CrossMark

**RESEARCH ARTICLE**

# Extreme learning machines: a new approach for modeling dissolved oxygen (DO) concentration with and without water quality variables as predictors

Salim Heddam[1] · Ozgur Kisi[2]

**Abstract** In this paper, several extreme learning machine (ELM) models, including standard extreme learning machine with sigmoid activation function (S-ELM), extreme learning machine with radial basis activation function (R-ELM), online sequential extreme learning machine (OS-ELM), and optimally pruned extreme learning machine (OP-ELM), are newly applied for predicting dissolved oxygen concentration with and without water quality variables as predictors. Firstly, using data from eight United States Geological Survey (USGS) stations located in different rivers basins, USA, the S-ELM, R-ELM, OS-ELM, and OP-ELM were compared against the measured dissolved oxygen (DO) using four water quality variables, water temperature, specific conductance, turbidity, and pH, as predictors. For each station, we used data measured at an hourly time step for a period of 4 years. The dataset was divided into a training set (70%) and a validation set (30%). We selected several combinations of the water quality variables as inputs for each ELM model and six different scenarios were compared. Secondly, an attempt was made to predict DO concentration without water quality

variables. To achieve this goal, we used the year numbers, 2008, 2009, etc., month numbers from (1) to (12), day numbers from (1) to (31) and hour numbers from (00:00) to (24:00) as predictors. Thirdly, the best ELM models were trained using validation dataset and tested with the training dataset. The performances of the four ELM models were evaluated using four statistical indices: the coefficient of correlation ($R$), the Nash-Sutcliffe efficiency (NSE), the root mean squared error (RMSE), and the mean absolute error (MAE). Results obtained from the eight stations indicated that: (i) the best results were obtained by the S-ELM, R-ELM, OS-ELM, and OP-ELM models having four water quality variables as predictors; (ii) out of eight stations, the OP-ELM performed better than the other three ELM models at seven stations while the R-ELM performed the best at one station. The OS-ELM models performed the worst and provided the lowest accuracy; (iii) for predicting DO without water quality variables, the R-ELM performed the best at seven stations followed by the S-ELM in the second place and the OP-ELM performed the worst with low accuracy; (iv) for the final application where training ELM models with validation dataset and testing with training dataset, the OP-ELM provided the best accuracy using water quality variables and the R-ELM performed the best at all eight stations without water quality variables. Fourthly, and finally, we compared the results obtained from different ELM models with those obtained using multiple linear regression (MLR) and multilayer perceptron neural network (MLPNN). Results obtained using MLPNN and MLR models reveal that: (i) using water quality variables as predictors, the MLR performed the worst and provided the lowest accuracy in all stations; (ii) MLPNN was ranked in the second place at two stations, in the third place at four stations, and finally, in the fourth place at two stations, (iii) for predicting DO without water quality variables, MLPNN is ranked in the second place at five stations, and ranked in the third, fourth,

✉ Salim Heddam
heddamsalim@yahoo.fr

Ozgur Kisi
ozgur.kisi@iliauni.edu.ge

[1]  Faculty of Science, Agronomy Department, Hydraulics Division University, 20 Août 1955, Route El Hadaik, BP 26 Skikda, Algeria

[2]  School of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia

and fifth places in the remaining three stations, while MLR was ranked in the last place with very low accuracy at all stations. Overall, the results suggest that the ELM is more effective than the MLPNN and MLR for modelling DO concentration in river ecosystems.

**Keywords** Modeling · Dissolved oxygen · Extreme learning machine · OP-ELM · OS-ELM · S-ELM · R-ELM · MLPNN · MLR

## Introduction

Dissolved oxygen concentration (DO) is the amount of molecular oxygen dissolved in water and plays a crucial role in determination of water quality. Among all other water quality variables that must be regularly measured and monitored, DO is practically the most important. The importance given to DO concentration is not only for itself, but also for its influence on so many other water quality variables and several other processes. DO provides regular information about water quality in rivers, streams, and lakes and is required for the survival of the aerobic aquatic organisms. Understanding the variation of DO is therefore a major priority for water resource managers. In addition to the traditional methods used for measuring DO concentration, previous attempts have been conducted to propose alternative methods to estimate DO concentrations mainly based on mathematical modeling.

To date, several researchers worldwide have attempted to develop effective and robust models for DO concentration. These studies can be divided into two categories: (i) estimation models and (ii) forecasting models. For the first category, the following three subgroups exist: (1) artificial neural network (ANNs) models, including multilayer perceptron neural network (MLPNN), radial basis function neural network (RBFNN), generalized regression neural network (GRNN), cascade correlation artificial neural networks (CCNN), and Elman neural networks (ELNN) (Akkoyunlu et al. 2011; Ay and Kisi 2012; Heddam 2014a, 2016a; Diamantopoulou et al. 2007; Liu et al. 2012); (2) fuzzy logic and neuro-fuzzy models (Ranković et al. 2012; Heddam 2014b; Ay and Kisi 2016); (3) evolutionary models (Kisi et al. 2013; Heddam 2014c, 2016b); (4) artificial neural network and wavelet decomposition conjunction models (Xu and Liu 2013; Evrendilek and Karakaya 2014a, b, 2015); (5) support vector machines (SVM) (Liu et al. 2013; Liu et al. 2014). For the second category, the objective is the development of models capable estimating DO concentrations using only antecedent values of DO without inclusion of water quality variables as inputs (Alizadeh and Kavianpour 2015; Wang et al. 2013; Faruk 2010; Heddam 2016b).

Akkoyunlu et al. (2011) compared MLPNN and RBFNN to estimate DO at different depths in Lake Ecosystem in Turkey. Ay and Kisi (2012) attempted to estimate DO at the

upstream station of Foundation Creek, Colorado, USA, using data from the downstream and vice versa, using two models; MLPNN and RBFNN. Heddam (2014a) demonstrated that GRNN model was a good and powerful tool for estimating hourly DO concentration. Heddam (2016a) applied RBFNN for estimating DO using four water quality variables. Diamantopoulou et al. (2007) firstly presented a CCNN model for predicting DO at a monthly time step. The application of the ELNN was reported by Liu et al. (2012). In all the above studies, the authors selected several water quality variables as inputs at different time steps, from hourly to monthly. Adaptive neuro-fuzzy inference system (ANFIS) is another well-known model that widely reported as a powerful tool for modeling DO concentration. Ranković et al. (2012) were firstly demonstrated that the ANFIS model could be successfully applied for modeling DO concentration using monthly data from Gruža Reservoir, Serbia. Heddam (2014b) and Ay and Kisi (2016) compared two ANFIS models, (i) grid partition-based fuzzy inference system, named ANFIS_GRID, and (ii) subtractive clustering-based fuzzy inference system, named ANFIS_SUB, for estimating DO, at hourly and monthly time steps, respectively.

Few studies in the literature have focused on the application of the evolutionary models for modeling DO concentration. Kisi et al. (2013) demonstrated that gene expression programming (GEP) provided better results compared to the standard MLPNN and ANFIS models. Heddam (2014c) applied two dynamic evolving neural-fuzzy inference systems, (i) offline-based system named DENFIS-OF and (ii) online-based system, named DENFIS-ON, for modeling DO at an hourly time step, and demonstrated that the DENFIS were the best in comparison to the MLPNN and multiple linear regression (MLR) models. Heddam (2016b) proposed evolving fuzzy neural network (EFuNN) for estimating DO and demonstrated that the EFuNN was more accurate than the MLR model.

Combined wavelet transforms (WT) and ANNs and sometimes MLR models, is another kind of approach adopted in the literature for estimating DO concentration. Meanwhile, Xu and Liu (2013) applied a wavelet neural network (WNN) conjunction models for DO estimation using monthly data from Jishan Lake in Duchang County, Jiangxi Province, China. Evrendilek and Karakaya (2014a) selected the orthogonal wavelet families for denoising time series of water quality variables, and applied three ANN models for estimating DO, at Yenicaga Lake, Bolu, in Turkey. Evrendilek and Karakaya (2014b, 2015) applied DWT and MLR conjunction models for estimating DO concentration using data from Lake Abant, Turkey. Support vector machines (SVM) based on kernel function is largely used in the area of water science; however, it is relatively a new form of models used for modeling DO. Liu et al. (2013) compared least squares support vector regression (LSSVR), SVM and MLPNN for modeling DO in river crab culture ponds,

China, and demonstrated that LSSVR was the best. In another study, Liu et al. (2014) demonstrated that particle swarm optimization (PSO) significantly improves the predictive capabilities of the LSSVR model in predicting DO. A particular type of model called the extended stochastic harmonic analysis (ESHA) was developed by Abdul-Aziz and Ishtiaq (2014) and compared with the classical harmonic analysis (CHA) for estimating hourly DO concentration, and found that the ESHA model was better when compared to CHA model.

Forecasting DO concentration models were rarely presented in the literature. This kind of models adopt antecedent values as input to forecast the future DO value at a specific horizon. Faruk (2010) investigated the capabilities of the MLPNN in forecasting monthly DO concentration in the Buyuk Menderes basin, Turkey. Wang et al. (2013) applied three models, namely, bootstrapped wavelet neural network (BWNN), MLPNN, and WNN, for forecasting monthly DO. Alizadeh and Kavianpour (2015) compared two forecasting models: MLPNN and WNN models in DO forecasting. The authors demonstrated that using four antecedent values of DO as inputs, the model provided better results at hourly time step when compared to the model at daily time step. Heddam (2016c) used the OP-ELM for forecasting DO in 7 days ahead, using previous values of DO measured at an hourly time step.

An analysis of the works reported above revealed that most of the works focused on the prediction of DO without water quality variables (category of forecasting models), is rarely adopted. Recently, Heddam (2016d) proposed a new kind of model for predicting DO without water variables. The author compared RBFNN and MLR for modeling DO using the components of the Gregorian calendar that are: (i) year number, (ii) number of months from 1 to 12, (iii) number of days from 1 to 31, and (iv) number of hours from 0:00 to 24:00. To the best of the authors' knowledge, there is no such study in the literature which examines the potential of extreme learning machines (ELM)-based approach, for modeling DO concentration with water quality variables as predictor. Hence, this study aims to introduce new kind of models for estimating DO concentration in river. The proposed ELM models were first applied for estimating DO using four water quality variables as inputs, and second, they were applied for forecasting DO without water quality variables and using only the component of the Gregorian calendar as input. Finally, we investigated whether the proposed ELM models can outperform the well-known multilayer perceptron neural network (MLPNN) and multiple linear regression (MLR) approaches for modeling DO concentration with and without water quality variables as predictors.

## Methods

ELM is a new data-driven model, first presented by Huang et al. (2004) and developed later by Huang et al. (2006a, b).

ELM has been introduced to improve the learning ability of the standard single hidden layer feedforward neural networks (SLFN). As a result, several attempts have been made over the past few years, with success, to improve the original algorithm; consequently, some changes were made and several types of ELM have been introduced. The present work aims to compare four ELM categories: the standard ELM with sigmoid activation function called S-ELM, the ELM with radial basis activation function (R-ELM), the online sequential extreme learning machine (OS-ELM), and the optimally pruned extreme learning machine (OP-ELM) for the first time. The four ELM models were also compared with the standard multilayer perceptron neural network (MLPNN) and multiple linear regression (MLR) models. Detailed descriptions of the proposed models are presented below.

### Extreme learning machines

In order to improve the performance and the learning ability of the SLFN, Huang et al. (2006a, b) introduced the ELM as a new learning algorithm. In ELM, parameters of hidden nodes are randomly selected and the output weights are analytically determined using least squares method (Huang et al. 2006a, b; Huang et al. 2012). Details of the mathematical formulation of the ELM can be summarized as follows (Fig. 1):

For $N$ arbitrary training samples represented by $(x_i, y_i)$, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{iD}]^T \in R^D$ and $[y_{i1}, y_{i2}, \ldots, y_{iD}]^T \in R^D$; The output of an SLFN with $N$ hidden nodes can be represented by (Huang et al. 2006a, b):

$$f_N(x) = \sum_{i=1}^{N} \beta_i f_i(x_j) = \sum_{i=1}^{N} \beta_i f_i(w_i \cdot x_j + b_i)$$
$$= o_j \quad x \in R^n, w_i \in R^n \tag{1}$$

where $w_i$ is the weight vector connecting the input layer to the $i$th hidden node, $\beta_i$ is the weight connecting the $i$th hidden node to the output node, and $f$ is the hidden node's activation function. Consequently, $(w_i \times x)$ denotes the inner product of
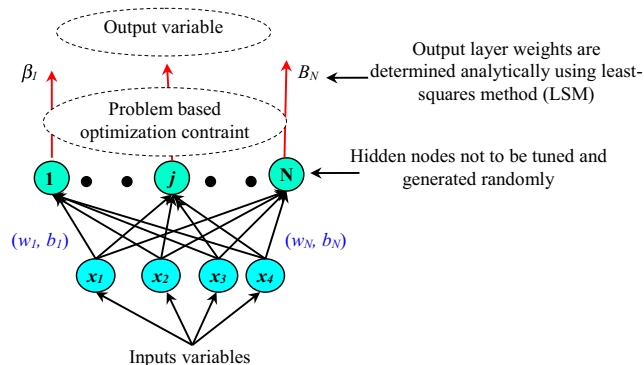


**Fig. 1** ELM structure (Huang 2015)

vectors $w_i$ and the input $x$. Here, $w_i = [w_{i1}, w_{i2}, \ldots, w_{iD}]^T$ is the weight vector connecting the $i$th hidden node to the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{ik}]^T$ is the weight vector connecting the $i$th hidden node to the output nodes, and $b_i$ is the threshold of the $i$th hidden node. The standard SLFN with N hidden nodes with activation function $f(x)$ can be written as (Huang et al. 2015):

$$H\beta = T \tag{2}$$

where

$$H = \begin{bmatrix} f_1(w_1 \cdot x_1 + b_1) & \cdots & f_N(w_N \cdot x_1 + b_N) \\ \vdots & \vdots & \vdots \\ f_1(w_1 \cdot x_N + b_1) & \cdots & f_N(w_N \cdot x_N + b_N) \end{bmatrix} \tag{3}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix} \tag{4}$$

ELM simply solves the function by:

$$\beta = H^+ T \tag{5}$$

where $H^+$ is the Moore-Penrose generalized inverse of matrix $H$. In the present study, we applied two types of ELM; S-ELM with sigmoid activation function and R-ELM with radial basis activation function. Two activation functions are presented below (Huang 2015; Liang et al. 2006).

