Title

Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: a case study of Wen-Rui Tang River, China.

Journal

Environmental science and pollution research international, 24(19)

ISSN

0944-1344

Authors

Ji, Xiaoliang
Shang, Xu
Dahlgren, Randy A
et al.

Publication Date

2017-07-01

DOI

10.1007/s11356-017-9243-7

Peer reviewed

Environ Sci Pollut Res
DOI 10.1007/s11356-017-9243-7

RESEARCH ARTICLE

Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: a case study of Wen-Rui Tang River, China

Xiaoliang Ji[1] · **Xu Shang**[1] · **Randy A. Dahlgren**[1,2] · **Minghua Zhang**[1,2]

Received: 17 December 2016 / Accepted: 9 May 2017

Abstract Accurate quantification of dissolved oxygen (DO) is critically important for managing water resources and controlling pollution. Artificial intelligence (AI) models have been successfully applied for modeling DO content in aquatic ecosystems with limited data. However, the efficacy of these AI models in predicting DO levels in the hypoxic river systems having multiple pollution sources and complicated pollutants behaviors is unclear. Given this dilemma, we developed a promising AI model, known as support vector machine (SVM), to predict the DO concentration in a hypoxic river in southeastern China. Four different calibration models, specifically, multiple linear regression, back propagation neural network, general regression neural network, and SVM, were established, and their prediction accuracy was systemically investigated and compared. A total of 11 hydro-chemical variables were used as model inputs. These variables were measured bimonthly at eight sampling sites along the rural-suburban-urban portion of Wen-Rui Tang River from 2004 to 2008. The performances of the established models were assessed through the mean square error (MSE), determination coefficient (R^2), and Nash-Sutcliffe (NS) model efficiency. The results indicated that the SVM model was superior to other models in predicting DO concentration in Wen-Rui Tang River. For SVM, the MSE, R^2, and NS values for the testing subset were 0.9416 mg/L, 0.8646, and 0.8763, respectively. Sensitivity analysis showed that ammonium-nitrogen was the most significant input variable of the proposal SVM model. Overall, these results demonstrated that the proposed SVM model can efficiently predict water quality, especially for highly impaired and hypoxic river systems.

Keywords Dissolved oxygen · Hypoxic river systems · Support vector machine · Artificial neural networks · Water quality prediction · Wen-Rui Tang River

Responsible editor: Marcus Schulz

✉ Minghua Zhang
mhzhang@ucdavis.edu

Xiaoliang Ji
jixiao556677@126.com

[1] Zhejiang Province Key Laboratory of Watershed Science and Health, Southern Zhejiang Water Research Institute (iWATER), Wenzhou Medical University, Wenzhou 325035, China

[2] Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA

Introduction

As a crucial global environmental issue, surface water quality impairment degrades the health of aquatic ecosystems, threatens drinking water availability, and induces harmful algal blooms and hypoxia in many freshwater and coastal ecosystems (Bowes et al. 2010; Gao and Zhang 2010; Morse and Wollheim 2014). The water quality deterioration can be attributed to urbanization, population growth, excessive water consumptions, industrial wastewater discharge, and agricultural activities, while the system lacks adequate wastewater treatment facilities (Chen et al. 2016; Gupta 2008; Singh et al. 2011). The increasingly serious pollution causes low dissolved oxygen (DO) levels and worsens life conditions in aquatic systems.

DO is a critical water quality indicator that can be used to assess the health of aquatic ecosystems (Basant et al. 2010; Wen et al. 2013). Oxygen-producing processes (e.g., reaeration from the atmosphere and photosynthesis) and oxygen-consuming processes (e.g., respiration by aquatic organisms, chemical oxidation, and sediment oxygen demand) both determine DO

Published online: 23 May 2017

concentrations in surface waters (Kuo et al. 2007; Quinn et al. 2005). Sufficient DO in a water body is essential for supporting aquatic life and basic organic oxygen demand (e.g., decomposition of organic matter). Quantification of DO levels in waters has become a major concern for water resource managers and government agencies. Accurately predicting DO levels in hypoxic river systems remains a challenge for water managers. This is because anthropogenic activities in these river systems usually couple with non-linear, dynamic, and complex biochemical processes. To date, the complex mechanistic models, which entail a full understanding of the underlying physical relationships in water column, such as River and Stream Water Quality Model (QUAL2K), Environmental Fluid Dynamics Code (EFDC), and Water Quality Analysis Simulation Program (WASP6), are widely applicable for DO prediction (Chapra and Pellettier 2003; Tetratech, Inc, 2002; Wool et al. 2006). Most of these mechanistic models require a large amount of input data. Sometimes these required data may not be readily available which hinders the implementation of the model outputs for many watersheds, especially those in the developing countries. Moreover, a particular degree of expertise and experience is indispensable to build such a model, making these models overly complex and difficult to implement for prediction.

To overcome these limitations, researchers have developed artificial intelligence (AI) models, such as artificial neural networks (ANNs) and support vector machine (SVM), to model water quality parameters. These models have the advantages of requiring less input data and solving non-linear problems (Chau 2006; He et al. 2014; Kisi et al. 2012). During the past decade, numerous reports have cited that AI models successfully simulate DO content or other water quality parameters. For example, Singh et al. (2009) used back propagation neural network (BPNN) and feed-forward neural network (FFNN) models to compute DO and biochemical oxygen demand (BOD) concentrations in Gomti River (India), and the results indicated that these models were capable of capturing long-term trends observed for DO and BOD. Antanasijević et al. (2013a) employed BPNN, general regression neural network (GRNN), and recurrent neural network (RNN) to estimate the DO content in Danube River (North Serbia), demonstrating the efficiency of the RNN model for DO prediction. Chen and Liu (2014) developed BPNN and adaptive neuro-fuzzy inference system (ANFIS) to simulate the DO levels of Feitsui Reservoir (China), and found that the latter was able to accurately predict DO concentration. Nevertheless, ANNs have some shortcomings inherent in its architecture, namely, overfitting, poor reproducibility, and local minima. The SVM method, proposed by Vapnik (1995), is a promising method in the data-driven prediction field. Recently, SVM has been extensively applied to address various water resource problems, such as stream flow forecasting, water-level prediction, and water quality parameter simulation (Khan and Coulibaly 2006; Lin et al. 2006; Noori et al. 2015; Yoon et al. 2011). However, studies related to DO prediction through the SVM model are few. More importantly, DO simulation in most previous research was performed only under high DO levels in surface waters (see Appendix Table 4). Thus, whether the SVM model can be successfully used for DO prediction in highly impaired river systems, which often tend to be hypoxic, remains unclear.

The Wen-Rui Tang River is a major river system across the rural-suburban-urban interface in Wenzhou, southeastern China. The water quality is severely degraded since large amounts of untreated industrial effluents, domestic sewage, and agricultural wastewater were often directly discharged into the river system. Currently, the water quality of Wen-Rui Tang River does not meet the lowest water quality classification (class V) in China (State Environment Protection Bureau of China 2002a). Moreover, many portions of urban waterways of the river are considered dead zones due to persistent hypoxia. Therefore, the Wen-Rui Tang River, a representative example of the hypoxia river systems, is worthy of our exploration.

The applicability of SVM model in estimating DO content in Wen-Rui Tang River was investigated in this research, which was aimed at overcoming the difficulties associated with DO prediction of hypoxic river systems. The performance of SVM for DO estimation was discussed and compared with BPNN, GRNN, and multiple linear regression (MLR). BPNN is by far the most popular among ANN paradigms (Chau and Wu 2010). GRNN is characterized by fast learning speed, easy-to-use, and convergence to the optimal regression surface. In addition, BPNN and GRNN both have shown great potential in predicting DO concentration (Antanasijević et al. 2013a, 2014a; Wen et al. 2013). MLR is a well-known statistical technique for establishing linear links between input and output variables and is the minimum standard for model intercomparison (Abrahart et al. 2012). We expect that the proposed SVM model will be an efficient tool for water quality management and pollution control of Wen-Rui Tang River and other hypoxic river systems.

