

Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China

Xiaohu Wen · Jing Fang · Meina Diao ·
Chuanqi Zhang

Received: 15 February 2012 / Accepted: 6 September 2012 / Published online: 22 September 2012
© Springer Science+Business Media B.V. 2012

Abstract Identification and quantification of dissolved oxygen (DO) profiles of river is one of the primary concerns for water resources managers. In this research, an artificial neural network (ANN) was developed to simulate the DO concentrations in the Heihe River, Northwestern China. A three-layer back-propagation ANN was used with the Bayesian regularization training algorithm. The input variables of the neural network were pH, electrical conductivity, chloride (Cl^-), calcium (Ca^{2+}), total alkalinity, total hardness, nitrate nitrogen

($\text{NO}_3\text{-N}$), and ammonical nitrogen ($\text{NH}_4\text{-N}$). The ANN structure with 14 hidden neurons obtained the best selection. By making comparison between the results of the ANN model and the measured data on the basis of correlation coefficient (r) and root mean square error (RMSE), a good model-fitting DO values indicated the effectiveness of neural network model. It is found that the coefficient of correlation (r) values for the training, validation, and test sets were 0.9654, 0.9841, and 0.9680, respectively, and the respective values of RMSE for the training, validation, and test sets were 0.4272, 0.3667, and 0.4570, respectively. Sensitivity analysis was used to determine the influence of input variables on the dependent variable. The most effective inputs were determined as pH, $\text{NO}_3\text{-N}$, $\text{NH}_4\text{-N}$, and Ca^{2+} . Cl^- was found to be least effective variables on the proposed model. The identified ANN model can be used to simulate the water quality parameters.

X. Wen (✉) · M. Diao · C. Zhang
Key Laboratory of Coastal Zone Environmental Processes,
Yantai Institute of Coastal Zone Research, Chinese
Academy of Sciences,
Chunhui Rd 17,
Yantai 264003 Shandong Province, China
e-mail: xhwen@yic.ac.cn

X. Wen · M. Diao · C. Zhang
Shandong Provincial Key Laboratory of Coastal Zone
Environmental Processes, Yantai Institute of Coastal Zone
Research, Chinese Academy of Sciences,
Chunhui Rd 17,
Yantai 264003 Shandong Province, China

J. Fang
Cold and Arid Regions Environmental and Engineering
Research Institute, Chinese Academy of Sciences,
West Donggang Rd 320,
Lanzhou 730000 Gansu Province, China

M. Diao · C. Zhang
University of Chinese Academy of Sciences,
19A Yuquanlu,
100049 Beijing, China

Keywords Artificial neural network · Dissolved oxygen · Modeling · Heihe River

Introduction

The surface water quality is one of the major issues today because of its effects on human health and aquatic ecosystems. With the increase in population, there is increasing pressure on water resources. Surface water quality in a region is largely determined both by natural processes including the lithology of the basin, atmospheric inputs, and climatic conditions, and by anthropogenic inputs

such as municipal and industrial wastewater discharge. On the other hand, rivers play a major role in assimilating or transporting municipal and industrial wastewater and runoff from agricultural land. Municipal and industrial wastewater constitutes a constant pollution source, whereas surface runoff is a seasonal phenomenon, largely affected by climate within the basin (Singh et al. 2004).

Dissolved oxygen (DO) is one of the important water quality parameters of an aquatic ecosystem and a significant status indicator for the aquatic ecosystems. The sources of DO in a water body include reaeration 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). Identification and quantification of DO profiles of river is one of the primary concerns for water resources managers.

Several DO models such as deterministic and stochastic models have been developed in order to manage the best practices for conserving the DO in water bodies (Ansa-Ansare et al. 2000; García et al. 2002; Wang et al. 2003; Hull et al. 2008; Shukla et al. 2008). Most of these models are complex and need several different input data which are not easily accessible, making it a very expensive and time-consuming process (Suen et al. 2003). Artificial neural networks (ANNs) are flexible modeling tools with the capability of learning the mathematical mapping between input and output variables of nonlinear systems and generalizing the processes of control, classification, and prediction. They are capable of providing a neuron computing approach to solve complex problems. In the last decade, ANNs have been widely successfully applied to various water resources problems, such as hydrological processes (Nayak et al. 2004; Sahoo et al. 2005; Dastorani et al. 2010; Guo et al. 2011; Wu and Chau 2011; Senkal et al. 2012), water resources management (Kralisch et al. 2003; Sreekanth and Datta 2010), groundwater problems (Daliakopoulos et al. 2005; Dixon 2005; Garcia and Shigidi 2006; Nayak et al. 2006; Ghose et al. 2010; Banerjee et al. 2011), and water quality (Ha and Stenstrom 2003; Kuo et al. 2006; Anctil et al. 2009; da Costa et al. 2009; Dogan et al. 2009; Chang et al. 2010; He et al. 2011). ANNs also have been used for modeling and forecasting DO (Kuo et al. 2007; Singh et al. 2009; Ranković et al. 2010; Najah et al. 2011). Furthermore, some intelligence algorithm such as genetic algorithm

was used with ANN model for the management of the watershed water quality problem (Kuo et al. 2006). ANNs, as effective tool for the computation of water quality, can be regarded as a powerful predictive alternative to traditional modeling techniques.

In the arid northwest of China, water resources play a dominant role in the development of the economy. Careful management is important for ecological and environmental protection (Wen et al. 2007). Due to extensive use of surface water, the quality of the surface water has also impacted for the last few decades (Wang et al. 1999). Predicting the water quality evolution of surface water in these arid regions can enhance understanding for river water systems and help decision makers effectively manage water resources. The main purpose of this study is to analyze and discuss the performances of ANNs in modeling of DO in the Heihe River.