The sigmoid function:

$$f(w, b, x) = \frac{1}{1 + e^{-(w \cdot x + b)}} \tag{6}$$

The radial basis functions with the popular Gaussian function:

$$f(w, b, x) = e^{-b\|x - w\|^2} \tag{7}$$

ELM has been successfully applied for modeling reference evapotranspiration ($ET_0$) at the north, mid, and southern parts of Iraq (Abdullah et al. 2015); prediction of monthly effective drought index in eastern Australia (Deo and Şahin 2015); simulation of monthly mean streamflow water level in eastern Queensland (Deo and Şahin 2016); forecasting monthly streamflow discharge rates in the Tigris River, Iraq (Yaseen et al. 2016); and prediction of daily water level in the Urmia Lake, Iran (Shiri et al. 2016).

## Online sequential extreme learning machine

Online sequential extreme learning machine (OS-ELM) was first proposed by Liang et al. (2006) as a variant ELM algorithm. OS-ELM is a modified version of the original ELM, developed for single hidden layer feedforward networks (SLFNs) with additive or radial basis function (RBF) hidden nodes (Liang et al. 2006). In the original ELM, the learning algorithm starts only if all the training samples are ready, consequently, when new data are available, the ELM must be retrained (Liang et al. 2006). Contrary to the ELM, OS-ELM is capable of learning the available data one-by-one or "chunk-by-chunk" with a fixed or varying chunk size (Liang et al. 2006; Huang et al. 2011); when the learning procedure for any particular observation is completed, the observation is discarded from the pattern (Liang et al. 2006; Huang et al. 2011). The OS-ELM algorithm is implemented in two stages: (i) the initialization and (ii) the online sequential learning phase. The learning algorithm is summarized as follows (Singh et al. 2016): (i) Training data are sequential one-by-one or block-by-block, (ii) only a new data sample is presented to the learning algorithm for training; the data samples are not used for training again and again; (iii) once finished with the data sample, it will be discarded from the pattern; and (iv) prior knowledge about the length of the data samples to be presented to the learning algorithm is not necessary (Singh et al. 2016).

If a new chunk of data is received, it results in a problem of minimizing (Liang et al. 2006):

$$\left\| \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} w - \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \right\| \tag{8}$$

Consequently, the connecting weight $w$ becomes:

$$w^{(1)} = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \tag{9}$$

where

$$K_1 = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = K_0 + H_1^T H_1 \tag{10}$$

In order to increase the efficiency of sequential learning, $w^{(1)}$ must be calculated as a function of $w^{(0)}$, $K_1$, $H_1$, and $T_1$:

$$\begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} = H_0^T T_0 + H_1^T T_1$$
$$= K_1 w^{(0)} - H_1^T H_1 w^{(0)} + H_1^T T_1 \tag{11}$$

$w^{(1)}$ can be expressed as follows by combining (10) and (11):

$$w^{(1)} = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} = w^{(0)} + K_1^{-1} H_1^T \left( T_1 - H_1 w^{(0)} \right) \tag{12}$$

Iteratively, when the $(k + 1)$ new chunk of data arrives, the recursive method is implemented for acquiring the updated solution. $w^{(k+1)}$ can be updated by:

$$w^{(k+1)} = w^{(k)} + K_{k+1}^{-1} H_{k+1}^T \left( T_{k+1} - H_{k+1} w^{(k)} \right) \tag{13}$$

with

$$K_{k+1}^{-1} = K_k^{-1} - K_k^{-1} H_{k+1}^T \left( I + H_{k+1} K_k^{-1} H_{k+1}^T \right)^{-1} \times H_{k+1} K_k^{-1} \tag{14}$$

OS-ELM has been successfully applied for many engineering problems (Lima et al. 2016; Yadav et al. 2016; Wang and Han 2014).

## Optimally pruned extreme learning machine

Optimally pruned extreme learning machine (OP-ELM) was first proposed by Miche et al. (2008a, b) and presented later by Miche et al. (2010). OP-ELMs accepts multiple types of functions, i.e., Gaussian, sigmoid, and linear (Pouzols and Lendasse 2010a, b). Using a real-world data set on a toy example (Miche et al. 2008b) demonstrated in a clear and effective manner that the original ELM suffers from an important drawback that lead to a decrease in the accuracy of the algorithm, when irrelevant or correlated variables are presented in the training data set (Sovilj et al. 2010). OP-ELM is presented as a good solution for pruning away irrelevant variables. (Miche et al. 2008a, 2010). The OP-ELM is achieved in three steps (Moreno et al. 2014; Grigorievskiy et al. 2014): (i) an over-parameterized ELM is built with initially large number of neurons; (ii) an appropriate ranking of the hidden neurons is made upon based on their contribution to the linear explanation of the ELM output, using the multi-response sparse regression (MRSR) (Similä and Tikka 2005) or least angle regression (LARS) (Efron et al. 2004) if the output is one-dimensional; (iii) leave-one-out (LOO) validation is used to decide how many neurons to prune. Presented as a good alternative to solve complex problems, OP-ELM has been successfully applied for many engineering problems (Sorjamaa et al. 2008; Sovilj et al. 2010; Moreno et al. 2014; Grigorievskiy et al. 2014; Akusok et al. 2015; Heddam 2016c).

## Multilayer perceptron neural network

Multilayer perceptron neural network (MLPNN) (Rumelhart et al. 1986) is a well-known ANNs model and mainly adopted for solving problems in environmental science. MLPNN is a universal approximator (Hornik et al. 1989; Hornik 1991). MLPNN is composed of layers of units called neurons. MLPNN has three types of layers: an input layer, one or more hidden layers, and an output layer. In the present study, we have used a model with only one hidden layer with sigmoid activation function and a linear activation function also called identity function in the output layer. The number of neurons in the hidden layer is determined with trial and error. The input layer contains the predictors (pH, TE, SC, and TU) for models with water quality variables, and the year, month, day, and hour numbers for the models without water quality variables. Finally, the output layer contains only one neuron; the DO concentration. More details about MLPNN can be found in Haykin (1999).

## Multiple linear regression)

Multiple linear regression (MLR) is a method adopted to provide a simple linear relation between a set of predictors ($x_i$), and one variable called response ($Y$); the model can be presented as follow:

$$Y = f(x_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_i x_i \quad (15)$$

where $Y$ is the response (DO); $x_i$ is the water quality variables (pH, TE, SC, and TU); and $\beta_i$ is the parameters of the models.

## Case studies

DO modeling in river ecosystems with and without water quality variables was demonstrated using eight different datasets from different basins in USA: Clackamas, Willamette, Umpqua, Tualatin, Delaware, and Cumberland basins. Detailed presentation of the rivers with the identification code of each station (ID) at different basins is reported in Table 1. For the eight stations, data are available at the United States Geological Survey (USGS) website:http://or.water.usgs.gov/cgi-bin/grapher/table_setup.pl?site_id. The water quality data used in this study consisted of measured water temperature (TE, °C), turbidity (TU FNU), pH (std. unit), specific conductance (SC, μS/cm) and dissolved oxygen (DO, mg/L). Water quality variables were measured at 1-h time interval. For each station, we selected a period of record for four calendar years. Detailed descriptions of the data set with the period of records are reported in Table 2. Each data set is divided into two subsets: (i) a training subset (70%) and (ii) a validation subset (30%); results for each station are reported in Table 3. Table 4 represents the statistical parameters of water quality variables in the study area. In the table, the terms $X_{mean}$, $X_{max}$, $X_{min}$, $S_x$, $C_v$, and $R$ denote the mean, maximum, minimum, standard deviation, coefficient of variation, and the coefficient of correlation between the variable and the DO, respectively.

## Application and results

The aim of this study is to demonstrate the capability and usefulness of the ELM models for the prediction of DO concentration in river ecosystems with and without water quality variables as inputs. The investigation was conducted in three phases. First, by selecting four water quality variables (TE, pH, SC, and TU), several combinations were selected and six scenarios (Table 5) (i) TE, pH, SC, and TU; (ii) TE, pH, and SC; (iii) TE, SC, and TU; (iv) TE and SC; (v) TE and pH; and (vi) TE and TU were compared. The selection of the six combinations is mainly based on the correlation coefficient. Second, an attempt is made to predict DO concentration

**Table 1** Geographical location of the selected stations

| Site | Basin | Station ID |
|---|---|---|
| Clackamas River near Oregon City | Clackamas Basin, Oregon | 14211010 |
| Willamette River at Portland | Willamette Basin, Oregon | 14211720 |
| N. Umpqua River at Soda Springs, nr Toketee Falls | Umpqua Basin, Oregon | 14316460 |
| N. Umpqua River near Idleyld Park | Umpqua Basin, Oregon | 14317450 |
| Tualatin River at Oswego Diversion Dam | Tualatin Basin, Oregon | 14207200 |
| Tualatin River at Hwy 219 at Jackson Bottom | Tualatin Basin, Oregon | 14206241 |
| Delaware River at Trenton, NJ | Delaware basin Commission | 1463500 |
| Mill Creek at Lebanon Road at Nashville, TN | Cumberland Basin, Tennessee | 3431083 |

without water quality variables as input. Within this context, we have selected four inputs variables that are (i) year number, (ii) month number from 1 to 12, (iii) day number of the month from 1 to 31, and (iv) hour number from 0:00 to 24:00. For example, the inputs of the pattern that correspond to 15 December 2010 at 22:00 are described as below:

1. Input 1: year number, i.e., "2010".
2. Input 2: month number, i.e., "12".
3. Input 3: day number, i.e., "15".
4. Input 4: hour number, i.e., "22.00".

Flowchart of the proposed approaches with and without water quality variables is presented in Fig. 2.

Similar to the models with water quality variables, the models without water quality always have four input variables. Third, the best ELM models are evaluated by training with validation dataset and testing with training dataset, with and without water quality variables.

In order to remove the effect of different units of measurement, the data were standardized using the $Z$ score method (Gulgundi and Shetty 2016):

$$x_{ni,k} = \frac{x_{i,k} - m_k}{Sd_K} \qquad (16)$$

where $x_{ni,\,k}$ is the normalized value of the variable $k$ (input or output) for each sample $i$. $x_{i,k}$ is the original value of the variable $k$. $m_k$ and $S_{dk}$ are the mean value and standard deviation of the variable $k$. According to Gulgundi and Shetty (2016), the $Z$ score method is mainly adopted in the studies based on statistical analysis; furthermore, the effect of different units of measurement was completely eliminated. It has been demonstrated that the $Z$ score method (Gulgundi and Shetty 2016; Heddam 2016c, d, e) improves the performances and the accuracy of the models. In this study, four statistical indices were selected to evaluate the different ELM models. These indices are the coefficient of correlation ($R$), the Nash-Sutcliffe efficiency (NSE), the root mean squared error (RMSE), and the mean absolute error (MAE), calculated as follows (Legates and McCabe 1999):

$$R = \left[ \frac{\frac{1}{N}\sum (O_i - O_m)(P_i - P_m)}{\sqrt{\frac{1}{N}\sum_{i=1}^{n}(O_i - O_m)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{n}(P_i - P_m)^2}} \right] \qquad (17)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}[O_i - P_i]^2}{\sum_{i=1}^{N}[O_i - O_m]^2} \qquad (18)$$

**Table 2** Data set presentation for all stations

| Station ID | Latitude | Longitude | Period of record | Total pattern | Incomplete pattern | Final pattern |
|---|---|---|---|---|---|---|
| 14211010 | 45° 22′ 46″ | 122° 34′ 34″ NAD27 | 01 January 2010–31 December 2013 | 35,064 | 364 | 34,700 |
| 14211720 | 45° 31′ 03″ | 122° 40′ 09″ NAD83 | 01 January 2010–31 December 2013 | 35,064 | 2258 | 32,806 |
| 14316460 | 43° 18′ 22″ | 122° 30′ 42″ NAD27 | 01 January 2010–31 December 2013 | 35,064 | 1366 | 33,698 |
| 14317450 | 43° 19′ 29″ | 122° 59′ 55″ AD27 | 01 January 2009–31 December 2012 | 35,064 | 1237 | 33,827 |
| 14207200 | 45° 21′ 24″ | 122° 41′ 02″ NAD27 | 01 January 2010–31 December 2013 | 35,064 | 4780 | 30,284 |
| 14206241 | 45° 30′ 01″ | 122° 59′ 24″ NAD27 | 01 January 2010–31 December 2013 | 35,064 | 1888 | 33,176 |
| 1463500 | 40° 13′ 18″ | 74° 46′ 41″N AD83 | 01 January 2007–31 December 2010 | 35,064 | 3327 | 31,737 |
| 3431083 | 36° 09′ 50″ | 86° 41′ 57″ NAD27 | 01 January 2010–31 December 2013 | 35,064 | 4690 | 30,374 |

45° 22′ 46″: 45°, 22 min, and 46 s; 35,064 = 8784 + (8760 × 3); 8760 = 24 h × 365 days (2008 and 2012 are leap years); one pattern corresponds to 1 h in a day

**Table 3** The number of data used in training (70%) and validation (30%) subsets for all stations

| Models | Stations | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 14,211,010 | 14,211,720 | 14,316,460 | 14,317,450 | 14,207,200 | 14,206,241 | 1,463,500 | 3,431,083 |
| All data | 34,700 | 32,806 | 33,698 | 33,827 | 30,284 | 33,176 | 31,737 | 30,374 |
| Training | 24,290 | 22,965 | 23,589 | 23,679 | 21,199 | 23,224 | 22,216 | 21,262 |
| Validation | 10,410 | 9841 | 10,109 | 10,148 | 9085 | 9952 | 9521 | 9112 |

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (O_i - P_i)^2} \qquad (19)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |O_i - P_i| \qquad (20)$$

where $N$ is the number of data points, $O_i$ is the measured value, and $P_i$ is the corresponding model prediction. $O_m$ and $P_m$ are the average values of $O_i$ and $P_i$. For applying the ELM models, we used the following Matlab codes: (i) for OP-ELM model, we used the code available at http://research.ics.aalto.fi/eiml/software.shtml and for the OS-ELM, S-ELM, and R-ELM we used the code available at http://www.ntu.edu.sg/home/egbhuang/elm_codes.html.

## Predicting DO using water quality variables as predictors

Table 6 shows the evaluation of the OP-ELM1, OS-ELM1, S-ELM1, R-ELM1, MLPNN1, and MLR1 models in training and validation phases for each station. It should be noted that Table 6 contains only the best models. The complete results, including the best and the worst models developed are reported in Tables A1–A6 as supplementary materials.

