, MA, pp 318–362
- Shi IJ (2000) Reducing prediction error by transforming input data for neural networks. *J Comput Civil Eng* 14:109–116
- Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality, a case study. *Ecol Model* 220:888–895
- Soyupak S, Karaer F, Gürbüz H, Kivrak E, Sentürk E, Yazici A (2003) A neural network-based approach for calculating dissolved oxygen profiles in reservoirs. *Neural Comput Appl* 12:166–172
- Talib A, Amat MI (2012) Prediction of chemical oxygen demand in Dondang River using artificial neural network. *Int J Inf Educat Technol* 2(3):259–261
- Venkatesan P, Anitha S (2006) Application of a radial basis function neural network for diagnosis of diabetes mellitus. *Curr Sci* 91:1195–1199
- Verma AK, Singh TN (2013) Prediction of water quality from simple field parameters. *Environ Earth Sci* 69(3):821–829. doi:[10.1007/s12665-012-1967-6](https://doi.org/10.1007/s12665-012-1967-6)
- Wang J, Sui J, Guo L, Karney BW, Jüpner R (2010) Forecast of water level and ice jam thickness using the back propagation neural network and support vector machine methods. *Int J Environ Sci Technol* 7(2):215–224
- Wasserman PD (1989) Neural computing: theory and practice. Van Nostrand Reinhold, New York
- Xiang SL, Liu ZM, Ma LP (2006) Study of multivariate linear regression analysis model for ground water quality prediction. *Guizhou Sci* 24:60–62
- Yilmaz I, Kaynar O (2011) Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Syst Appl* 38:5958–5966
- Zhang H, Li S (2010) Effects of physical and biochemical processes on the dissolved oxygen budget for the Pearl River Estuary during summer. *J Mar Syst* 79(2):65–88

