MGPB

A software sensor for in-situ monitoring of the 5-day biochemical oxygen demand

The Mining-Geology-Petroleum Engineering Bulletin
UDC: 502:550.8
DOI: 10.17794/rgn.2018.1.3

Preliminary communication

Rana Kasem[1]; Dimah ALabdeh[1]; Roohollah Noori[2*]; Abdulreza Karbassi[3]

[1]*Graduate Faculty of Environment, University of Tehran, Tehran, Iran, PhD Candidate in Environmental Engineering-Water Resources*
[2]*Graduate Faculty of Environment, University of Tehran, Tehran, Iran, Assistant Professor of Environmental Engineering*
[3]*Graduate Faculty of Environment, University of Tehran, Tehran, Iran, Associate Professor of Environmental Engineering*

Abstract

Due to the time-consuming procedure for determining the 5-day biochemical oxygen demand (BOD_5), the present study developed two software sensors based on artificial intelligence techniques. It is aimed to estimate this indicator instantaneously. For this purpose, feed-forward and radial basis function neural networks (FFANN and RBFANN, respectively) were used. FFANN and RBFANN were employed to estimate the maximum values of BOD_5 ($BOD_{5(max)}$) as a function of average, maximum and minimum dissolved oxygen in the Sefidrood River. Also, Levenberg-Marquardt (LM), resilient backpropagation, and scaled conjugate gradient algorithms were used to optimize the FFANN parameters. The results demonstrated that the performance of the LM algorithm in tuning the FFANN was better than the others, in the verification step. Furthermore, the performance of each model was evaluated according to the mean square error, correlation coefficient and developed discrepancy ratio. The results showed that the performance of both FFANN and RBFANN models for the prediction of the $BOD_{5(max)}$ were approximately the same.

Keywords

FFANN, RBFANN, Dissolved oxygen, Calibration, BOD_5.

1. Introduction

Rivers are one of the main sources of water supply for diverse uses including industry, drinking, recreation and agriculture (**Fan et al., 2008**). Due to the discharge of municipal-industrial wastewaters and agricultural drainages into rivers, evaluating the quality of these valuable water bodies is essential. In this direction, predicting the water quality parameters (WQP) such as the 5-day biochemical oxygen demand (BOD_5), as an indicator of organic load, can provide the proper framework for managers to improve the water quality in the rivers.

Numerical and data-driven models are usually used for the simulation of WQP based on the existing facilities as well as the user's needs. Numerical models have always been a useful tool for the simulation of WQP and were used especially for the simulation of BOD_5. **Ning et al. (2001)** discussed pollution prevention actions in the river Kao-Ping in Taiwan. They used QUAL2E, a steady state one-dimensional model that applies a finite difference scheme for solving the governing equations on pollutants' transport in rivers. Subsequently, simulation of BOD_5, dissolved oxygen (DO) and total phosphorous (TP) was carried out by Ning et al. (2001). Simulated results indicated that using economic instruments is required to reduce and control the pollution load of BOD_5 in the river. **Park and Lee (2002)** compared the performance of two models QUAL2E and QUAL2K (a developed version of QUAL2E that includes some advantages such as the conversion of algae death to BOD_5 and denitrification) for the simulation of BOD_5 in the Nakdong River, Korea. The researchers concluded that the QUAL2K performance was better than the QUAL2E. **Fan et al. (2008)** combined two models QUAL2K and HEC-RAS (a software developed by the U.S. Army Corps of Engineers) to assess WQP in the Keelung River, Taiwan. They used QUAL2K for the simulation of BOD_5. HEC-RAS was used for simulating the hydraulic constants of atmospheric reaeration and water level profile variations. The results showed that the combination of two models provides a good tool for water quality simulation in tidal rivers. **Sharma et al. (2015)** modelled BOD_5 using QUAL2Kw (a developed version of QUAL2K that uses the genetic algorithm for automatic calibration) and reported that the model performance was satisfactory.

It should be pointed out that data-driven models are more flexible than numerical ones. Also, due to the proper performance of a set of these models called intelligent techniques, their application for BOD_5 prediction has increased in recent years. **Chaves and Chang (2006)** used

Corresponding author: Roohollah Noori
noor@ut.ac.ir

a feed-forward artificial neural network (FFANN) to simulate WQP in the Shihmen Reservoir located on the Tahan River in Taiwan. For this purpose, they used the water quality data for 47 days at five points in three different depths. Hydro-meteorological data and WQP were defined as input and outputs, respectively. They concluded that FFANN was well capable of predicting WQP. **Dogan et al. (2007)** applied FFANN to estimate the BOD_5 concentration in the Melen-River, Turkey. Inputs (including chemical oxygen demand (COD), temperature, DO, water flow, chlorophyll-a, ammonia, nitrite and nitrate) were measured at 11 sites in the Melen-River during 2001-2002. A comparison of the modelled values with the measured ones showed that the FFANN model performed well for the BOD_5 prediction with a correlation coefficient (R) equal to 0.93. **Singh et al. (2009)** applied the FFANN model to predict BOD_5 and DO in the Gomti River, India. The results revealed that the predicted and measured DO and BOD_5 were in close agreement. The values of R (in case of BOD_5) for calibration and verification steps were 0.92 and 0.88, respectively.