According to the obtained results, some important conclusions can be drawn. First, among the six OP-ELM models, the OP-ELM1 is always the best followed by the OP-ELM2 in the second place, the OP-ELM3 and OP-ELM4 ranked as third. Second, it is clear from Table A1 that inclusion of turbidity in the models does not help to significantly improve the performances of the proposed models, and the accuracy of the models is slightly improved. However, the TU is not suitable as input variable and may not be a predictor of DO concentration. Third, water TE and pH are the most important predictors of DO. According to Table A1, the $R$ varies with different combinations in the training phase. The best $R$ value (0.990) was obtained using OP-ELM1 at two stations: USGS 14211720 and USGS 14317450. Similarly, the best NSE value (0.980) was obtained using OP-ELM1 at the USGS 14317450 station. In terms of the RMSE and MAE, the best performances were obtained using OP-ELM1, with RMSE equal to 0.152 mg/L and MAE equal to 0.110 mg/L, at the USGS 14316460 station. In addition, the results for the six OP-ELM models in the validation phase are also summarized in Table A1. As shown in Table A1, the OP-ELM1 and OP-ELM2 provide the best accuracy at eight stations, with high $R$ and NSE, and low RMSE and MAE. Furthermore, the results obtained in the validation phase indicate that: (i) inclusion of the TU as input does not help to improve the accuracies of the proposed models (ii) OP-ELM5 model with TE and pH as inputs provides good results at four stations with $R$ values ranging from 0.914 to 0.987; NSE values range from 0.835 to 0.975. The validation results for the best OP-ELM1 models show smaller RMSE and MAE with values equal to 0.172 mg/L and 0.116, respectively. Measured and predicted DO concentrations with OP-ELM1 at the eight stations in the validation phase, using water quality variables as inputs are presented in Fig. 3. The scatterplots apparently illustrate that the estimates of the OP-ELM1 model closely follow the corresponding measured DO values (especially the extreme values).

Results obtained using the six OS-ELM models in the training and validation phases at all eight stations are summarized in Table A2. It should be observed here that the six OS-ELM models give high accuracy and the OS-ELM1 performs better than the other models by showing lower RMSE and MAE values, and higher NSE and $R$ values. The results indicate that the OS-ELM1 and OS-ELM2 are the best among the all six models. These models relatively give the high $R$ and NSE, and low RMSE and MAE values. It can be observed that the OS-ELM models yield results worse than those obtained using the OP-ELM models. When comparing these two models taking into account the increasing importance of the input variables from two to four inputs, the OP-ELM models perform superior to the OS-ELM models. According to Table A2, the R and NSE values obtained using OS-ELM1 model in the validation phase, lie in an interval between 0.923 and 0.987 and 0.852 and 0.974, respectively. Similarly, the validation results of the best OS-ELM1 models, respectively, show smaller RMSE and MAE with values equal to 0.179 and 0.134 mg/L when compared to other models. The

**Table 4** Statistical parameters of the used data sets for all stations

| Station | Data set | Unit | $X_{mean}$ | $X_{max}$ | $X_{min}$ | $S_x$ | $C_v$ | $R$ |
|---|---|---|---|---|---|---|---|---|
| USGS ID 14211010 | TE | °C | 10.37 | 22.30 | 0.00 | 5.02 | 0.48 | −0.91 |
| | SC | uS/cm | 50.66 | 76.00 | 26.00 | 10.18 | 0.20 | −0.65 |
| | pH | / | 7.59 | 9.20 | 6.70 | 0.32 | 0.04 | -0.15 |
| | TU | FNU | 3.76 | 520.00 | 0.00 | 12.21 | 3.25 | 0.18 |
| | DO | mg/L | 11.34 | 15.20 | 7.20 | 1.39 | 0.12 | 1.00 |
| USGS ID 1463500 | TE | °C | 13.42 | 30.90 | 0.00 | 9.19 | 0.69 | −0.89 |
| | SC | uS/cm | 191.68 | 305.00 | 85.00 | 39.89 | 0.21 | −0.33 |
| | pH | / | 7.87 | 9.50 | 6.60 | 0.51 | 0.06 | -0.01 |
| | TU | FNU | 6.61 | 410.00 | 0.00 | 17.01 | 2.57 | −0.06 |
| | DO | mg/L | 10.99 | 16.90 | 5.70 | 2.25 | 0.21 | 1.00 |
| USGS ID 3431083 | TE | °C | 16.76 | 32.90 | 0.20 | 7.58 | 0.45 | −0.76 |
| | SC | uS/cm | 504.23 | 821.00 | 126.00 | 77.62 | 0.15 | 0.26 |
| | pH | / | 8.07 | 9.00 | 7.00 | 0.28 | 0.03 | 0.58 |
| | TU | FNU | 12.95 | 620.00 | 0.20 | 28.45 | 2.20 | −0.12 |
| | DO | mg/L | 9.35 | 18.90 | 2.40 | 2.67 | 0.29 | 1.00 |
| USGS ID 14206241 | TE | °C | 10.83 | 22.00 | 1.20 | 4.10 | 0.38 | −0.90 |
| | SC | uS/cm | 91.84 | 134.00 | 53.00 | 12.68 | 0.14 | −0.30 |
| | pH | / | 7.20 | 7.80 | 6.70 | 0.20 | 0.03 | -0.41 |
| | TU | FNU | 10.73 | 220.00 | 0.00 | 11.49 | 1.07 | 0.18 |
| | DO | mg/L | 10.07 | 13.50 | 7.30 | 1.01 | 0.10 | 1.00 |
| USGS ID 14207200 | TE | °C | 13.11 | 23.90 | 1.80 | 5.61 | 0.43 | −0.87 |
| | SC | uS/cm | 198.34 | 380.00 | 67.00 | 82.00 | 0.41 | −0.81 |
| | pH | / | 7.02 | 7.40 | 6.70 | 0.11 | 0.02 | 0.07 |
| | TU | FNU | 9.79 | 120.00 | 0.10 | 11.37 | 1.16 | 0.46 |
| | DO | mg/L | 8.59 | 12.90 | 4.30 | 1.67 | 0.19 | 1.00 |
| USGS ID 14211720 | TE | °C | 12.24 | 24.00 | 1.60 | 5.43 | 0.44 | −0.95 |
| | SC | uS/cm | 72.08 | 98.00 | 48.00 | 9.35 | 0.13 | −0.53 |
| | pH | / | 7.27 | 8.10 | 6.60 | 0.18 | 0.02 | -0.52 |
| | TU | FNU | 9.33 | 170.00 | 0.80 | 13.12 | 1.41 | 0.53 |
| | DO | mg/L | 11.57 | 14.90 | 6.00 | 1.76 | 0.15 | 1.00 |
| USGS ID14316460 | TE | °C | 7.36 | 15.20 | 0.80 | 3.29 | 0.06 | −0.96 |
| | SC | uS/cm | 54.08 | 69.00 | 27.00 | 5.89 | 0.11 | −0.16 |
| | pH | / | 7.68 | 8.90 | 7.00 | 0.18 | 0.02 | -0.48 |
| | TU | FNU | 1.61 | 990.00 | 0.00 | 7.53 | 4.68 | 0.06 |
| | DO | mg/L | 11.83 | 14.70 | 10.00 | 0.83 | 0.07 | 1.00 |
| USGS ID 14317450 | TE | °C | 9.33 | 20.50 | 0.00 | 4.34 | 0.46 | −0.98 |
| | SC | uS/cm | 53.02 | 71.00 | 28.00 | 8.84 | 0.17 | −0.47 |
| | pH | / | 7.69 | 8.70 | 6.70 | 0.27 | 0.04 | -0.50 |
| | TU | FNU | 3.04 | 330.00 | 0.00 | 8.16 | 2.68 | 0.15 |
| | DO | mg/L | 11.57 | 14.80 | 8.90 | 1.17 | 0.10 | 1.00 |

$X_{mean}$ mean; $X_{max}$ maximum; $X_{min}$ minimum; $S_x$ standard deviation; $C_v$ coefficient of variation; $R$ coefficient of correlation with DO, *TE* water temperature, *DO* dissolved oxygen, *SC* specific conductance, *TU* turbidity, *uS/cm* micro Siemens per centimeter, *mg/L* milligram per liter, *FNU* Formazin Nephelometric Unit

RMSE and MAE accuracies of the best OS-ELM1 model were increased by 25.40 and 26.60% using OP-ELM1, respectively. Measured and predicted DO concentrations with OS-ELM1, at the eight stations in the validation phase, using water quality variables as inputs are presented in Fig. 4. Scatterplots apparently illustrate that the estimates of the OS-ELM1 model closely follows the corresponding measured DO values.

Table A3 shows the model evaluation of the six S-ELM models in the training and validation phases at all eight

**Table 5** The input combinations of different models

| Models | | | | | | Inputs combinations |
|---|---|---|---|---|---|---|
| OP-ELM | OS-ELM | S-ELM | R-ELM | MLPNN | MLR | |
| OP-ELM1 | OS-ELM1 | S-ELM1 | R-ELM1 | MLPNN1 | MLR1 | TE, SC, pH, TU |
| OP-ELM2 | OS-ELM2 | S-ELM2 | R-ELM2 | MLPNN2 | MLR2 | TE, SC, pH |
| OP-ELM3 | OS-ELM3 | S-ELM3 | R-ELM3 | MLPNN3 | MLR3 | TE, SC,TU |
| OP-ELM4 | OS-ELM4 | S-ELM4 | R-ELM4 | MLPNN4 | MLR4 | TE, SC |
| OP-ELM5 | OS-ELM5 | S-ELM5 | R-ELM5 | MLPNN5 | MLR5 | TE, pH |
| OP-ELM6 | OS-ELM6 | S-ELM6 | R-ELM6 | MLPNN6 | MLR6 | TE, TU |

*OP-ELM* optimally pruned extreme learning machine, *OS-ELM* online sequential extreme learning machine, *S-ELM* standard extreme learning machine, *R-ELM* radial basis function extreme learning machine, *MLR* multiple linear regression, *MLPNN* multilayer perceptron neural network

stations. According to Table A3, in the validation phase, the S-ELM1 is the best model for the all stations, with the highest $R$ and NSE values, and the lowest RMSE and MAE values. The $R$ and NSE values of the best S-ELM1 model are 0.988, and 0.977, respectively. Based on the RMSE and MAE as seen in Table A3, the S-ELM1 is selected as the best model in approximating hourly DO with the values equal to 0.179 and 0.140 mg/L, respectively. Furthermore, it can be observed that the S-ELM1 models yield results worse than those obtained using the OP-ELM1 model and yield results better than those obtained using the OS-ELM1 models. In the validation phase, the OP-ELM1 improves the S-ELM1 model with 18.60 and 57.14% reductions in RMSE and MAE, respectively. By contrast, the accuracy of the S-ELM1 model in respect to the RMSE and MAE was increased by 21.83 and 59.40%, respectively, against the OS-ELM1 model. Among the two-input models (combination 5, TE and pH inputs), the S-ELM5 is the best model with R, NSE, RMSE, and MAE values equal to 0.987, 0.974, 0.192, and 0.152 mg/l, respectively. Measured and predicted DO concentrations with S-ELM1 models at eight stations in the validation phase, using water quality

variables as predictors are presented in Fig. 5. It is obvious from the parameters of the fit line equations in scatterplots that the S-ELM1 model estimates are close to corresponding DO values.

Table A4 shows the model evaluation of the six R-ELM models in the training and validation phases at all eight stations. Similar to the results obtained using OP-ELM, OS-ELM, and S-ELM, the R-ELM1 is the best model at the all stations, with the highest $R$ and NSE values, and the lowest RMSE and MAE values. The best results is obtained using the R-ELM1 at the USGS 14317450 with $R$, NSE, RMSE, and MAE values equal to 0.988 and 0.976, respectively. In addition, R-ELM1 has the minimum value of RMSE (0.189 mg/L) and MAE (0.150 mg/L) values. As seen in Table A4, the R-ELM5 with only two water quality variables as inputs (TE and pH) is the best model with R, NSE, RMSE, and MAE values equal to 0.987, 0.974, 0.189, and 0.150 mg/L, respectively. The R-ELM1 model clearly performs worse than the OP-ELM1 and S-ELM1 models; however, the R-ELM1 provides better results than the OS-ELM1 model. In the validation phase, OP-ELM1 improves the OS-ELM1, S-ELM1, and R-ELM1 models, with the 10.41,



**Fig. 2** Flowchart of the two proposed modeling scenarios: scenario 1 using water quality variables as predictor and scenario 2 without water quality variables as predictors
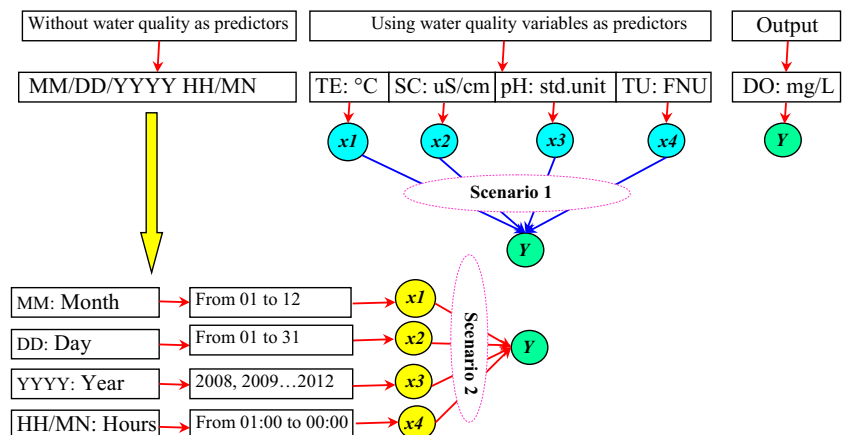
**Table 6** Performances of different models in modeling DO at all stations using water quality variables as predictors