Material and methods

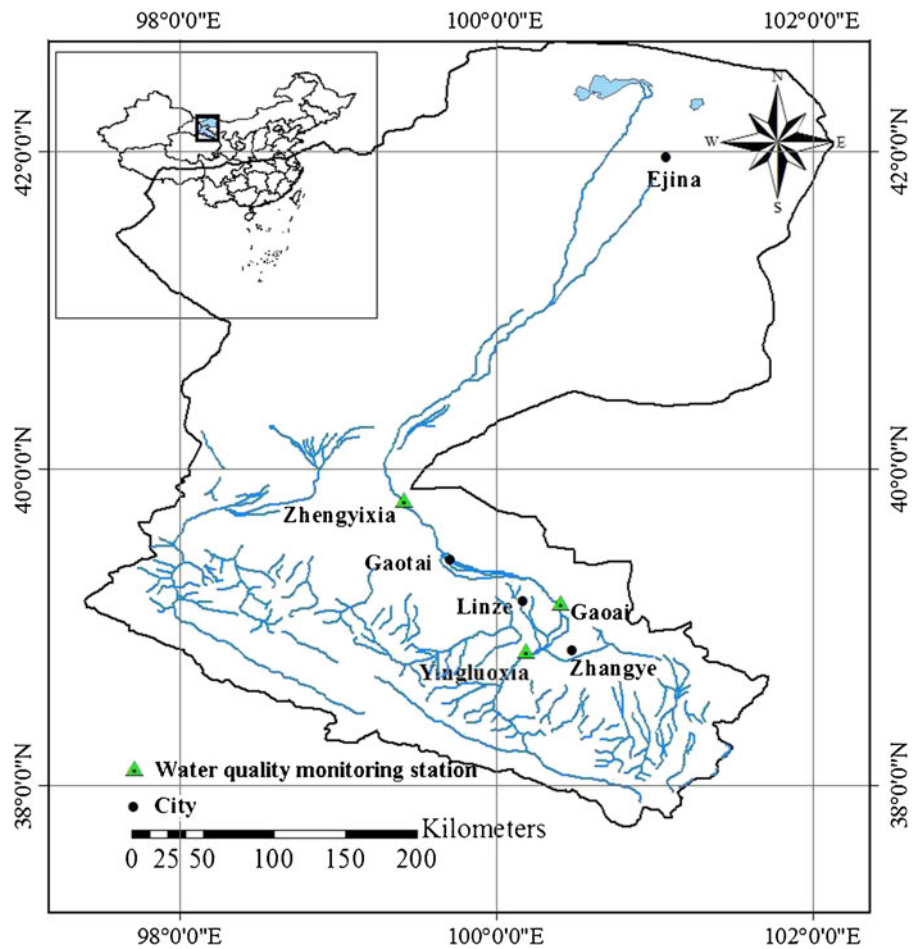
Study area

The Heihe River in northwestern China is one of the largest inland rivers in China, covering an area of $1.3 \times 10^5 \text{ km}^2$. It originates from the Qilian Mountains, flowing through the Zhangye basin and the lower reaches (also known as the Ejina Basin) (Fig. 1). The middle reaches of Heihe River, from the mountain outlet (Yingluoxia) to the end of the middle reaches of the Heihe River (Zhengyixia), is 185 km in length with an average slope of 2 %, covering an area of $1.08 \times 10^4 \text{ km}^2$, including Zhangye City, Linze County, and Gaotai County (Fig. 1). This area has an arid continental climate with a mean annual temperature of 3–7 °C. The average annual precipitation ranges from 50 to 150 mm, with the majority (~80 %) falling from June to September. The average annual potential evaporation is 2,000–2,200 mm (Gao 1991).

Water sampling procedure

The data collected from three water quality monitoring stations, including Yingluoxia, Gaoai, and Zhengyixia (Fig. 1). The Yingluoxia is located in the entrance of the middle reaches of Heihe River, and the stations are situated at upstream sites of the study area. This station receives pollution from nonpoint sources, i.e., mostly from agricultural activities with relatively low

Fig. 1 Location of the study area and the water quality monitoring station



river pollution; the Gaoai and Zhengyixia are located in the central part and the end of the middle reaches of Heihe River, respectively, representing the high pollution of the river. These stations are situated at downstream sites of the study area. These stations receive pollution from point and nonpoint sources, i.e., agricultural and livestock farms, domestic wastewater, and surface runoff from villages. The river quality was monitored monthly at three different sites over 6 years (2003–2008) comprising nine water quality parameters. Although more than 20 water quality parameters were available, only nine parameters were selected due to their continuity in measurement at all selected water quality monitoring stations. The selected water quality parameters included pH, electrical conductivity (EC, microsiemens per centimeter), chloride (Cl^- , milligrams per liter), calcium (Ca^{2+} , milligrams per liter), total alkalinity (TA, milligrams per liter), total hardness (TH, milligrams per liter), nitrate nitrogen ($\text{NO}_3\text{-N}$, milligrams per liter), ammoniacal nitrogen

($\text{NH}_4\text{-N}$, milligrams per liter), and dissolved oxygen (DO, milligrams per liter). DO, EC, and pH were measured in the field by the portable multi-parameter water quality analyzer, and the analytical precision of DO was within $\pm 2\%$. Other water quality parameters were determined using Standard Methods (APHA 1995).

The independent water quality parameters showed a coefficient of variation between 2.69 and 151.17% (Table 1). Such variability among the samples may be attributed to the large geographical variations in climate and seasonal influences in the study area. Parameter pH showed lowest variation. Compared to the natural origin parameters, water quality parameters of anthropogenic origin showed high variations due to the buffering capacity of the river. The correlation coefficient between DO and the input parameters was calculated and presented in Table 1.