The relation between BOD_5 and DO is rather complex. In other words, many other parameters such as toxic substances and algae respiration could greatly influence the relationship between BOD_5 and DO. Thus, it is difficult to use simple data-driven models such as linear regression to construct a software-sensor for BOD_5 prediction. In the present investigation, two software sensors were applied by ANN models to predict BOD_5 as a function of DO in the Sefidrood River. Due to the little attention to the simulation of BOD_5 with radial basis function neural network (RBFANN), FFANN was also used to simulate the maximum value of BOD_5 ($BOD_{5(max)}$) in the Sefidrood River. Also, this research aimed to compare the performance of FFANN and RBFANN models for BOD_5 prediction.

2. Methods and materials

2.1. Case study and data

For the in-situ measurement of BOD_5 with the aid of FFANN and RBFANN models, the Sefidrood River Basin, located in the northwest of Iran, was selected. The river basin covers an area of 59196 square kilometres, located between the Alborz and the Zagros Mountains. Pollution sources in the Sefidrood River Basin are urban, agricultural and industrial (**Noori et al., 2013a and 2015**). A field study along the river was performed to detect appropriate sampling stations. Firstly, 94 monitoring stations were selected in such a way to cover the whole length of the river and its tributaries (see **Figure 1**). Then, sampling was carried out for each season throughout a one year period. Dissolved oxygen (DO) was measured in-situ. BOD_5 samples were carried to the laboratory and analysed within 24 hours after collection

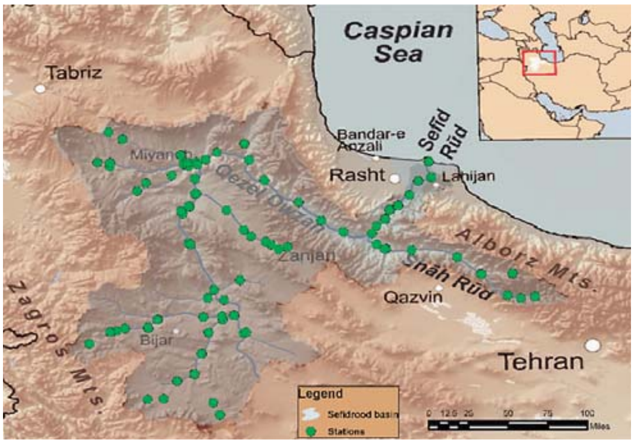

Figure 1: The Sefidrood River Basin and sampling points

(**APHA, 1995**). Thereafter, the data was checked for outliers.

One needs five days to determine BOD_5 using conventional methods whereas in-situ measurement of this parameter would be very significant for water quality managers. Therefore, this study aimed to apply a software-sensor for in-situ measurement of $BOD_{5(max)}$ as a function of WQP such as DO. To achieve this goal, the minimum, average and maximum values of DO, (DO_{min}, DO_{avg} and DO_{max}) measured online in four seasons, were used as suggested by **Noori et al. (2012)**. **Figure 2** shows the correlation matrix between DO parameters and $BOD_{5(max)}$.

Generally, **Figure 2** indicates a weak correlation between DO and $BOD_{5(max)}$. This can be justified by many factors such as discharge of industrial effluents without proper treatment that cause a negative influence on microbial activates (**Noori et al., 2013b**). In other words, the presence of toxic substances, along with algae respiration due to the eutrophic status in some parts of the river can influence microbial activities to various degrees. Based on **Figure 2**, the largest value of the coefficient of determination (R^2) is between DO_{max} and $BOD_{5(max)}$ and the smallest one is between DO_{min} and $BOD_{5(max)}$. Considering this correlation matrix, it is not possible to predict $BOD_{5(max)}$ based on DO because there is no linear relation between the two variables whereas it is clear that a complicated relation exists between them. Therefore, it is difficult to provide a software-sensor for modelling BOD_5 as a function of DO by using simple data-driven models such as linear regression. So, it is necessary to use a strong tool such as ANN.

Figure 3 shows the applied methodology in this research for developing a software-sensor based on FFANN and RBFANN models step-by-step to predict $BOD_{5(max)}$ in the Sefidrood River. According to this figure, after the collection of data, inputs (DO_{min}, DO_{avg} and DO_{max}) and target ($BOD_{5(max)}$) were imported to MATLAB software for the application of FFANN and RBFANN models. The data entered to MATLAB were standardized to be limited in the range [-1, +1]. Then,

