

## Integrated versus isolated scenario for prediction dissolved oxygen at progression of water quality monitoring stations

A. Najah<sup>1</sup>, A. El-Shafie<sup>2</sup>, O. A. Karim<sup>2</sup>, and O. Jaafar<sup>2</sup>

<sup>1</sup>PhD Candidate, Dept. Civil & Structural Eng, Universiti Kebangsaan Malaysia, UKM, Malaysia

<sup>2</sup>Lecturers, Dept. Civil & Structural Eng, Universiti Kebangsaan Malaysia, UKM, Malaysia

Received: 15 June 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 23 June 2011

Revised: 22 August 2011 – Accepted: 22 August 2011 – Published: 29 August 2011

**Abstract.** This study examined the potential of Multi-layer Perceptron Neural Network (MLP-NN) in predicting dissolved oxygen (DO) at Johor River Basin. The river water quality parameters were monitored regularly each month at four different stations by the Department of Environment (DOE) over a period of ten years, i.e. from 1998 to 2007. The following five water quality parameters were selected for the proposed MLP-NN modelling, namely; temperature (Temp), water pH, electrical conductivity (COND), nitrate ( $\text{NO}_3$ ) and ammonical nitrogen ( $\text{NH}_3\text{-NL}$ ). In this study, two scenarios were introduced; the first scenario (Scenario 1) was to establish the prediction model for DO at each station based on five input parameters, while the second scenario (Scenario 2) was to establish the prediction model for DO based on the five input parameters and DO predicted at previous station (upstream). The model needs to verify when output results and the observed values are close enough to satisfy the verification criteria. Therefore, in order to investigate the efficiency of the proposed model, the verification of MLP-NN based on collection of field data within duration 2009–2010 is presented. To evaluate the effect of input parameters on the model, the sensitivity analysis was adopted. It was found that the most effective inputs were oxygen-containing ( $\text{NO}_3$ ) and oxygen demand ( $\text{NH}_3\text{-NL}$ ). On the other hand, Temp and pH were found to be the least effective parameters, whereas COND contributed the lowest to the proposed model. In addition, 17 neurons were selected as the best number of neurons in the hidden layer for the MLP-NN architecture. To evaluate the performance of the proposed model, three statistical indexes were used, namely; Coefficient of Efficiency (CE), Mean Square Error (MSE) and Coefficient of Correlation (CC). A relatively low correlation between the observed and predicted values in the testing data set was obtained in

Scenario 1. In contrast, high coefficients of correlation were obtained between the observed and predicted values for the test sets of 0.98, 0.96 and 0.97 for all stations after adopting Scenario 2. It appeared that the results for Scenario 2 were more adequate than Scenario 1, with a significant improvement for all stations ranging from 4 % to 8 %.

### 1 Introduction

Water is a vital resource necessary for all aspects of human and ecosystem survival and health. In addition to drinking and personal hygiene, water is needed for agricultural production, industrial and manufacturing processes, hydroelectric power generation, waste assimilation, recreation, navigation, enhancement of fish and wildlife and a variety of other purposes (Biswas, 1981). The term water quality is used to describe the condition of water, including its chemical, physical and biological characteristics. Water quality is one of the main characteristics of a river, which purpose is not only for human water supply (Dogan et al., 2009; Lopes et al., 2005).

Dissolved oxygen (DO) is one of the important water quality parameters for the survival of aquatic life. It is a critical parameter used frequently and continuously to determine the water quality of rivers. The sources of DO in a water body include re-aeration from the atmosphere, photosynthetic oxygen production and DO loading. The sinks include oxidation of carbonaceous and nitrogenous materials, sediment oxygen demand, and respiration by aquatic plants (Kuo et al., 2007). The problems associated with low concentrations of DO in rivers have been recognized for over a century and the impacts of low DO concentrations or, at the extreme, anaerobic conditions in a normally well-oxygenated river system, are an unbalanced ecosystem with fish mortality, odours and other aesthetic nuisances. When DO concentrations are reduced, aquatic animals are forced to alter their breathing patterns or lower their level of activity. Both of these actions will



Correspondence to: A. Najah  
(ali\_najah@ymail.com)

retard their development, and can cause reproductive problems (such as increased egg mortality and defects) and/or deformities (Kalff, 2002; Cox, 2003)

The simplicity of measurement of dissolved oxygen obscures the fact that a number of physical and chemical processes within the water body contribute to the dissolved oxygen level within the stream (Lopes et al., 2005). Notably, the Winkler method is the most reliable technique used to measure dissolved oxygen in freshwater systems. This is a multi-step chemical method, where the test is performed on-site, as delays between sample collections and testing may result in an alteration in oxygen content (Sengorur et al. 2006).

Water quality modelling is the basis of water pollution control project. It predicts the water quality tendency of varieties according to the current water environment quality condition, transfer and transformation rule of the pollutants in the river basin (Najah et al., 2010). In addition, several water quality models, such as deterministic and stochastic models have been developed in order to manage the best practices for conserving water quality (Hull et al., 2008; Einax et al., 1999). Most of these models are very complex and require a significant amount of field data to support the analysis. Furthermore, many statistical-based water quality models, which assume the relationship between response variable and prediction variable, are linear and distributed normally (Ansare et al., 2000; Garcia et al., 2002). However, as water quality can be affected by so many factors, traditional data processing methods are no longer efficient enough for solving the problem (Xiang et al., 2006), as such factors show a complicated non-linear relation to the variables of water quality forecast. Therefore, utilizing statistical approaches usually does not possess high precision (Rankovic et al., 2010).

Recently, the neural networks approach has been applied to many branches of science. There are a number of studies in which neural networks are used to address water resources problems (Alvisi et al., 2006; El-shafie et al., 2008; Najah et al., 2009; Akhtar et al., 2009; Hung et al., 2009; Ming et al., 2010; Najah et al., 2010a, b; El-shafie et al., 2011). In water quality issues, artificial Neural Networks (ANNs) were first applied by French and Recknagel (1994) to the task of learning to predict algal blooms from water quality databases. In their application, a feed-forward ANN was trained to make predictions of abundance of species of phytoplankton in Saldenbach Reservoir, Germany. Similarly, Yabunaka et al. (1997) also applied ANNs to predict algal bloom by simulating the future growth of five phytoplankton species and the chlorophyll-*a* (Chl-*a*) concentrations in the same lake.

Hence, motivated by the successful applications in modelling non-linear system behaviours in a wide range of areas, ANNs are used to predict water quality parameters in complex systems. The literature offers some recent successful ANNs applications related to water quality predictions. The main intentions were to minimize fieldwork and improve the

accuracy of prediction. For instance, Hatzikos et al. (2005) utilized neural networks with active neurons as a modelling tool for predicting seawater quality indicators, such as water temperature, pH, dissolved oxygen and turbidity.

The water quality of Johor River is deteriorating due to increasing levels of various pollutants. It continues to be silted and contaminated by wastes due to the lack of enforcement by local authorities. These contaminants eventually flow into the estuaries of Johor River, which are rich habitats that provide spawning and feeding areas for fish and poultry.

This study demonstrated the application of Artificial Neural Network to predict water quality parameters in terms of dissolved oxygen (DO), having the dynamic processes hidden in the measured data itself. The use of Multi-Layer Perceptron Neural Network (MLP-NN) model in water quality prediction in Johor River could be complementary in capturing patterns of historical data set and improving the prediction accuracy.

## 2 Methods and materials

### 2.1 Study area data analysis

Johor is the second largest state in the Malaysia Peninsular, with an area of 18 941 km<sup>2</sup>. The Johor River and its tributaries are important sources of water supply, not only for the state of Johor but also for Singapore. The river comprises 122.7 km long drains, covering an area of 2636 km<sup>2</sup>. It originates from Mount Gemuruh and flows through the south-eastern part of Johor and finally into the Straits of Johor. The catchment is irregular in shape. The maximum length and breadth are 80 km and 45 km, respectively. About 60 % of the catchment comprises undulating highlands rising to a height of 366 m, while the remainder encompasses lowland and swampy areas.

The station's location map is provided in Fig. 1a and b. This station includes four locations along the main stream of the river, which are near to the mouth of the major tributaries. The proposed models in this research were constructed under the assumption that land use/cover has remained unchanged during the study period. However, land use/cover is an important factor for the prediction of water quality parameters. A more precise prediction of water quality parameters could be achieved by adding variables representing the land use/cover status into the model.

The selection of appropriate input parameters is a very important aspect for the neural network modelling. In order to use the MLP-NN structures effectively, the input parameters must be selected with great care. This highly depends on better understanding of the problem. The choice of input parameters based on statistical correlation analysis is the most popular analytical technique for selecting input. The drawback of cross-correlation is that only able to capture linear dependence between two variables. Consequently, it can lead



to omission of important inputs that are nonlinearity related to the output. To evaluate the effect of input parameters on the model, two evaluation processes were used. First, a priori knowledge supported by statistical correlation analysis. The second assessment process was based on the prediction accuracy of water quality parameters. Moreover, in the literature, various input parameters have been used to create the model for predicting dissolved oxygen (Table 1).