The available data are generally divided into training, validation, and testing subsets to develop an ANN model. The training set is used to estimate the

Table 1 Basic statistics of the measured water quality parameters in Heihe River

	Unit	Min	Max	Average	SD	CV (%)	Correlation
pH	–	7.40	8.70	8.18	0.22	2.69	0.26
EC	μS/cm	374.00	3050.00	890.60	613.94	68.94	–0.25
Ca ²⁺	mg/l	31.90	127.00	60.61	12.87	21.23	0.23
Cl [–]	mg/l	3.94	387.00	55.97	84.61	151.17	–0.27
TH	mg/l	183.00	838.00	333.09	145.64	43.72	–0.18
TA	mg/l	138.00	337.00	204.33	40.93	20.03	–0.08
NH ₄ -N	mg/l	0.00	4.52	0.62	0.83	133.87	–0.28
NO ₃ -N	mg/l	0.08	3.28	1.15	0.64	55.65	0.10
DO	mg/l	3.60	11.80	7.84	1.72	21.94	1.00

SD standard deviation, *CV* coefficient of variation

unknown connection weights; the validation set is used to decide when to stop training in order to avoid overfitting and/or which network structure is optimal; and the test set is used to assess the generalization ability of the trained model (Maier et al. 2010). In this study, the complete river water quality data set (164 samples × 8 variables) was randomly divided into three sections including training, validation, and test sets comprised of 100 (60 %), 32 (20 %), and 32 (20 %) samples, respectively. The output variables (DO) corresponding to the input variables belonged to the same water sample, thus measured in the same time and space.

In view of the requirements of the neural computation algorithm, the raw data of both the input and output variables were normalized to an interval by transformation. All the variables were normalized ranging from –1 to 1 as follow equation:

$$x_n = 2 \times \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} - 1 \quad (1)$$

where x_n and x_i represent the normalized and original training, test, and validation data; x_{\min} and x_{\max} denote the minimum and maximum of the training, test, and validation data.

Artificial neural network modeling

An artificial neural network (ANN) is a mathematical structure designed to mimic the information processing functions of a network of neurons in the brain (Hinton 1992; Jensen 1994). ANNs are highly parallel systems that process information through many interconnected units that respond to inputs through modifiable weights, thresholds, and mathematical transfer functions. Each unit processes the pattern of activity it receives from

other units and then broadcasts its response to still other units. ANNs are particularly well suited for problems in which large data sets contain complicated nonlinear relations among many different inputs. ANNs are able to find and identify complex patterns in data sets that may not be well described by a set of known processes or simple mathematical formulae.

Multilayer perceptron neural network

Among the various types of ANNs that have been developed over the years, the multilayer perceptron (MLP) neural network structure is the most commonly used and well-researched class of ANNs (Ouarda and Shu 2009). A feed forward MLP network consists of an input layer which receives the values of the input variables, an output layer which provides the model output, and one or more hidden layers. Nodes in each layer are interconnected through weighted acyclic arcs from each preceding layer to the following, without lateral or feedback connections (Shu and Ouarda 2007). Principe et al. (2000) emphasize that the main advantage is in being easy to use, and the key disadvantages are that they train slowly and require a large amount of training data, and easily to get stuck in a local minimum. However, MLP with a sufficient number of hidden units can approximate any continuous function to a prespecified accuracy; in other words, MLP networks are universal approximations (Cherkassky and Mulier 1998).

Back-propagation neural network and leaning algorithm

It has been well recognized that a neural network with one hidden layer is capable of approximating any finite

nonlinear function with high accuracy and three more hidden layered systems are known to cause unnecessary computational overload (Kim and Gilley 2008). Hence, an MLP network with one hidden layer trained by back-propagation (BP) neural network was used to build the ANN model for modeling of the river water DO with eight input variables as shown in Fig. 2. The activation function consists of a tan-sigmoid function in the hidden layer and a linear function in the output layer. The mathematical expression of the MLP is as follows:

$$\bar{x}_j = \sum_i w_{ij}x_i + w_j \tag{2}$$

$$x_j = f(\bar{x}_j) = \frac{1}{1 + e^{-\bar{x}_j}} \tag{3}$$

where x_i is the output of node i located in any one of the previous layers, w_{ij} the weight associated with the link connecting nodes i and j , and w_j the bias of node j .

Since the weights w_{ij} are actually internal parameters associated with each node i , changing the weights of a node will alter the behavior of the node and in turn alter the behavior of the whole back-propagation MLP. First, a squared error measure for the p th input–output pair is defined as:

$$E_p = \sum_k (d_k - x_k)^2 \tag{4}$$

$$e_j = \begin{cases} -2(d_j - x_j) \frac{\partial x_j}{\partial \bar{x}_j} = -2(d_j - x_j)x_j(1 - x_j) & \text{If node } j \text{ is a output node} \\ \frac{\partial x_j}{\partial \bar{x}_j} \sum_{k_j < k} \frac{\partial E_p}{\partial \bar{x}_k} \frac{\partial \bar{x}_k}{\partial x_j} = x_j(1 - x_j) \sum_{k_j < k} e_k w_{jk} & \text{Otherwise} \end{cases} \tag{6}$$

where w_{jk} is the connection weight from node j to k and w_{jk} is zero if there is no direct connection. Then, the weight update Δw_{jk} for off-line learning is:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \sum_p \frac{\partial E_p}{\partial w_{jk}} \tag{7}$$

where η is a learning rate that affects the convergence speed and stability of the weights during learning. In vector form,

$$\Delta w = -\eta \nabla_w E \tag{8}$$

where $E = \sum_p E_p$. This corresponds to a way of using the true gradient direction based on the entire data set.