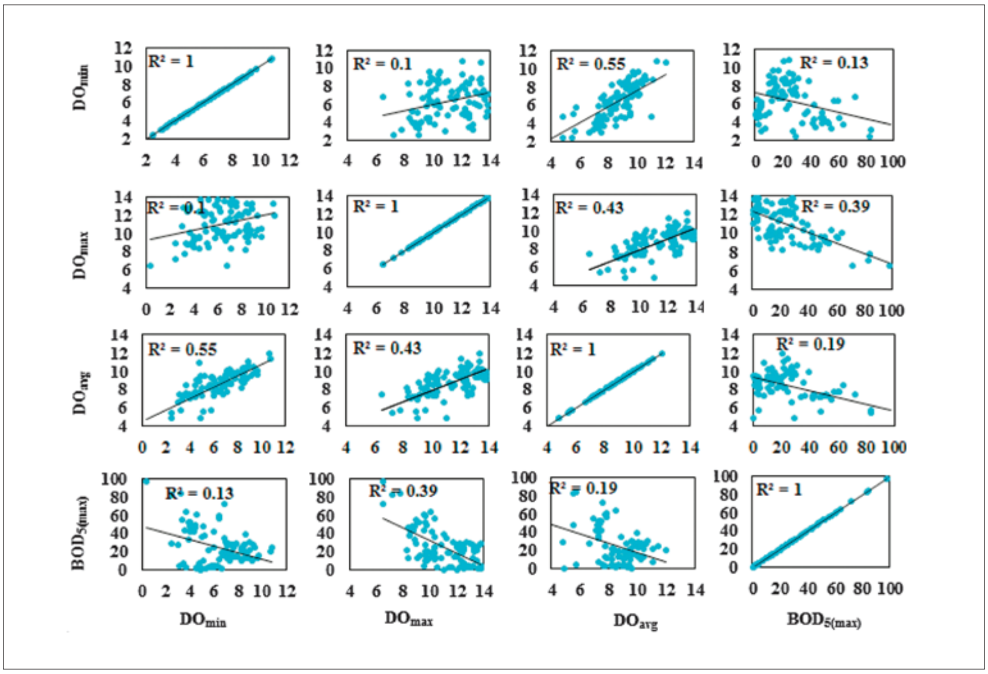

Figure 2: The correlation matrix between DO and $BOD_{5(max)}$ parameters

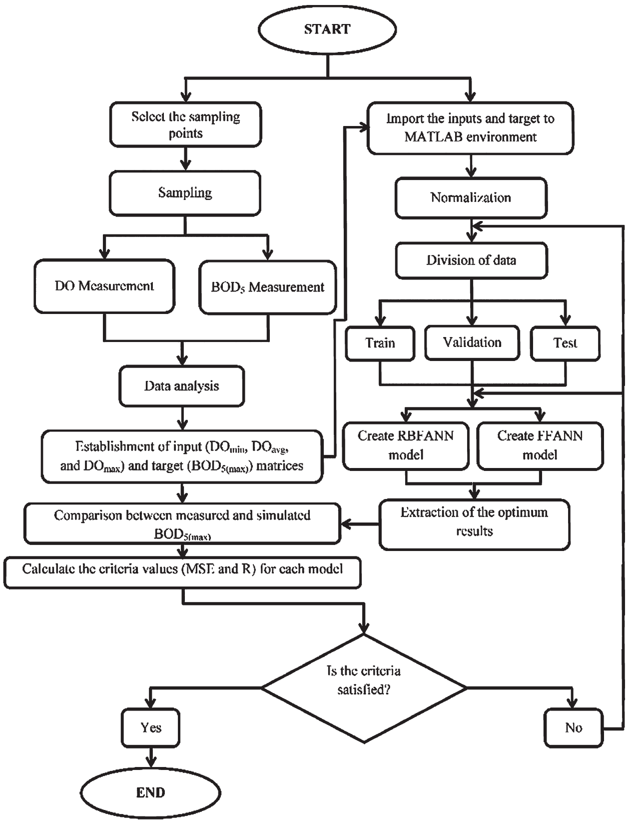

Figure 3: Applied methodology in this research for developing a software-sensor based on FFANN and RBFANN models step-by-step to predict $BOD_{5(max)}$ in the Sefidrood River

data was divided into three groups (train, validation, and verification). Subsequently, the two models (FFANN and RBFANN) were set and the modelled $BOD_{5(max)}$ values were compared with measured ones. Finally, the performance of each model was evaluated by the mean square error (MSE), R and developed discrepancy ratio (DDR).

2.2. Feed-forward artificial neural network

The FFANN is comparable to the natural nervous system. It has neurons as processing elements and links which represent the connections among the neurons (**Dogan et al., 2007**). Every link has a weight parameter associated with itself (**Maier and Dandy, 2000**). The neurons are located in layers and each layer has a specific transfer function. FFANN consists of an input-layer, an output-layer, and several hidden layers, but one hidden-layer is enough to estimate each complex parameter (**Noori et al., 2010b; Singh and Gupta, 2014**). **Figure 4** shows an applied FFANN model in which DO_{min}, DO_{avg} and DO_{max} are as inputs and $BOD_{5(max)}$ is the target. ANN was applied to the data set collected from 94 sample points. The inputs were standardized to be limited in a range between -1 and 1 (**Basant et al., 2010**) and were divided into three categories: training, validation and verification. The laboratory data set from the first to 65th station, 66th to 80th station, and 81st to the 94th were selected as training, validation, and verification sets, respectively. Also, to optimize the network weights, Levenberg-Marquardt (LM), scaled conjugate gradient (SCG) and resilient backpropagation (RP), training functions were used. LM and SCG have a good efficacy and RP training function has a high speed in training procedure. Therefore, these functions have been utilized in the present research. The SCG, against the LM algorithm, avoids the time-consuming calculation of the

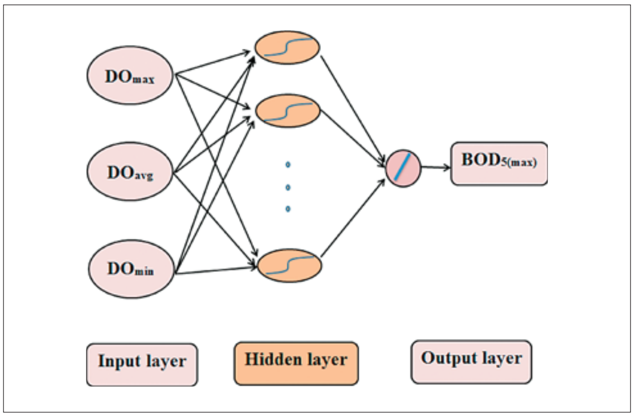

Figure 4: The applied three-layer FFANN in this study

Hessian matrix in the modelling process. It is the main advantage of this algorithm. Also, SCG merges between two approaches: the model-trust region and the conjugate gradient (CG). More information on SCG, LM and RP algorithms can be found in **Noori et al. (2009a; 2010a; 2010b)**.