Based on the literature, existing measured values and statistical analyses, the following five water quality parameters were selected for the ANN modelling in this study, namely; temperature (Temp), water pH, electrical conductivity (COND), nitrate ( $\text{NO}_3$ ) and ammoniacal nitrogen ( $\text{NH}_3\text{-NL}$ ). The river water quality parameters were monitored regularly each month at four different stations by the Department of Environment (DOE) over a period of ten years, i.e. from 1998 to 2007. The analysed results of the study sites are given in Table 2.

The water quality data that used in this study was collected within the Johor River. Both in-situ measurements and laboratory analysis were conducted. Four observation points were selected based on the location of the stations. At each point samples were taken at three different depths. These depths are surface, middle and bottom. All these parameters are measured during sampling using a water quality checker which is known as Multi-parameter YSI 550A with five sensors. The YSI was re-calibrated daily to ensure data accuracy.

pH is the indicator for acidic and alkaline conditions of water status. Notably, the INWQS threshold range of pH for Malaysian rivers is 5.00 to 9.00 (DOE, 1994). From the results, the mean pH of the Johor River varied from 6.22 to 6.39. At all stations, pH was almost equal and did not show a statistically significant difference. Basically, the pH value is controlled by the dissolved carbon dioxide ( $\text{CO}_2$ ), which forms carbonic acid in water (Hem, 1985). The main source of such chemical should be urban runoff or industrial wastewater.

On the other hand, electrical conductivity (COND) is associated with major water quality parameters due to the dilution effect of stream flow and can be used as a general water quality indicator. The mean COND of Johor River varied from 20.01  $\mu\text{S}$  to 22.14  $\mu\text{S}$  except at Station 3, which was 53.8  $\mu\text{S}$ . This change in COND might be an indicator of a discharge or some other source of pollution that entered the stream.

Meanwhile,  $\text{NO}_3$  ion is usually derived from anthropogenic sources like agricultural fields, domestic sewage and other waste effluents containing nitrogenous compounds (Das and Acharya, 2003). It appeared that the mean values among all locations did not vary greatly.

Ammoniacal nitrogen ( $\text{NH}_3\text{-N}$ ) is used to measure the amount of ammonia, i.e. a toxic pollutant often found in agriculture fertilizer and domestic sewage.  $\text{NH}_3\text{-N}$  has been promoted as a tool to define the status of surface water quality in Malaysia (DOE, 2003). The mean  $\text{NH}_3$  of the Johor River

varied from 0.1  $\text{mg l}^{-1}$  to 0.15  $\text{mg l}^{-1}$ . At all stations,  $\text{NH}_3$  was almost equal and did not show a statistically significant difference.

The coefficient of variation (CV) is employed to measure the data statistical dispersion, which is the mean normalized standard deviation of the given data set (Singh et al., 2009).

$$cv(\%) = \frac{SD}{Mean} * 100 \quad (1)$$

All parameters showed a coefficient of variation between 3.08 % and 214.96 %. Such variability among the samples might be due to the large geographical variations in climate influences in the study area. Temperature showed the lowest variation which might be due to the buffering capacity of the river. The correlation coefficient between DO and the input parameters was calculated and presented in Table 3.

## 2.2 Artificial Neural Network (ANN)

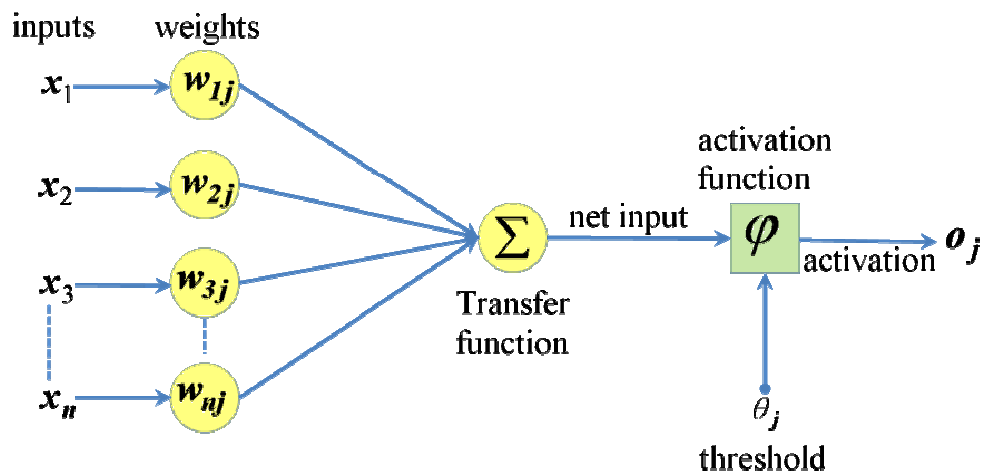
An artificial neural network (ANN) is tailored to mimic natural neural networks using a computing process (Haykin, 1999). Among many types of ANNs, the most widely used is the feed-forward neural network such as multi-layer perceptron (MLP) network with back-propagation training algorithm. The MLP is organized as layers of computing elements, known as neurons, which are connected between layers via weights. Apart from the input layer receiving inputs from the environment and the output layer generating the network's response, one or more intermediate hidden layers also exist. For brevity, we refrain from discussing the details of the neural network methodology and instead refer the reader to the papers written by Lek et al. (1996b), as well as Olden and Jackson (2001) for more comprehensive treatments.

Generally, forecasting models can be divided into statistical and physical based approaches. Statistical approaches determine the relationships between historical data sets, whereas physical based approaches model the underlying processes directly. MLP networks are closely related to statistical models and are the type of ANN most suited to forecasting applications (Rumelhart et al., 1986). When using ANNs for forecasting, the modelling philosophy employed is similar to that used in traditional statistical approaches. In both cases, the unknown model parameters (i.e. the connection weights in the case of ANNs) are adjusted to obtain the best match between the historical set of model inputs and the corresponding outputs.

These neural networks are commonly used in ecological studies because they are believed to be universal approximates of any continuous function (Hornik and White, 1989). A neural network consists of at least three or more layers, which comprise an input layer, an output layer and a number of hidden layers, as shown in Fig. 2. Each neuron in one layer is connected to the neurons in the next layer, but there are no connections between the units of the same layer (Kasabov and Foundations, 1996). The number of neurons in each

**Table 1.** Input parameters used in previous studies for the ANN model.

| Author(s)and year      | Input Variable   | Location(s)                             |
|------------------------|--|---|
| Sengorur et al. (2006) | BOD, Temp, Water discharge, NO2-N, NO3-N                           | N/A                                     |
| Kuo et al. (2007)      | pH, Chl-a, NH4N, No3N, temp, month                                 | Te-Chi Reservoir, Taiwan                |
| Ying et al. (2007)     | Turbidity, Temp, pH, Hardness, Alkalinity, Chloride, NH4-N, NO2-N  | Yuqiao reservoir, china                 |
| Palani et al. (2008)   | DO, Chl-a, temp  | Singapore coastal, Singapore            |
| Zaqoot et al. (2009)   | Conductivity, Turbidity, Temp, PH, Wind speed                      | Mediterranean Sea along Gaza, Palestine |
| Singh et al. (2009)    | pH, TS, T-ALK, T-Hard, CL, PO4, K, Na, NH4N, NO3N, COD             | Gomti, India                            |
| Rankovic et al. (2010) | Temp, PH, Chloride, Conductivity, phosphate, NH3, Mg, No2, Fe, No3 | Gruža reservoir, Serbia                 |



**Fig. 2.** A typical multi-layer perceptron neural network architecture.

layer may vary depending on the problem. The weighted sum of the input components is calculated as follows (Freeman and Skapura, 1991):

$$Net_j = \sum_{i=1}^n W_{ij} + \theta_j \tag{2}$$

where  $Net_j$  is the weighted sum of the  $j$ th neuron for the input data received from the preceding layer with  $n$  neurons,  $W_{ij}$  is the weight between the  $j$ th neuron and the  $i$ th neuron in the preceding layer and  $\theta_j$  is the bias term of the  $j$ th neuron. The output of the  $j$ th neuron  $out_j$  is calculated with a sigmoid function as follows:

$$out_j = f(Net_j) = \frac{1}{1 + e^{-Net_j}} \tag{3}$$

The network is trained by adjusting the weights. The training process is done with a large number of training sets and training cycles (epochs). The main goal of the learning procedure is to find the optimal set of weights, which can ideally produce the correct output for the relative input. The output of the network is compared with the desired response to determine the error. The performance of the MLP is measured in terms of a desired signal and the criterion for con-

vergence. For one sample, it is determined by the sum square error (SSE), expressed as follows:

$$SSE = \sum_{i=1}^m (T_i - out_i)^2 \tag{4}$$

where  $T_i$  and  $out_i$  are the desired (target) output and output of the neural network, respectively, for the  $i$ th output neuron, and  $m$  is the number of neurons in the output layer.