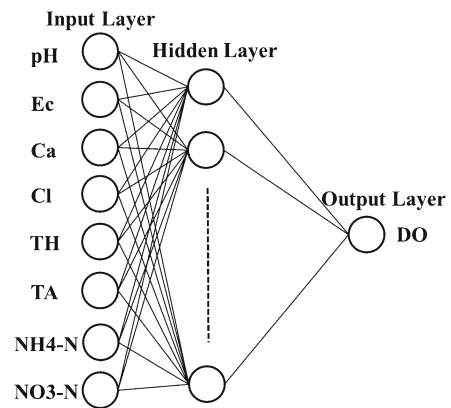


Fig. 2 General conceptual neural network for the DO in the Heihe River

where d_k is the desired output for node k , and x_k is the actual output for node k when the input part of the p th data pair is presented. To find the gradient vector, an error term e_j for node i is defined as:

$$e_j = \frac{\partial E_p}{\partial \bar{x}_j} = \frac{\partial \sum_k (d_k - x_k)^2}{\partial \bar{x}_j} \tag{5}$$

By the chain rule, the recursive formula for e_j can be written as:

The way we adapt to speed-up training is to use the momentum term:

$$\Delta w = -\eta \nabla_w E + \alpha \Delta w_{prev} \tag{9}$$

where Δw_{prev} is the previous update amount and α is the momentum constant. As for the detail of the backpropagation MLP, interested readers can refer to any literatures addressing neural network theory for more information (Freeman and Skapura 1991; Jang et al. 1997; Kuo et al. 2006).

There are several optimization methods to improve the convergence speed and the performance of network training. In this paper, the Bayesian regularization BP algorithm was selected. The Bayesian regularization is an algorithm that automatically sets optimum values for

the parameters of the objective function. In the approach used, the weights and biases of the network are assumed to be random variables with specified distributions. To estimate regularization parameters which are related to the unknown variances, statistical techniques are used. The advantage of this algorithm is that whatever the size of the network, the function will not be over-fitted. Bayesian regularization has been effectively used (Porter et al. 2000; Coulibaly et al. 2001a, b; Anctil et al. 2004; Krishna et al. 2008). A more detailed discussion of the Bayesian regularization can be found in the literature (MacKay 1992).

Determining the number of neurons in the hidden layer is an important task when designing an ANN (Shu and Ouarda 2007). Too many hidden nodes may lead to the problem of overfitting. Too few nodes in the hidden layer may cause the problem of underfitting. The appropriate number of nodes in a hidden layer was recommend ranging from $(2n^{1/2}+m)$ to $(2n+1)$, where n was the number of input nodes and m is the number of output nodes (Fletcher and Goss 1993). In this study, a trial and error procedure for the hidden node selection was carried out by gradually varying the number of nodes in the hidden layer.

During the training processes, there are three factors that are associated with the weight optimization algorithms. These are: (1) initial weight matrix, (2) learning rate, and (3) stopping criteria such as (a) fixing the number of epoch size, (b) setting a target error goal, and (c) fixing minimum performance gradient. The initial weights are randomly generated between -1 and 1 with a random number generator. The value of the learning parameter is not fixed. Maier and Dandy (1998, 2000) reported that optimization of learning parameter was highly problem dependent and should be selected so that oscillation in error surface can be avoided. Hagan et al. (1996) demonstrated that the learning became unstable for higher values (>0.035). Thus, the learning rate was set as 0.01 .

The mean square error (MSE) can be used to determine how well the network output fits the desired output. MSE, the smaller values ensuring the better performance, is defined as follows

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2 \quad (10)$$

where n is the number of input samples, and O_i and P_i are the measured and network output value from the

i th elements, respectively. The maximum numbers of epochs, target error goal MSE, and the minimum performance gradient were set as 10^5 , 10^{-5} , and 10^{-5} , respectively. Training stops when any of these conditions occur. All the computations were performed using MATLAB software (MathWorks, Inc., Natwick, MA).

Statistical forecasting of ANN model

The performance of developed models can be evaluated using several statistical tests that describe the errors associated with the model. The MSE, the coefficient of correlation (r), and the root mean square error (RMSE) were used to provide an indication of goodness of fit between the measured and modeled values.

Coefficient of correlation is defined as the degree of correlation between the measured and modeled values:

$$r = \frac{\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2 \sum_{i=1}^n (O_i - \bar{O})^2}} \quad (11)$$

The RMSE can be calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (12)$$

where n is the number of input samples; and O_i and P_i are the measured and network output value from the i th elements, respectively. \bar{O} and \bar{P} are their average, respectively.

Results and discussion

DO model result

The optimum number of neurons was determined based on the minimum MSE value of the training data set. The training of the BP MLP-NN was performed with a variation of 5–17 neurons. Each architecture configuration was trained 50 times with different initializations, and then, the best network was retrained to calculate the overall accuracy. Determined by the relationship between the numbers of neurons versus MSE during training, the MSE value decreased to

