Artificial neural network modeling of the effluent quality index for municipal wastewater treatment plants using quality variables: south of Tehran wastewater treatment plant

Maliheh Falah Nezhad, Naser Mehrdadi, Ali Torabian
and Sadegh Behboudian

ABSTRACT

Using treated wastewater could be a realistic solution for the scarcity of fresh water in Iran; therefore, effluent quality evaluation is an essential task. In the present study, the effluent quality index (EQI) for south of Tehran municipal wastewater treatment plant was found considering eight quality parameters and sub-index formulae method. Further, with the help of an artificial neural network (ANN), design and applicability of feed-forward, three-layer perceptron neural network model for computing the EQI has been assessed. In this study, ANN modeling is done by ANN toolbox in MATLAB 2013. The modeling efforts showed that the optimal network architecture was 8–7–1 and that the best EQI predictions were associated with the Levenberg–Marquardt back propagation training algorithm and Tansig transfer function. The EQI predictions of this model had significant and very high correlation (R = 0.96, MSE = 0.1) with the measured EQI values. The ANN approach which is used in this article suggested powerful tool to EQI computation and prediction, since it reduced lengthy computations and using various sub-index formula for each value.

Key words | artificial neural networks, back-propagation, effluent quality index, municipal wastewater treatment plant, sub-index

Maliheh Falah Nezhad (corresponding author)
Naser Mehrdadi
Ali Torabian
Department of Environmental Engineering,
    Graduate Faculty of Environment,
University of Tehran,
Ghods Street,
Tehran,
Iran
E-mail: *m_falahnezhad@yahoo.com*

Sadegh Behboudian
College of Engineering, School of Civil Engineering,
University of Tehran,
PO Box 11155-4563,
Tehran,
Iran

INTRODUCTION

Water quality management has become an important issue throughout the world (Najufpoor *et al.* 2007). Most countries have planned water quality monitoring and assessment of surface water in terms of their physical, chemical, biological and nutrient constituents (Sapkal & Valunjkar 2013).

Water quality index (WQI) is a description of biological, chemical, and physical characteristics of water in connection with intended use(s) and a set of standards (Liou *et al.* 2004; Boyacioglu 2007; Khalil *et al.* 2011; Gazzaz *et al.* 2012). Since no individual parameter can express water quality sufficiently, water quality indices have been developed to integrate measurements of a set of parameters into a single index (Zandbergen & Hall 1998).

Due to the scarcity of fresh water resources in the Middle East regions, wastewater reuse could be a realistic option to compensate the lack of fresh water and until now, the largest and most popular wastewater reuse has been in the agricultural irrigation field (Palese *et al.* 2009; Alfarra *et al.* 2011).

Wastewater treatment plant (WWTP) effluents must respect all the quality standards and cannot be described by purely theoretical models (Verlicchi *et al.* 2011).

Effluent quality index (EQI), which is able to rapidly evaluate whether an effluent is appropriate for agricultural, recreational, industrial or disposal purposes, could be of great help to managers and decision makers in water resources planning. Moreover, this index is a useful tool for comparing different wastewater treatment processes.

The index is an appropriate tool for:

(a) quick comparison of water quality;
(b) rapid evaluation of treatment options;
(c) rapid evaluation of the improvement in effluent quality.

Although constructing an EQI has many advantages, the calculations are complicated and unintentional errors make it difficult and inaccurate. Thus, an alternative direct and quick means of computing and forecasting EQI values based on artificial neural network (ANN) modeling is suggested. This model has the potential to reduce computation time and errors. Therefore, this study illustrates the design of a neural network model as an alternative to EQI computation methods.

The ANNs are popular tools for modeling highly complicated relationships, processes, and phenomena. They have been widely used in environmental modeling and predicting such as river flow forecasting (e.g. Shamseldin *et al.* 2002; Teschl & Randeu 2006); rainfall-runoff modeling (e.g. Riad *et al.* 2004; Wu & Chau 2011); water demand forecasting (Behboudian *et al.* 2013); solid waste forecasting (Abdoli *et al.* 2011) and WQ forecasting (e.g. Palani *et al.* 2008; Singh *et al.* 2009; Khalil *et al.* 2011). However, the specific use of ANNs to develop predictive models for EQI has not been yet investigated by researchers. Therefore, the main objectives of this study were to: (1) construct the EQI with traditional methods; (2) demonstrate the potential of the ANN for producing models capable of efficient forecasting of the EQI; (3) illustrate the general framework for ANN model design (e.g. selection of network type; determination of appropriate input variables and number of hidden neurons; and specification of the optimum settings of the network training parameters).