Materials and methods

Study area

The Wen-Rui Tang River watershed, located in southeastern China, covers an area of 353 km^2 (Fig. 1). The river originates from the Lishui mountain stream, meanders through forests, agriculture areas, urban zone with a population of ~7.2 million inhabitants, and agricultural areas, from northwest to southeast, and joins the Fei-Yun River, which eventually

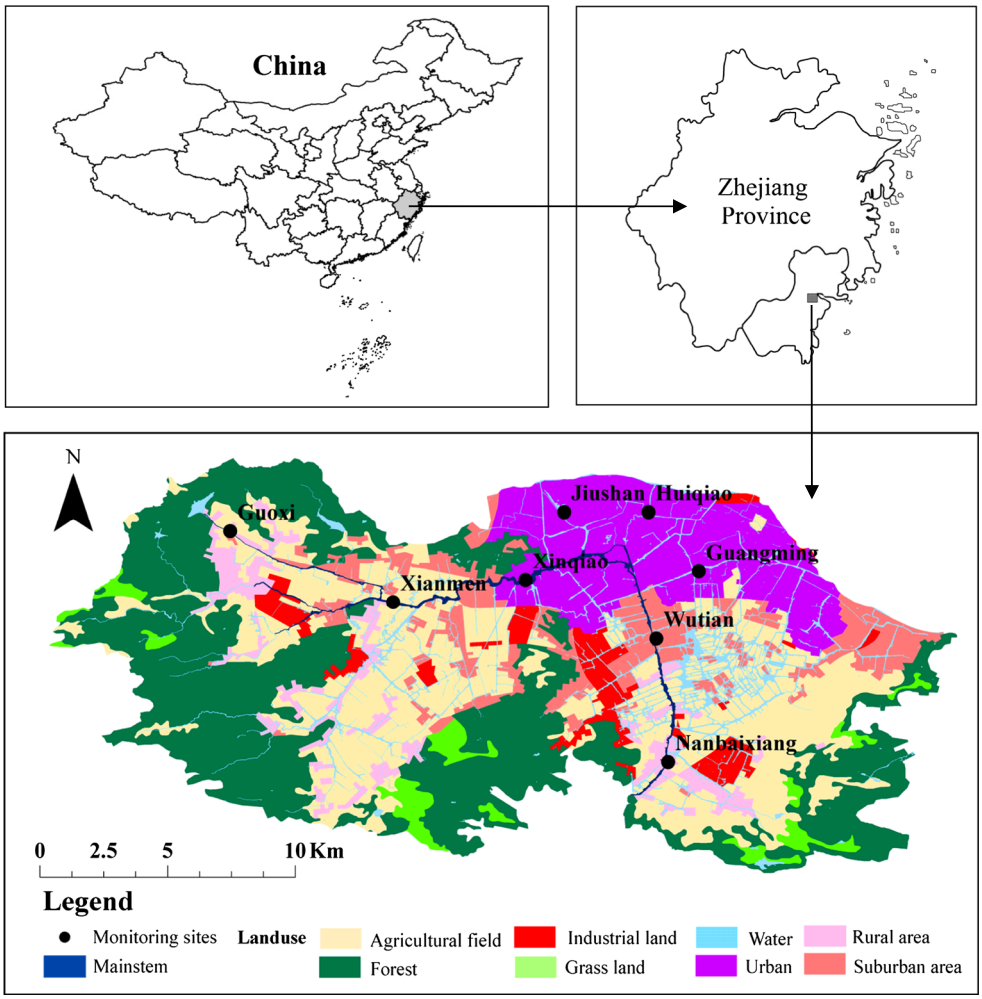

Fig. 1 Land use map and monitoring sites of the Wen-Rui Tang River watershed

drains into the East China Sea. The dominant land use categories are agriculture, hilly forest, and urban, accounting for 39.5, 38, and 14.5% of the entire watershed, respectively. The width of the urban portion of the river varies from 13 to 150 m, with an average value of ~50 m (Mei et al. 2011). This region has a subtropical oceanic climate with hot and humid summers and mild and dry winters. The long-term average annual rainfall for the watershed is 1800 mm with ~70% precipitation occurring from April to September. Since the rapid urbanization and economic development in the 1980s, the water quality of the river has been severely impaired and experienced recurring hypoxia. On the basis of the water functions of different land use zones, the desired quality standards were set to be class IV water quality standard (water suitable for industrial uses) in urban and suburban zones and class V water quality standard (water suitable for agricultural purposes) in rural zone (Yang et al. 2013). Correspondingly, the expected DO concentration is higher than 3 mg/L in urban and suburban zones and 2 mg/L in rural zone. Most DO concentrations from 2004 to 2008 are less than the minimum allowable value, indicating that the Wen-Rui Tang River is a typical hypoxic river system (Fig. 2).

Data

Water quality data were measured bimonthly over a period of 5 years (2004–2008) at eight monitoring sites. Data were obtained from the Wenzhou Environmental Protection Bureau (WEPB). The eight sites belonged to different land use zones, including two rural area sites (Guoxi and Nanbaixiang), four urban area sites (Xinqiao, Jiushan, Huiqiao, and Guangming), and two suburban area sites (Xianmen and Wutian). Water quality parameters included water temperature (Temp), pH, potassium permanganate-chemical oxygen demand index (COD_{Mn}), 5-day biochemical oxygen demand (BOD_5), ammonium-nitrogen (NH_4^+-N), petroleum (Petrol), total phosphorus (TP), cadmium-chemical oxygen demand index (COD_{Cr}), fluoride (F), total nitrogen (TN), electrical conductivity (EC), and dissolved oxygen (DO). Water samples were analyzed according to the National Quality Standards for Surface Waters (State Environment Protection Bureau of China 2002b). The particular set of parameters was used here because they would directly or indirectly affect DO concentrations in aquatic ecosystem. The statistical summary of these parameters is

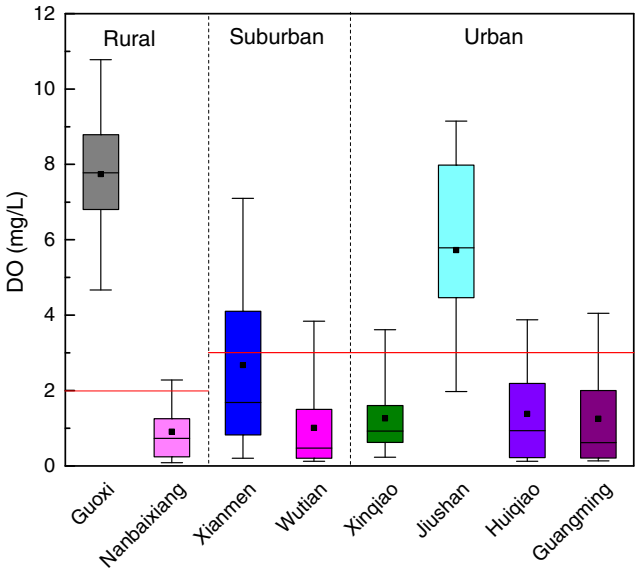

Fig. 2 DO concentrations of Wen-Rui Tang River watershed from 2004 to 2008 (*box plot* denotes 25th, 50th, and 75th percentiles; the *whiskers* indicate the 5th and 95th percentiles; the *small black boxes* denote mean value; *red lines* represent the minimum allowable value)

shown in Table 1. For DO, the concentrations ranged from 0.04 to 10.80 mg/L, and 65% of the samples are worse than the minimum allowable value.

Artificial neural networks

ANNs (e.g., BPNN and GRNN) emulate information processing functions of neural networks in the brain. ANNs are suitable for modeling complex non-linear relationships in different data sets that cannot be described by linear mathematical formulas. The brief theory of BPNN and GRNN is provided in this section.

Back propagation neural network

BPNN is a traditional and powerful non-linear regression tool based on the back propagation of the error gradient (Hagan et al. 1996; Haykin 1999). In this supervised learning algorithm, the input vectors and corresponding target vectors are employed to train a network until a termination criterion, such as a specified maximum number of epochs or a minimum error, is achieved (Chen and Liu 2014). Figure 3a illustrates the BPNN architecture. The basic BPNN structure usually comprises three distinctive layers, namely one input layer, one hidden layer, and one output layer (Wu et al. 2009; Xie et al. 2006). Each layer consists of one or more basic elements, called neurons or nodes, and connection pathways that link them. The details of the BPNN approach are as follows: (i) initialize the network. The number of neurons in three layers and the weight values between the input and hidden layers (w_{ij}) and between the hidden and output layers (w_{jk}) need to be initialized. (ii) Obtain the output results of the hidden layer. The linear and sigmoid functions can be used as the transfer functions between the input and hidden layers. The former can express a linear relationship between these layers. The latter is suitable for highly non-linear systems and contains logistic sigmoid (logsig) and hyperbolic tangent sigmoid (tansig) functions. The output values of the logsig function are within 0 and 1, and those of the tansig function are within −1 and 1. The tansig function has better sensitivity and accuracy than the logsig function. Hence, the tansig function was selected and described by the following equation:

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

(iii) Obtain the results of the output layer. The transfer function between the hidden and output layers is a linear function. (iv) Calculate the mean square error (MSE), which is an evaluation criterion of network performance. (v) Update weight and bias values according to the calculated MSE. Levenberg-Marquardt algorithm, which has good performance and learning speed, is used to complete the process. The procedure is reiterated until the results satisfy the stopping criteria earlier mentioned.