Table 2 Performance parameters of the artificial neural network model

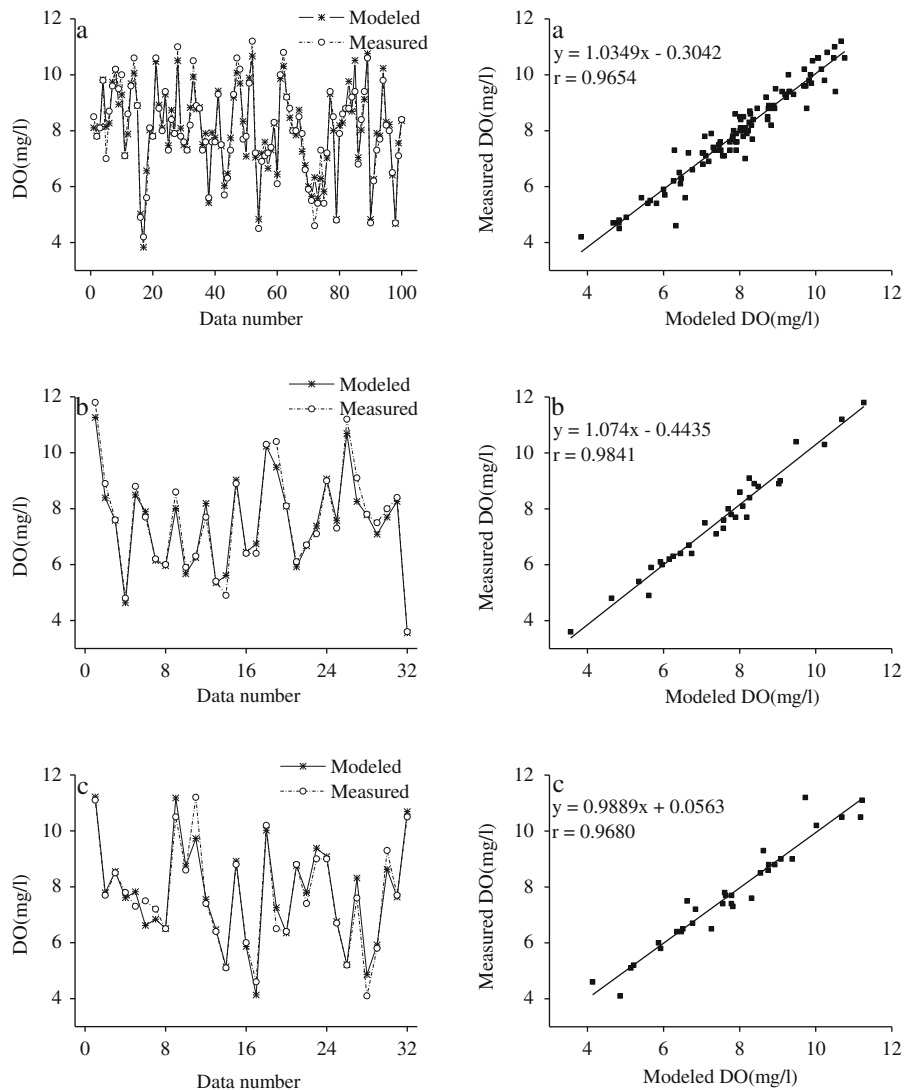
Model	Structure		RMSE	<i>r</i>
DO	8-14-1	Training	0.4272	0.9654
		Validation	0.3667	0.9841
		Test	0.4570	0.9680

0.1825 when 14 neurons were used. Thus, 14 neurons were selected as the best number of neurons.

The selected ANN for the DO model was composed of one input layer with eight input variables, one hidden layer with 14 neurons, and one output layer with one output variable. The coefficient of

correlation (*r*) and RMSE were computed for the training. Validation and test data sets used for the DO model were presented in Table 2. Figure 3 showed the fittings between measured and modeled values of DO in training, validation, and testing sets. The coefficient of correlation (*r*) values for the training, validation, and test sets were 0.9654, 0.9841, and 0.9680, respectively. The respective values of RMSE for the training, validation, and test sets were 0.4272, 0.3667, and 0.4570, respectively. A closely followed pattern of variation by the measured and modeled DO concentrations in the Heihe River was shown in Fig. 2, with coefficient of correlation (*r*) and RMSE values suggesting a good-fit of the DO model to the data set.

Fig. 3 Comparison of the measured and modeled DO values in **a** training, **b** validation, and **c** testing sets



Sensitivity analysis

To evaluate the effect of input variables on the DO model, two evaluation processes were used. Firstly, the performance evaluation of various combinations of the parameters was investigated utilizing the coefficient of correlation (r) and RMSE approaches to determine the most effective variables on the output. The optimal network architecture of the various combinations of the parameters was selected based on the one with minimum of MSE using the 14 neurons. Overall, nine networks were compared as shown in Table 3. Each one demonstrated the extents which the eliminated variable would affect the network accuracy. Apparently, the precision of model became higher if Cl^- was eliminated from the input variables to the model, where minimum RMSE and coefficient of correlation (r) were determined to be 0.4712 and 0.9691 for the testing data set, respectively. Therefore, Cl^- could be excluded. Conversely, the coefficient of correlation (r) reduced if other input parameters was removed, which reduced the ability of ANN in the capability modeling. Furthermore, DO was found to be sensitive to the pH, $\text{NH}_4\text{-N}$, and $\text{NO}_3\text{-N}$ variables.

Secondly, the neural net weight matrix was used to assess the relative importance of the input variables (Garson 1998; Elmolla et al. 2010). In this study, the proposed network consisted of eight variables. Assuming the connection weights from the input nodes to the hidden nodes demonstrate the relative predictive importance of the independent

variable, the importance of each input variable can be expressed as follows:

$$I_j = \frac{\sum_{m=1}^{m=Nh} \left(\left(|W_{jm}^{ih}| / \sum_{k=1}^{Ni} |W_{km}^{ih}| \right) \times |W_{mn}^{ho}| \right)}{\sum_{k=1}^{k=Ni} \left\{ \sum_{m=1}^{m=Nh} \left(\left(|W_{jm}^{ih}| / \sum_{k=1}^{Ni} |W_{km}^{ih}| \right) \times |W_{mn}^{ho}| \right) \right\}} \quad (6)$$

where I_j is the relative importance of the j th input variable on the output variable; Ni and Nh are the number of input and hidden neurons, respectively; W is connection weight; the superscripts i , h , and o refer to input, hidden, and output layers, respectively; and subscripts k , m , and n refer to input, hidden, and output neurons, respectively.