### 2.3 Dissolved Oxygen (DO) prediction with MLP-NN

In fact, the prediction procedure is, by detention, an operation through which the future dissolved oxygen pattern can be provided. In this study, the ANN with its non-linear and stochastic modelling capabilities was utilized to develop a prediction model that mimicked the DO pattern at the Johor River based on the five input parameters (Scenario 1) mentioned earlier, which can be expressed as follows:

$$DO_N = f_{MLP-NN}(pH_N, Temp_N, NO_{3N}, NH_{3N}, COND_N) \tag{5}$$

$$N = 1, 2, 3, 4$$

where  $DO_N$  is the dissolved oxygen at station  $N$ , and  $f_{MLP-NN}(\cdot)$  is the non-linear function predictor constructed

**Table 2.** Basic statistics of the measured water quality parameters in Johor River.

| Sampling Site |      | COND ( $\mu\text{s}$ ) | PH – | NH3-NL ( $\text{mg l}^{-1}$ ) | TEMP ( $^{\circ}\text{C}$ ) | NO3 ( $\text{mg l}^{-1}$ ) | DO ( $\text{mg l}^{-1}$ ) |
|---------------|------|------------------------|------|-------------------------------|-----------------------------|----------------------------|---------------------------|
| DO-1          | Mean | 22.14                  | 6.39 | 0.14                          | 27.03                       | 0.64                       | 6.25                      |
|               | Min  | 3.00                   | 5.49 | 0.01                          | 24.08                       | 0.01                       | 2.91                      |
|               | Max  | 56.00                  | 7.83 | 1.07                          | 30.33                       | 2.93                       | 7.34                      |
|               | SD   | 6.67                   | 0.45 | 0.18                          | 0.83                        | 0.36                       | 0.59                      |
|               | CV   | 30.13                  | 7.07 | 129.30                        | 3.08                        | 56.00                      | 9.52                      |
| DO-2          | Mean | 53.80                  | 6.22 | 0.10                          | 27.20                       | 0.66                       | 6.15                      |
|               | Min  | 42.00                  | 5.43 | 0.01                          | 25.34                       | 0.27                       | 4.84                      |
|               | Max  | 69.60                  | 7.28 | 0.38                          | 29.82                       | 1.28                       | 6.96                      |
|               | SD   | 7.03                   | 0.33 | 0.08                          | 0.87                        | 0.20                       | 0.37                      |
|               | CV   | 13.06                  | 5.39 | 81.43                         | 3.19                        | 29.40                      | 6.10                      |
| DO-3          | Mean | 20.22                  | 6.36 | 0.15                          | 27.31                       | 0.76                       | 5.79                      |
|               | Min  | 1.00                   | 5.67 | 0.01                          | 23.35                       | 0.01                       | 4.43                      |
|               | Max  | 45.00                  | 8.41 | 2.46                          | 31.93                       | 4.70                       | 6.78                      |
|               | SD   | 6.98                   | 0.44 | 0.31                          | 1.20                        | 0.61                       | 0.66                      |
|               | CV   | 34.51                  | 6.98 | 214.96                        | 4.41                        | 80.34                      | 11.44                     |
| DO-4          | Mean | 20.01                  | 6.29 | 0.14                          | 27.42                       | 0.64                       | 5.55                      |
|               | Min  | 9.00                   | 5.59 | 0.01                          | 24.58                       | 0.01                       | 4.41                      |
|               | Max  | 38.00                  | 8.09 | 0.83                          | 29.78                       | 3.22                       | 7.53                      |
|               | SD   | 4.54                   | 0.38 | 0.15                          | 1.03                        | 0.35                       | 0.61                      |
|               | CV   | 22.66                  | 6.10 | 106.55                        | 3.74                        | 53.69                      | 11.05                     |

**Table 3.** The correlation coefficient between DO and the input parameters.

|      | COND     | pH       | NH3-NL   | TEMP     | NO3      |
|------|----------|----------|----------|----------|----------|
| DO-1 | 0.041324 | 0.461876 | 0.057188 | 0.351994 | 0.016422 |
| DO-2 | -0.09643 | 0.061553 | -0.21588 | 0.355585 | -0.25981 |
| DO-3 | 0.17228  | -0.26767 | -0.04009 | 0.197713 | -0.40085 |
| DO-4 | -0.0658  | -0.18141 | -0.33326 | -0.06502 | -0.32051 |

by the MLP-NN network. Hence, a total of four models for DO prediction were constructed.

Most of the recent studies attempted to predict the concentrations of DO at each station. Generally, the water pollution of a downstream station is affected by the discharge of local area from the upstream station (Zaqoot et al., 2009). Hence, it was required to consider the effect of DO at the upstream station in the proposed model. Therefore, second scenario (Scenario 2) was formed to establish the model prediction for DO at each station based on the five input parameters. The predicted DO at the previous station (upstream) can be expressed following Eq. (6). This procedure of using the predicted DO can be repeated for the third and fourth stations at downstream. The schematic representation of the proposed networks for Scenario 2 is shown in Fig. 3.

$$\text{DO}_{N+1} = f_{\text{MLP-NN}}(\text{pH}_{N+1}, \text{Temp}_{N+1}, \text{NO}_{3N+1}, \text{NH}_{3N+1}, \text{COND}_{N+1}, \text{DO}_{pN}) \quad (6)$$

The ANN models were established using the above two equations. The architecture of the networks consisted of an input layer of five and six neurons for Scenario 1 and Scenario 2, respectively. The hyperbolic tangent sigmoid transfer function was employed between the input and the hidden layers. Moreover, a linear transfer function was employed between the hidden and output layers (corresponding to the predicted DO). Finally, the optimal ANN, together with a flowchart of the algorithm's procedure, is shown in Fig. 4.

It is important to divide the data set in such a way that the training, validation and test data sets are statistically comparable. In this study, the water quality data were divided into three sets. The first set contained 60 % of the data set used as the training set; the second test contained 25 % of the

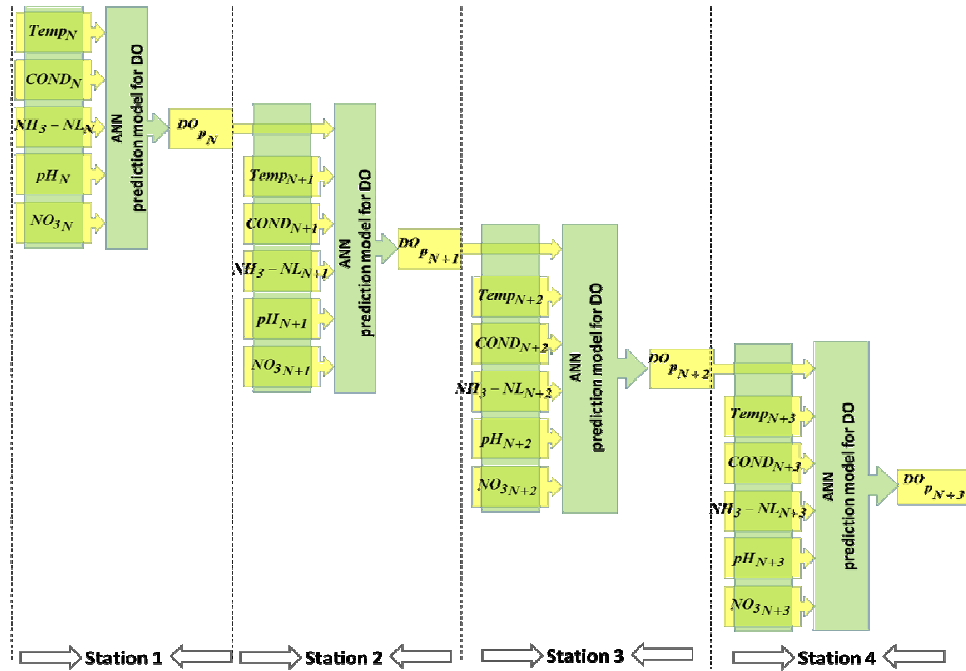


Fig. 3. A schematic representation for Scenario 2.

data set used as the validation set and 15 % of the data set, which the network had never seen before, was used as the testing set. The statistical properties (i.e. mean, standard deviation, range) from them were compared, as shown in Fig. 5. According to the statistical properties of those data sets, no significant differences between the divisions of the data were observed. All samples were normalized in the [0 1] range. Thus, all of the data sets ( $X_i$ ) from the training, validation and test sets were scaled to a new value  $x_i$  as follows:

$$x_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (7)$$

2.4 Selection of back propagation training algorithm

The back propagation (BP) learning algorithm (Rumelhart et al., 1986) is a method conventionally used to perform the training of Artificial Neural Networks for adjusting weighted connections. Standard back propagation is a gradient descent algorithm in which the network weights are moved along the negative of the gradient of the performance function. Although traditional BP uses a gradient descent algorithm to determine the weights in the network, it computes rather slowly due to linear convergence.