The backpropagation algorithm is used for training the perceptron network. This supervised learning algorithm is carried out through backpropagation, a generalization of the least mean squares algorithm in the linear perceptron (**Singh et al., 2009**). The error (e) in output-node j in the n^{th} data point can be calculated by **Equation 1** as follows:

$$e_j(n) = t_j(n) - y_j(n) \tag{1}$$

where:

t – the target values,

y – the output values.

The weights of the nodes are based on the corrections which minimize the error in the entire output, given by **Equation 2**.

$$\varepsilon(n) = \frac{1}{2}\sum_j e_j^2(n) \tag{2}$$

Using gradient descent (GD), the change in each weight (Δw) is written by **Equation 3**

$$\Delta w_{ji}(n) = -\sigma \frac{\partial \varepsilon(n)}{\partial \gamma_j(n)}\, y_i(n) \tag{3}$$

where:

y_i – the output of the previous neuron,

σ – the learning rate.

The calculation of derivative counts on the induced local field γ_j. For an output-node, the derivative is calculated as **Equation 4.**

$$-\frac{\partial \varepsilon(n)}{\partial \gamma_j(n)} = e_j(n)\,\phi\,\gamma_j(n) \tag{4}$$

where:

ϕ – the derivative of the transfer function.

The analysis is more difficult for the change in weights to a hidden-node, but it is shown as **Equation 5.**

$$-\frac{\partial \varepsilon(n)}{\partial \gamma_j(n)} = \phi\left(\gamma_j(n)\right)\sum_k -\frac{\partial \varepsilon(n)}{\partial \gamma_k(n)}\, w_{kj}(n) \tag{5}$$

This counts on the change in weights of the k^{th} nodes, which represents the output-layer.

The linear and sigmoidal transfer functions are very popular in the application of ANN models. In some investigations, a sigmoidal transfer function was used in place of a linear function in the output-layer of FFANN (**Nilsson et al., 2006; Sahoo and Ray, 2006**). This can limit the outputs of FFANN to a small range (**Haykin, 1994**). In the present study, for the FFANN model with one hidden-layer, the transfer functions of the hidden-layer and output-layer were selected as tangent-sigmoid and linear functions. To achieve the best architecture, the number of the hidden-layer neurons was identified using the trial-error procedure.

2.3. Radial basis function artificial neural network

The RBFANN was firstly introduced by Lowe and Broomhead (1988) for the application to problems of supervised learning (**Orr, 1996**). This network has a faster learning process in comparison to other neural networks (**Han et al., 2012**). The basic RBFANN structure contains three layers (input, output and hidden layers). The hidden-layer functions transfer the nonlinear input-space to the linear hidden-space (**Liu et al., 2004**; **Singh et al., 2013**). In this study, to construct a software-sensor for the online prediction of $BOD_{5\,(max)}$ using a RBFANN model, input vectors DO_{min}, DO_{avg} and DO_{max} were selected (see **Figure 5**).

A basic RBFANN with one output and k hidden-layer nodes is represented by **Equation 6**

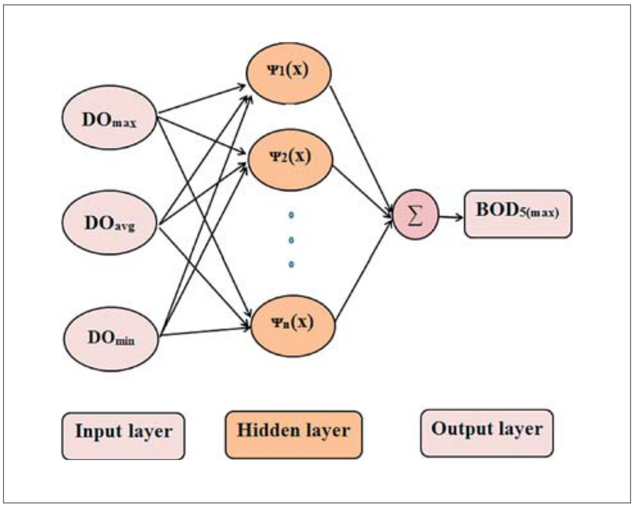

Figure 5: The structure of RBFANN in this study

$$y = \sum_{k=1}^{k} w_k \psi_k(x) \tag{6}$$

where:

x – the input of RBFANN,

w_k – the connecting weights between k^{th} hidden-node and the output-layer,

ψ_k – the output-value of the k^{th} hidden-node which usually based on Gaussian function as **Equation 7**.

$$\psi_k(x) = \exp\left(-\left\|x-\mu_k\right\| / 2\rho_k^2\right) \tag{7}$$

where:

μ_k – the centre vector of the k^{th} hidden-node,

$\left\|x-\mu_k\right\|$ – the Euclidean distance between x and μ_k,

ρ_k – the width of the k^{th} hidden-node.