| Models | Stations | Training (70%) | | | | Validation (30%) | | | | Ranks |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NSE | $R$ | RMSE | MAE | NSE | $R$ | RMSE | MAE | |
| OP-ELM1 | 14,211,010 | 0.968 | 0.984 | 0.235 | 0.179 | 0.966 | 0.983 | 0.240 | 0.183 | 1 |
| | 14,211,720 | 0.950 | 0.990 | 0.405 | 0.186 | 0.933 | 0.989 | 0.472 | 0.198 | 1 |
| | 14,316,460 | 0.965 | 0.982 | 0.152 | 0.110 | 0.962 | 0.981 | 0.158 | 0.116 | 1 |
| | 14,317,450 | 0.980 | 0.990 | 0.169 | 0.130 | 0.979 | 0.989 | 0.172 | 0.133 | 1 |
| | 14,207,200 | 0.954 | 0.977 | 0.360 | 0.272 | 0.952 | 0.976 | 0.363 | 0.278 | 1 |
| | 14,206,241 | 0.944 | 0.971 | 0.225 | 0.169 | 0.941 | 0.970 | 0.231 | 0.174 | 3 |
| | 1,463,500 | 0.934 | 0.967 | 0.583 | 0.435 | 0.930 | 0.965 | 0.596 | 0.446 | 1 |
| | 3,431,083 | 0.857 | 0.926 | 1.014 | 0.763 | 0.853 | 0.924 | 1.014 | 0.767 | 1 |
| OS-ELM1 | 14,211,010 | 0.959 | 0.979 | 0.265 | 0.206 | 0.958 | 0.979 | 0.268 | 0.209 | 5 |
| | 14,211,720 | 0.933 | 0.981 | 0.467 | 0.273 | 0.919 | 0.981 | 0.520 | 0.280 | 5 |
| | 14,316,460 | 0.949 | 0.974 | 0.183 | 0.135 | 0.951 | 0.975 | 0.179 | 0.134 | 5 |
| | 14,317,450 | 0.974 | 0.987 | 0.193 | 0.153 | 0.974 | 0.987 | 0.192 | 0.152 | 5 |
| | 14,207,200 | 0.939 | 0.969 | 0.414 | 0.323 | 0.938 | 0.969 | 0.413 | 0.323 | 4 |
| | 14,206,241 | 0.945 | 0.972 | 0.223 | 0.166 | 0.941 | 0.970 | 0.232 | 0.172 | 2 |
| | 1,463,500 | 0.918 | 0.958 | 0.653 | 0.500 | 0.915 | 0.956 | 0.660 | 0.505 | 5 |
| | 3,431,083 | 0.857 | 0.926 | 1.015 | 0.761 | 0.852 | 0.923 | 1.016 | 0.767 | 2 |
| S-ELM1 | 14,211,010 | 0.961 | 0.980 | 0.258 | 0.202 | 0.960 | 0.980 | 0.260 | 0.203 | 4 |
| | 14,211,720 | 0.935 | 0.981 | 0.463 | 0.266 | 0.920 | 0.981 | 0.518 | 0.275 | 4 |
| | 14,316,460 | 0.955 | 0.977 | 0.173 | 0.127 | 0.956 | 0.978 | 0.170 | 0.127 | 4 |
| | 14,317,450 | 0.977 | 0.988 | 0.179 | 0.139 | 0.977 | 0.988 | 0.179 | 0.140 | 3 |
| | 14,207,200 | 0.936 | 0.968 | 0.422 | 0.329 | 0.936 | 0.968 | 0.419 | 0.327 | 5 |
| | 14,206,241 | 0.915 | 0.956 | 0.276 | 0.207 | 0.914 | 0.956 | 0.280 | 0.211 | 5 |
| | 1,463,500 | 0.924 | 0.961 | 0.629 | 0.475 | 0.921 | 0.960 | 0.636 | 0.482 | 4 |
| | 3,431,083 | 0.838 | 0.915 | 1.081 | 0.808 | 0.837 | 0.915 | 1.067 | 0.805 | 5 |
| R-ELM1 | 14,211,010 | 0.968 | 0.984 | 0.235 | 0.180 | 0.966 | 0.983 | 0.240 | 0.184 | 2 |
| | 14,211,720 | 0.940 | 0.984 | 0.443 | 0.243 | 0.925 | 0.984 | 0.502 | 0.251 | 3 |
| | 14,316,460 | 0.962 | 0.981 | 0.158 | 0.115 | 0.959 | 0.979 | 0.165 | 0.120 | 2 |
| | 14,317,450 | 0.977 | 0.988 | 0.180 | 0.140 | 0.976 | 0.988 | 0.182 | 0.142 | 4 |
| | 14,207,200 | 0.950 | 0.975 | 0.375 | 0.285 | 0.948 | 0.974 | 0.376 | 0.287 | 2 |
| | 14,206,241 | 0.959 | 0.979 | 0.192 | 0.142 | 0.951 | 0.975 | 0.210 | 0.156 | 1 |
| | 1,463,500 | 0.930 | 0.964 | 0.603 | 0.453 | 0.926 | 0.962 | 0.616 | 0.465 | 2 |
| | 3,431,083 | 0.871 | 0.933 | 0.962 | 0.721 | 0.859 | 0.927 | 0.992 | 0.746 | 3 |
| MLPNN1 | 14,211,010 | 0.964 | 0.982 | 0.248 | 0.192 | 0.963 | 0.981 | 0.252 | 0.195 | 3 |
| | 14,211,720 | 0.970 | 0.985 | 0.309 | 0.227 | 0.969 | 0.984 | 0.314 | 0.232 | 2 |
| | 14,316,460 | 0.959 | 0.979 | 0.166 | 0.121 | 0.959 | 0.979 | 0.164 | 0.122 | 3 |
| | 14,317,450 | 0.978 | 0.989 | 0.176 | 0.138 | 0.978 | 0.989 | 0.177 | 0.138 | 2 |
| | 14,207,200 | 0.945 | 0.972 | 0.391 | 0.303 | 0.944 | 0.972 | 0.393 | 0.303 | 3 |
| | 14,206,241 | 0.928 | 0.963 | 0.254 | 0.193 | 0.926 | 0.962 | 0.259 | 0.197 | 4 |
| | 1,463,500 | 0.926 | 0.962 | 0.618 | 0.470 | 0.923 | 0.961 | 0.626 | 0.476 | 3 |
| | 3,431,083 | 0.847 | 0.921 | 1.048 | 0.792 | 0.846 | 0.920 | 1.038 | 0.791 | 4 |
| MLR1 | 14,211,010 | 0.947 | 0.973 | 0.302 | 0.243 | 0.944 | 0.972 | 0.308 | 0.247 | 6 |
| | 14,211,720 | 0.951 | 0.975 | 0.395 | 0.291 | 0.952 | 0.975 | 0.393 | 0.294 | 6 |
| | 14,316,460 | 0.933 | 0.966 | 0.210 | 0.159 | 0.931 | 0.965 | 0.213 | 0.159 | 6 |
| | 14,317,450 | 0.966 | 0.983 | 0.218 | 0.174 | 0.966 | 0.983 | 0.217 | 0.174 | 6 |
| | 14,207,200 | 0.841 | 0.917 | 0.667 | 0.524 | 0.842 | 0.918 | 0.657 | 0.519 | 6 |
| | 14,206,241 | 0.843 | 0.919 | 0.375 | 0.288 | 0.848 | 0.921 | 0.372 | 0.287 | 6 |
| | 1,463,500 | 0.893 | 0.945 | 0.746 | 0.563 | 0.888 | 0.942 | 0.757 | 0.571 | 6 |
| | 3,431,083 | 0.726 | 0.852 | 1.405 | 1.062 | 0.731 | 0.855 | 1.370 | 1.050 | 6 |

Due to lack of space the results for all six combinations (1 to 6) can be found in the supplementary material. The table contains only the best models

3.91, and 5.49% reductions in RMSE, respectively. In addition, the accuracy of the OP-ELM1 model in respect to the MAE is increased by 12.50, 5.00, and 6.33% against the OS-ELM1, S-ELM1, and R-ELM1, respectively. Measured and predicted DO concentrations with R-ELM1 models at eight stations in the validation phase, using water quality variables as predictors are presented in Fig. 6. It is obvious from the parameters of the fit line equations in scatterplots that the DO estimates by the R-ELM1 models closely follow the measured values.

Table A5 shows the model evaluation of the six MLPNN models in the training and validation phases

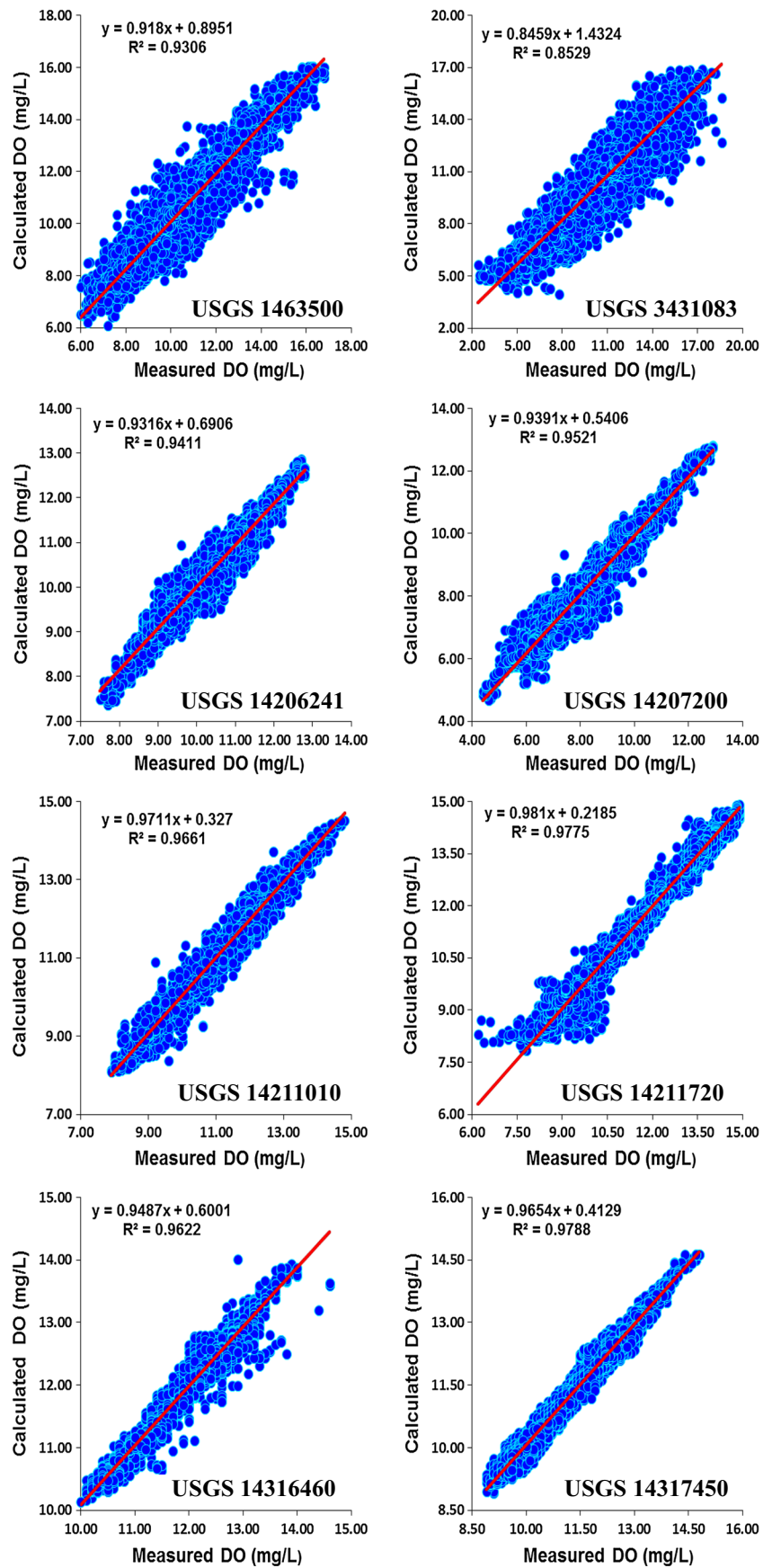**Fig. 3** Scatterplot of measured vs calculated DO using OP-ELM1 models in the validation period

**Fig. 4** Scatterplot of measured vs calculated DO using OS-ELM1 models in the validation period
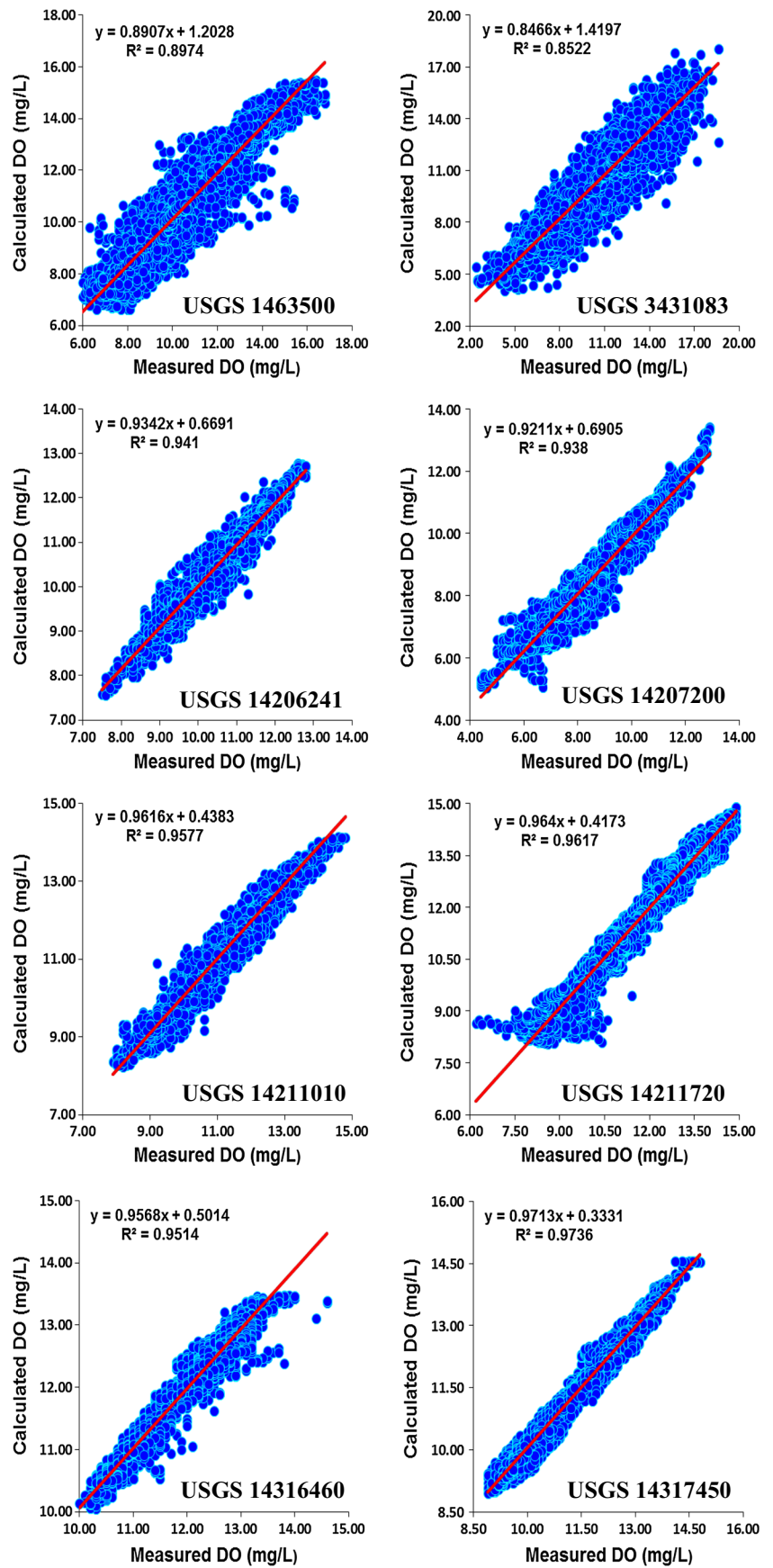
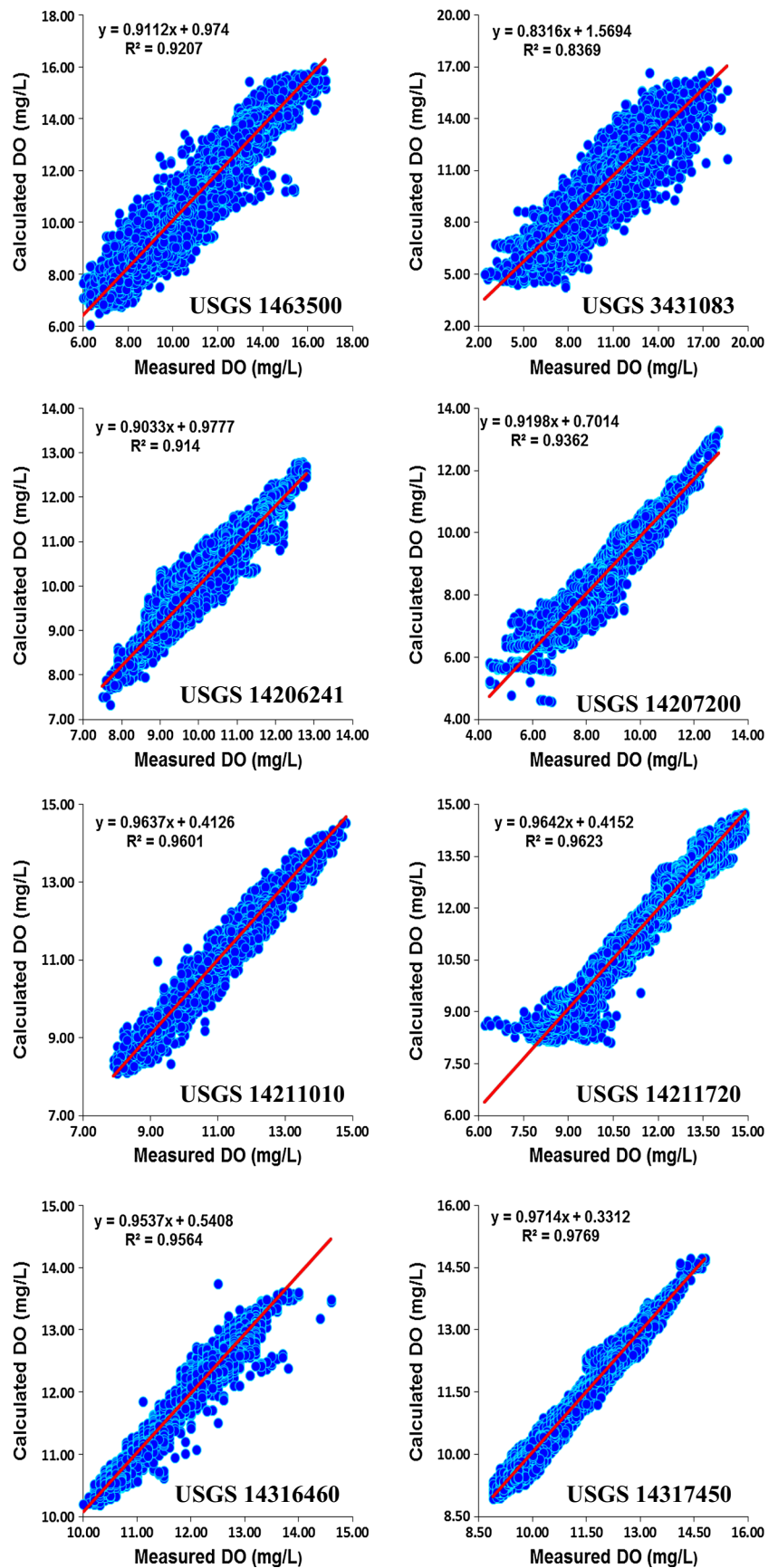**Fig. 5** Scatterplot of measured vs calculated DO using S-ELM1 models in the validation period