Table 4 showed the connection weight values for the proposed model. The relative importance of each of the input variables as computed by Eq. (6) was shown in Fig. 4, illustrating the significance of a variable compared with the others in the model. Although the network did not necessarily represent physical meaning through the weights, it suggested that all the variables had strong effects on the prediction of DO (Singh et al. 2009), where the predictor contributions ranged from 7.4 to 18.9 % and pH, $\text{NO}_3\text{-N}$, $\text{NH}_4\text{-N}$, and Ca^{2+} had relatively high contributions Fig. 3. In addition, pH and $\text{NO}_3\text{-N}$ were the high influential variables with relative importance of 18.9 and 15.8 %. It was obvious that the most effective inputs were those which included oxygen containing ($\text{NO}_3\text{-N}$) and oxygen demanding ($\text{NH}_4\text{-N}$). Moreover, Cl^- revealed the least contribution on the proposed model.

Table 3 Modeling accuracy of the input variables was eliminate from the model

Model	Combination	Structure	RMSE			r		
			Training	Validation	Test	Training	Validation	Test
1	All	8-14-1	0.4272	0.3667	0.4570	0.9654	0.9841	0.9680
2	Eliminate pH	7-14-1	0.5019	0.4232	0.7179	0.9448	0.9806	0.9168
3	Eliminate EC	7-14-1	0.6779	0.9484	0.4933	0.9160	0.8407	0.9632
4	Eliminate Ca^{2+}	7-14-1	0.5005	0.5691	0.5209	0.9600	0.9437	0.9472
5	Eliminate Cl^-	7-14-1	0.3857	0.4493	0.4712	0.9738	0.9570	0.9691
6	Eliminate TH	7-14-1	0.5120	0.6524	0.6157	0.9519	0.9274	0.9454
7	Eliminate TA	7-14-1	0.5522	0.6251	0.5970	0.9471	0.9388	0.9393
8	Eliminate NH_4^+	7-14-1	0.9775	1.6073	0.8595	0.8084	0.5515	0.8914
9	Eliminate NO_3^-	7-14-1	0.7021	0.7035	0.6812	0.9118	0.9301	0.9260

Table 4 Connection weights between input and hidden layers (W1) and weights between hidden and output layers (W2)

Neuron	W1								W2
	Input variables								Target
	pH	EC	Ca ²⁺	Cl ⁻	TH	TA	NH ₄ -N	NO ₃ -N	DO
1	2.9030	0.5507	0.3237	0.1295	-0.7388	0.5463	-0.4320	-1.1923	1.8273
2	-1.0740	-0.7635	-1.7234	-0.4427	-0.5311	-0.6613	1.1612	0.8638	1.9211
3	1.0478	-0.4759	-0.9306	-0.4848	0.1470	-0.4110	-0.2323	0.2586	1.4489
4	1.5011	-1.2504	0.9808	-0.2178	0.4498	0.9944	0.1010	0.5281	1.6881
5	-0.9063	-0.3498	-0.2671	-0.4753	-0.5861	-1.0770	-0.3427	1.0466	1.8861
6	-0.6823	0.3497	-0.0643	0.1412	-1.0480	-0.4506	0.9581	-1.4659	-1.5894
7	1.1961	0.6122	-0.6630	0.4970	-0.0106	0.1695	-0.0721	0.4011	-1.3215
8	-2.9759	-0.6781	-1.3443	0.8044	-1.1719	0.5936	1.0460	0.4538	1.9414
9	-0.0614	1.2704	1.6516	0.9095	0.9652	0.6873	-0.3609	-1.4873	1.7432
10	-0.7089	-0.2954	1.2947	-1.1460	-0.3713	0.3468	1.9001	0.7855	-2.1393
11	-1.4200	0.5245	-0.5357	0.3485	0.5622	-0.5899	2.2096	-1.0692	2.0094
12	-0.9548	-0.3865	-0.5079	0.1309	0.2503	0.5959	1.1128	1.9301	1.5296
13	-0.4341	-1.0879	0.0636	-0.2805	-1.2685	-0.6450	-0.3724	0.0146	-1.5940
14	-0.1735	0.6558	-1.3277	-0.0438	-0.1984	0.3388	-0.2157	1.5416	-1.6090

These relationships represented that high levels of dissolved organic matter consume large amounts of oxygen, which underwent anaerobic fermentation processes leading to formation of ammonia and organic acids. Hydrolysis of these acidic materials causes a decrease of water pH values (Vega et al. 1998; Singh et al. 2004).

Conclusion

An artificial neural network was developed to simulate the dissolved oxygen (DO) concentration in the Heihe

River (Northwestern China). A three-layer (one input layer, one hidden layer, and one output layer) BPNN was used with the Bayesian regularization training algorithm. Water quality variables such as pH, electrical conductivity (EC), chloride (Cl⁻), calcium (Ca²⁺), total alkalinity (TA), total hardness (TH), nitrate nitrogen (NO₃-N), and ammonical nitrogen (NH₄-N) were used as the input data to obtain the output of the neural network, DO. Fourteen neurons were selected as the best number of neurons based on the minimum value of MSE of the training data set. A well-trained ANN produced results with the coefficient of correlation (*r*) of 0.9654, 0.9841, and 0.9680, and the RMSE of 0.4272, 0.3667, and 0.4570 for the training, validation, and test sets, respectively, with good match between the measured and modeled DO. The sensitivity analysis showed that the input variables such as pH, NO₃-N, NH₄-N, and Ca²⁺ had strong effect on DO. In addition, pH and NO₃-N were the high influential parameters with relative importance of 18.9 and 15.8 %, while Cl⁻ revealed the least contribution on the proposed model and can be excluded. The result demonstrated that the proposed ANN model was a better choice for modeling DO levels with limited knowledge of the water quality parameters.