There are a number of variations on the basic algorithm that are based on other standard optimization techniques. In this paper, Levenberg-Marquardt algorithm (LMA) was used, which appears to be the fastest method for training moderate-sized feed-forward neural networks (Demuth et al., 2008). The LMA is a very simple but robust method, which provides a numerical solution to the problem of minimizing

a function over a space of parameters for the function. Principally, it involves in solving the following equation:

$$\delta = (\mathbf{J}^t \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^t \mathbf{E} \quad (8)$$

where  $\mathbf{I}$  is the identity matrix,  $\mathbf{J}$  is the Jacobian matrix for the system,  $\lambda$  is the Levenberg’s non-negative damping factor,  $\delta$  is the weight update vector that we want to find and  $\mathbf{E}$  is the error vector containing the output errors for each input vector used in training the network. The  $\delta$  tells us by how much we should modify our network weights to reach a better solution.  $\lambda$  is adjusted at each iteration. If the reduction of  $E$  is rapid, a smaller value can be used, bringing the algorithm closer to the Gauss-Newton algorithm, whereas if iteration gives insufficient reduction in the residual,  $\lambda$  can be increased, giving a step closer to the gradient descent direction. In that way, LMA is considered as a hybrid between the classical Newton and steepest descent algorithms (Souza et al., 2009). The Jacobian matrix can be created by taking the partial derivatives of each output in respect to each weight and has the following form:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial F(x_1, \omega)}{\partial \omega_1} & \dots & \frac{\partial F(x_1, \omega)}{\partial \omega_w} \\ \vdots & \ddots & \vdots \\ \frac{\partial F(x_N, \omega)}{\partial \omega_1} & \dots & \frac{\partial F(x_N, \omega)}{\partial \omega_w} \end{bmatrix} \quad (9)$$

2.5 Performance criteria

Due to the fact that water parameters had been truthfully monitored over the 10-yr period, the performances of the

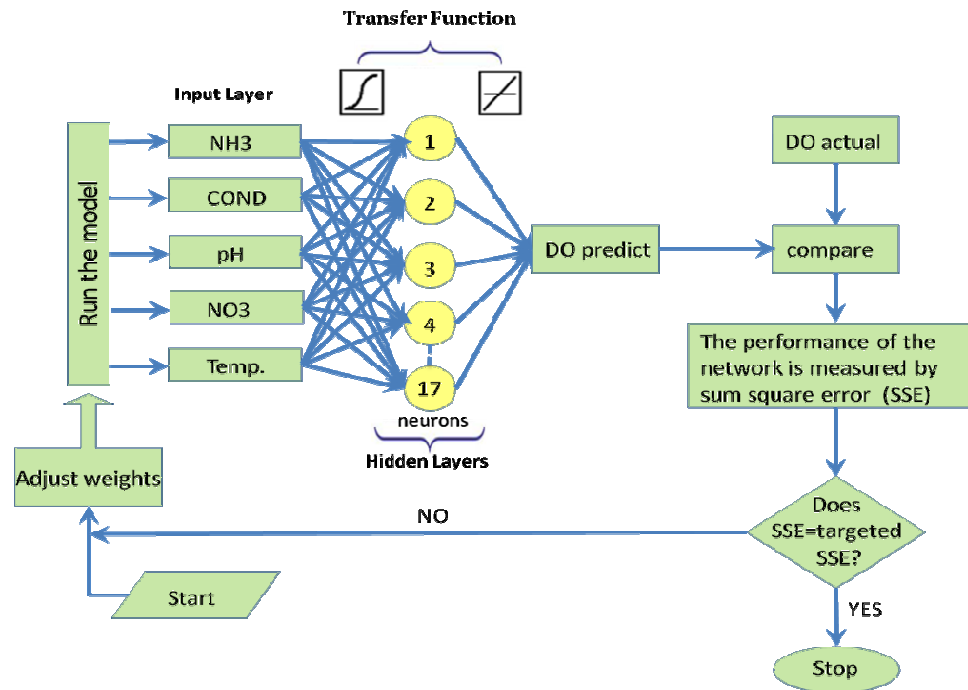


Fig. 4. An optimal architecture of ANN, together with a flowchart of the algorithm's procedure.

proposed models could be examined and evaluated. The performances of the models were evaluated according to three statistical indexes. Coefficient of Efficiency (CE) is often used to evaluate the model performance, introduced by Nash and Sutcliffe (1970).

$$CE = 1 - \frac{\sum_{i=1}^n (DO_m - DO_p)^2}{\sum_{i=1}^n (DO_m - \overline{DO_m})^2} \quad (10)$$

where  $n$  is the number of observations,  $DO_p$  and  $DO_m$  are the predicted and measured dissolved oxygen, respectively, and  $\overline{DO_m}$  is the average of measured dissolved oxygen.

The Mean Square Error (MSE) can be used to determine how well the network output fits the desired output. The smaller values of MSE ensure better performance. It is defined as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (DO_m - DO_p)^2 \quad (11)$$

The coefficient of correlation (CC) is often used to evaluate the linear relationship between the predicted and measured dissolved oxygen. It is defined as follows:

$$CC = \frac{\sum_{i=1}^n (DO_m - \overline{DO_m})(DO_p - \overline{DO_p})}{\sqrt{\sum_{i=1}^n (DO_m - \overline{DO_m})^2 \sum_{i=1}^n (DO_p - \overline{DO_p})^2}} \quad (12)$$

### 3 Results and discussion

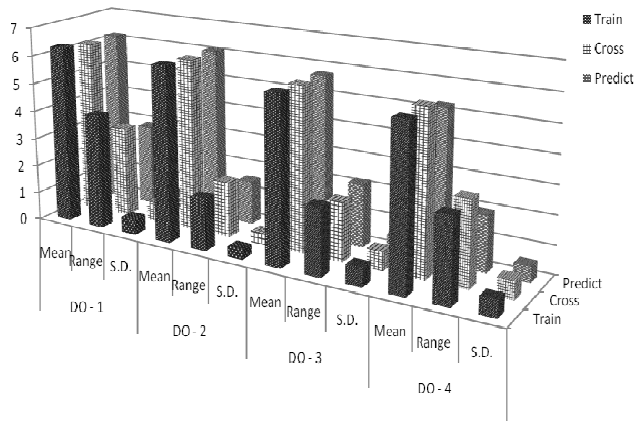
#### 3.1 Optimizations of the Neurons Number

One of the most important characteristics of MLP-NN is the number of neurons in the hidden layer. If an insufficient number of neurons are used, the network will be unable to model the complex data and the resulting fit will be poor. On the contrary, if too many neurons are used, the training time may become excessively long and the network may over fit the data. In this study, the number of neurons needed in the hidden layers to achieve the precision criteria was generally determined by trial and error approach. The optimum number of neurons was determined based on the minimum value of Mean Square Error (MSE) of the training data set. The training of the MLP-NN was performed with a variation of 1–20 neurons. Figure 6 shows the relationship between the numbers of neurons versus MSE during training. It was obvious from the figure that the MSE equalled to 0.2761 when one neuron was used and decreased to 0.0310 when 17 neurons were used. Enlarging the neurons more than 17 did not significantly decrease MSE. Thus, 17 neurons were selected as the best number of neurons.

#### 3.2 Test and validation of the model

Figure 4 shows the proposed architecture used to predict the dissolved oxygen at Station 1 (DONN1), which was developed according to the procedure discussed in the previous





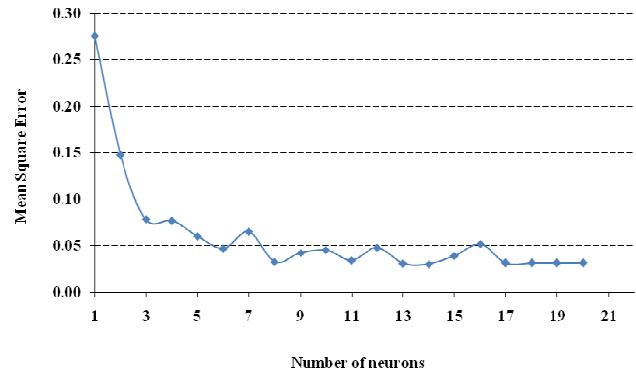
**Fig. 5.** Statistical properties (i.e. mean, standard deviation, range) for training, validation and testing data sets.

section. Training, validation and testing processes of the MLP-NN model were performed to minimize the Mean Square Error (MSE) between the output and the desired response, as shown in Fig. 7. It was apparent that the performance goal was achieved in less than 15 iterations (epochs). On the other hand, Fig. 8 illustrates the comparison between the predicted versus observed DO using 45° line of graph and two deviation lines of ±15 % deviation from the 45° line for both validation and testing data sets. It was obvious from Fig. 8 that DONN1 can predict the DO with a high level of accuracy, whereby the error for majority of the records did not reach 15 %, while the error of a few records fell within 15 %.

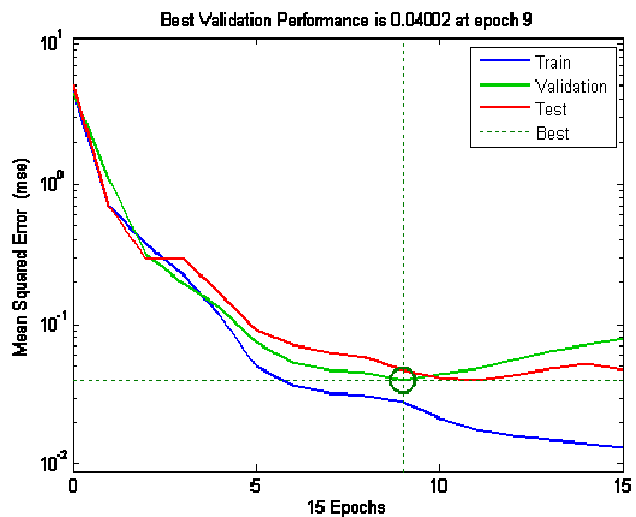
### 3.3 Sensitivity analysis

To evaluate the effect of input parameters on the model, two evaluation processes were used. First, the performance evaluation of various possible combinations of the parameters was investigated utilizing Coefficient of Efficiency (CE) and Mean Square Error (MSE) approaches to determine the most effective parameters on the output. Overall, six networks were compared, as shown in Table 4. Each one demonstrated how significant the eliminated parameter would affect the network accuracy. Apparently, the precision of MLP-NN became higher if all the suggested parameters were used as the input to the model, where minimum MSE and CE were determined to be 0.05 and 0.95 for the testing data set, respectively. Meanwhile, the level accuracy of the second network slightly decreased (MSE = 0.07, CE = 0.91) when COND was eliminated.