**Fig. 6** Scatterplot of measured vs calculated DO using R-ELM1 models in the validation period
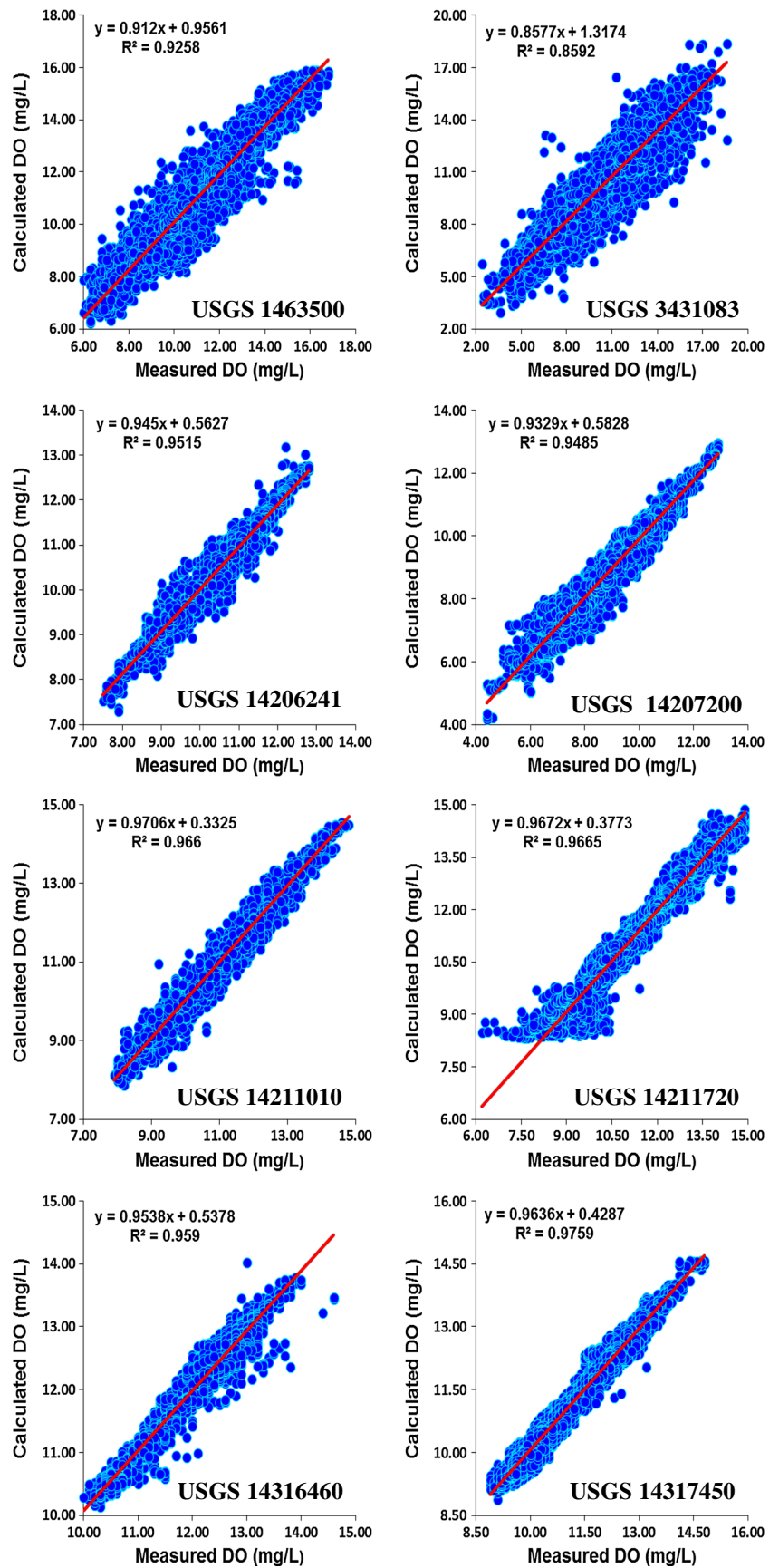
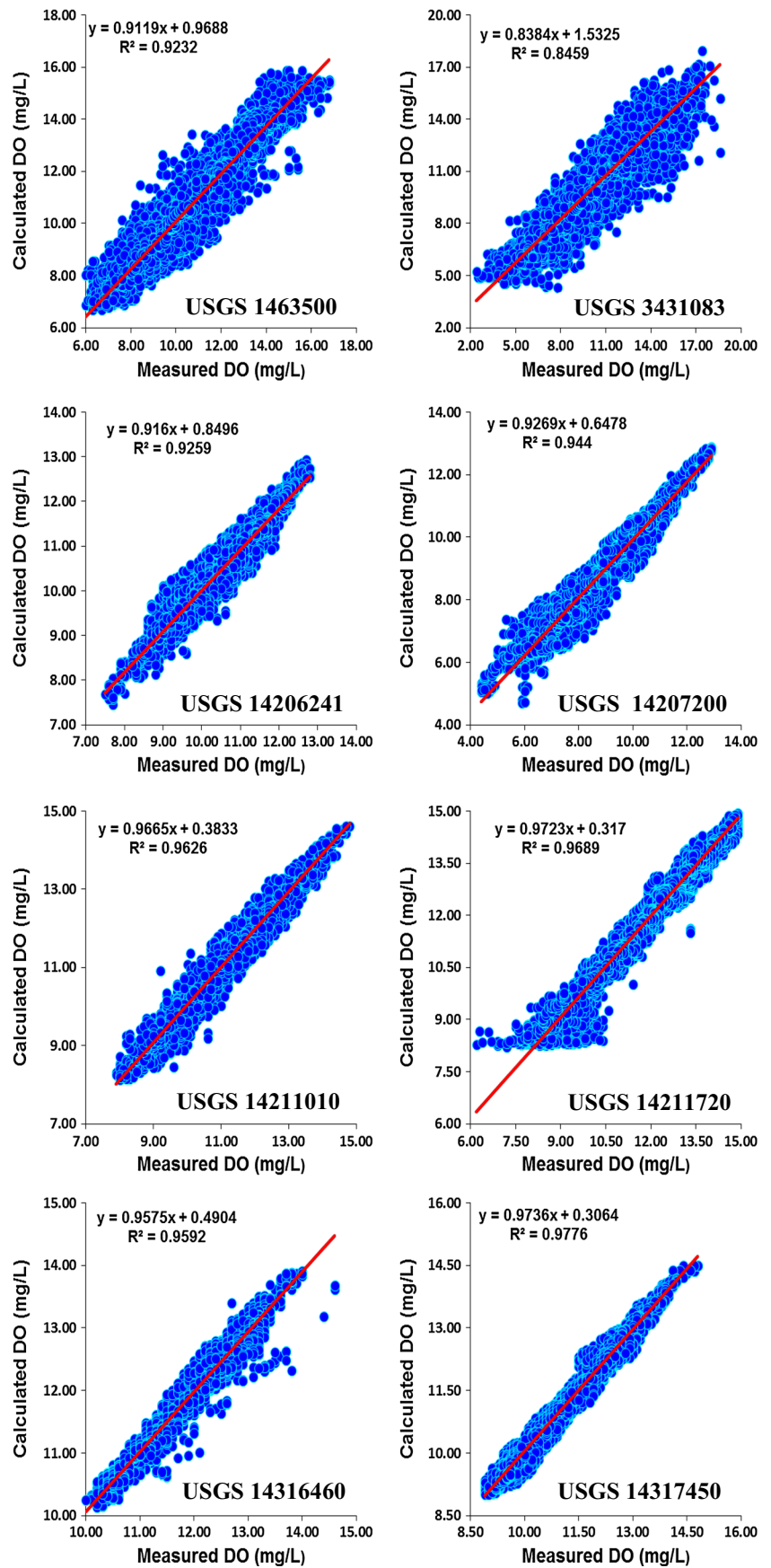**Fig. 7** Scatterplot of measured vs calculated DO using MLPNN1 models in the validation period

**Fig. 8** Scatterplot of measured vs calculated DO using MLR1 models in the validation period
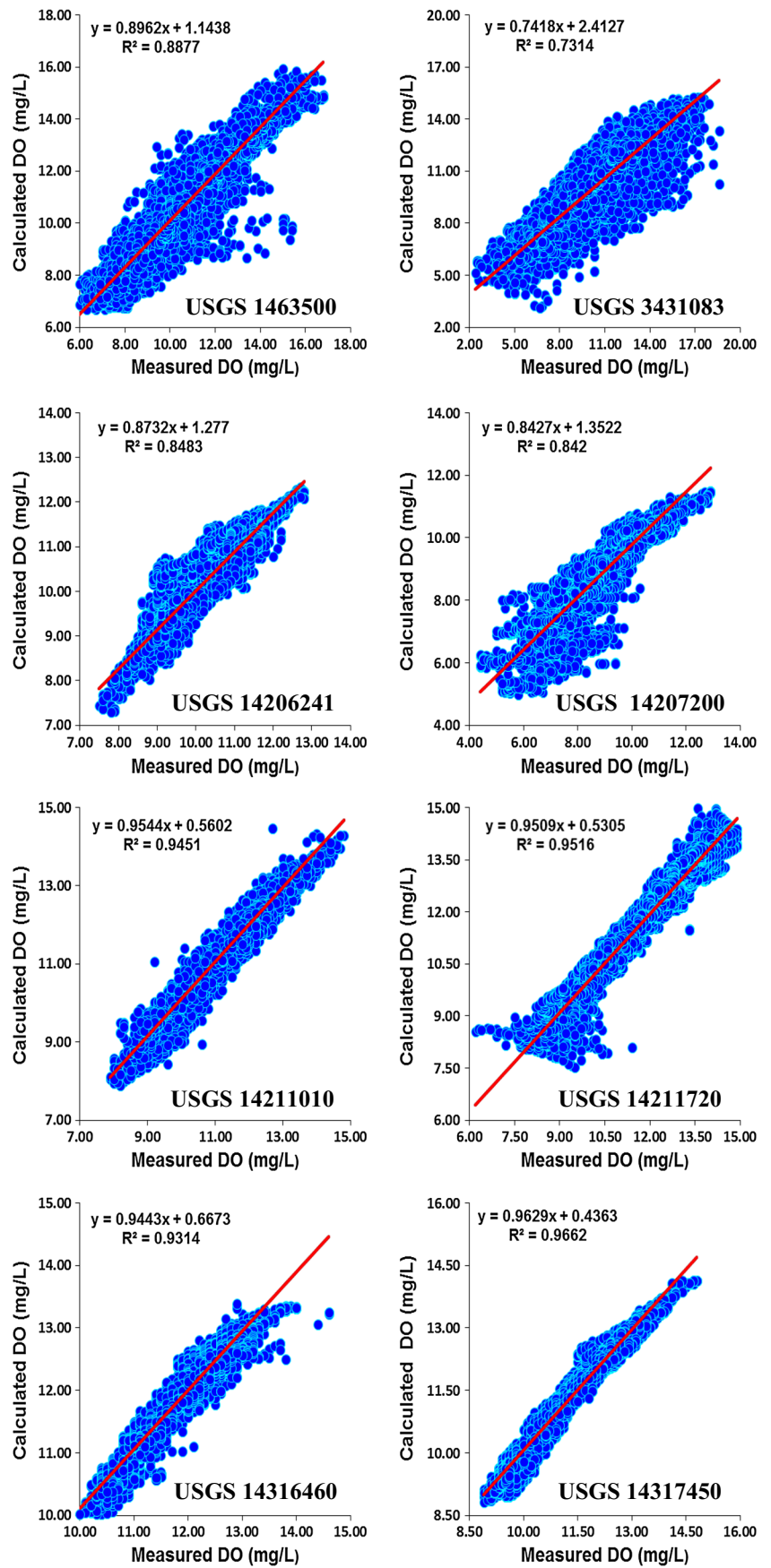
**Table 7** Ranking and comparison of different six models in modeling DO at all stations using water quality variables as predictors

| Stations | Comparison |
|---|---|
| 14211010 | OP-ELM1 > R-ELM1 > MLPNN1 > S-ELM1 > OS-ELM1 > MLR1 |
| 14211720 | OP-ELM1 > MLPNN1 > R-ELM1 > S-ELM1 > OS-ELM1 > MLR1 |
| 14316460 | OP-ELM1 > R-ELM1 > MLPNN1 > S-ELM1 > OS-ELM1 > MLR1 |
| 14317450 | OP-ELM1 > MLPNN1 > S-ELM1 > R-ELM1 > OS-ELM1 > MLR1 |
| 14207200 | OP-ELM1 > R-ELM1 > MLPNN1 > OS-ELM1 > S-ELM1 > MLR1 |
| 14206241 | R-ELM1 > OS-ELM1 > OP-ELM1 > MLPNN1 > S-ELM1 > MLR1 |
| 1463500 | OP-ELM1 > R-ELM1 > MLPNN1 > S-ELM1 > OS-ELM1 > MLR1 |
| 3431083 | OP-ELM1 > OS-ELM1 > R-ELM1 > MLPNN1 > S-ELM1 > MLR1 |