The RBFANN training process involves determining the centre vectors of the nodes in the hidden-layer, the width of the hidden-layer nodes, and the connecting weights between the two layers (hidden and output layers).

The RBFANN, compared to the FFANN, needs more neurons but its design and training process are faster. In this research, the *newrb* function was used in the MATLAB environment to design and tune the RBFANN model for $BOD_{5(max)}$ estimation. This function creates a RBFANN that includes one neuron the first time. Neurons are added to the network until the sum of the squared error reaches the error goal or the number of the hidden-layer neurons reaches the maximum number (**Demuth and Beale, 2004**).

2.4. Models' evaluation

To estimate the performance of FFANN and RBFANN models, three statistical indices; R, MSE, and DDR were used. R denotes the strength and the direction of a linear relationship between the measured and predicted values. R values range from -1 to +1. The (+) and (-) signs are used for positive and negative linear correlation, respectively. If the observed and predicted values have a strong positive linear correlation, R is close to +1. When the correlation is greater than 0.8, it means that there is a strong fit between the observed and predicted values. R can be represented by **Equation 8**.

$$\mathrm{R} = \frac{n\sum fp - \left(\sum f\right)\left(\sum p\right)}{\sqrt{n\left(\sum f^2\right) - \left(\sum f\right)^2}\sqrt{n\left(\sum p^2\right) - \left(\sum p\right)^2}} \tag{8}$$

where:

f – modelled values,

p – observed values,

n– the number of values.

In statistics, MSE estimates the quality of a model for predicting values. MSE values close to zero are better. It is calculated by **Equation 9** (**Dogan et al., 2016**)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(p_i - f_i)^2 \tag{9}$$

In addition, the DDR statistic that introduces a graphical view for the models' performance is calculated as shown in **Equation 10**.

$$\mathrm{DDR} = \frac{f_i}{p_i} - 1 \tag{10}$$

DDR values must be standardized and then, the normalized value of DDR (Q_{DDR}) is calculated using the Gaussian function. Finally, Q_{DDR} via standardized DDR values must be illustrated. In the obtained figure, more tendencies to the centre line and also, a bigger value of the maximum Q_{DDR} indicates more accuracy. More information on DDR can be found in **Noori et al. (2010a)**.

3. Results and discussion

3.1. Results of feed-forward artificial neural network

In this study, the unknown network parameters, i.e. weights and biases, were optimized in the model using LM, SCG, and RP algorithms, respectively. The results revealed that the best FFANN performance was achieved by the application of 20 neurons in the hidden-layer of the network. The results for the application of each training algorithm based on statistical characteristics, i.e. MSE and R, have been shown in **Table 1**.

Table 1: Calibration-verification results for FFANN and RBFANN

Model	Optimizing algorithm	Calibration		Verification	
		MSE	R	MSE	R
FFANN	LM	0.035	0.89	0.040	0.89
	RP	0.065	0.78	0.057	0.85
	SCG	0.064	0.81	0.054	0.82
RBF	GD	0.041	0.87	0.039	0.90

Based on **Table 1**, it is observed that the tuned model by LM algorithm includes more desirable R and MSE values during both the calibration (training and validation) and verification steps. Therefore, it can be noticed that LM performance is better than SCG and RP algorithms. Besides, **Figure 6** shows a scatter diagram of the predicted $BOD_{5(max)}$ vs measured ones during both calibration-verification steps. The results of the predicted $BOD_{5(max)}$ vs the measured ones and also an error diagram for the FFANN model in each station can be seen in **Figure 7**. This figure shows that the FFANN model tuned by the LM algorithm worked well during both the calibration-verification steps. However, the error diagram illustrates that the FFANN model in some stations was faced with high error values that influence the application of this model as a management tool for BOD_5