Compared with the results reported in the previous research, Rankovic et al. (2010) developed a separate neural network model for each independent input variable in order to determine the most effective variable. The correlation was 0.4802 between the conductivity and DO for the testing data set. The negative correlation of DO with electrical conduc-



**Fig. 6.** Relationship between the number of neurons and MSE.



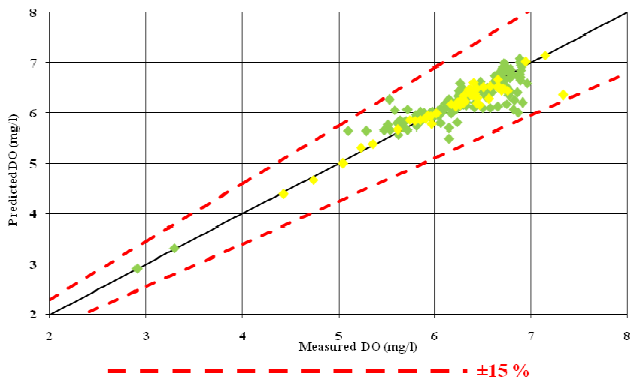
**Fig. 7.** Training, validation and test mean squared errors for the developed model.

tivity was well documented (Zaqoot et al., 2009). Although it was evident that COND was less effective on the DO, the level of accuracy increased 3 % when we considered it in the model input, as shown in the first network in Table 4. Therefore, COND would not be neglected in this study. Conversely, the Coefficient of Efficiency reduced gradually if any one of the input parameters was removed, which reduced the ability of ANN in the capability prediction. Furthermore, DO was found to be sensitive to the NO<sub>3</sub> parameter, where the level accuracy of model six (eliminate NO<sub>3</sub>) decreased (MSE = 0.26, CE = 0.6) for the testing data set.

Singh et al. (2009) computed the DO levels in the Gomti River (India) using three-layer feed-forward neural networks with back propagation learning. The sensitivity analysis revealed that NO<sub>3</sub> provided relatively higher contributions to the network. Figure 9 shows the effect on predictive accuracy if any one of the input parameters was removed from the model. For each of the five input parameters, removing the

**Table 4.** Predictive accuracy if any one of the input parameters was removed from model.

| Model | Combination  | Structure | MSE      |      | CE       |      |
|-------|--|-----------|----------|------|----------|------|
|       |  |           | Validate | Test | Validate | Test |
| 1     | pH,Temp,COND,NH <sub>3</sub> -NL,No <sub>3</sub><br>(all parameters) | 5-17-1    | 0.03     | 0.05 | 0.94     | 0.95 |
| 2     | pH,Temp,NH <sub>3</sub> -NL,No <sub>3</sub><br>(eliminate COND)      | 4-17-1    | 0.01     | 0.07 | 0.98     | 0.91 |
| 3     | Temp,COND,NH <sub>3</sub> -NL,No <sub>3</sub><br>(eliminate pH)      | 4-17-1    | 0.02     | 0.09 | 0.94     | 0.88 |
| 4     | pH,COND,NH <sub>3</sub> -NL,No <sub>3</sub><br>(eliminate Temp)      | 4-17-1    | 0.026    | 0.18 | 0.96     | 0.84 |
| 5     | pH,Temp,COND,No <sub>3</sub><br>(eliminate NH <sub>3</sub> -NL)      | 4-17-1    | 0.025    | 0.19 | 0.96     | 0.76 |
| 6     | pH,Temp,COND,NH <sub>3</sub> -NL<br>(eliminate No <sub>3</sub> )     | 4-17-1    | 0.16     | 0.26 | 0.81     | 0.6  |



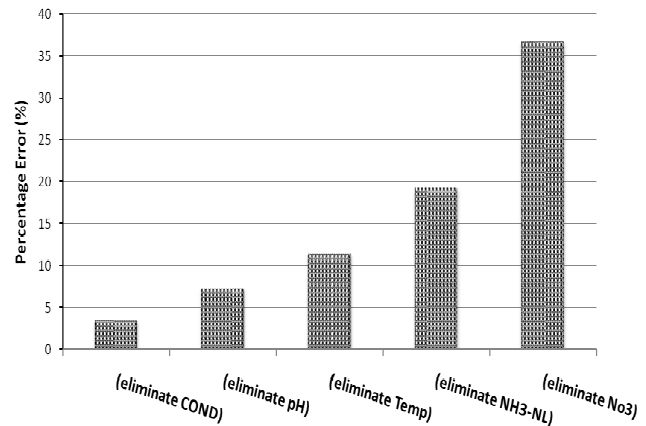
**Fig. 8.** A scatter diagram of the predicted versus observed DO for validation (in green dot) and testing (in yellow dot).

parameter dramatically increased the predictive error. Thus, the five input parameters were essential to the model.

The second assessment process was based on partitioning the neural network connection weights in order to determine the relative importance of each input parameter in the network (Garson, 1998; Emad et al., 2010). In this study, the proposed network consisted of five environmental parameters. Assuming the connection weights from the input nodes to the hidden nodes demonstrate the relative predictive importance of the independent parameter, the importance of each input parameter can be expressed as follows:

$$I_j = \frac{\sum_{m=1}^{m=N_h} \left( \left( \frac{|w_{jm}^{ih}|}{\sum_{k=1}^{N_i} |w_{km}^{ih}|} \right) \times |w_{mn}^{ho}| \right)}{\sum_{k=1}^{k=N_i} \left\{ \sum_{m=1}^{m=N_h} \left( \left( \frac{|w_{jm}^{ih}|}{\sum_{k=1}^{N_i} |w_{km}^{ih}|} \right) \times |w_{mn}^{ho}| \right) \right\}} \quad (13)$$

where  $I_j$  is the relative importance of  $j$ th input parameter on the output parameter,  $N_i$  and  $N_h$  are the numbers of input and hidden neurons, respectively, and  $W$  is the connection weight. Meanwhile, superscripts “ $i$ ”, “ $h$ ” and “ $o$ ” refer



**Fig. 9.** Percentage error of the model if any one of the input parameters was removed.

to the input, hidden and output layers, respectively, whereas subscripts “ $k$ ”, “ $m$ ” and “ $n$ ” refer to the input, hidden and output neurons, respectively.

Table 5 shows the connection weights values for the proposed model. It is important to note that Garson’s algorithm uses the absolute values of the connection weights when calculating parameter contributions. The relative importance of each of the input parameters as computed by Eq. (13) is shown in Fig. 10. The relative importance showed the significance of a parameter compared with the others in the model. Although the network did not necessarily represent physical meaning through the weights, it suggested that all the parameters had strong effects on the prediction of DO, where the predictor contributions ranged from 15 to 25 %. It was obvious that the most effective inputs were those which included oxygen containing (NO<sub>3</sub>) and oxygen demanding (NH<sub>3</sub>-NL). On the other hand, Temp and pH were found to be the least

**Table 5.** Connection weights between the input and hidden layers ( $W_1$ ) and weights between hidden and output layers ( $W_2$ ).

| Neuron | $W_1$            |         |                     |         |                 | $W_2$   |
|--------|------------------|---------|---------------------|---------|-----------------|---------|
|        | Input parameters |         |                     |         |                 | Target  |
|        | COND             | pH      | NH <sub>3</sub> -NL | Temp    | No <sub>3</sub> | Do      |
| 1      | -0.3496          | -0.5559 | 0.3896              | -2.5327 | 0.8759          | -0.8541 |
| 2      | -0.4448          | 1.079   | 1.0898              | -1.1716 | -1.1379         | 0.8674  |
| 3      | -1.0216          | 0.669   | 0.5456              | 1.8672  | 1.428           | -0.3121 |
| 4      | 1.426            | 0.0858  | -1.5348             | 1.4329  | -1.5608         | -2.0752 |
| 5      | -1.3073          | -1.3222 | -0.7894             | -1.5151 | -1.1319         | 0.8688  |
| 6      | -0.0383          | -1.0967 | 2.0217              | -2.6691 | -2.6661         | -1.4198 |
| 7      | -0.2122          | 0.3851  | 1.6225              | -1.4701 | -1.791          | 1.939   |
| 8      | -1.1948          | 1.4197  | 0.5578              | 0.3979  | -1.936          | -1.0546 |
| 9      | 1.3426           | -0.2442 | -1.8487             | 1.1279  | 2.1421          | 0.8314  |
| 10     | -0.5551          | -2.713  | -0.7412             | -0.4472 | 1.2528          | 1.191   |
| 11     | -1.1745          | -0.3326 | -1.6775             | -2.2023 | 1.7435          | -1.9648 |
| 12     | -0.5291          | 2.4511  | -1.3133             | -0.2754 | 0.1694          | 1.0078  |
| 13     | 0.4766           | -1.0507 | -1.6751             | 0.5298  | -0.8397         | 0.9967  |
| 14     | 1.5033           | -0.0905 | -1.6797             | -0.2412 | -0.8862         | 0.3342  |
| 15     | 0.5713           | 0.7769  | -0.2276             | 1.4275  | 2.2699          | -1.1364 |
| 16     | -1.7212          | -1.2637 | -1.1528             | 0.9447  | -1.015          | -1.0251 |
| 17     | 0.8254           | 0.3256  | 2.8875              | -0.5906 | 1.2672          | -1.5751 |