General regression neural network

GRNN, a variation of radial basis function (RBF) networks, is designed for function regression and approximation. The main difference between GRNN and traditional BPNN is the fixed architecture of the former for a given input-output data set and the requirement of determining an optimal number of hidden neurons in the latter (Yaseen et al. 2016). GRNN can be

Table 1 Statistical summary of water quality parameters

Parameters	*N*	Unit	Minimum	Maximum	Mean	SD	Cv (%)
Temp	240	°C	8.80	33.00	21.40	7.39	34.53
pH	240	–	6.64	9.26	7.38	0.35	4.80
COD_{Mn}	240	mg/L	0.74	17.20	6.46	2.66	41.14
BOD_5	240	mg/L	0.63	46.30	7.47	5.84	78.25
NH_4^+-N	240	mg/L	0.04	19.00	5.78	4.34	75.08
Petrol	240	mg/L	0.03	0.66	0.09	0.08	81.39
TP	240	mg/L	0.03	1.77	0.47	0.32	69.01
COD_{Cr}	240	mg/L	7.90	88.00	26.62	13.87	52.10
F	240	mg/L	0.07	1.65	0.43	0.20	45.94
TN	240	mg/L	0.55	19.90	6.83	4.34	63.51
EC	240	mS/m	6.10	173.00	38.37	21.84	56.93
DO	240	mg/L	0.04	10.80	2.74	2.90	105.76

SD standard deviation, *Cv* coefficient of variation

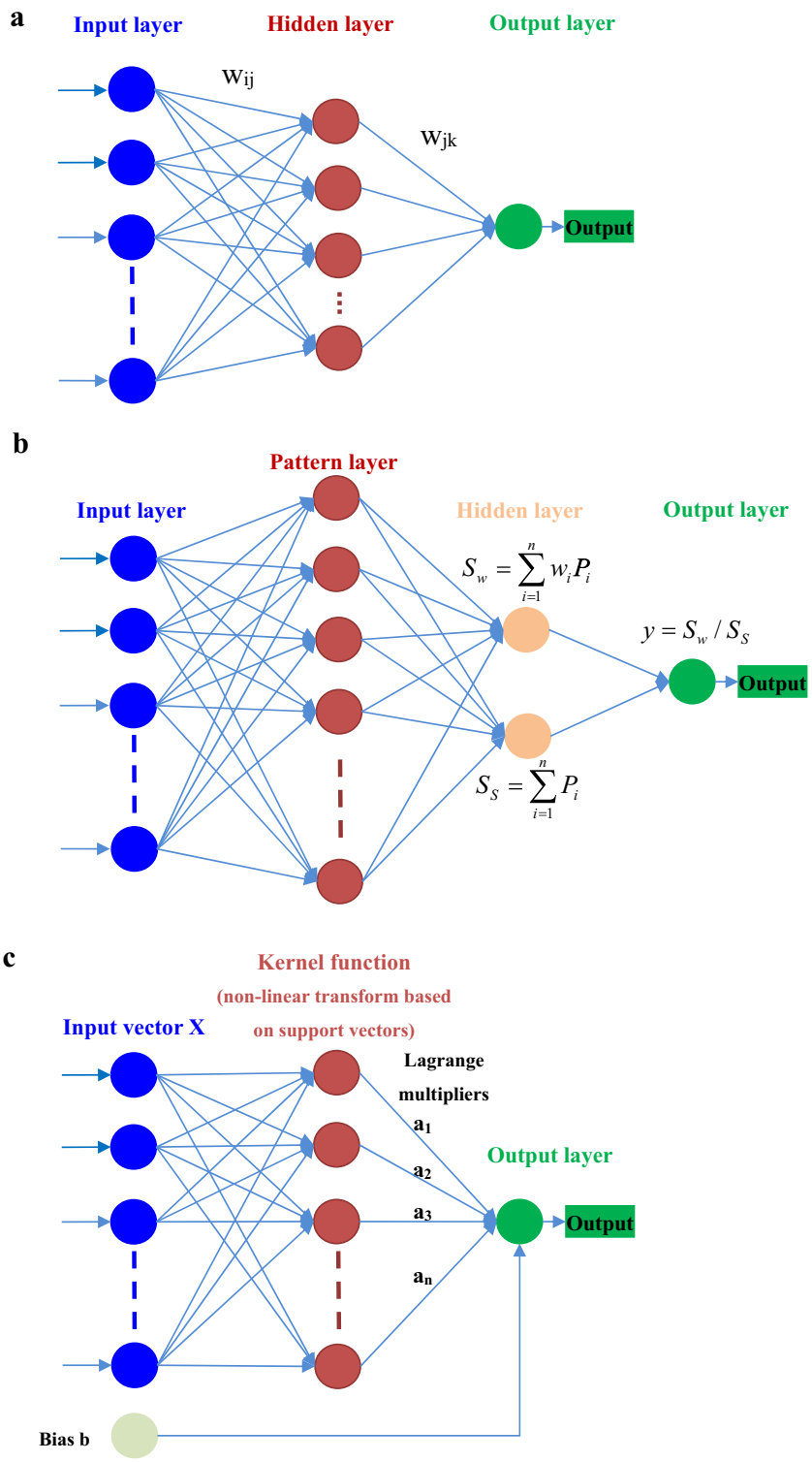

Fig. 3 Schematic diagram of the models used in this study: **a** BPNN, **b** GRNN, and **c** SVM

successfully applied to deal with various linear and non-linear problems and can accurately perform predictions without requiring large samples (Alilou and Yaghmaee 2015). Any smooth function approximation problem can be addressed by GRNN model (Kim and Kim 2008). Also, the GRNN model is capable of generating consistent forecasts such that when the training data set size becomes large and the estimation error approaches zero, with only mild restrictions on the function (Cigizoglu and Alp 2006). This learning algorithm with fast learning speed and convergence demonstrated good results in environmental modeling (Antanasijević et al. 2013b, c, 2014b; Heddam 2014a, b).

GRNN consists of four layers, namely, the input, pattern, summation, and output layers (Fig. 3b). In the present case, the numbers of neurons in the input and output layers were 11 and 1, respectively, and each neuron represented an independent variable. The pattern layer contained 192 neurons, which was equal to the number of data patterns used for model training. The summation layer included two neurons, which was equivalent to the number of output variables plus one. Upon GRNN implementation, the neurons in the pattern layer (P_i) are able to memorize the relationship between the neurons in the input layer and proper response in the pattern layer. The Gaussian function of the pattern P_i is defined as follows:

$$P_i = \exp\left(-\frac{(X-X_i)^T(X-X_i)}{2\sigma^2}\right) \tag{2}$$

where X indicates the input variable, X_i indicates a training sample of the ith neuron in pattern layer, and σ is the smoothing factor, which represents the width of the Gaussian curve and thus determines the accuracy of GRNN (Antanasijević et al. 2013b).

Subsequently, the two neurons in summation layer, namely S_s and S_w, can be expressed as follows:

$$Ss = \sum_{i=1}^{n} Pi \tag{3}$$

$$Sw = \sum_{i=1}^{n} w_i P_i \tag{4}$$

where w_i is the weight used to connect the ith neuron to the summation layer.

Finally, the predicted values are determined in the output layer. The output, denoted by y, can be derived using the following:

$$y = S_w/S_s \tag{5}$$

Support vector machine

The SVM method, which is developed based on statistical learning theory, has been successfully applied to classification and regression problems (Mohammadpour et al. 2015). The basic concept behind SVM is to map the original data sets to higher dimensional features of space and construct an optimal separating plane (SP), from which the distance to all the data points is minimal (Lin et al. 2008; Pan et al. 2008; Qu and Zuo 2010).

Detailed expression of SVM has been extensively reported in numerous studies (Cristianine and Taylor 2000; Raghavendra and Deka, 2014; Vapnik 1998). Consequently, a brief description of this model was provided in the present study, and the schematic diagram of SVM is illustrated in Fig. 3c. For training data set $\{(x_i, y_i), i = 1,\ldots, n\}$, $x \in R^m$, $y \in R$, where n is the total number of data patterns, x is the input vector of m components, and y is the corresponding output value, the SVM regression function can be expressed as follows:

$$f(x) = w \cdot \phi(x) + b \tag{6}$$

where w is the weight vector, b is the bias, and $\phi(x)$ indicates the non-linear transfer function. The parameters w and b, which define the location of SP, can be determined by minimizing the following regularized risk function:

$$\text{Minimize}: \frac{1}{2}\|w\|^2 + c\sum_{1}^{n}\left(\xi_i + \xi_i^*\right) \tag{7}$$

Subject to $y_i - w\cdot\phi(x) - b \le \varepsilon + \xi_i$ $w\cdot\phi(x) + b - y_i \le \varepsilon + \xi_i^*$ $\xi_i \ge 0, \xi_i \ge 0$

where C is the regularization parameter, ξ_i and ξ_i^* are slack variables, and Eq. (7) is solved in a dual form using the Lagrangian multipliers.

$$\text{Maximize}: -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_i - a_i^*\right)\left(a_j - a_j^*\right)K\left(x_i, x_j\right) - \sum_{i=1}^{n}\left(a_i - a_i^*\right) + \sum_{i=1}^{n}\left(a_i - a_i^*\right)y_i \tag{8}$$

Subject to $\sum_{i=1}^{n}\left(a_i - a_i^*\right)$ $0 \le a_i \le C$ $0 \le a_i^* \le C$

where $K(x_i, x)$ is the kernel function.

By imposing the Karush-Kuhn-Tucker (KKT) optimality condition, w^* is obtained, that is,

$$w^* = \sum_{i=1}^{n}\left(a_i - a_i^*\right)\cdot K\left(x_j, x\right) \tag{9}$$

Finally, the SVM is expressed as follows:

$$f(x) = \sum_{i=1}^{n}\left(a_i - a_i^*\right)\cdot K(x_i, x) + b \tag{10}$$

For SVM, four options, namely, polynomial, sigmoid, linear, and RBF, can be used for the kernel function. In this paper, RBF was adopted as the kernel function of SVM because of the following reasons: First, the RBF kernel map input vectors into a high-dimensional feature space in a non-linear fashion. Hence, it has the ability to model complicated non-linear relationships that the linear kernel function does not have. Second, available adjustable parameters in RBF are fewer than those in polynomial and sigmoid kernels, and thus, RBF is easier to use than the polynomial and sigmoid kernels (Keerthi and Lin 2001). Third, many studies have demonstrated excellent performance of RBF

(Dibike et al. 2001; Keerthi and Lin 2001). RBF is defined as follows:

$$K(x_i, x) = \exp\left(-g\|x_i - x\|^2\right) \tag{11}$$

where g is the adjustable kernel parameter.