> superior

at all eight stations. According to Table A5, MLPNN1 have the best accuracy and performed the best in comparison to the all other five models. In the validation phase, the $R$ varies with different combinations. The best $R$ and NSE values (0.989 and 0.978) were obtained using MLPNN1 at the USGS 14317450station. In terms of the RMSE and MAE, the best performances were obtained using MLPNN1, with RMSE equal to 0.177 mg/L and MAE equal to 0.138 mg/L. Although the results obtained using MLPNN at the eight stations are very satisfactory, we also conclude that OP-ELM1 is performing better compared to the MLPNN1 in all stations, and the MLPNN1 ranks only second, third, and fourth, in terms of the performance criteria. Measured and predicted DO concentrations with MLPNN1 models at eight stations in the validation phase, using water quality variables as predictors are presented in Fig. 7.

Results using MLR are presented herein to draw a robust conclusion about the results obtained using the ELM and MLPNN models. Results using MLR are tabulated in Table A6. According to Table A6, both MLR1 and MLR2 perform well for modeling DO in all stations. On the other hand, MLR1 provided the best results compared to all other models. By comparing the results obtained using MLR1 and all ELM and MLPNN1 models, MLR1 presents the worst overall performances and rank in the last place. Measured and predicted DO concentrations with MLR1 models at eight stations in the validation phase, using water quality variables as predictors are presented in Fig. 8.

Finally, we compared the results obtained using the OP-ELM1, OS-ELM1, S-ELM1, R-ELM1, MLPNN1, and MLR1 models. The performances of the six models are tabulated in Table 6. As shown in Table 6, the six models are ranked according to the performances of each model in the validation phase at all eight stations. According to Table 6, OP-ELM1 model provides the best performance and is ranked in the first place at seven stations and in the second place at one station.

R-ELM1 is ranked in the first place in one station, in the second place in four stations, and it is third at only one station. MLPNN is ranked in the second place at two stations, in the third place in four stations, and finally, in the fourth place in two stations. OS-ELM1 and S-ELM1models are ranked either third, fourth, and fifth, except at two stations where OS-ELM1 model is ranked in the second place, as reported in Table 6. Finally, according to Table 6, we can conclude that the MLR1 is the worst model among all the compared models, and ranked in the sixth place. Details comparison and ranking of the six models is reported in Table 7.

Finally, in order to illustrate the capabilities and usefulness of the best obtained models OP-ELM1, OS-ELM1, S-ELM1, R-ELM1, MLPNN1, and MLR1 having TE, pH, SC, and TU variables as inputs, the models were also evaluated by interchanging the training and validation datasets by setting the training dataset (30%) and validation dataset (70%) and the results are summarized in Table 8. As seen from the table, the results are always good and the models provide high accuracy, with the exception of the R-ELM1 at the USGS 3431083 station, in which the performances of the model are dramatically decreased. For the 30% training and 70% validation case, the major parts of the results approximate the results obtained at 70% training and 30% validation case, which may mean that the models are capable to predict DO with a high accuracy. Comparing the performances of the four models, OP-ELM1 yields the best results and ranks in the first place at six stations, while the MLR1 presents the worst overall performances.

## Predicting DO without water quality variables as predictors

Table 9 shows the model evaluation of the four ELM, MLPNN, and MLR models without water quality variables

**Table 8** Evaluation of the optimal models trained with validation dataset and tested with training dataset by using water quality variables as predictors

| Station | Models | Training (30%) | | | | Validation (70%) | | | | Ranks |
|---------|--------|------|------|------|------|------|------|------|------|-------|
| | | NSE | R | RMSE | MAE | NSE | R | RMSE | MAE | |
| 14211010 | OP-ELM1 | 0.967 | 0.984 | 0.235 | 0.181 | 0.967 | 0.983 | 0.239 | 0.183 | 1 |
| | OS-ELM1 | 0.952 | 0.976 | 0.284 | 0.223 | 0.953 | 0.976 | 0.284 | 0.222 | 5 |
| | S-ELM1 | 0.958 | 0.979 | 0.268 | 0.212 | 0.959 | 0.979 | 0.266 | 0.210 | 4 |
| | R-ELM1 | 0.967 | 0.984 | 0.235 | 0.180 | 0.966 | 0.983 | 0.240 | 0.183 | 2 |
| | MLPNN1 | 0.964 | 0.982 | 0.246 | 0.191 | 0.965 | 0.982 | 0.246 | 0.189 | 3 |
| | MLR1 | 0.945 | 0.972 | 0.306 | 0.246 | 0.947 | 0.974 | 0.301 | 0.242 | 6 |
| 14211720 | OP-ELM1 | 0.932 | 0.989 | 0.476 | 0.202 | 0.947 | 0.988 | 0.418 | 0.201 | 1 |
| | OS-ELM1 | 0.917 | 0.980 | 0.526 | 0.284 | 0.932 | 0.980 | 0.473 | 0.277 | 5 |
| | S-ELM1 | 0.920 | 0.981 | 0.517 | 0.275 | 0.934 | 0.981 | 0.464 | 0.269 | 4 |
| | R-ELM | 0.924 | 0.984 | 0.504 | 0.247 | 0.939 | 0.984 | 0.448 | 0.243 | 2 |
| | MLPNN1 | 0.966 | 0.983 | 0.328 | 0.249 | 0.967 | 0.983 | 0. 454 | 0.246 | 3 |
| | MLR1 | 0.952 | 0.975 | 0.394 | 0.294 | 0.931 | 0.975 | 0.496 | 0.291 | 6 |
| 14316460 | OP-ELM1 | 0.967 | 0.983 | 0.148 | 0.109 | 0.953 | 0.977 | 0.177 | 0.128 | 1 |
| | OS-ELM1 | 0.948 | 0.974 | 0.185 | 0.137 | 0.945 | 0.972 | 0.192 | 0.140 | 4 |
| | S-ELM1 | 0.958 | 0.979 | 0.165 | 0.123 | 0.944 | 0.972 | 0.193 | 0.133 | 3 |
| | R-ELM1 | 0.964 | 0.982 | 0.153 | 0.113 | 0.914 | 0.959 | 0.238 | 0.153 | 5 |
| | MLPNN1 | 0.961 | 0.980 | 0.161 | 0.119 | 0.955 | 0.977 | 0.173 | 0.127 | 2 |
| | MLR1 | 0.934 | 0.967 | 0.208 | 0.158 | 0.933 | 0.966 | 0.211 | 0.159 | 6 |
| 14317450 | OP-ELM1 | 0.980 | 0.990 | 0.166 | 0.128 | 0.979 | 0.989 | 0.173 | 0.132 | 1 |
| | OS-ELM1 | 0.975 | 0.987 | 0.188 | 0.147 | 0.975 | 0.987 | 0.189 | 0.147 | 5 |
| | S-ELM1 | 0.977 | 0.989 | 0.178 | 0.138 | 0.977 | 0.988 | 0.180 | 0.140 | 4 |
| | R-ELM1 | 0.978 | 0.989 | 0.177 | 0.137 | 0.977 | 0.988 | 0.180 | 0.139 | 3 |
| | MLPNN1 | 0.978 | 0.989 | 0.176 | 0.137 | 0.977 | 0.989 | 0.179 | 0.139 | 2 |
| | MLR1 | 0.966 | 0.983 | 0.219 | 0.175 | 0.966 | 0.983 | 0.217 | 0.174 | 6 |
| 14207200 | OP-ELM1 | 0.955 | 0.977 | 0.350 | 0.267 | 0.950 | 0.975 | 0.374 | 0.282 | 1 |
| | OS-ELM1 | 0.907 | 0.953 | 0.504 | 0.396 | 0.905 | 0.952 | 0.515 | 0.402 | 5 |
| | S-ELM1 | 0.939 | 0.969 | 0.410 | 0.318 | 0.937 | 0.968 | 0.418 | 0.323 | 4 |
| | R-ELM1 | 0.950 | 0.975 | 0.369 | 0.281 | 0.947 | 0.973 | 0.386 | 0.292 | 2 |
| | MLPNN1 | 0.945 | 0.972 | 0.388 | 0.300 | 0.942 | 0.971 | 0.402 | 0.310 | 3 |
| | MLR1 | 0.842 | 0.918 | 0.657 | 0.520 | 0.841 | 0.917 | 0.667 | 0.526 | 6 |
| 14206241 | OP-ELM1 | 0.944 | 0.972 | 0.226 | 0.168 | 0.938 | 0.969 | 0.236 | 0.175 | 4 |
| | OS-ELM1 | 0.945 | 0.972 | 0.223 | 0.166 | 0.943 | 0.971 | 0.227 | 0.169 | 2 |
| | S-ELM1 | 0.916 | 0.957 | 0.277 | 0.207 | 0.913 | 0.955 | 0.280 | 0.208 | 5 |
| | R-ELM1 | 0.960 | 0.980 | 0.190 | 0.139 | 0.943 | 0.972 | 0.225 | 0.158 | 1 |
| | MLPNN1 | 0.927 | 0.963 | 0.257 | 0.194 | 0.925 | 0.962 | 0.259 | 0.195 | 3 |
| | MLR1 | 0.848 | 0.921 | 0.372 | 0.285 | 0.843 | 0.919 | 0.376 | 0.286 | 6 |
| 1463500 | OP-ELM1 | 0.932 | 0.965 | 0.590 | 0.442 | 0.931 | 0.965 | 0.598 | 0.445 | 1 |
| | OS-ELM1 | 0.895 | 0.946 | 0.730 | 0.551 | 0.898 | 0.948 | 0.727 | 0.552 | 5 |
| | S-ELM1 | 0.922 | 0.960 | 0.632 | 0.479 | 0.923 | 0.961 | 0.631 | 0.476 | 4 |
| | R-ELM1 | 0.931 | 0.965 | 0.594 | 0.444 | 0.930 | 0.965 | 0.600 | 0.448 | 2 |
| | MLPNN1 | 0.928 | 0.963 | 0.607 | 0.456 | 0.928 | 0.963 | 0.610 | 0.457 | 3 |
| | MLR1 | 0.888 | 0.942 | 0.756 | 0.571 | 0.893 | 0.945 | 0.746 | 0.564 | 6 |
| 3431083 | OP-ELM1 | 0.857 | 0.926 | 1.000 | 0.758 | 0.847 | 0.920 | 1.050 | 0.788 | 2 |
| | OS-ELM1 | 0.862 | 0.928 | 0.981 | 0.744 | 0.854 | 0.924 | 1.025 | 0.770 | 1 |
| | S-ELM1 | 0.840 | 0.916 | 1.059 | 0.801 | 0.834 | 0.913 | 1.091 | 0.818 | 5 |
| | R-ELM1 | 0.877 | 0.936 | 0.928 | 0.702 | 0.795 | 0.896 | 1.214 | 0.799 | 4 |
| | MLPNN1 | 0.843 | 0.918 | 1.049 | 0.792 | 0.840 | 0.916 | 1.073 | 0.802 | 3 |
| | MLR1 | 0.732 | 0.855 | 1.368 | 1.044 | 0.725 | 0.852 | 1.406 | 1.058 | 6 |

as predictors, at all eight stations. In the validation phase, as seen in Table 9, all the ELM models provide good approximates with respect to four statistical indices. R-ELM is the best model and ranks first at seven stations while the MLPNN ranks in the second place at five stations, followed by the S-ELM and the OS-ELM in the third and fourth places, respectively. Without water quality variables as predictors,

MLR presents the worst overall performances. Using MLR, the R and NSE values never exceed 0.44 and 0.186, respectively. Table 9 also shows that RMSE and MAE were found to be very high, and in some case, they reached to 2.573 and 2.104 mg/L, respectively. The R-ELM model yields low RMSE and MAE values (0.223 and 0.171 mg/L) at the USGS 14316460 and it is better than the all other models.