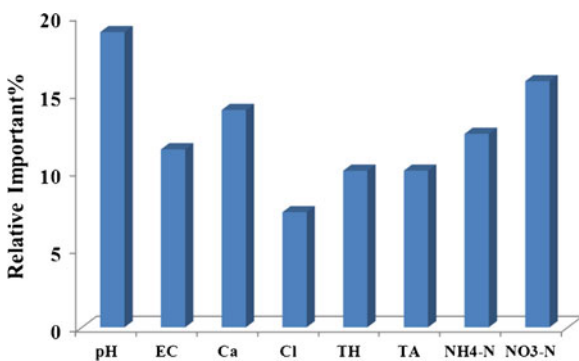


Fig. 4 The relative importance of the input variables to DO ANN model for Heihe River

Acknowledgments This work was supported by the One Hundred Person Project of the Chinese Academy of Sciences (29Y127D01), National Natural Science Foundation of China (41171026, 91025024). The author wishes to thank the anonymous reviewers for their reading of the manuscript, and for their suggestions and critical comments.

References

- Ancil, F., Perrin, C., & Andreassian, V. (2004). Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environmental Modelling and Software*, *19*, 357–368.
- Ancil, F., Filion, M., & Tournebize, J. (2009). A neural network experiment on the simulation of daily nitrate-nitrogen and suspended sediment fluxes from a small agricultural catchment. *Ecological Modelling*, *220*, 879–887.
- Ansa-Ansare, O. D., Marr, I. L., & Cresser, M. S. (2000). Evaluation of modelled and measured patterns of dissolved oxygen in a freshwater lake as an indicator of the presence of biodegradable organic pollution. *Water Research*, *34*, 1079–1088.
- APHA. (1995). *Standard methods for the examination of water and wastewater* (19th ed.). Washington: American Public Health Association.
- Banerjee, P., Singh, V. S., Chattopadhyay, K., Chandra, P. C., & Singh, B. (2011). Artificial neural network model as a potential alternative for groundwater salinity forecasting. *Journal of Hydrology*, *398*, 212–220.
- Chang, F. J., Kao, L. S., Kuo, Y. M., & Liu, C. W. (2010). Artificial neural networks for estimating regional arsenic concentrations in a blackfoot disease area in Taiwan. *Journal of Hydrology*, *388*, 65–76.
- Cherkassky, V., & Mulier, F. (1998). *Learning from data: Concepts, theory, and methods*. U.S.A.: Wiley.
- Coulibaly, P., Ancil, F., Aravena, R., & Bobée, B. (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, *37*, 885–896.
- Coulibaly, P., Bobee, B., & Ancil, F. (2001). Improving extreme hydrologic events forecasting using a new criterion for artificial neural network selection. *Hydrological Processes*, *15*, 1533–1536.
- da Costa, A. O., Silva, P. F., Sabara, M. G., & Da Costa, E. F., Jr. (2009). Use of neural networks for monitoring surface water quality changes in a neotropical urban stream. *Environmental Monitoring and Assessment*, *155*, 527–538.
- Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, *309*, 229–240.
- Dastorani, M. T., Moghadamnia, A., Piri, J., & Rico-Ramirez, M. (2010). Application of ANN and ANFIS models for reconstructing missing flow data. *Environmental Monitoring and Assessment*, *166*, 421–434.
- Dixon, B. (2005). Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: a GIS-based sensitivity analysis. *Journal of Hydrology*, *309*, 17–38.
- Dogan, E., Sengorur, B., & Koklu, R. (2009). Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. *Journal of Environmental Management*, *90*, 1229–1235.
- Elmolla, E. S., Chaudhuri, M., & Eltoukhy, M. M. (2010). The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. *Journal of hazardous materials*, *179*, 127–34.
- Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. *Information Management*, *24*, 159–167.
- Freeman, J. A., & Skapura, D. M. (1991). *Neural networks: algorithms, applications, and programming techniques*. MA: Addison–Wesley, Reading.
- Gao, Q. (1991). *Development and utilization of water resources in the Heihe River catchment* (p. 205). Lanzhou: Gansu Science and Technology Press. In Chinese.
- Garcia, L., & Shigidi, A. (2006). Using neural networks for parameter estimation in ground water. *Journal of Hydrology*, *318*, 215–231.
- García, A., Revilla, J. A., Medina, R., Alvarez, C., & Juanes, J. A. (2002). A model for predicting the temporal evolution of dissolved oxygen concentration in shallow estuaries. *Hydrobiologia*, *475/476*, 205–211.
- Garson, G. D. (1998). *Neural networks: an introductory guide for social scientists*. California: Sage Publications.
- Ghose, D. K., Panda, S. S., & Swain, P. C. (2010). Prediction of water table depth in western region. Orissa using BPNN and RBFN neural networks. *Journal of Hydrology*, *394*, 296–304.
- Guo, J., Zhou, J., Qin, H., Zou, Q., & Li, Q. (2011). Monthly streamflow forecasting based on improved support vector machine model. *Expert Systems with Applications*, *38*, 13073–13081.
- Ha, H., & Stenstrom, M. K. (2003). Identification of land use with water quality data in stormwater using a neural network. *Water Research*, *37*, 4222–4230.
- Hagan, M. T., Delmuth, H. B., & Beale, M. (1996). *Neural network design*. MA: PWS Publishing Company.
- He, B., Oki, T., Sun, F., Komori, D., Kanae, S., Wang, Y., Kim, H., & Yamazaki, D. (2011). Estimating monthly total nitrogen concentration in streams by using artificial neural network. *Journal of Environmental Management*, *92*, 172–177.
- Hinton, G. E. (1992). How neural networks learn from experience. *Scientific American*, *267*, 144–151.
- Hull, V., Parrella, L., & Falcucci, M. (2008). Modelling dissolved oxygen dynamics in coastal lagoons. *Ecological Modelling*, *2*, 468–480.
- Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). *Neuro-fuzzy and soft computing*. Englewood Cliffs: Prentice-Hall.
- Jensen, B. A. (1994). *Expert systems—neural networks, Instrument Engineers' Handbook* (3rd ed., pp. 48–54). Radnor: Chilton.
- Kim, M., & Gilley, J. E. (2008). Artificial neural network estimation of soil erosion and nutrient concentrations in runoff from land application areas. *Computers and Electronics in Agriculture*, *64*(2), 268–275.
- Kralisch, S., Fink, M., Flügel, W. A., & Beckstein, C. (2003). A neural network approach for the optimisation of watershed management. *Environmental Modelling and Software*, *18*, 815–823.
- Krishna, B., Satyaji Rao, Y. R., & Vijaya, T. (2008). Modelling groundwater levels in an urban coastal aquifer using artificial neural networks. *Hydrological Processes*, *22*, 1180–1188.