The generalization performance of SVM depends on the regularization parameter C in combination with RBF kernel parameter g. The regularization parameter C controls the trade-off between minimizing training error and minimizing model complexity. An extremely small C value can cause an insufficient fitting problem, while an extremely large C value can cause the algorithm to overfit the training data (Wang et al. 2007). The RBF kernel parameter g defines the width of the kernel. Finding the optimal parameters of SVM has no specific strategy. Usually, a simple trial-and-error method is employed to adjust these parameters. In recent years, optimization algorithms, including exhaustive grid search algorithm, genetic algorithm (GA), and particle swarm optimization (PSO) algorithm, have been used to perform the optimization of C and g. Although GA and PSO are both novel simulated evolutionary and random optimization algorithms, they can be easily be trapped in a local optimum and are thus prone to significant errors in quantitative calculation. The grid search algorithm prevents local optimum and presents high precision solution and quick convergence. Accordingly, an exhaustive grid search technique coupled with cross validation (i.e., CV) was employed to determine the optimal combination of C and g (Hsu et al. 2007).

Model performance assessment

The model performance can be evaluated through various statistical measures, including determination coefficient (R^2), efficiency coefficient (CE), agreement index (d), Nash-Sutcliffe (NS) model efficiency, MSE, mean absolute error (MAE), standard error of prediction (SEP), and mean absolute relative error (MARE). Legates and McCabe (1999) suggested that a good examination of model performance should contain at least one goodness-of-fit or relative error measure (e.g., NS) and at least one absolute error measure (e.g., MSE). In addition, R^2 is the widely applied statistical score metric and is often useful in estimating the model performance. A model can be adequately assessed by R^2, NS, and MSE, and thus, the performances of the established models in this study were evaluated using these performance indexes. MSE measures the discrepancy between the predicted and observed values. R^2 indicates the degree of correlation among the predicted and observed values. NS is an indicator of the predictive power of model (Borah and Bera 2004). Basically, MSE = 0, R^2 = 1, and NS = 1 represent the best fit between observed and predicted values. The mathematical expressions of the performance indexes are given by the following:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2 \tag{12}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}\left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(P_i - \overline{P}\right)^2 \sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2}}\right)^2 \tag{13}$$

$$\mathrm{NS} = 1 - \frac{\sum_{i-1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}\left(Oi - \overline{O}\right)^2} \tag{14}$$

where n represents the number of data points, P_i and O_i are the predicted and observed ith values of the DO concentration, respectively, and $\overline{P}$ and $\overline{O}$ are the average values of predicted and observed DO concentrations, respectively.

Table 2 Spearman correlation coefficients of the water quality parameters

	Temp	PH	COD_{Mn}	BOD_5	NH_4^+-N	Petrol	TP	COD_{Cr}	F	TN	EC	DO
Temp	1											
PH	0.014	1										
COD_{Mn}	0.111*	−0.028	1									
BOD_5	−0.237**	−0.059	0.721**	1								
NH_4^+-N	−0.082	−0.188**	0.785**	0.711**	1							
Petrol	0.049	−0.145*	0.426**	0.351**	0.359**	1						
TP	0.102	−0.115*	0.682**	0.561**	0.711**	0.302**	1					
COD_{Cr}	−0.129*	0.035	0.792**	0.753**	0.669**	0.339**	0.515**	1				
F	0.096	−0.042	0.447**	0.268**	0.576**	0.210**	0.412**	0.295**	1			
TN	−0.095	−0.182**	0.791**	0.729**	0.967**	0.372**	0.720**	0.696**	0.546**	1		
EC	−0.157**	0.242**	0.417**	0.446**	0.447**	0.141**	0.373*	0.429**	0.441**	0.429**	1	
DO	0.069	0.342**	−0.537**	−0.531**	−0.719**	−0.366**	−0.617**	−0.479**	−0.513**	−0.690**	−0.382**	1

*$p < 0.05$ level; **$p < 0.01$ level (one-tailed)

Table 3 Comparison of results based on different regression models

Model	Training			Testing		
	MSE (mg/L)	R^2	NS	MSE (mg/L)	R^2	NS
MLR	2.9022	0.6797	0.6732	2.0652	0.6780	0.7286
BPNN	1.4254	0.8421	0.8395	1.6658	0.7854	0.7811
GRNN	0.9857	0.9012	0.8890	1.0959	0.8341	0.8560
SVM	0.6603	0.9262	0.9256	0.9416	0.8646	0.8763

Results and discussion

In this work, MLR, BPNN, GRNN, and SVM were used to predict DO content in Wen-Rui Tang River, and their performances were systemically investigated and compared. Of these models, the MLR model was implemented using the SPSS 17.0 software (SPSS Inc., Chicago, USA). BPNN, GRNN, and SVM model construction were conducted using the MATLAB 2014a software (MathWorks Inc., Natick, USA) installed in Windows XP.

It is essential to select appropriate input variables for data-driven model development (Muttil and Chau 2007). In previous studies, many different variables have been employed as model inputs for DO estimation (Antanasijević et al. 2013a; Najah et al. 2014; Singh et al. 2009). Correlation analysis (e.g., Pearson and Spearman method) can be performed to select initial input variables. In the present study, the data for all variables were not normally distributed (as determined by the Kolmogorov-Smirnov test), and therefore, the Spearman method was used to select initial input variables for these regression models. Significant correlation was observed between DO and other water quality parameters except for Temp (Table 2). However, water temperature should also be considered, since it can influence DO content in a water body (Cox 2003). Hence, all 11 water quality parameters were used as inputs for modeling DO concentration.

The change of water quality through time illustrates the seasonal variations. Therefore, a full hydrologic year or more should be considered for testing the features of DO changes. The collected water quality data were divided into two subsets, namely, training subset, which contained data from 2004 to 2007 (192 samples), and testing subset, which contained data from 2008 (48 samples). The training subset was used to establish the model, and the testing subset was used to test the accuracy of the established model.

Results of MLR model

The MLR, a classic linear regression method, is often used as a reference to evaluate other non-linear models. A linear regression model is composed of regression variables, regression coefficients, and dependent variables (Awchi 2014). Equation (15) was obtained and then utilized to predict the DO content in Wen-Rui Tang River.

$$\begin{aligned} DO = {} & -11.293 - 0.065Temp + 0.2669pH - 0.069COD_{Mn} \\ & - 0.034BOD_5 - 0.241NH_4^+ + 1.614Petrol - 0.643TP \\ & + 0.01COD_{Cr} - 2.868F - 0.021TN - 0.026EC \end{aligned} \quad (15)$$

The obtained minimum DO concentration may approach zero but cannot have a negative value. Unfortunately, some negative values were present in the outputs of the regression models (i.e., MLR, BPNN, GRNN, and SVM). These values serve no purpose as predictions. Here, the predicted negative values were adjusted to the minimum observed value (0.04 mg/L for training data set and 0.2 mg/L for testing data set). After changing the negative values, the performance of different regression models is presented in Table 3.

As shown in Table 3, the performance of the MLR model was insufficient for prediction purposes, with MSE = 2.9022 mg/L, $R^2 = 0.6797$ ($p < 0.01$), and NS = 0.6732 for the training set and MSE = 2.0652 mg/L, $R^2 = 0.6780$ ($p < 0.01$), and NS = 0.7286 for the testing set. Figure 4 presents the observed and predicted DO values obtained through the MLR model during the testing phase. It was clear that the observed and predicted values were not well superposed, and the differences between the predictions

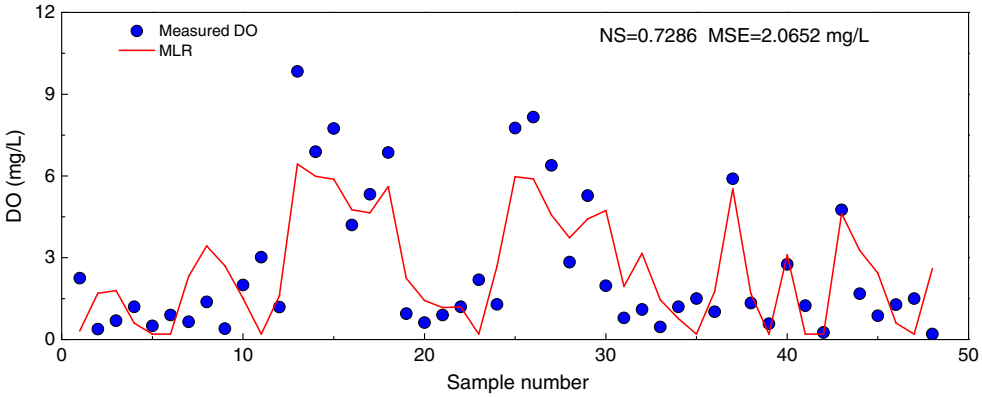

Fig. 4 Observed and predicted DO values by MLR model during testing stage

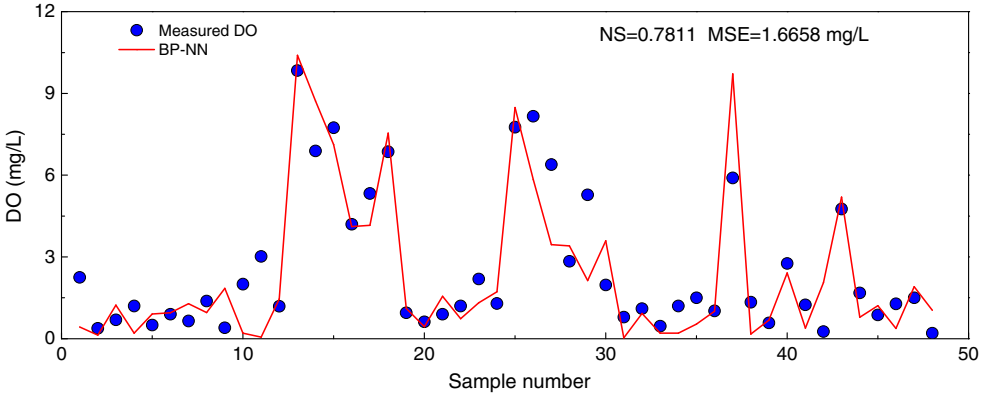

Fig. 5 Observed and predicted DO values by BPNN model during testing stage

and observed values were large. The overall results showed that the prediction accuracy of MLR model requires further improvement.