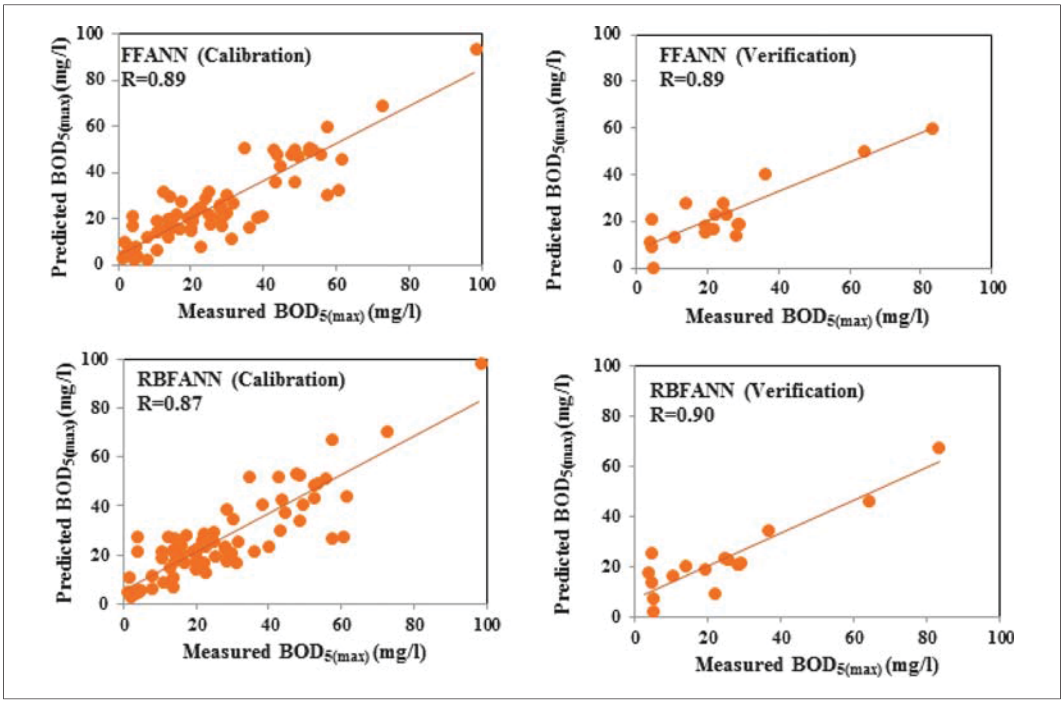

Figure 6: Scatter diagrams of the predicted $BOD_{5(max)}$ by FFANN and RBFANN models vs observed ones for calibration-verification steps

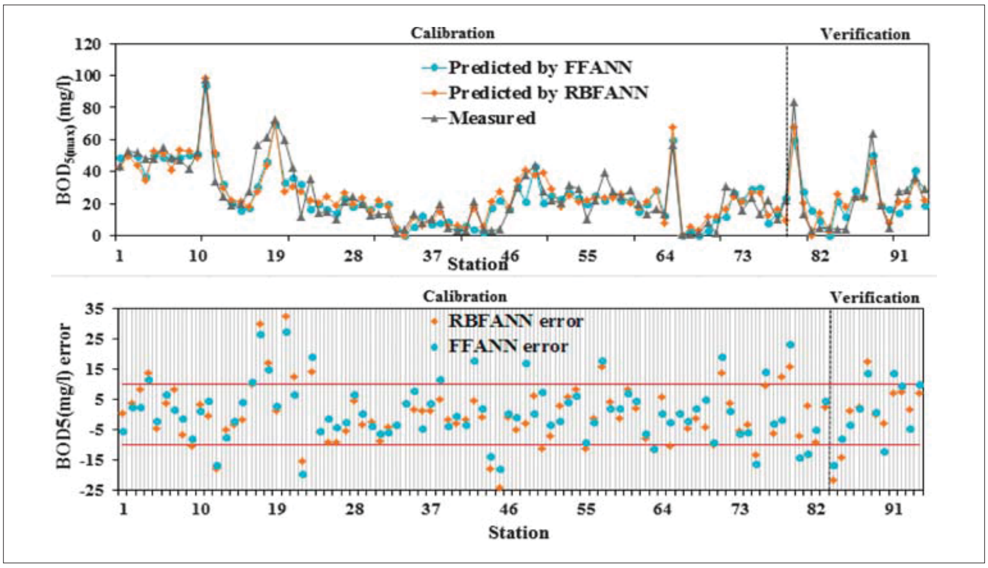

Figure 7: The predicted $BOD_{5(max)}$ vs measured ones and the error diagrams for FFANN and RBFANN models in stations

monitoring in the Sefidrood River. In **Figure 7**, the two red lines indicate the range of acceptable error which is less than 10%. This figure shows that the error for approximately 25% of the points for FFANN model is out of this range.

3.2. Results of radial basis function artificial neural network

In RBFANN the width of the hidden nodes was optimized through trial and error. The maximum number of the hidden-layer neurons was selected with consideration of the inputs. Finally, the width value and the maximum hidden-layer's neurons were considered within ranges 0 to 4 and 5 to 30, respectively. The RBFANN weights were updated depending on the GD approach. The obtained results by the trial and error procedure specified that the best RBFANN performance was in the application of 20 neurons in the hidden-layer of the network and considering the width of the hidden nodes equal to 2. MSE and R indices for the calibration-verification steps of RBFANN model have been shown in **Table 1**. This table shows that the tuned model by RBFANN includes proper R and MSE values during the calibration-verification steps. Also, the scatter and trend diagrams of the predicted $BOD_{5(max)}$ vs measured ones for RBFANN model have been shown in **Figure 6**. Moreover, the error diagram (see **Figure 7**) clearly shows the poor performance of the RBFANN model in some points so that the error of 23% of the points is more than 10%.

However, a visual inspection of the results indicates that the tuned models by RBFANN and FFANN have approximately the same performance. To have a better evaluation regarding the performance of FFANN and RBFANN models, the Q_{DDR} via standardized DDR values for FFANN and RBFNN in the verification step have been shown in **Figure 8**. Based on this figure, the tendency to the centre line in RBFANN is as same as that in FFANN model. Besides, the maximum of Q_{DDR} for the both models is close to 0.4. Therefore, the DDR analysis reveals that the performance of RBFANN is practically same as the FFANN model.