- Kuo, J. T., Wang, Y. Y., & Lung, W. S. (2006). A hybrid neural-genetic algorithm for reservoir water quality management. *Water Research*, *40*, 1367–1376.
- Kuo, J., Hsieh, M., Lung, W., & She, N. (2007). Using artificial neural network for reservoir eutrophication prediction. *Ecological Modelling*, *200*, 171–177.
- MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation*, *4*, 415–447.
- Maier, H. R., & Dandy, G. C. (1998). The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study. *Environmental Modelling and Software*, *13*, 193–209.
- Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling and Software*, *15*, 101–124.
- Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions. *Environmental Modelling and Software*, *25*, 891–909.
- Najah, A., El-Shafie, A., Karim, O. A., & Jaafar, O. (2011). Integrated versus isolated scenario for prediction dissolved oxygen at progression of water quality monitoring stations. *Hydrology and Earth System Sciences*, *15*, 2693–2708.
- Nayak, P., Sudheer, K., Rangan, D., & Ramasastri, K. (2004). A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, *291*, 52–66.
- Nayak, P. C., Rao, Y. R. S., & Sudheer, K. P. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, *20*, 77–90.
- Ouarda, T. B. M. J., & Shu, C. (2009). Regional low-flow frequency analysis using single and ensemble artificial neural networks. *Water Resources Research*, *45*, w11428.
- Porter, D. W., Gibbs, P. G., Jones, W. F., Huyakorn, P. S., Hamm, L. L., & Flach, G. P. (2000). Data fusion modeling for groundwater systems. *Journal of Contaminant Hydrology*, *42*, 303–335.
- Principe, J. C., Euliano, N. R., & Lefebvre, C. W. (2000). *Neural and adaptive systems: fundamentals through simulations*. New York: Wiley.
- Ranković, V., Radulović, J., Radojević, I., Ostojić, A., & Čomić, L. (2010). Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. *Ecological Modelling*, *221*, 1239–1244.
- Sahoo, G. B., Ray, C., Wang, J. Z., Hubbs, S. A., Song, R., Jasperse, J., & Seymour, D. (2005). Use of artificial neural networks to evaluate the effectiveness of riverbank filtration. *Water Research*, *39*, 2505–2516.
- Senkal, O., Yildiz, B. Y., Sahin, M., & Pestemalci, V. (2012). Precipitable water modelling using artificial neural network in Cukurova region. *Environmental Monitoring and Assessment*, *184*, 141–147.
- Shu, C., & Ouarda, T. B. M. J. (2007). Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resources Research*, *43*, W07438.
- Shukla, J. B., Misra, A. K., & Chandra, P. (2008). Mathematical modeling and analysis of the depletion of dissolved oxygen in eutrophied water bodies affected by organic pollutants. *Nonlinear Analysis-Real*, *9*, 1851–1865.
- Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Research*, *38*, 3980–3992.
- Singh, K. P., Basant, A., Malik, A., & Jain, G. (2009). Artificial neural network modeling of the river water quality—a case study. *Ecological Modelling*, *220*, 888–895.
- Sreekanth, J., & Datta, B. (2010). Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models. *Journal of Hydrology*, *393*, 245–256.
- Suen, J. P., Eheart, J. W., & Asce, M. (2003). Evaluation of neural networks for modeling nitrate concentration in rivers. *Journal of Water Resources Planning and Management—ASCE*, *129*, 505–510.
- Vega, M., Pardo, R., Barrado, E., & Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of River water by exploratory data analysis. *Water Research*, *32*, 3581–3592.
- Wang, G., Cheng, G., & Yang, Z. (1999). The utilization of water resources and its influence on eco-environment in the northwest arid area of China. *Journal of Natural Resources*, *14*, 109–116 (In Chinese).
- Wang, H., Hondzo, M., Xu, C., Poole, V., & Spacie, A. (2003). Dissolved oxygen dynamics of streams draining an urbanized and an agricultural catchment. *Ecological Modelling*, *160*, 145–161.
- Wen, X. H., Wu, Y. Q., Lee, L. J. E., Su, J. P. & Wu, J. (2007). Groundwater flow modeling in the Zhangye basin, Northwestern China. *Environmental Geology*, *53*, 77–84.
- Wu, C. L., & Chau, K. W. (2011). Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology*, *399*, 394–409.