Results of BPNN model

In BPNN, the learning rate, target error goal MSE, and maximum numbers of epochs were set as 0.01, 0.0001, and 1000, respectively. The appropriate neuron number in the hidden layer must be selected in advance for model construction, because too many neurons may result in overfitting problem and insufficient neurons may lead to inadequate information capture by the model (Chen and Liu 2014; Wen et al. 2013). Fletcher and Goss (1993) suggested that the appropriate number of neurons in the hidden layer ranges from $(a + 2n^{1/2})$ to $(2n + 1)$, where a is the output neuron number and n indicates the input neuron number. Gradually varying the number of neurons in the hidden layer from 5 to 23 through trial and error was employed to seek the optimal number. Within this range, a BPNN model, having one input layer with 11 input variables, one hidden layer with 10 neurons, and one output layer with 1 output variable, was the optimal network for DO prediction.

The results of BPNN model are listed in Table 3. The BPNN model was superior to the MLR model in terms of predicting the DO values. MSE decreased from 2.0652 mg/L in the MLR model to 1.6658 mg/L in the BPNN model, whereas the R^2 and NS increased from 0.6780 and 0.7286 in the MLR model to 0.7854 and 0.7811 in the BPNN model, respectively. Figure 5 shows the observed and predicted DO values obtained through the BPNN model during the testing phase. The comparison results between Figs. 4 and 5 indicated that the BPNN model outperformed the MLR model in the prediction of DO content. The conclusion was consistent with that of previous research, in which the accuracy of the MLR model for predicting DO was inferior to the ANN models (Akkoyunlu et al. 2011; Antanasijević et al. 2013a; Wen et al. 2013). The findings can be explained as follows: the correlation among water quality parameters tends to be non-linear. The greatest problem in MLR methodology is the assumption of a linear input-output relationship, but this assumption is unacceptable for complex system. Conversely, a great advantage of ANNs or SVM models is their ability to model non-linear relationships.

Results of GRNN model

The smoothing factor, which has a significant effect on the prediction ability of the model, is the only required parameter to be determined in the GRNN model. An extremely large smoothing factor induces oversmoothing and will typically cause most of the input patterns to appear similar, and an extremely small smoothing factor cannot provide a smooth regression surface (Yaseen et al. 2016). In our study, an

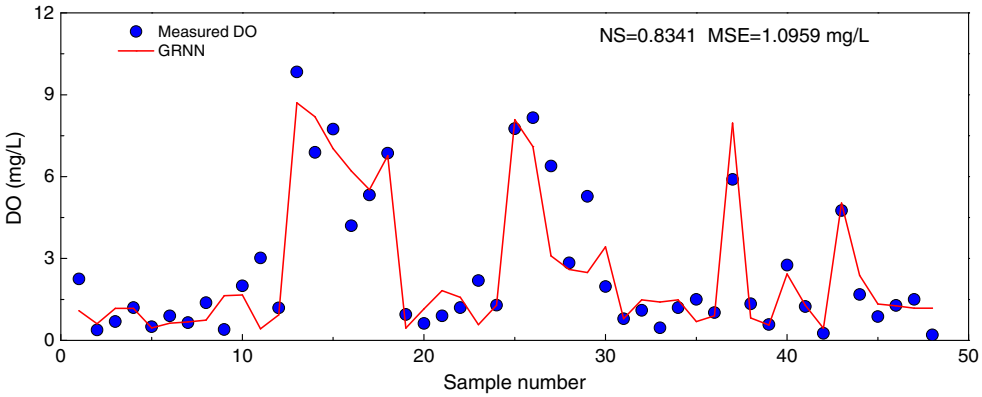

Fig. 6 Observed and predicted DO values by GRNN model during testing stage

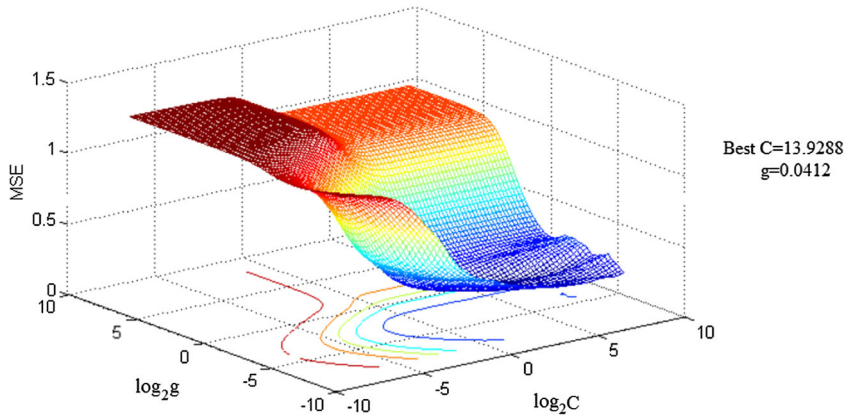

Fig. 7 Surface plot of optimization of SVM tuning parameters during grid search

interactive algorithm was used to determine the intermediate value of the smoothing factor within the range of 0.01–1 (Šiljić et al. 2015). Finally, the smoothing factor was determined as 0.7, and the results of GRNN model for DO prediction are listed in Table 3. Compared with the BPNN model, the GRNN model achieved better performance in predicting DO content. MSE decreased from 1.6658 mg/L in the BPNN model to 1.0959 mg/L in the GRNN model. By contrast, the R^2 increased from 0.7854 in the BPNN model to 0.8341 in the GRNN model, and NS increased from 0.7811 in the BPNN model to 0.8560 in the GRNN model. A visual assessment of the observed and predicted DO values in the testing phase depicts that the predictions of GRNN were closer to the measured DO values than those of BPNN (Figs. 5 and 6).

Results of SVM model

The SVM method proposed by Vapnik can simultaneously minimize estimation errors and model dimensions, has superior generalization and accurate prediction capabilities, and can prevent overfitting problems (Borin et al. 2006; Durand et al. 2007; Langeron et al. 2007; Pierna et al. 2004). During SVM model development, the determination of the optimal combination of C and g is greatly important in constructing high-performance regression models. C is the regularization parameter that controls the degree of empirical error in optimization problem, and g is the RBF kernel parameter that significantly affects the generalization ability of SVM (Noori et al. 2011; Singh et al. 2011). In the present study, the fivefold CV and grid search method were employed to determine the optimal pairwise C and g during the construction of the SVM model. Fivefold CV was used since it could yield lowest MSE (see Appendix Fig. 13). In fivefold CV, the samples of training data set are divided into five subsets of equal size. Then, four subsets are employed to construct the model, and the remaining subset is used for validation. Consequently, each instance of training data set is predicted once (Singh et al. 2014). The CV procedure can prevent overtraining problems, and the MSE value was served as the criterion for parameter optimization. The grid search algorithm divides the search scope of the parameters for optimization into grids and traverses all grid points to obtain the optimal value. The accuracy of grid search optimization depends on parameter range and interval size. Increase in parameter range and decrease in step size can enhance optimization accuracy (Wang et al. 2007). For this reason, a grid search over a space of $C \times g = \{(C, g) | C \in (2^{-8}, 2^{8}), g \in (2^{-8}, 2^{8})\}$ with step size of $2^{0.2}$ was used to perform this optimization. Figure 7 shows the three-dimensional (3D) view of the optimization results for pairwise parameters C and g obtained through the grid search method with fivefold CV. The combination of