However, in comparison with previous studies, it can be found that the applied FFANN and RBFANN models in this research have a good performance for $BOD_{5(max)}$ prediction. Regarding the FFANN model, **Onkal-Engin et al. (2005)** tuned a FFANN model with R equal to 0.93 for the prediction of BOD_5. **Singh et al. (2009)** calibrated different FFANN models to predict BOD_5. They reported that the models had a good performance with R ranged from 0.70 to 0.85 between the measured and the modelled BOD_5. Also, the R value for the presented FFANN model by **Dogan et al. (2007)** was 0.93. In another work, **Noori et al. (2013)** tuned a FFANN model with R equal to 0.94 for the prediction of BOD_5. Noted although there are some studies that aimed to predict BOD_5 using FFANN, an application of the RBFANN model, which in this case, is rare. Thus, this study aimed to investigate the RBFANN performance for BOD_5 prediction in rivers. Also, artificial intelligence techniques such as FFANN and RBFANN models are sensitive to the case study and selection of inputs. So the models' performance is highly influenced by the selected inputs.

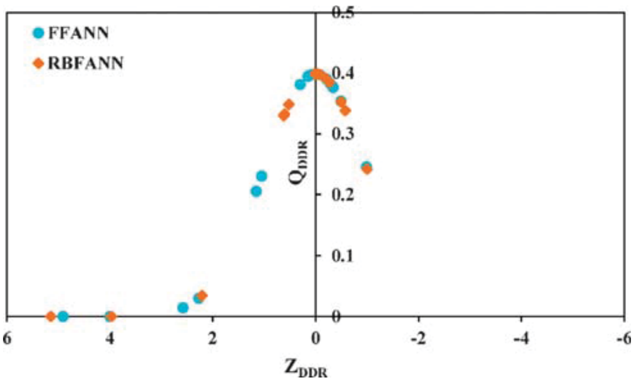

Figure 8: Q_{DDR} via standardized DDR values for FFANN and RBFNN in testing step

In this regard, this study aimed to tune a model whereas faced with some financial constraints. Therefore, only the main effective parameter on BOD_5 variations, i.e. DO was measured and other effective parameters such as nutrients, industrial effluents and flow discharge were ignored. By considering the facts, the obtained results are acceptable and could be applied to understanding the pollution trends in the Sefidrood River.

4. Conclusions

This study aimed to estimate the $BOD_{5(max)}$ as a function of DO_{min}, DO_{avg} and DO_{max} in the Sefidrood River Basin, Iran, using FFANN and RBFANN models. The performance of the tuned models was assessed through R, MSE, and DDR indices. Calibration-verification results specified that FFANN and RBFANN models worked well and predicted $BOD_{5(max)}$ with R values close to 0.89 and 0.90 in the verification step, respectively. For FFANN, it was observed that the tuned model by the LM algorithm resulted in more desirable R and MSE values. In addition, R and MSE values for RBFANN and FFANN models were approximately acceptable. According to DDR analysis, the results demonstrated that the performance of both models was approximately the same and the maximum values of Q_{DDR} were practically the same for both networks. However, because of insufficient data about some effective parameters on BOD_5 such as algae respiration and toxic substances, both the tuned models did not have excellent performance in some stations. Therefore, it is recommended that the models are again tuned using more data if the decision makers aim to apply the results for management of water quality in the Sefidrood River.

5. References

APHA (American Public Health Association) (1995): Standard Methods for the Examination of Water and Wastewater. (19th ed.), Washington.

Basant, N., Gupta, S., Malik, A. and Singh, K.P. (2010): Linear and nonlinear modelling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water—a case study. Chemometrics And Intelligent Laboratory Systems, 104, 172-180. DOI: 10.1016/j.chemolab.2010.08.005.

Chavez, P. and Chang, F.J. (2006): Simulation of Multiple Water Quality Parameters Using Artificial Neural Networks. 7th International Conference on Hydro information, Nice, France.

Demuth, H.B. and Beale, M.H. (2004): Neural network toolbox for use with MATLAB: User guide. Math Works.

Dogan, E., Kocamaz, U.E., Utkucu, M. and Yıldırım, E. (2016): Modelling daily water level fluctuations of Lake Van (Eastern Turkey) using Artificial Neural Networks. Fundamental And Applied Nematology, 187, 177-189.

Dogan, E., Koklu, R. and Sengorur, B. (2007): Estimation of Biological Oxygen Demand Using Artificial Neural Network. International Earthquake Symposium, Kocaeli.

Dogan, E., Yuksel, I. and Kişi, O. (2007): Estimation of total sediment load concentration obtained by experimental study using artificial neural networks. Environmental Fluid Mechanics, 7, 271-288, DOI: 10.1007/s10652-007-9025-8.

Fan, C., Ko, C.H. and Wang, W.S. (2008): An innovative modeling approach using Qual2K and HEC-RAS integration to assess the impact of tidal effect on River Water quality simulation. Journal of Environmental Management, 90(5), 1824-1832.DOI: 10/.1016/j. jenvman.2008.11.011.

Han, H.G., Qiao, J.F. and Chen, Q.I. (2012): Model predictive control of dissolved oxygen concentration based on a self-organizing RBF neural network. Control Engineering Practice, 20(4), 465-476. DOI:10.1016/j.conengprac. 2012.01.001.

Haykin, S. (1994): Neural Networks: A Comprehensive Foundation. (2nd ed.), New York, Macmillan.