**Table 9** Performances of the proposed models in modeling DO at all stations without water quality variables as predictors

| Station | Models | Training (30%) | | | | Validation (70%) | | | | Ranks |
|---------|--------|-----|---|------|-----|-----|---|------|-----|-------|
| | | NSE | R | RMSE | MAE | NSE | R | RMSE | MAE | |
| 14211010 | OP-ELM | 0.857 | 0.926 | 0.494 | 0.387 | 0.853 | 0.923 | 0.500 | 0.394 | 4 |
| | OS-ELM | 0.851 | 0.923 | 0.504 | 0.407 | 0.849 | 0.922 | 0.505 | 0.408 | 5 |
| | S-ELM | 0.899 | 0.948 | 0.416 | 0.329 | 0.896 | 0.947 | 0.420 | 0.332 | 3 |
| | R-ELM | 0.951 | 0.975 | 0.289 | 0.223 | 0.939 | 0.969 | 0.322 | 0.251 | 1 |
| | MLPNN | 0.925 | 0.961 | 0.359 | 0.279 | 0.924 | 0.961 | 0.360 | 0.280 | 2 |
| | MLR | 0.177 | 0.423 | 1.188 | 0.961 | 0.186 | 0.438 | 1.175 | 0.949 | 6 |
| 14211720 | OP-ELM | 0.855 | 0.940 | 0.691 | 0.482 | 0.840 | 0.939 | 0.731 | 0.494 | 4 |
| | OS-ELM | 0.852 | 0.939 | 0.698 | 0.499 | 0.838 | 0.938 | 0.736 | 0.509 | 5 |
| | S-ELM | 0.874 | 0.952 | 0.644 | 0.446 | 0.856 | 0.949 | 0.694 | 0.462 | 3 |
| | R-ELM | 0.934 | 0.982 | 0.465 | 0.265 | 0.889 | 0.968 | 0.610 | 0.343 | 1 |
| | MLPNN | 0.925 | 0.962 | 0.489 | 0.386 | 0.923 | 0.960 | 0.496 | 0.393 | 2 |
| | MLR | 0.107 | 0.329 | 1.691 | 1.393 | 0.112 | 0.329 | 1.688 | 1.394 | 6 |
| 14316460 | OP-ELM | 0.856 | 0.925 | 0.309 | 0.238 | 0.849 | 0.921 | 0.316 | 0.244 | 5 |
| | OS-ELM | 0.859 | 0.927 | 0.306 | 0.238 | 0.858 | 0.926 | 0.306 | 0.239 | 4 |
| | S-ELM | 0.887 | 0.942 | 0.274 | 0.212 | 0.887 | 0.942 | 0.273 | 0.213 | 3 |
| | R-ELM | 0.925 | 0.962 | 0.223 | 0.171 | 0.915 | 0.957 | 0.236 | 0.183 | 1 |
| | MLPNN | 0.889 | 0.943 | 0.271 | 0.205 | 0.889 | 0.943 | 0.270 | 0.208 | 2 |
| | MLR | 0.069 | 0.267 | 0.785 | 0.669 | 0.066 | 0.268 | 0.785 | 0.668 | 6 |
| 14317450 | OP-ELM | 0.850 | 0.921 | 0.461 | 0.361 | 0.849 | 0.921 | 0.460 | 0.357 | 5 |
| | OS-ELM | 0.909 | 0.953 | 0.358 | 0.287 | 0.906 | 0.952 | 0.363 | 0.289 | 1 |
| | S-ELM | 0.904 | 0.950 | 0.368 | 0.294 | 0.901 | 0.949 | 0.373 | 0.295 | 3 |
| | R-ELM | 0.951 | 0.975 | 0.262 | 0.208 | 0.867 | 0.934 | 0.432 | 0.338 | 4 |
| | MLPNN | 0.903 | 0.950 | 0.370 | 0.287 | 0.904 | 0.951 | 0.366 | 0.289 | 2 |
| | MLR | 0.085 | 0.274 | 1.137 | 0.927 | 0.092 | 0.297 | 1.128 | 0.923 | 6 |
| 14207200 | OP-ELM | 0.777 | 0.882 | 0.789 | 0.607 | 0.771 | 0.878 | 0.791 | 0.611 | 5 |
| | OS-ELM | 0.815 | 0.903 | 0.718 | 0.563 | 0.812 | 0.902 | 0.716 | 0.562 | 4 |
| | S-ELM | 0.850 | 0.922 | 0.647 | 0.507 | 0.844 | 0.919 | 0.652 | 0.511 | 3 |
| | R-ELM | 0.908 | 0.953 | 0.507 | 0.382 | 0.882 | 0.939 | 0.568 | 0.433 | 1 |
| | MLPNN | 0.874 | 0.935 | 0.594 | 0.466 | 0.868 | 0.932 | 0.602 | 0.469 | 2 |
| | MLR | 0.141 | 0.375 | 1.549 | 1.254 | 0.145 | 0.381 | 1.529 | 1.240 | 6 |
| 14206241 | OP-ELM | 0.697 | 0.835 | 0.522 | 0.413 | 0.696 | 0.834 | 0.525 | 0.415 | 5 |
| | OS-ELM | 0.766 | 0.875 | 0.458 | 0.362 | 0.768 | 0.877 | 0.459 | 0.362 | 3 |
| | S-ELM | 0.875 | 0.936 | 0.334 | 0.261 | 0.776 | 0.886 | 0.451 | 0.343 | 2 |
| | R-ELM | 0.885 | 0.941 | 0.322 | 0.252 | 0.857 | 0.926 | 0.360 | 0.281 | 1 |
| | MLPNN | 0.752 | 0.867 | 0.472 | 0.365 | 0.754 | 0.868 | 0.473 | 0.364 | 4 |
| | MLR | 0.166 | 0.423 | 0.865 | 0.699 | 0.169 | 0.425 | 0.868 | 0.700 | 6 |
| 1463500 | OP-ELM | 0.794 | 0.891 | 1.032 | 0.809 | 0.783 | 0.885 | 1.052 | 0.826 | 5 |
| | OS-ELM | 0.850 | 0.922 | 0.882 | 0.670 | 0.844 | 0.919 | 0.893 | 0.681 | 4 |
| | S-ELM | 0.886 | 0.941 | 0.770 | 0.584 | 0.880 | 0.938 | 0.782 | 0.595 | 2 |
| | R-ELM | 0.925 | 0.962 | 0.623 | 0.459 | 0.908 | 0.953 | 0.686 | 0.513 | 1 |
| | MLPNN | 0.870 | 0.933 | 0.821 | 0.625 | 0.864 | 0.930 | 0.832 | 0.634 | 3 |
| | MLR | 0.156 | 0.406 | 2.090 | 1.790 | 0.162 | 0.413 | 2.066 | 1.767 | 6 |
| 3431083 | OP-ELM | 0.688 | 0.830 | 1.497 | 1.153 | 0.676 | 0.822 | 1.505 | 1.166 | 4 |
| | OS-ELM | 0.717 | 0.848 | 1.426 | 1.116 | 0.705 | 0.840 | 1.436 | 1.127 | 2 |
| | S-ELM | 0.701 | 0.837 | 1.465 | 1.132 | 0.690 | 0.831 | 1.471 | 1.143 | 3 |
| | R-ELM | 0.809 | 0.899 | 1.172 | 0.899 | 0.781 | 0.884 | 1.235 | 0.954 | 1 |
| | MLPNN | 0.684 | 0.827 | 1.506 | 1.169 | 0.670 | 0.819 | 1.518 | 1.186 | 5 |
| | MLR | 0.069 | 0.263 | 2.587 | 2.114 | 0.051 | 0.229 | 2.573 | 2.104 | 6 |

As can be seen from the comparison made in Table 9, the models' accuracies in estimating hourly DO values decrease by using the components of the Gregorian calendar as predictors instead of water quality parameters. Ranking accuracy of the four models is reported in Table 9 according to the performances of each model in the validation phase at all eight stations.

Table 10 shows the model evaluation of the four ELM models trained using validation dataset (30%) and tested using training dataset (70%) without water quality variables as inputs. A first analysis of the obtained results reveals that: (i) the performance accuracy of the developed models was decreased and (ii) the R-ELM provides the best results in all training and validation phases at all eight stations, followed by the

**Table 10**  Evaluation of the optimal models trained with validation dataset and tested with training dataset without water quality variables as predictors

| Station | Models | Training (30%) | | | | Validation (70%) | | | | Ranks |
|---|---|---|---|---|---|---|---|---|---|---|
| | | NSE | R | RMSE | MAE | NSE | R | RMSE | MAE | |
| 14211010 | OP-ELM | 0.839 | 0.916 | 0.522 | 0.414 | 0.824 | 0.908 | 0.548 | 0.434 | 5 |
| | OS-ELM | 0.861 | 0.928 | 0.485 | 0.391 | 0.859 | 0.927 | 0.491 | 0.396 | 4 |
| | S-ELM | 0.894 | 0.946 | 0.424 | 0.337 | 0.889 | 0.943 | 0.436 | 0.345 | 3 |
| | R-ELM | 0.954 | 0.977 | 0.280 | 0.215 | 0.919 | 0.961 | 0.372 | 0.282 | 1 |
| | MLPNN | 0.912 | 0.955 | 0.387 | 0.302 | 0.912 | 0.955 | 0.389 | 0.303 | 2 |
| | MLR | 0.187 | 0.436 | 1.174 | 0.947 | 0.175 | 0.419 | 1.189 | 0.960 | 6 |
| 14211720 | OP-ELM | 0.841 | 0.941 | 0.729 | 0.487 | 0.843 | 0.934 | 0.719 | 0.505 | 5 |
| | OS-ELM | 0.833 | 0.936 | 0.747 | 0.506 | 0.842 | 0.934 | 0.719 | 0.510 | 4 |
| | S-ELM | 0.869 | 0.955 | 0.662 | 0.436 | 0.878 | 0.953 | 0.633 | 0.439 | 3 |
| | R-ELM | 0.916 | 0.982 | 0.529 | 0.275 | 0.920 | 0.975 | 0.513 | 0.317 | 1 |
| | MLPNN | 0.944 | 0.972 | 0.423 | 0.334 | 0.944 | 0.972 | 0.422 | 0.334 | 2 |
| | MLR | 0.110 | 0.329 | 1.690 | 1.395 | 0.104 | 0.326 | 1.695 | 1.394 | 6 |
| 14316460 | OP-ELM | 0.870 | 0.933 | 0.292 | 0.228 | 0.861 | 0.928 | 0.304 | 0.237 | 5 |
| | OS-ELM | 0.879 | 0.938 | 0.282 | 0.220 | 0.869 | 0.932 | 0.294 | 0.229 | 4 |
| | S-ELM | 0.889 | 0.943 | 0.270 | 0.210 | 0.879 | 0.938 | 0.283 | 0.219 | 3 |
| | R-ELM | 0.929 | 0.964 | 0.217 | 0.168 | 0.911 | 0.954 | 0.243 | 0.186 | 1 |
| | MLPNN | 0.905 | 0.951 | 0.251 | 0.193 | 0.900 | 0.949 | 0.257 | 0.199 | 2 |
| | MLR | 0.072 | 0.277 | 0.782 | 0.665 | 0.073 | 0.275 | 0.784 | 0.667 | 3 |
| 14317450 | OP-ELM | 0.875 | 0.935 | 0.419 | 0.335 | 0.867 | 0.930 | 0.434 | 0.344 | 5 |
| | OS-ELM | 0.907 | 0.952 | 0.361 | 0.290 | 0.897 | 0.946 | 0.382 | 0.305 | 4 |
| | S-ELM | 0.906 | 0.951 | 0.364 | 0.290 | 0.896 | 0.946 | 0.384 | 0.305 | 3 |
| | R-ELM | 0.955 | 0.977 | 0.252 | 0.198 | 0.922 | 0.961 | 0.331 | 0.264 | 1 |
| | MLPNN | 0.917 | 0.958 | 0.340 | 0.267 | 0.913 | 0.955 | 0.352 | 0.274 | 2 |
| | MLR | 0.094 | 0.297 | 1.127 | 0.919 | 0.086 | 0.273 | 1.137 | 0.924 | 6 |
| 14207200 | OP-ELM | 0.691 | 0.832 | 0.919 | 0.718 | 0.681 | 0.826 | 0.944 | 0.737 | 5 |
| | OS-ELM | 0.819 | 0.905 | 0.704 | 0.550 | 0.813 | 0.902 | 0.722 | 0.564 | 4 |
| | S-ELM | 0.848 | 0.921 | 0.645 | 0.506 | 0.842 | 0.918 | 0.665 | 0.520 | 3 |
| | R-ELM | 0.911 | 0.955 | 0.492 | 0.370 | 0.861 | 0.931 | 0.624 | 0.474 | 1 |
| | MLPNN | 0.856 | 0.925 | 0.627 | 0.496 | 0.856 | 0.925 | 0.634 | 0.500 | 2 |
| | MLR | 0.145 | 0.381 | 1.529 | 1.240 | 0.140 | 0.375 | 1.549 | 1.254 | 6 |
| 14206241 | OP-ELM | 0.772 | 0.879 | 0.455 | 0.358 | 0.755 | 0.869 | 0.469 | 0.369 | 5 |
| | OS-ELM | 0.778 | 0.882 | 0.449 | 0.356 | 0.765 | 0.875 | 0.459 | 0.366 | 4 |
| | S-ELM | 0.776 | 0.886 | 0.451 | 0.343 | 0.875 | 0.936 | 0.334 | 0.261 | 2 |
| | R-ELM | 0.896 | 0.946 | 0.308 | 0.240 | 0.850 | 0.922 | 0.367 | 0.282 | 1 |
| | MLPNN | 0.814 | 0.902 | 0.411 | 0.326 | 0.804 | 0.897 | 0.419 | 0.333 | 3 |
| | MLR | 0.169 | 0.424 | 0.868 | 0.699 | 0.165 | 0.420 | 0.866 | 0.699 | 6 |
| 1463500 | OP-ELM | 0.823 | 0.907 | 0.949 | 0.741 | 0.822 | 0.907 | 0.961 | 0.753 | 5 |
| | OS-ELM | 0.848 | 0.921 | 0.881 | 0.678 | 0.848 | 0.921 | 0.887 | 0.681 | 4 |
| | S-ELM | 0.883 | 0.940 | 0.772 | 0.584 | 0.881 | 0.939 | 0.785 | 0.597 | 2 |
| | R-ELM | 0.924 | 0.961 | 0.621 | 0.456 | 0.900 | 0.949 | 0.721 | 0.540 | 1 |
| | MLPNN | 0.863 | 0.929 | 0.836 | 0.641 | 0.866 | 0.931 | 0.833 | 0.640 | 3 |
| | MLR | 0.161 | 0.410 | 2.069 | 1.767 | 0.152 | 0.403 | 2.094 | 1.793 | 6 |
| 3431083 | OP-ELM | 0.679 | 0.824 | 1.497 | 1.168 | 0.660 | 0.813 | 1.563 | 1.215 | 5 |
| | OS-ELM | 0.714 | 0.845 | 1.413 | 1.095 | 0.705 | 0.840 | 1.456 | 1.120 | 2 |
| | S-ELM | 0.691 | 0.831 | 1.468 | 1.129 | 0.687 | 0.829 | 1.500 | 1.147 | 3 |
| | R-ELM | 0.808 | 0.899 | 1.158 | 0.890 | 0.772 | 0.879 | 1.281 | 0.983 | 1 |
| | MLPNN | 0.675 | 0.822 | 1.506 | 1.176 | 0.675 | 0.822 | 1.529 | 1.182 | 4 |
| | MLR | 0.053 | 0.230 | 2.572 | 2.104 | 0.067 | 0.261 | 2.590 | 2.118 | 6 |

MLPNN in the second place, S-ELM in the third place, OS-ELM in the fourth place, OP-ELM in the fifth place, and finally, MLR with the lowest accuracy ranks in the last place. In the validation phase, R-ELM yields the lowest RMSE (0.186 mg/L) and MAE (0.243 mg/L), and the highest $R$ (0.975) and NSE (0.920) indices. Based on these results, it is clear that none of the proposed models are able to provide the performances achieved using the R-ELM.