Among black-box models, ANN is a prevalent modeling method with growing applications (Banerjee *et al.* 2011). The most common application of ANN is to map a set of input vectors to the specified output ones. In the case of EQI modeling, the input vectors are quality parameters such as biochemical oxygen demand (BOD), total dissolved solids (TDS), fecal coliform (FC), etc. and the output vector is the EQI. Thus, in order to model EQI two steps must be followed: (1) EQI construction for a set of observed data; and (2) ANN modeling in order to predict EQI.

MATERIALS AND METHODS

Developing a new index for Iran's effluent of municipal wastewater treatment plants includes the following two sections.

EQI formula

Determining the significant parameters is the first step in developing the EQI. Initially, 15 parameters were selected including BOD, chemical oxygen demand (COD), total suspended solids (TSS), TDS, pH, total P, total N, NO_3, NH_3, NO_2, T, dissolved oxygen (DO), fecal coliform, total coliform, and PO_4. Based on the Delphi method 100 experts filled the questionnaires and the average weight of each parameter is calculated. Finally, the following parameters were selected: BOD, COD, TSS, PO_4, NH_4, fecal coliform, pH, TDS.

In the next step, a weight is assigned to each selected parameter based on the result of the Fuzzy Topsis method. The parameter which has adverse effects on human health has complicated and expensive treatment process and more aesthetically unpleasant is assigned higher weight in questionnaires so that it increases EQI. Afterward, the parameter weights are defined using the Fuzzy Topsis decision making method. The related formula and description about this method can be found in Hwang & Yoon (1981) and Wu & Chen (2007).

The third step involved determination of quality function (curve), i.e. sub-index for each selected parameter. Sub-indices are calculated by converting the value of each selected variable into a non-dimensional, scaled value through a sub-index rating curve where each variable has its own rating curve on a scale of deteriorating EQI, mostly from 0 to 100 (Liou *et al.* 2004; Kaurish & Younos 2007). Basically, these rating curves were developed in this study from statistical data analysis of raw data acquired by long-term treated wastewater analysis with the assistance of wastewater quality expert

opinions. The final step is the aggregation and calculation of the EQI which is the most important step in calculating EQI. In order to minimize ambiguity and eclipsing, it is necessary to identify an appropriate function for aggregation (Obaid & Al-Shujairi 2013). In this study, the multiplicative product method of aggregation is used. Overall effluent quality index is determined by the formula given in Equation (1):

$$EQI = \sum_{n=1}^{n=8} w_n I_n \tag{1}$$

where I_n is the sub-index relating to the parameters COD, BOD, TSS, NH_4, PO_4, Fc and W_n is the corresponding weight for each parameter.

Developing artificial neural networks

ANNs are non-linear processing systems operating in parallel (Sarkar *et al.* 2009). The concept of ANNs was developed based on the real biological human brain. Analogous to the brain, ANNs are composed of nodes and connections. The nodes are called neurons, which are the simplest and the major computational units and the connections are network weights, which are responsible for establishing several types of ANN proposed so far. However, two kinds of networks, i.e. back-propagation neural network (BPNN) and radial basis function (RBF), are the most common network structures used for engineering purposes.

The feed forward back propagated MLP is the default network type for most MLPs. It has multiple layers of neurons with nonlinear transfer functions that allow the network to learn nonlinear and linear relationships between input and output vectors. Different training algorithms can be used in the training stage: gradient descent, gradient descent with momentum, gradient descent which has a variable learning rate, gradient descent with momentum, which has a variable learning rate, and RP are the most famous training functions. Conjugate gradient (CG) algorithms are the other algorithms for the training procedure. The standard back propagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient), where the performance function is decreasing most rapidly. For faster convergence than steepest descent

directions in the CG algorithms, a search is performed along conjugate directions. The Levenberg–Marquardt algorithm is also a training algorithm designed to approach second order training speed without having to compute the Hessian matrix.

Feed forward back-propagation neural network

BPNN is the most popular network, especially in the case of non-linear approximation (Kashaninejad *et al.* 2009). Each BPNN consists of an input layer, an output layer and one or more hidden layers. Back-propagation is used as the learning algorithm, and it involves working backward from the output layer to adjust the weights accordingly, and reducing the average error across all layers. This process is repeated until the weights reach their optimal values and the error between the output of the network and the actually desired output is minimized (Aydiner *et al.* 2005).