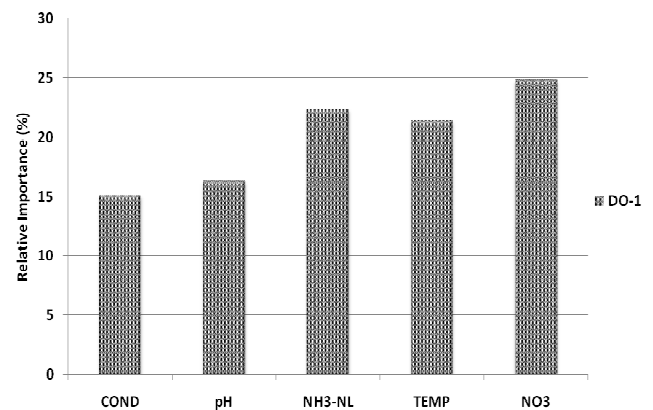
effective parameters. Moreover, COND revealed the lowest contribution on the proposed model. These findings agreed with those found in previous evaluation (the combinations of parameters).

### 3.4 Performance of the proposed scenarios

Considering the same architecture that was used to predict dissolved oxygen at Station 1 (DONN1), the DO at Stations 2, 3 and 4 was predicted. Figure 11 demonstrates the performance of the proposed models. Apparently, the scatter plot of the three models showed that the error approximately fell on the ideal line except three records, which remarkably exceeded 15 % and were also found in the third model that was used to predict the DO at Station 3. These three records were more deviated from the observed value attributes due to the fact that the extreme values were found in the samples which were polluted by noise signals owing to systematic and random errors.

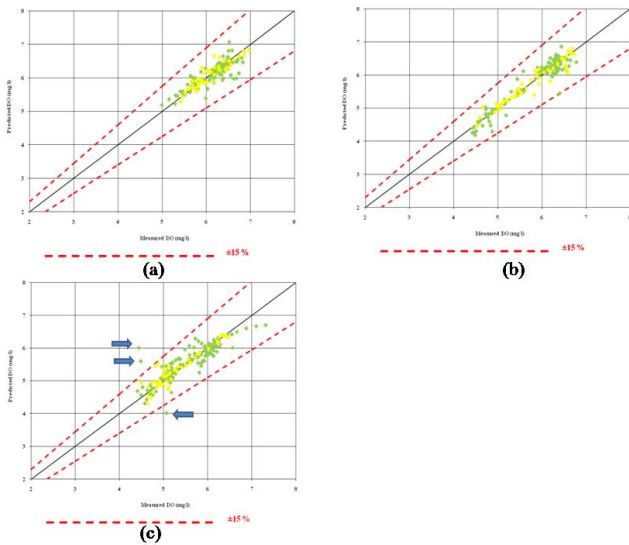
The Coefficient of Efficiency (CE) and MSE computed for validation and testing data sets used for the three stations are presented in Table 6. The CE values for the three stations for the validation data set were controlled within an acceptable range, i.e. 0.9–0.97, while the CE values for the three stations for the test data set were 0.83, 0.89 and 0.86, respectively. The respective MSE values of validation and testing data sets for the three locations were 0.03–0.05 for DONN2, 0.03–0.10 for DONN3 and 0.06–0.10 for DONN4.

The relatively low correlation between the observed and predicted values in the testing phase was perhaps due to the



**Fig. 10.** The relative importance of each input parameter at Station 1.

non-homogenous nature of water quality parameters. Moreover, Ying et al. (2007) showed that the selection of affecting factors (input parameters) plays a key role since these factors have great impact on the forecast results. Thus, it was evident that the low correlation in this study was attributed to the fact that, the input parameters did not include all the relevant parameters. In addition, water pollution at the downstream station was related to the discharge from the upstream station. Hence, to overcome the problem, this study introduced another approach (Scenario 2) so that a high level of accuracy could be reached. This approach was related to the



**Fig. 11.** Scatter diagrams of the predicted versus observed DO for validation (in green dot) and testing (in yellow dot); (a) DO-NN2, (b) DO-NN3 and (c) DO-NN4.

**Table 6.** Coefficient of Efficiency (CE) and Mean Square Error (MSE) associated with MLP-NN models for each station at Johor River.

| Model | MSE        |         | CE         |         |
|-------|------------|---------|------------|---------|
|       | Validation | Testing | Validation | Testing |
| DONN2 | 0.03       | 0.05    | 0.90       | 0.83    |
| DONN3 | 0.03       | 0.10    | 0.97       | 0.89    |
| DONN4 | 0.06       | 0.10    | 0.92       | 0.86    |

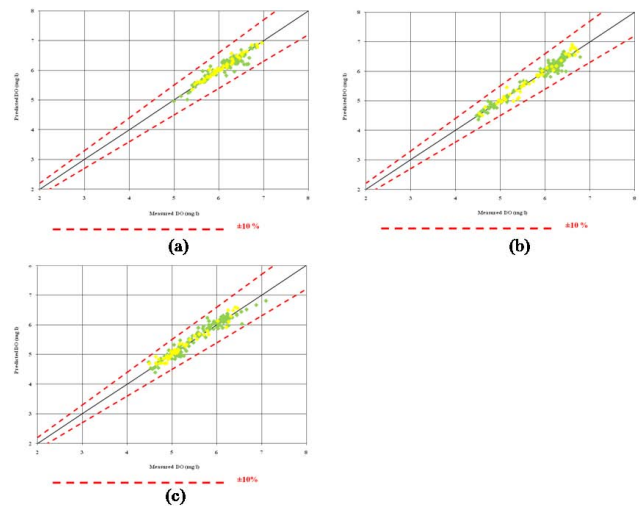
prediction of the DO, with consideration of the predicted DO at the upstream station as the input to the model, as expressed by Eq. (6).

Fig. 12 demonstrates the performance of the proposed models (Scenario 2). Apparently, the scatter plot of the three models showed that the error approximately fell on the ideal line for both validation and testing data sets.

In comparison between Scenario 1 and Scenario 2, Scenario 2 was able to achieve a high level of accuracy in simulating the magnitude and patterns of DO at all stations and reducing the deviation error from ±15 % that reached by Scenario 1 to ±10 %.

For further analysis, we adopted the accuracy improvement (AI) index for correlation coefficient statistical index to measure the significance of the proposed Scenario 2 over Scenario 1, expressed as follows:

$$AI(\%) = \left( \frac{CC_{Scen2} - CC_{Scen1}}{CC_{Scen2}} \right) * 100 \quad (14)$$



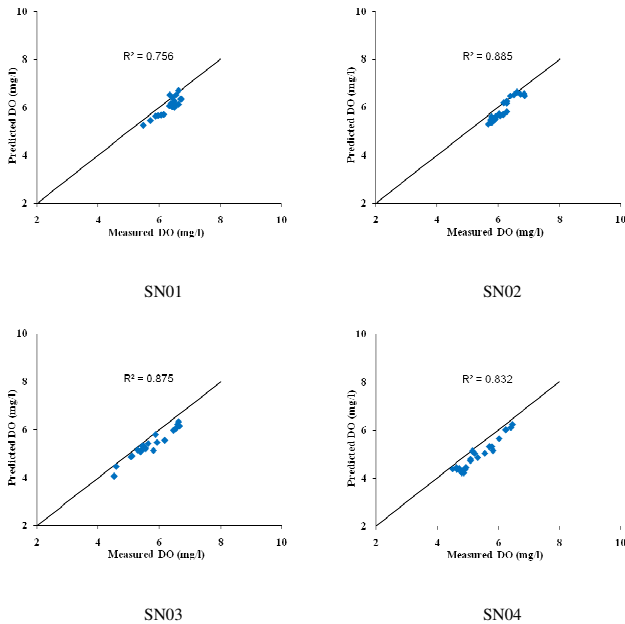
**Fig. 12.** Scatter diagrams of the predicted versus observed DO for validation (in green dot) and testing (in yellow dot); (a) DO-NN2, (b) DO-NN3 and (c) Do-NN4.