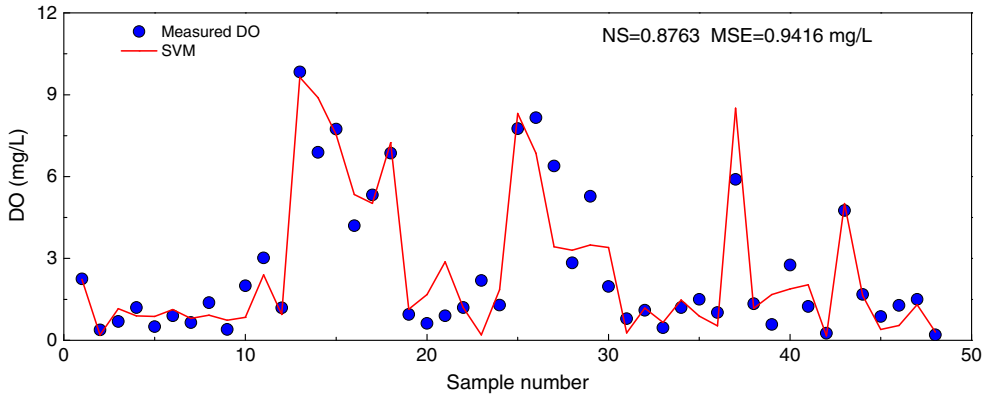

Fig. 8 Observed and predicted DO values by SVM model during the testing stage

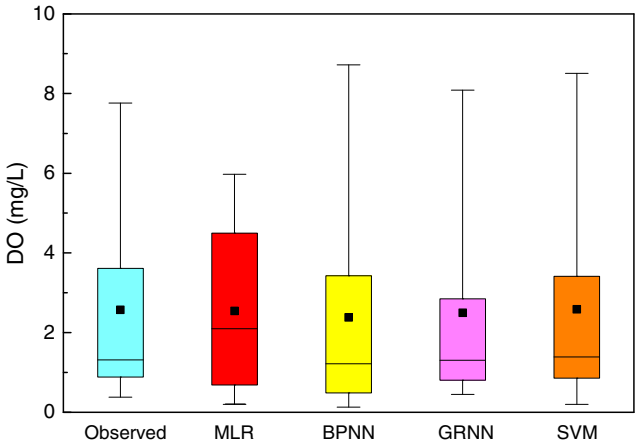

Fig. 9 Box plots of the observed DO compared with predicted DO from MLR, BPNN, GRNN, and SVM (*box plot* denotes 25th, 50th, and 75th percentiles; the *whiskers* indicate the 5th and 95th percentiles; the *small black boxes* denote mean value)

13.9288 (C) and 0.0412 (g), which had the lowest MSE, was determined as the optimal value. The SVM model was subsequently constructed with the optimal pairwise C and g.

Compared with the GRNN model, the SVM model performed better in the prediction of DO concentration in the training and testing data sets (Table 3). For the testing data set, MSE decreased from 1.0959 mg/L in the GRNN model to 0.9416 mg/L in the SVM model. By contrast, the R^2 increased from 0.8341 in the GRNN model to 0.8646 in the SVM model, and the NS increased from 0.8560 in the GRNN model to 0.8763 in the SVM model. The observed and predicted DO values obtained through the SVM model in the testing phase are illustrated in Fig. 8. It was clear from the figure that the SVM predictions were in fairly good agreement with the observations. By comparing Figs. 4, 5, 6, and 8, it appeared that the SVM estimates were closer to the corresponding observed DO values than those of the other models, indicating that the SVM model outperformed other models in terms of prediction accuracy.

With respect to MSE and NS values, the quantitative models can be arranged in the following order: SVM > GRNN > BPNN > MLR. The AI models are able to deal with highly non-linear, dynamic, and complex DO balance process, which always occurs in hypoxic river systems, and thus are superior to conventional linear statistical models (e.g., MLR). Among the three AI models, the SVM model has the best performance because of the following advantages and characteristics: (1) simultaneous minimization of prediction error and model complexity by introducing a kernel trick to the SVM, (2) prevention of overfitting problem based on the principle of structural risk minimization (SRM), and (3) obtainable satisfactory performance from sparse data such as the bimonthly water quality data used in this study.

To demonstrate how closely the SVM model predictions compared with the observed values and other regression models, the box plots were utilized (Fig. 9). According to the

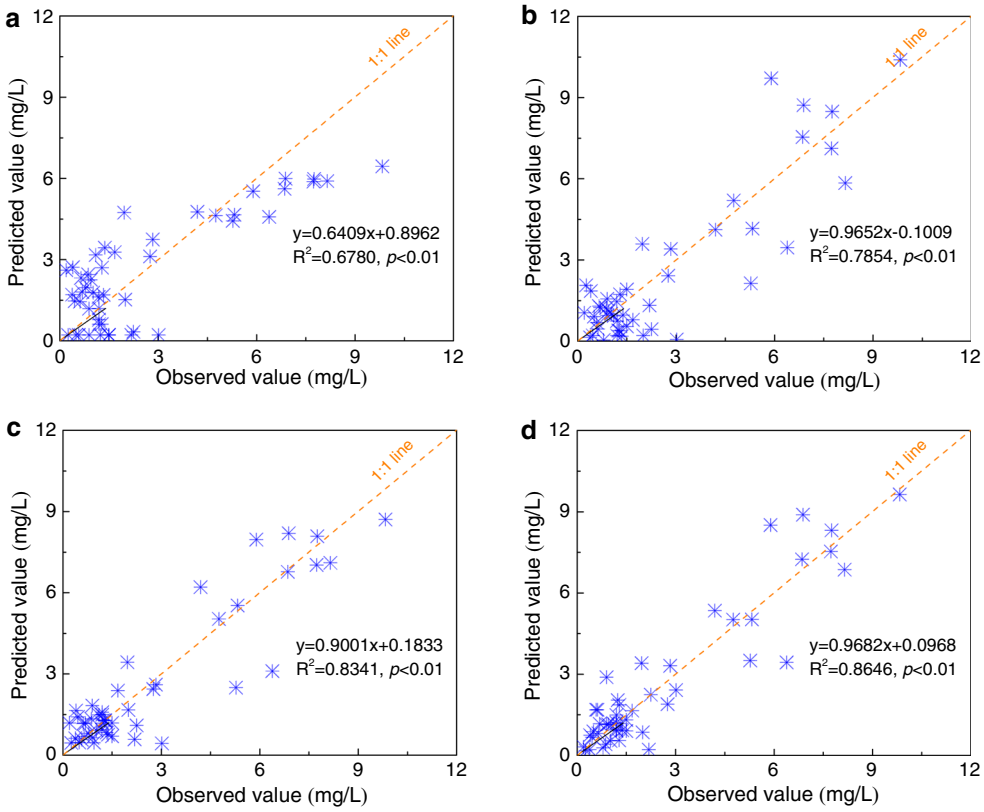

Fig. 10 Scatter plots of observed and predicted values obtained through **a** MLR, **b** BPNN, **c** GRNN, and **d** SVM models in the testing data set

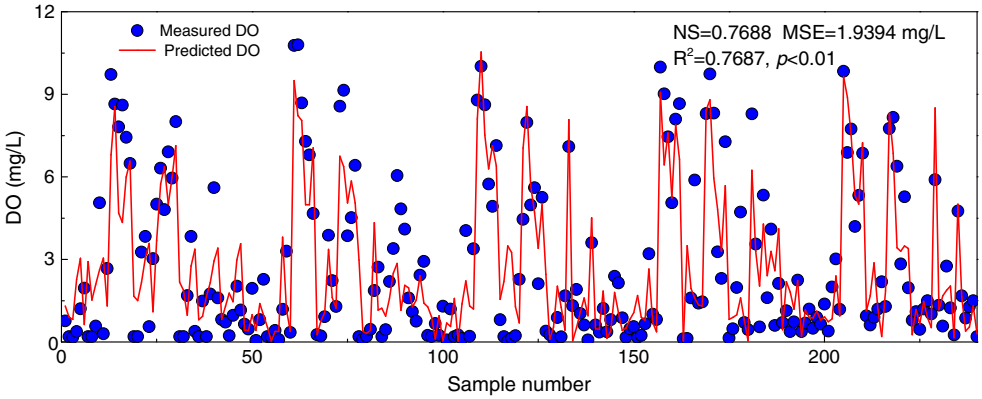

Fig. 11 Observed and predicted DO values by SVM model during 2004–2008

respective quartile values and whiskers, the degree of overall spread in the observed and predicted values can be obtained. From the figure, the spread of SVM-based predictions closely resembled observed DO values. A similar conclusion can also be drawn from Fig. 10, which shows the scatter plots of observed and predicted values obtained through MLR, BPNN, GRNN, and SVM in the testing data set. The 1:1 line denotes ideal results. In particular, the closer the data points are to this line, the better the result of the model produces (Yang et al. 2017). Compared with MLR, BPNN, and GRNN, SVM had the best fitting effect since all the points in this model clustered closely to the 1:1 line. Overall, the SVM model can be used as a superior alternative to ANNs (e.g. BPNN and GRNN) and conventional linear model (e.g., MLR) for the prediction of DO concentration in Wen-Rui Tang River.

The results of the SVM model will vary if the training and testing data sets are changed. Within the water quality data from 2004 to 2008, each year was selected as the testing set and the other 4 years were divided into individual training sets. Consequently, five different input and output combinations were used for the development of SVM and each year was predicted once. Figure 11 illustrates the observed and predicted DO values during 2004–2008. The small MSE value (1.9394 mg/L) as well as the high values of R^2 (0.7687) and NS (0.7688) implied that the SVM model can provide a genuinely good performance for DO simulation in the Wen-Rui Tang River.