Liu, Y., Zheng, Q., Shi, Z. and Chen, J. (2004): Training radial basis function networks with particle swarms. In International Symposium on Neural Networks. Springer Berlin Heidelberg, 317-322. DOI: 10.1007/978-3-540-28647-9_54.

Maier, H.R. and Dandy, G.C. (2000): Neural Networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environmental Modelling and Software, 15, 101-124. DOI: 10.1016/S1364-8152(99)00007-9.

Nilsson, P., Uvo, C.B. and Ronny, B. (2006): Monthly runoff simulation: Comparing and combining conceptual and neural network models. Journal of Hydrology, 321, 344-363. DOI: 10.1016/j.jhydrol.2005.08.007.

Ning, S.K., Chang, N.B., Yang, L., Chen, H.V. and Hsu, H.Y. (2001): Assessing pollution prevention program by QUAL2E simulation analysis for the Kao-Ping River Basin, Taiwan. Journal of Environmental Management, 61(1), 61-76. DOI:10.1006/jema.2000. 0397.

Noori, R., Hoshiyaripour, G.A. Ashrafi, K. and Arabi, B.A. (2010b): Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. Atmospheric Environment, 44(4), 476-482. DOI:10.1016/j.atmosenv.2009.11.005.

Noori, R., Karbassi, A., Farokhnia, A. and Dehghani, M. (2009a): Predicting the longitudinal dispersion coefficient using support vector machine and adaptive neuro-fuzzy inference system techniques. Environmental Engineering Science, 26, 1503-1510. DOI: 10.1089/ees.2008.0360.

Noori, R., Khakpour, A., Omidvar, B. and Farokhnia, A. (2010a): Comparison of ANN and principal component analysis multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statist. Expert Systems with Applications, 37(8), 5856-5862. DOI:10.1016/j.eswa.2010.02.020.

Noori, R., Karbassi, A., Ashrafi, K., Ardestani, M., Mehrdadi, N. and Bidhendi, G.R.N. (2012): Active and online prediction of BOD5 in river systems using reduced-order support vector machine. Environmental Earth Sciences, 67(1), 141-149. DOI:10.1007/s12665-011-1487-9

Noori, R., Safavi, S. and Shahrokni, S.A.N. (2013b): A reduced-order adaptive neuro-fuzzy inference system model as a software sensor for rapid estimation of five-day biochemical oxygen demand. Journal of Hydrology, 495, 175-185. DOI:10.1016/j.jhydrol.2013.04.052

Noori, R., Ashrafi, K., Karbassi, A.R., Ardestani, M. and Mehrdadi, N. (2013a): Development and application of reduced-order neural network model based on proper orthogonal decomposition for BOD_5 monitoring: Uncertainty analysis. Environmental Progress & Sustainable Energy, 32(2), 344-349. DOI: 10.1002/ep.11610

Noori, R., Yeh, H.D., Abbasi, M., Kachoosangi, F.T. and Moazami, S. (2015): Uncertainty analysis of support vector machine for online prediction of five-day biochemical oxygen demand. Journal of Hydrology, 527, 833-843. DOI:10.1016/j.jhydrol.2015.05.046

Onkal-Engin, G., Demira, L. and Enginb, S.N. (2005): Determination of the relationship between sewage odour and BOD by neural networks. Environmental Modelling and Software, 20, 843-850. DOI: 10.1016/j.envsoft.2004.04.012.

Orr, M.J. 1996. Introduction to Radial Basis Function Networks, University of Edinburgh, Scotland. URL: http://www.cc.gatech.edu/~isbell/tutorials/rbf-intro.pdf.

Park, S.S. and Lee, Y.S. (2002): A water quality modeling study of the Nakdong River, Korea. Ecological Modelling, 152(1), 65-75. DOI: 10.1016/S0304-3800(01)00489-6.

Sahoo, G.B. and Ray, C. (2006): Flow forecasting for a Hawaii stream using rating curves and neural networks. Journal of Hydrology, 317, 63-80. DOI: 10.1016/j.jhydrol.2005.05.008.

Sharma, D., Kansal, A. and Pelletier, G. (2015): Water quality modeling for urban reach of Yamuna River, India (1999–2009), using QUAL2Kw. Applied Water Science, DOI 10.1007/s13201-015-0311-1.

Singh, K.P., Basant, A., Malik, A. and Jain, G. (2009): Artificial neural network modeling of the river water quality – a case study. Ecological Modelling, 220(6), 888-895. DOI: 0.1016/j.ecolmodel.2009.01.004.

Singh, K.P. and Gupta, S. (2012): Artificial intelligence based modeling for predicting the disinfection by-products in water. Chemometrics and Intelligent Laboratory Systems, 114, 122-131. DOI:10.1016/j.chemolab.2012.03.014.

Singh, K.P., Gupta, S., Ojha, P. and Rai, P. (2013): Predicting adsorptive removal of chlorophenol from aqueous solution using artificial intelligence based modeling approaches. Environmental Science and Pollution Research, 20, 2271-2287. DOI: 10.1007/s11356-012-1102-y.

Singh, K.P., Ojha, P., Malik, A. and Jain, G. (2009): Partial least squares and artificial neural networks modeling for predicting chlorophenol removal from aqueous solution. Chemometrics and Intelligent Laboratory Systems, 99, 150-160. DOI: 10.1016/j.chemolab.2009.09.004.