**Table 7.** A summary of correlation coefficient for Scenario 1 and Scenario 2 and the AI %.

| Model | Scenario 1 | Scenario 2 | AI     |
|-------|------------|------------|--------|
|       | CC         | CC         | CC (%) |
| DONN2 | 0.91       | 0.98       | 8      |
| DONN3 | 0.91       | 0.96       | 6      |
| DONN4 | 0.93       | 0.97       | 4      |

Where  $CC_{Scen2}$  is the value of the correlation coefficient for Scenario 2, while  $CC_{Scen1}$  is the same statistical index for Scenario 1. Examining Table 7 carefully, it can be observed that Scenario 2 was more adequate than Scenario 1, with a significant improvement for all stations ranging from 4 % to 8 %. Prediction accuracy was significantly improved after introducing Scenario 2 for all stations.

For further assessment, the proposed models were compared with the results reported in the literature. Soyupak et al. (2003) employed the ANN modelling approach to calculate the pseudo steady state time and space dependent DO concentrations in three different reservoirs, with entirely different properties. The correlation coefficients between neural network estimates and field measurements were higher than 0.95. In addition, Sengorur et al. (2006) employed the feed-forward (FF) type ANN for computing the monthly values of DO. The findings demonstrated that the ANN results were very close to the observed values of DO where the correlation coefficient equalled to 0.9186. Ying et al. (2007) adopted the BP neural network to forecast water quality at Yuqiao Reservoir. The correlation between the forecast and actual measured of DO values was 0.9418. Likewise, Kuo

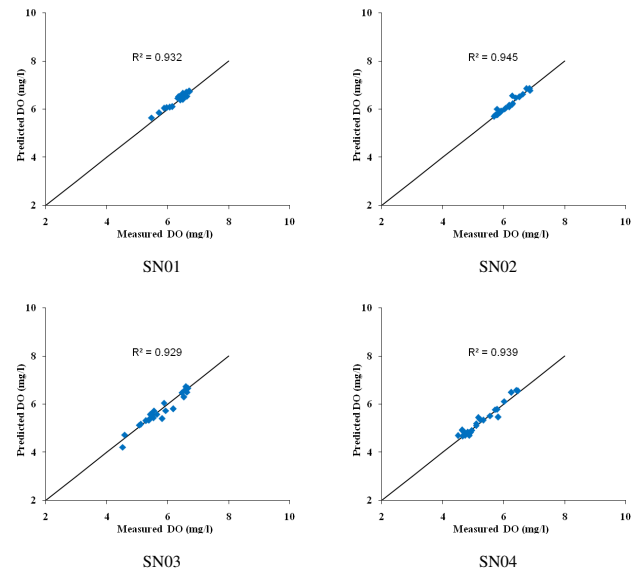


**Fig. 13.** Scatter plots between the observed and predicted value for each of the DO at each stations (Scenario 1).

et al. (2007) used the back-propagation neural network for predicting the dissolved oxygen in the Te-Chi Reservoir in Taiwan. The correlation coefficients between the predicted values and measured data of DO were above 0.7 for training and testing data sets. Meanwhile, Zaqoot et al. (2009) used the ANNs-Multilayer Perceptron (MLP) network to predict the next fortnight's dissolved oxygen concentrations in the water of Mediterranean Sea along Gaza. The coefficient of determination between the measured and model computed values of DO was 0.996. On the other hand, Singh et al. (2009) constructed an artificial neural network (ANN) model to predict the water quality at Gomti River, India. The coefficients of determination between the measured and model computed values of DO for the training, validation and test sets were 0.70, 0.74 and 0.76, respectively. Furthermore, Rankovic et al. (2010) developed a feed-forward neural network (FNN) model to predict the dissolved oxygen in Gruza Reservoir, Serbia. The correlation coefficients between the predicted values and measured values of DO were 0.974 and 0.8738 for training and testing data sets, respectively.

The developed model was compatible with the results of other researchers/authors. High coefficients of correlation were obtained between the observed and predicted values for the test sets of 0.98, 0.96 and 0.97 for all stations. These results revealed that the input parameters selected in this study had direct relevance with the target (DO). The selection of input parameters might affect the model output remarkably (Singh et al., 2009).

The results also indicated that the proposed model was basically an attractive alternative, offering a relatively fast al-



**Fig. 14.** Scatter plots between the observed and predicted value for each of the DO at each stations (Scenario 2).

gorithm with good theoretical properties to predict the dissolved oxygen where the performance goal was achieved in less than 15 iterations (epochs) and can be extended to predict different water quality parameters.

The model needs to verify when output results and the observed values are close enough to satisfy the verification criteria. Therefore, in order to investigate the efficiency of the proposed model, the verification of MLP-NN based on collection of field data within duration 2009–2010 is presented. Scatter plots between the observed and predicted value for each of the DO for both scenarios are presented in Figs. 13 and 14.

The comparison of the proposed model (Scenario 2) and Scenario 1 for prediction water quality parameters shows that the network output of Scenario 2 could depict the behaviour of the water quality parameters pattern more accurately than Scenario 1. It can be seen that most of the predicted water quality parameter values are close to the actual observation. A value of  $R^2$  should be close to 1,  $R^2$  more than 0.9 indicates a very satisfactory model performance, a value between 0.6–0.9 indicates a fairly good performance, and values below 0.5 indicate unsatisfactory performance. The proposed model showed efficiency in predicting the concentration of water quality parameters in the Johor River, and it was compatible with the results of other researchers/authors. The results also indicated that the proposed model was basically an attractive alternative, offering a relatively fast algorithm with good theoretical properties to predict the water quality parameters and can be extended to predict different water quality parameters.

Nevertheless, due to the fact that water quality forecast can be easily affected by external environment, the obtained

model sometimes produces results much deviated from the actual values, therefore further study needs to be done in future work to identify the suitable forecast model, understand its laws of changes and solve the problem of forecast deviation. In general, this research work has managed to integrate several analytical and modeling methods that would prove to be useful for various institutions that are directly involved in the management of river basin in Malaysia. Moreover, the tools used in this work could form a basis for a more effective decision making process on the part of the policy makers in order to help maintain and improve the management of river basins.

However, it should be emphasized that there are no structured methods today to identify what network structure can be best approximate the function, mapping the inputs to outputs. In addition, pre-processing for the data is essential step for water quality prediction model and required more survey and analysis that could lead to better accuracy in their application. Moreover, the optimal selection of the key parameter still required to be achieved by augmenting the AI model with other optimization model such as genetic algorithm or particle swarm optimization methods. On the other hand, the variable selection (input pattern) in AI model is always a challenging task due to the complexity of the hydrologic process. Some other advanced ANN model, namely; Dynamic Neural Network DNN could be investigated and might provide better predict model. The investigation and application of more robust input pattern selection approaches, such as systematic searching of optimal or near optimal variable combination in DNN with ensemble procedure, would be desirable in future water quality prediction research studies.

#### 4 Conclusions

In this study, model based on artificial neural networks was developed for prediction of the DO concentration in water of the Johor river (Malaysia). Among many types of ANNs, the most widely used is the three-layer feed-forward neural network which is introduced in this study. LMA algorithm is used for its faster convergence speed and lower error rate to overcome the shortcomings of traditional BP algorithm as slow to converge and easy to reach extreme minimum value. The architecture of the proposed model giving the optimal result was multi-layer perceptron neural network (MLP-NN) with The hyperbolic tangent sigmoid transfer function between the input and the hidden layers, linear transfer function between the hidden and output layers. The number of neuron is varying in the range 1–20 neurons. 17 neurons were selected as the best number of neurons based on the minimum value of MSE of the training data set. The sensitivity analysis showed that all studied input parameters (temperature (Temp), water pH, electrical conductivity (COND), nitrate ( $\text{NO}_3$ ) and ammonical nitrogen ( $\text{NH}_3\text{-NL}$ )) have strong effect on dissolved oxygen. In addition,  $\text{NO}_3$  is the most influ-

ential parameter with relative importance of 25 %. Two scenarios were introduced; the first scenario (Scenario 1) was to establish the prediction model for water quality parameters at each station based on five input parameters, while the second scenario (Scenario 2) was to establish the prediction model for water quality parameters based on the five input parameters and the value of DO at previous station (upstream). In comparison between Scenario 1 and Scenario 2, Scenario 2 was able to achieve a high level of accuracy in simulating the magnitude and patterns of all water quality parameters at all stations where the accuracy improvement percentage (AI%) ranging from 4 % to 8 % after utilizing Scenario 2. The verification of the developed model based on collection of field data within duration 2009–2010 showed that the model perform very satisfactorily to predict DO concentration ( $R^2$  values equal or bigger than 0.9) for all stations. The result showed that the proposed model can be applied successfully and contribute to enhance the accuracy of water quality prediction.

*Acknowledgements.* The authors wish to thank the Department of Environment for providing the required data for this research. This research was supported by the research grant given to the second and third authors from University Kebangsaan Malaysia; UKM-GUP-PLW-08-13-308 and FRGS Fund UKM-KK-02-FRGS0125-2009.