Sensitivity analysis

To evaluate the contribution of each independent input parameter to the DO, sensitivity analysis was performed. In this study, sensitivity analysis was carried out by constructing 11 SVM models each use single water quality parameter as input variable (Dogan et al. 2009). The effects of each parameter were assessed on the basis of the MSE during the test stage, and the results are shown in Fig. 12. This analysis indicated that NH_4^+-N with the lowest MSE had the greatest influence on DO levels in Wen-Rui Tang River. Other input variables, which provided comparably high contributions to DO, were TN, TP, COD_{Cr}, BOD_5, and COD_{Mn}. These findings can be explained according to the following: First, decomposition of large amounts of nitrogen compounds from untreated domestic sewage and industrial wastewater consumes large amounts of DO and produces NH_4^+-N. Second, nutrients (e.g., NH_4^+-N, TN, and TP) significantly affect aquatic plants, algae, and microorganisms, which closely relate with DO concentration in water body through photosynthesis, respiration, and decomposition. Besides, DO consumption from organic matter

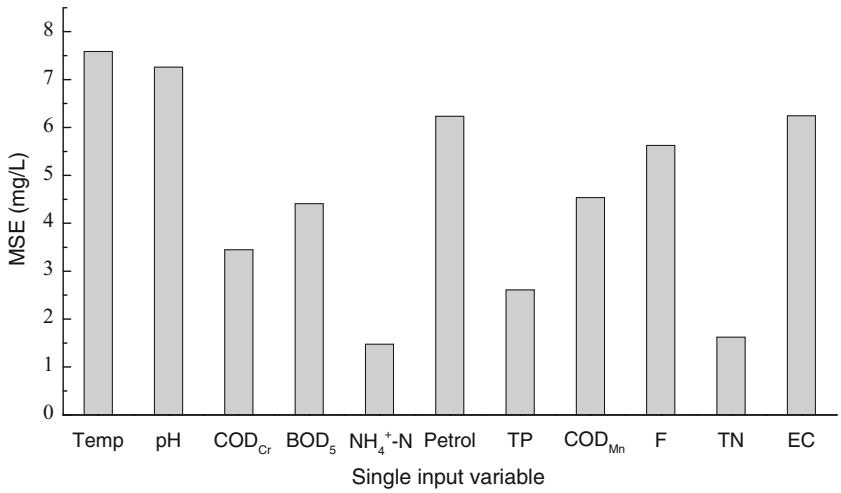

Fig. 12 SVM model efficiency for single input

decomposition is commonly measured as BOD and COD, so COD_{Cr}, BOD_5, and COD_{Mn} have high contributions to DO. Based on these results, the reduction of the concentration of organic pollution parameters (e.g., NH_4^+-N, COD_{Cr}, BOD_5, and COD_{Mn}) is essential for the improvement and remediation of DO level in Wen-Rui Tang River. These desired results may be achieved through the establishment of water treatment plants and sewage pipeline systems in urban and suburban areas and application of best management practices (e.g., constructed wetland, riverbank stabilization, and buffer strips) in rural areas.

Conclusions

In this paper, we investigated the performance of the SVM model in estimating DO concentration in a hypoxic river system. The prediction accuracy of SVM was compared with those of MLR, BPNN, and GRNN models. All the prediction models were established on the basis of the hydro-chemical data of Wen-Rui Tang River for the 2004–2008 period. MSE, R^2, and NS were used as criteria to evaluate the performance of each model. Among the four models, SVM exhibited the optimal performance in predicting DO concentration. Compared with MLR, BPNN, and GRNN models, the MSE of SVM model decreased by 54, 43, and 14%, respectively. Sensitivity analysis showed that input variables with significant effects on the proposed SVM model were in the following descending order: NH_4^+-N, TN, TP, COD_{Cr}, BOD_5, and COD_{Mn}. The results of this study can serve as a reference for the estimation of other water quality parameters (e.g., COD and BOD) in hypoxic river systems.

Although excellent prediction accuracy was achieved in this study, it should be noted that only hydro-chemical data were considered as input variables. Other parameters regarding weather (e.g., precipitation and wind speed), hydrology (e.g., flow velocity and flux), and social economy (e.g., gross domestic product, population density, and municipal waste generation) were not taken into account here because the corresponding data were unavailable. In the future, various data will be required to reinforce the conclusions drawn from this paper. The sensitivity analysis results may be inconsistent for each single monitoring site which belongs to different land use zones. Thus, further investigations are necessary to explore the contributions of each input parameter on the developed model at each monitoring site. Furthermore, it is also recommended to determine the uncertainty of the SVM model for DO since the prediction results, in fact, are not necessarily certain due to the algorithm characteristics.

Acknowledgements This work was supported by the Science and Technology Department of Zhejiang Province (2008C03009). The authors would like to thank the Wenzhou Environmental Protection Bureau (WEPB) for the data provided for the Wen-Rui Tang River.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Funding This study was funded by the Science and Technology Department of Zhejiang Province (grant number 2008C03009).

Appendix

Table 4 DO concentration (mg/L) in previous studies

Author(s)	Locations	The period of water quality data	DO concentration (mean ± SD)
Singh et al. (2009)	Gomti (India)	1994.1–1999.12, 2002.1–2005.12	5.61 ± 2.66
Ranković et al. (2010)	Gruža Reservoir (Serbia)	2000–2003	5.01 ± 3.53
Antanasijević et al. (2013a)	Danube River (Europe)	2004–2009	Train data set, 10.9; test data set, 10.3
Wen et al. (2013)	Heihe River (China)	2003–2008	7.84 ± 1.72
Antanasijević et al. (2014a)	Danube River (Europe)	2002–2010	10.0 ± 2.1
Chen and Liu (2014)	Feitsui Reservoir, Taiwan (China)	1993–2011	7.14 ± 1.11
Najah et al. (2014)	Johor River (Malaysian)	1998–2007	Site 1, 7.34 ± 0.59
			Site 2, 6.96 ± 0.37
			Site 3, 6.78 ± 0.66
			Site 4, 7.53 ± 0.61
Nemati et al. (2015)	Tai Po River, Hong Kong (China)	1991–2011	8.18 ± 1.12

SD standard deviation

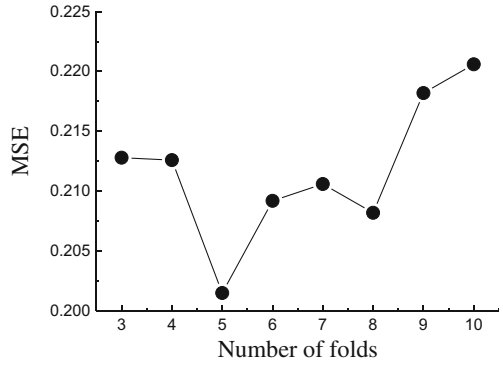

Fig. 13 SVM model efficiency with different numbers of folds in cross validation

References

Abrahart RJ, Anctil F, Coulibaly P, Dawson CW, Mount NJ, See LM, Shamseldin AY, Solomatine DP, Toth E, Wilby RL (2012) Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. Prog Phys Geogr 36:480–513

Akkoyunlu A, Altun H, Cigizoglu HK (2011) Depth-integrated estimation of dissolved oxygen in a lake. J Environ Eng 137(10):961–967

Alilou VK, Yaghmaee F (2015) Application of GRNN neural network in non-texture image inpainting and restoration. Pattern Recogn Lett 62:24–31

Antanasijević DZ, Pocajt VV, Povrenović DS, Perić-Grujić AA, Ristić MD (2013a) Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. Environ Sci Pollut Res 20(12):9006–9013

Antanasijević DZ, Pocajt VV, Povrenović DS, Ristić MD, Perić-Grujić AA (2013b) PM_{10} emission forecasting using artificial neural networks and genetic algorithm input variables optimization. Sci Total Environ 443:511–519

Antanasijević DZ, Ristić MD, Perić-Grujić AA, Pocajt VV (2013c) Forecasting human exposure to PM_{10} at the national level using an artificial neural network approach. J Chemom 27(6):170–177

Antanasijević DZ, Pocajt VV, Povrenović DS, Perić-Grujić AA, Ristić MD (2014a) Modelling of dissolved oxygen content Danube River using artificial neural networks and Monte Carlo simulation uncertainty analysis. J Hydrol 519:1895–1907

Antanasijević DZ, Ristić MD, Perić-Grujić AA, Pocajt VV (2014b) Forecasting GHG emissions using an optimized artificial neural network model based on correlations and principal component analysis. Int J Greenh Gas Con 20(5):244–253

Awchi TA (2014) River discharges forecasting in northern Iraq using different ANN techniques. Water Resour Manag 28(3):801–814

Basant N, Gupta S, Malik A, Singh KP (2010) Linear and non-linear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water—a case study. Chemometr Intell Lab 104(2):172–180

Borah DK, Bera M (2004) Watershed-scale hydrologic and nonpoint-source pollution models: reviews of application. T ASAE 47(3): 789–803

Borin A, Ferrão MF, Mello C, Maretto DA, Poppi RJ (2006) Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. Anal Chim Acta 579(1):25–32

Bowes MJ, Neal C, Jarvie HP, Smith JT, Davies HN (2010) Predicting phosphorus concentrations in British rivers resulting from the introduction of improved phosphorus removal from sewage effluent. Sci Total Environ 408:4239–4250

Chapra S, Pellettier G (2003) QUAL2K: a modeling framework for simulating river and stream water quality. Civil and environmental engineering dept., Tufts University, Medford

Chau KW (2006) A review on integration of artificial intelligence into water quality modeling. Mar Pollut Bull 52:726–733

Chau KW, Wu CL (2010) A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. J Hydroinf 12(4):458–473

Chen WB, Liu WC (2014) Artificial neural network modeling of dissolved oxygen in reservoir. Environ Monit Assess 186(2):1203–1217

Chen JB, Li FY, Fan ZP, Wang YJ (2016) Integrated application of multivariate statistical methods to source apportionment of watercourses in the Liao River Basin, Northeast China. Int J Environ Res Public Health 13(10):1035–1061