## Conclusions

In the present study, an attempt was made to estimate DO concentration using four different ELM models; also, the paper investigated whether ELM can outperform MLPNN and MLR, for modeling DO with and without water quality variables as predictors. Various water quality datasets obtained from eight different stations; USA were used for developing

the models. The novelty of the present work lies on one hand in the application of the ELM approaches at the first time, and in the other hand, an attempt to estimate DO concentration without water quality variables and using the components of the Gregorian calendar as inputs. From the results obtained, we can draw the following conclusions. First, in the first application using the four water quality variables as inputs, the OP-ELM was found to be more accurate than the OS-ELM, S-ELM, and R-ELM at seven stations and provided the best accuracy in the validation phase, with high $R$ and NSE values (from 0.924 to 0.989) and (from 0.853 to 0.933), respectively. Comparison of the best OP-ELM models suggests that the OP-ELM models are superior to the MLPNN and MLR at all stations, and the MLR is the worst models in all stations. Therefore, the OP-ELM model is a very promising tool for modeling DO concentration. Second, in the second application using the components of the Gregorian calendar as inputs, the R-ELM performed the best at seven stations and provided the best accuracy in the validation phase, with high $R$ and NSE values (from 0.884 to 0.969) and (from 0.781 to 0.939), respectively. The MLPNN is ranked in the second place at five stations, while the MLR results suggest that the linear model is very poor with very low accuracy, which leads to the conclusion that the MLR is insufficient for mapping a good relation between DO and the components of the Gregorian calendar; hence, the selection of nonlinear models like ELM and MLPNN are highly recommended. In addition, the important conclusion could be drawn that, in some cases, the inclusion of the turbidity as input did not improve the accuracies of the proposed models; on the contrary, it made them worse. The study also indicated that the hourly DO might be successfully estimated by the ELM models using only components of the Gregorian calendar as inputs when water quality data are missing or do not exist.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that there are no conflicts of interest regarding this manuscript.

# References

Abdul-Aziz OI, Ishtiaq KS (2014) Robust empirical modelling of dissolved oxygen in small rivers and streams: scaling by a single reference observation. J Hydrol 511:648–657. doi:10.1016/j.jhydrol.2014.02.022

Abdullah SS, Malek MA, Abdullah NS, Kisi O, Yap KS (2015) Extreme learning machines: a new approach for prediction of reference evapotranspiration. J Hydrol 527:184–195. doi:10.1016/j.jhydrol.2015.04.073

Akkoyunlu A, Altun H, Cigizoglu H (2011) Depth-integrated estimation of dissolved oxygen in a lake. ASCE J Environ Eng 137(10):961–967. doi:10.1061/(ASCE)EE.1943-7870.0000376

Akusok A, Veganzones D, Miche Y, Björk K-M, du Jardin P, Severin E, Lendasse A (2015) MD-ELM: originally mislabeled samples detection using OP-ELM model. Neurocomputing 159:242–250. doi:10.1016/j.neucom.2015.01.055

Alizadeh MJ, Kavianpour MR (2015) Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean. Mar Pollut Bull 98:171–178. doi:10.1016/j.marpolbul.2015.06.052

Ay M, Kisi O (2012) Modeling of dissolved oxygen concentration using different neural network techniques in Foundation Creek, El Paso County, Colorado. ASCE J Environ Eng 138(6):654–662. doi:10.1061/ (ASCE) EE.1943-7870.0000511

Ay M, Kisi O (2016) Estimation of dissolved oxygen by using neural networks and neuro fuzzy computing techniques. KSCE J Civ Eng 00(0):1–9. doi:10.1007/s12205-016-0728-6

Deo RC, Şahin M (2015) Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. Atmos Res 153:512–525. doi:10.1016/j.atmosres.2013.11.002

Deo RC, Şahin M (2016) An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queenslad. Environ Monit Assess 188:90. doi:10.1007/s10661-016-5094-9

Diamantopoulou MJ, Antonopoulos VZ, Papamichail DM (2007) Cascade correlation artificial neural networks for estimating missing monthly values of water quality parameters in rivers. Water Resour Manag 21:649–662. doi:10.1007/s11269-006-9036-0

Efron B, Hastie T, Johnstone I, Tibshirani R (2004) Least angle regression. Ann Stat 32:407–499. doi:10.1214/009053604000000067

Evrendilek F, Karakaya N (2014a) Monitoring diel dissolved oxygen dynamics through integrating wavelet denoising and temporal neural networks. Environ Monit Assess 186:1583–1591. doi:10.1007/s10661-013-3476-9

Evrendilek F, Karakaya N (2014b) Regression model-based predictions of diel, diurnal and nocturnal dissolved oxygen dynamics after wavelet denoising of noisy time series. Physica A 404:8–15. doi:10.1016/j.physa.2014.02.062

Evrendilek F, Karakaya N (2015) Spatiotemporal modeling of saturated dissolved oxygen through regressions after wavelet denoising of remotely and proximally sensed data. Earth Sci Inf 8:247–254. doi:10.1007/s12145-014-0148-4

Faruk DÖ (2010) A hybrid neural network and ARIMA model for water quality time series prediction. Eng Appl Artif Intell 23:586–594. doi:10.1016/j.engappai.2009.09.015

Grigorievskiy A, Miche Y, Ventelä AM, Séverin E, Lendasse A (2014) Long-term time series prediction using OP-ELM. Neural Netw 51:50–56. doi:10.1016/j.neunet.2013.12.002

Gulgundi MS, Shetty A (2016) Identification and apportionment of pollution sources to groundwater quality. Environ Process 3:451–461. doi:10.1007/s40710-016-0160-4

Haykin S (1999) Neural networks a comprehensive foundation. Prentice Hall, Upper Saddle River

Heddam S (2014a) Generalized regression neural network (GRNN) based approach for modelling hourly dissolved oxygen concentration in the upper Klamath River, Oregon, USA. Environ Techno 35(13):1650–1657. doi:10.1080/09593330.2013.878396

Heddam S (2014b) Modelling hourly dissolved oxygen concentration (DO) using two different adaptive neurofuzzy inference systems (ANFIS): a comparative study. Environ Monit Assess 186:597–619. doi:10.1007/s10661-013-3402-1

Heddam S (2014c) 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, Oregon, USA. Environ Sci Pollut Res 21:9212–9227. doi:10.1007/s11356-014-2842-7

Heddam S (2016a) Simultaneous modelling and forecasting of hourly dissolved oxygen concentration (DO) using radial basis function neural network (RBFNN) based approach: a case study from the Klamath River, Oregon, USA. Model Earth Syst Environ 2:135. doi:10.1007/s40808-016-0197-4

Heddam S (2016b) Fuzzy neural network (EFuNN) for modelling dissolved oxygen concentration (DO). In: Kahraman C, Sari IU (eds) Intelligence Systems in Environmental Management: Theory and Applications, Intelligent Systems Reference Library 113, pp 231–253. doi:10.1007/978-3-319-42993-9_11

Heddam S (2016c) Use of optimally pruned extreme learning machine (OP-ELM) in forecasting dissolved oxygen concentration (DO) several hours in advance: a case study from the Klamath River, Oregon, USA. Environ Process 3(4):909–937. doi:10.1007/s40710-016-0172-0

Heddam S (2016d) New modelling strategy based on radial basis function neural network (RBFNN) for predicting dissolved oxygen concentration using the components of the Gregorian calendar as inputs: case study of Clackamas River, Oregon, USA. Model. Earth Syst. Environ 2:167. doi:10.1007/s40808-016-0232-5

Heddam S (2016e) Secchi disk depth estimation from water quality parameters: artificial neural network versus multiple linear regression models? Environ Process 3(1):525–536. doi:10.1007/s40710-016-0144-4

Hornik K (1991) Approximation capabilities of multilayer feedforward networks. Neural Netw 4(2):251–257. doi:10.1016/0893-6080(91)90009-T

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal Approximators. Neural Netw 2:359–366. doi:10.1016/0893-6080(89)90020-8

Huang G (2015) What are extreme learning machines? Filling the gap between frank Rosenblatt's dream and John von Neumann's puzzle. Cogn Comput 7:263–278. doi:10.1007/s12559-015-9333-0

Huang GB, Zhu QY, Siew CK (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE Proceedings of International Joint Conference on Neural Networks, vol. 2, pp 985–990. doi:10.1109/IJCNN.2004.1380068

Huang GB, Chen L, Siew CK (2006a) Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans Neural Netw 17(4):879–892. doi:10.1109/TNN.2006.875977

Huang GB, Zhu QY, Siew CK (2006b) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501. doi:10.1016/j.neucom.2005.12.126

Huang GB, Wang DH, Lan Y (2011) Extreme learning machines: a survey. Int J Mach Learn Cybern 2:107–122. doi:10.1007/s13042-011-0019-y

Huang GB, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern Part B Cybern 42:513–529. doi:10.1109/TSMCB.2011.2168604

Huang G, Huang GB, Song S, You K (2015) Trends in extreme learning machines: a review. Neural Netw 61:32–48. doi:10.1016/j.neunet.2014.10.001

Kisi O, Akbari N, Sanatipour M, Hashemi A, Teimourzadeh K, Shiri J (2013) Modeling of dissolved oxygen in river water using artificial intelligence techniques. J Environ Inform 22(2):92–101. doi:10.3808/jei.201300248

Legates DR, McCabe GJ (1999) Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation. Water Resour Res 35:233–241. doi:10.1029/1998WR900018

Liang NY, Huang GB, Saratchandran P, Sundararajan N (2006) A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Trans Neural Netw 17:1411–1423. doi:10.1109/TNN.2006.880583

Lima AR, Cannon AJ, Hsieh WW (2016) Forecasting daily streamflow using online sequential extreme learning machines. J Hydrol 537:431–443. doi:10.1016/j.jhydrol.2016.03.017

Liu S, Yan M, Tai H, Xu L, Li D (2012) Prediction of dissolved oxygen content in aquaculture of hyriopsis cumingii using Elman neural network. Li D, Chen Y (eds) Computer and Computing Technologies in Agriculture V (CCTA) 2011, Part III. IFIP Advances in Information and Communication Technology vol. 370, pp 508–518. doi:10.1007/978-3-642-27275-2-57.

Liu S, Xu L, Li D, Li Q, Jiang Y, Tai H, Zeng L (2013) Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization. Comput Electron Agric 95:82–91. doi:10.1016/j.compag.2013.03.009

Liu S, Xu L, Jiang Y, Li D, Chen Y, Li Z (2014) A hybrid WA-CPSO-LSSVR model for dissolved oxygen content prediction in crab culture. Eng Appl Artif Intell 29:114–124. doi:10.1016/j.engappai.2013.09.019

Miche Y, Sorjamaa A, Lendasse A (2008a) OP-ELM: theory, experiments and a toolbox. In: Proceedings of the international conference on artificial neural networks. Lecture Notes in Computer Science, vol. 5163, Prague, pp 145–154. doi:10.1007/978-3-540-87536-9_16.

Miche Y, Bas P, Jutten C, Simula O, Lendasse A (2008b) A methodology for building regression models using extreme learning machine: OP-ELM. In: ESANN 2008, European Symposium on Artificial Neural Networks, Bruges

Miche Y, Sorjamaa A, Bas P, Simula O, Jutten C, Lendasse A (2010) OP-ELM: optimally pruned extreme learning machine. IEEE Trans Neural Netw 21(1):158–162. doi:10.1109/TNN.2009.2036259

Moreno R, Corona F, Lendasse A, Graña M, Galvão LS (2014) Extreme learning machines for soybean classification in remote sensing hyperspectral images. Neurocomputing 128:207–216. doi:10.1016/j.neucom.2013.03.057

Pouzols FM, Lendasse A (2010a) Evolving fuzzy optimally pruned extreme learning machine: a comparative analysis. IEEE Int Conf Fuzzy Syst (FUZZ):1–8. doi:10.1109/FUZZY.2010.5584327

Pouzols FM, Lendasse A (2010b) Evolving fuzzy optimally pruned extreme learning machine for regression problems. Evol Syst 1:43–58. doi:10.1007/s12530-010-9005-y

Ranković V, Radulović J, Radojević I, Ostojić A, Čomić L (2012) Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system. J Hydro Inform 14(1):167–179. doi:10.2166/hydro.2011.084

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland PDP, Research Group (eds) Parallel distributed processing: explorations in the microstructure of cognition. Foundations, vol. I. MIT Press, Cambridge, pp 318–362

Shiri J, Shamshirband S, Kisi O, Karimi S, Bateni SM, Nezhad SH, Hashemi A (2016) Prediction of water-level in the Urmia Lake using the extreme learning machine approach. Water Resour Manag. doi:10.1007/s11269-016-1480-x

Similä T, Tikka J (2005) Multiresponse sparse regression with application to multidimensional scaling. In: Artificial neural networks: formal models and their applications-ICANN 2005, vol. 3697/2005, pp. 97–102. doi:10.1007/11550907_16

Singh RP, Dabas N, Chaudhary V, Nagendra (2016) Online sequential extreme learning machine for watermarking in DWT domain. Neurocomputing 174:238–249. doi:10.1016/j.neucom.2015.03.115

Sorjamaa A, Miche Y, Weiss R, Lendasse A (2008) Long-term prediction of time series using NNE-based projection and OP-ELM. In: Proceedings of the IEEE international joint conference on neural networks (IJCNN), Hong Kong, pp 2674–2680. doi:10.1109/IJCNN.2008.4634173.

Sovilj D, Sorjamaa A, Yu Q, Miche Y, Séverin E (2010) OPELM and OPKNN in long-term prediction of time series using projected input data. Neurocomputing 73:1976–1986. doi:10.1016/j.neucom.2009.11.033

Wang X, Han M (2014) Online sequential extreme learning machine with kernels for nonstationary time series prediction. Neurocomputing 145:90–97. doi:10.1016/j.neucom.2014.05.068

Wang Y, Zheng T, Zhao Y, Jiang J, Wan YG, Guo L, Wang P (2013) Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China. Environ Sci Pollut Res 20:8909–8923. doi:10.1007/s11356-013-1874-8

Xu L, Liu S (2013) Study of short-term water quality prediction model based on wavelet neural network. Math Comput Model 58:807–813. doi:10.1016/j.mcm.2012.12.023

Yadav B, Ch S, Mathur S, Adamowski J (2016) Discharge forecasting using an online sequential extreme learning machine (OS-ELM) model: a case study in Neckar River, Germany. Measurement 92:433–445. doi:10.1016/j.measurement.2016.06.042

Yaseen ZM, Jaafar O, Deo RC, Kisi O, Adamowski J, Quilty J, El-shafie A (2016) Boost stream-flow forecasting model with extreme learning machine data-driven: a case study in a semi-arid region in Iraq. J Hydrol. doi:10.1016/j.jhydrol.2016.09.035