Edited by: A. D. Reeves

#### References

- Akhtar, M. K., Corzo, G. A., van Andel, S. J., and Jonoski, A.: River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin, *Hydrol. Earth Syst. Sci.*, 13, 1607–1618, doi:10.5194/hess-13-1607-2009, 2009.
- Alvisi, S., Mascellani, G., Franchini, M., and B'ardossy, A.: Water level forecasting through fuzzy logic and artificial neural network approaches, *Hydrol. Earth Syst. Sci.*, 10, 1–17, doi:10.5194/hess-10-1-2006, 2006.
- Ansa-Ansare, O. D., Marr, I. L., and Cresser, M. S.: Evaluation of modelled and measured patterns of dissolved oxygen in a freshwater lake as an indicator of the presence of biodegradable organic pollution, *Water Resour. Res.*, 34, 1079–1088, 2000.
- Biswas, A. K.: *Models for Water Quality Management*, New York: McGraw-Hill Inc, 1981.
- Cox, B A.: 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, 2003.
- Das, J. and Acharya, B. C.: Hydrology and Assessment of Lotic Water Quality in Cuttack City, India, *J. Water, Air Soil Pollut.*, 150, 163–175, 2003.
- Demuth, H., Beale, M., and Hagan, M.: *Neural Network Toolbox: User's Guide, Version 6*, The Mathworks, Inc, <http://www.mathworks.com>, 2008.

- Department of Environment (DOE): Classification of Malaysian Rivers – Methodology and Classification of Ten Rivers. Ministry of Science, Technology and Environment, 2, 4.101–4.104, 1994.
- Department of Environment (DOE): Water quality management in Malaysia. Kuala Lumpur: DOE Documents, 2003.
- Dogan, E., Sengorur, B., and Koklu, R.: Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique, *J. Environ. Manage.*, 90, 1229–1235, 2009.
- Einax, J. W., Aulinger, A., Tumpling, W. V., and Prange, A.: Quantitative description of element concentrations in longitudinal river profiles by multiway PLS models, *Fres. J. Anal. Chem.*, 363, 655–661, 1999.
- El-shafie, A., Noureldin, A., Taha, M. R., and Basri, H.: Neural network model for Nile River inflow forecasting analysis of historical inflow data, *J. Appl. Sci.*, 8(24), 4487–4499, 2008.
- El-shafie, A., Mukhlisin, M., Najah, A. A., and Taha, M. R.: Performance of artificial neural network and regression techniques for rainfall-runoff prediction, *Int. J. Phys. Sci.*, 6(8), 1997–2003, 18 April, 2011
- Elmolla, E. S., Chaudhuri, M., and Eltoukhy, M. M.: The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process, *J. Hazard. Mater.*, 179, 127–134, 2010.
- Freeman, J. A. and Skapura, B. M.: Neural networks algorithms, applications, and programming techniques, Addison-Wesley, 1991.
- French, M. and Recknagel, F.: Modeling algal blooms in freshwaters using artificial neural networks, in: *Computer Techniques in Environmental Studies V*, vol. II: Environment Systems, edited by: Zanetti, P., Computational Mechanics Publications, Boston, 87–94, 1994.
- Garcia, A., Revilla, J. A., Medina, R., Alvarez, C., and Juanes, J. A.: A model for predicting the temporal evolution of dissolved oxygen concentration in shallow estuaries, *Hydrobiology*, 475/476, 205–211, 2002.
- Garson, G. D.: *Neural Networks an Introductory Guide for Social Scientists*, Sage Publications, California, 1998.
- Hatzikos, E., Anastasakis, L., Bassiliades, N., and Vlahavas, I.: Simultaneous prediction of multiple chemical parameters of river water quality with tide, in: *Proceedings of the Second International Scientific Conference on Computer Science*, IEEE Computer Society, Bulgarian Section, 2005.
- Haykin, S.: *Neural Networks: Comprehensive Foundation*, Upper Saddle River, N.J, USA: Prentice-Hall, 1999.
- Hem, J. D.: *Study and Interpretation of the Chemical Characteristics of Natural Water*, 3rd Ed. United States Geological Survey, Washington, 1985.
- Hornik, K. and White, H.: Multilayer feedforward networks are universal approximators, *Neural Netw.*, 2, 359–366, 1989.
- Hull, V., Parrella, L., and Falcucci, M.: Modelling dissolved oxygen dynamics in coastal lagoons, *Ecol. Model.*, 2, 468–480, 2008.
- Hung, N. Q., Babel, M. S., Weesakul, S., and Tripathi, N. K.: An artificial neural network model for rainfall forecasting in Bangkok, Thailand, *Hydrol. Earth Syst. Sci.*, 13, 1413–1425, doi:10.5194/hess-13-1413-2009, 2009.
- Kalff, J.: *Limnology: Inland Water Ecosystems*, Prentice-Hall, New Jersey, USA, 592 pp., 2002.
- Kasabov, N. K.: *Foundations of neural networks, fuzzy systems, and knowledge engineering*, Cambridge: MIT Press, 1996.
- Kuo, J., Hsieh, M., Lung, W., and She, N.: Using artificial neural network for reservoir eutrophication prediction, *Ecol. Model.*, 200, 171–177, 2007.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., and Aulagnier, S.: Application of neural networks to modeling non-linear relationships in ecology, *Ecol. Model.*, 90, 39–52, 1996b.
- Lopes, J. F., Silva, C. I., and Cardoso, A. C.: The water quality of the Ria de Aveiro lagoon, Portugal: From the observations to the implementation of a numerical model, *Mar. Environ. Res.*, 60, 594–628, 2005.
- Najah, A., Elshafie, A., Karim, O. A., and Jaffar, O.: Prediction of Johor River water quality parameters using artificial neural networks, *Eur. J. Scienti. Res.*, 28(3), 422–435 ISSN 1450-216X, 2009.
- Najah A. A., El-Shafie A., Karim O. A., and Jaafar, O.: Water Quality Prediction Model Utilizing Integrated Wavelet-ANFIS Model with Cross Validation, *Neural Comput. Appl.*, in press, doi:10.1007/s00521-010-0486-1, 2010b.
- Najah, A., Elshafie, A., Karim, O. A., and Jaffar, O.: Evaluation the efficiency of radial basis function neural network for prediction of water quality parameters, *Eng. Int. Syst.*, 4, 221–231, 2010c.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models, Part 1: a discussion of principles, *J. Hydrol.*, 10(3), 282–290, 1970.
- Olden, J. D. and Jackson, D. A.: Fish-habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks, *Trans. Am. Fish Soc.*, 130, 878–897, 2001.
- Palani, S., Liong, S. Y., and Tkalich, P.: An ANN application for water quality forecasting, *Marine Pollut. Bull.*, 56, 1586–1597, 2008.
- Radwan, M., Willems, P., El-Sadek, A., and Berlamont, J.: Modelling of dissolved oxygen and biochemical oxygen demand in river water using a detailed and simplified model, *Int. J. River Basin Manage.*, 1(2), 97–103, 2003.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by backpropagating errors, *Nature*, 323, 533–536, 1986.
- Sengorur, B., Dogan, E., Koklu, R., and Samandar, A.: Dissolved oxygen estimation using artificial neural network for water quality control, *Fresen. Environ. Bull.*, 15(9a), 1064–1067, 2006.
- Singh, K. P., Basant, A., Malik, A., and Jain, G.: Artificial neural network modeling of the river water quality – a case study, *Ecol. Model.*, 220, 888–895, 2009.
- Souza, C. R.: *Neural Network Learning by the Levenberg-Marquardt Algorithm with Bayesian Regularization*, <http://crsouza.blogspot.com/2009/11/neural-network-learning-by-levenberg.html>, 2009.
- Soyupak, S., Karaer, F., Gurbuz, H., Kivrak, E., Senturk, E., and Yazici, A.: A neural networkbased approach for calculating dissolved oxygen profiles in reservoirs, *Neural Comput. Appl.*, 12, 166–172, 2003.
- Ranković, V., Radulović, J., Radojević, I., Ostojia, A., and Eomija, L.: Neural network modelling of dissolved oxygen in the Gruza reservoir, Serbia, *Ecol. Model.*, 221, 1239–1244, 2010.
- Xiang, S. L., Liu, Z. M., and Ma, L. P.: Study of multivariate linear regression analysis model for ground water quality prediction, *Guizhou Science*, 24(1), 60–62, 2006.
- Yabunaka, K.-I., Hosomi, M., and Murakami, A.: Novel application

- of a backpropagation artificial neural network model formulated to predict algal bloom, *Water Sci. Technol.*, 36(5), 89–97, 1997.
- Yen-Ming Chiang, Li-Chiu Chang, Meng-Jung Tsai, Yi-Fung Wang, and Fi-John Chang: Dynamic neural networks for real-time water level predictions of sewerage systems-covering gauged and ungauged sites, *Hydrol. Earth Syst. Sci.*, 14, 1309–1319, doi:10.5194/hess-14-1309-2010, 2010.
- Ying, Z., Jun, N., Fuyi, C., and Liang, G.: Water quality forecast through application of BP neural network at Yuquiao reservoir, *J. Zhejiang Univ. Sci.*, A8, 1482–1487, 2007.
- Zaqoot, H. A., Ansari, A. K., and Unar, M. A.: Prediction of dissolved oxygen in the Mediterranean Sea along Gaza, Palestine-an artificial neural network approach, *Water Sci. Technol.*, 60(12), 3051–3059, 2009.