Cigizoglu HK, Alp M (2006) Generalized regression neural network in modelling river sediment yield. Adv Eng Softw 37(2):63–68

Cox BA (2003) A review of currently available in-stream-water-quality models and their applicability for simulating dissolved oxygen in lowland rivers. Sci Total Environ 314–316:335–377

Cristianine N, Taylor JS (2000) An introduction to support vector machine and other kernel based learning methods. Cambridge University Press, Cambridge

Dibike YB, Velickov S, Solomatine DP, Abbott MB (2001) Model induction with support vector machines: introduction and applications. J Comput Civ Eng 15:208–216

Dogan E, Sengorur B, Koklu R (2009) Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. J Environ Manag 90:1229–1235

Durand A, Devos O, Ruckebusch C, Huvenne JP (2007) Genetic algorithm optimization combined with partial least squares regression and mutual information variable selection procedures in near-infrared quantitative analysis of cotton–viscose textiles. Anal Chim Acta 595(1–2):72–79

Fletcher D, Goss E (1993) Forecasting with neural networks: an application using bankruptcy data. Inform Manage-Amster 24:159–167

Gao C, Zhang TL (2010) Eutrophication in a Chinese context: understanding various physical and socio-economic aspects. Ambio 39: 385–393

Gupta DA (2008) Implication of environmental flows in river basin management. Phys Chem Earth 33(5):298–303

Hagan MT, Demuth HP, Beale M (1996) Neural network design. PWS Publishing, Boston, MA

Haykin S (1999) Neural networks: a comprehensive foundation. Macmillan, New York

He ZB, Wen XH, Liu H, Du J (2014) A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J Hydrol 509:379–386

Heddam S (2014a) Generalized regression neural network-based approach for modelling hourly dissolved oxygen concentration in the Upper Klamath River, Oregon, USA. Environ Technol 35(13): 1650–1657

Heddam S (2014b) Modelling hourly dissolved oxygen concentration (DO) using dynamic evolving neural-fuzzy inference system (DENFIS)-based approach: case study of Klamath River at Miller Island Boat Ramp OR USA. Environ Sci Pollut Res 21(15):9212–9227

Hsu CW, Chang CC, Lin CJ (2007) A practical guide to support vector classification. URL<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

Keerthi SS, Lin CJ (2001) Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Comput 15:1667–1689

Khan SM, Coulibaly P (2006) Application of support vector machine in lake water level prediction. J Hydrol Eng 11(3):199–205

Kim S, Kim HS (2008) Neural networks and genetic algorithm approach for non-linear evaporation and evapotranspiration modeling. J Hydrol 351(3–4):299–317

Kisi O, Ozkan C, Akay B (2012) Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm. J Hydrol 428–429:94–103

Kuo JT, Hsieh MH, Lung WS, She N (2007) Using artificial neural network for reservoir eutrophication prediction. Ecol Model 200(1–2):171–177

Langeron Y, Doussot M, Hewson DJ, Duchêne J (2007) Classifying NIR spectra of textile products with kernel methods. Eng Appl of Artif Intel 20(3):415–427

Legates DR, McCabe GJ Jr (1999) Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. Water Resour Res 35(1):233–241

Lin JY, Cheng TC, Chau KW (2006) Using support vector machines for long term discharge prediction. Hydrolog Sci J 51(4):599–612

Lin SW, Lee ZJ, Chen SC, Tseng TY (2008) Parameter determination of support vectormachines and feature selection using simulated annealing approach. Appl Soft Comput 8(4):1505–1512

Mei K, Zhu YL, Liao LL, Dahlgren RA, Shang X, Zhang MH (2011) Optimizing water quality monitoring networks using continuous longitudinal monitoring data: a case study of Wen-Rui Tang River, Wenzhou, China. J Environ Monit 13:2755–2762

Mohammadpour R, Shaharuddin S, Chang CK, Zakaria NA, Ghani AA, Chan NW (2015) Prediction of water quality index in constructed wetlands using support vector machine. Environ Sci Pollut Res 22: 6208–6219

Morse NB, Wollheim WM (2014) Climate variability masks the impacts of land use change on nutrient export in a suburbanizing watershed. Biogeochemistry 121(1):45–59

Muttil N, Chau KW (2007) Machine-learning paradigms for selecting ecologically significant input variables. Eng Appl Artif Intell 20(6):735–744

Najah A, El-Shafie A, Karim OA, El-Shafie AH (2014) Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring. Environ Sci Pollut Res 21(13):1658–1670

Nemati S, Fazelifard MH, Terzi Ö, Ghorbani MA (2015) Estimation of dissolved oxygen using data-driven techniques in the Tai Po River, Hong Kong. Environ Earth Sci 74(5):1–9

Noori R, Karbassi AR, Moghaddamnia A, Han D, Zokaei-Ashtiani MH, Farokhnia A, Ghafari Gousheh M (2011) Assessment of input variables determination on the SVM model performance using PCA, gamma test, and forward selection techniques for monthly stream flow prediction. J Hydrol 401:177–189

Noori R, Yeh HD, Abbasi M, Kachoosangi FT, Moazami S (2015) Uncertainty analysis of support vector machine for online prediction of five-day biochemical oxygen demand. J Hydrol 527(6):833–843

Pan Y, Jiang J, Wang R, Cao H (2008) Advantages of support vector machine in QSPR studies for predicting auto-ignition temperatures of organic compounds. Chemometr Intell Lab Syst 92(2):169–178

Pierna JAF, Baeten V, Renier AM, Cogdill RP, Dardenne P (2004) Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds. J Chemom 18(7–8):341–349

Qu J, Zuo MJ (2010) Support vector machine based data processing algorithm for wear degree classification of slurry pump systems. Measurement 43(6):781–791

Quinn NTW, Jacobs K, Chen KW, Stringfellow WT (2005) Elements of decision support system for real-time management of dissolved oxygen in the San Joaquin River deep water ship channel. Environ Model Softw 20(12):1495–1504

Raghavendra NS, Deka PC (2014) Support vector machine applications in the field of hydrology: a review. Appl Soft Comput 19:372–386

Ranković V, Radulović J, Radojević I, Ostojić A, Čomić L (2010) Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. Ecol Model 221(8):1239–1244

Šiljić A, Antanasijević D, Perić-Grujić A, Ristić M, Pocajt V (2015) Artificial neural network modeling of biological oxygen demand in rivers at the national level with input selection based on Monte Carlo simulations. Environ Sci Pollut Res 22(6):4230–4241

Singh KP, Basant A, Malik A, Jain G (2009) Artifical neural network modeling of the river water quality—a case study. Ecol Model 220(6):888–895

Singh KP, Basant N, Gupta S (2011) Support vector machines in water quality management. Anal Chim Acta 703(2):152–162

Singh KP, Gupta S, Rai P (2014) Predicting dissolved oxygen concentration using kernel regression modeling approaches with non-linear hydro-chemical data. Environ Monit Assess 186:2749–2765

State Environment Protection Bureau of China (2002a) Environmental quality standards for surface water. China Environmental Science Press, Beijing (in Chinese)

State Environment Protection Bureau of China (2002b) Water and wastewater analysis method. China Environmental Science Press, Beijing (in Chinese)

Tetratech, Inc (2002) Draft user's manual for environmental fluid dynamics code Hydro Version (EFDC-Hydro). U.S. Environmental Protection Agency, Atlanta, GA

Vapnik VN (1995) The nature of statistical learning theory. Springer, New York

Vapnik VN (1998) Statistical learning theory. Wiley, New York

Wang J, Du HY, Liu HX, Yao XJ, Hu ZD, Fan BT (2007) Prediction of surface tension for common compounds based on novel methods using heuristic method and support vector machine. Talanta 73(1): 147–156

Wen XH, Fang J, Diao MN, Zhang XQ (2013) Artificial neural networks modeling of dissolved oxygen in the Heihe River, Northwestern China. Environ Monit Assess 185(5):4361–4371

Wool TA, Ambrose RB, Martin JL, Comer EA (2006) Water quality analysis simulation program (WASP) version 6.0 draft: user's manual. US Environmental Protection Agency, Athens, GA.

Wu CL, Chau KW, Li YS (2009) Methods to improve neural network performance in daily flows prediction. J Hydrol 372(1–4):80–93

Xie JX, Cheng CT, Chau KW, Pei YZ (2006) A hybrid adaptive time-delay neural network model for multi-step-ahead prediction of sunspot activity. Int J Environ Pollut 28(3–4):364–381

Yang L, Mei K, Liu X, Wu L, Zhang M, Xu J, Wang F (2013) Spatial distribution and source apportionment of water pollution in different administrative zones of Wen-Rui-Tang (WRT) river watershed, China. Environ Sci Pollut Res 20(8):1–12

Yang Y, Liu XS, Li WL, Jin Y, Wu YJ, Zheng JY, Zhang WT, Chen Y (2017) Rapid measurement of epimedin A, epimedin B, epimedin C, icariin, and moisture in Herba Epimedii using near infrared spectroscopy. Spectrochim Acta Part A Mol Biomol Spectrosc 171:351–360

Yaseen ZM, Jaafar O, Deo RC, Kisi O, Adamowski J, Quilty J, EI-Shafie A (2016) Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. J Hydrol 542:603–614

Yoon H, Jun SC, Hyun Y, Bae GO, Lee KK (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J Hydrol 396(1–2):128–138