The sum of the weighted inputs forms the input to the transfer function. Neurons can use any differentiable transfer function to generate their output. In BPNN, a global function, like the tangent function or the logistic function, is often used as an activation function. The main task in designing a BPNN network is to find the appropriate number of layers and also the number of neurons in each layer in the way that the overall network error minimizes.

Model performance

To evaluate the performance of the ANN model, two indices are assessed: mean square error (MSE) and correlation coefficient (R). The accuracy of the model is determined according to these criteria (Abdoli *et al.* 2011). The MSE can be defined as follows:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_{i,actual} - y_{i,predicted})^2 \tag{2}$$

where $y_{i,actual}$ is the output of the process, $y_{i,predicted}$ is the ith output of the network corresponding to the $y_{i,actual}$ and N is the total number of data points.

In order to reduce error impacts, it is necessary to standardize the data, that is, to convert the data into a

non-dimensional form of uniform range of variability (Dawson & Wilby 2001; Özesmi *et al.* 2006). The variables input to the neural network (eight EQVs and the EQI) were standardized to the range of, namely, (0,1).

In order to ensure that the different EQVs would have equal weights in the analysis, the raw EQVs were standardized through z-scale transformation to a mean of 0.0 and variance of 1.0 by using the equation (Güler *et al.* 2002; Kowalkowski *et al.* 2006):

$$Z_{ij} = (X_{ij} - \mu_j)\big/_{\sigma_j} \tag{3}$$

where Z_{ij} is the standard score of the jth value of the measured variable i; x_{ij} is the jth observation of variable i; and l_i and r_i are the mean and standard deviation values, respectively, of the population of variable i.

Case study

South of Tehran, the WTP is located in the southeastern region of the city. It currently covers 525,000 inhabitants for each module and it contains eight modules. The average inlet flow rate of this WWTP is about 2,600 LPS and it is expected to provide about 3,700 LPS in future. Activated sludge is the biological wastewater treatment processes of this WWTP and it includes different units of screen bar unit, primary settling tank, selector, aerated tank, secondary settling tank, and UV units.

Data

In this study a data set from the south of Tehran wastewater treatment plant has been used. Daily measurements for 15 parameters including BOD, COD, TSS, TDS, pH, total P, total N, NO$_3$, NH$_3$, NO$_2$, T, DO, fecal coliform, total

coliform, PO$_4$ over a range of 6 months was conducted. However, not all of them were used in ANN modeling and the most significant variables are defined through the Delfi method. Finally, eight variables were defined and involved in the simulation. Values of some descriptive statistics for these variables are shown in Table 1.

RESULTS AND DISCUSSION

The aim of developing an effluent quality index is to improve water quality by reducing the concentrations of the main pollutant parameters to the levels required for its final destination (Verlicchi *et al.* 2011).

As mentioned above, the most significant parameters for discharging the domestic effluent to the surface bodies and for reusing in irrigation and agriculture, industrial uses or recreational activities based on the Delphi method are as follows: BOD$_5$, COD, TSS, PO$_4$, NH$_4$, fecal coliform, TDS, pH.

The Fuzzy Topsis method is utilized to estimate the rate of importance for each parameter based on the fulfilled questionnaires regarding the effect of each parameter on health, aesthetic effect, and treatment cost process. The related formula and extra descriptions about this method can be found in the literature (Hwang & Yoon 1981; Wu & Chen 2007).

A panel of experts of about 100 people was selected: wastewater treatment operators, university professors, wastewater consultant companies, wastewater managers. Decision makers completed the questionnaires and assigned a weight 1 to 5 for each parameter considering the adverse effect based on the criteria. In addition, each criterion itself is assigned a weight based on its significance by decision makers. Also, it is supposed that all the decision maker groups have similar weights.

Table 1 | Data for effluent of south of Tehran WTP

Variable	BOD (mg/L)	COD (mg/L)	TSS (mg/L)	Fecal coliform (CFU/100)	NH$_4$ (mg/L)	PO$_4$ (mg/L)	pH	TDS (mg/L)
Minimum	1.00	5.60	1.00	100.00	0.04	0.50	5.50	120.00
Maximum	120.00	210.00	100.00	5,320.00	55.30	18.00	9.50	1,600.00
Average	17.80	33.57	18.49	638.51	4.80	5.68	7.31	588.60
STD	24.70	43.42	23.29	686.99	10.46	3.70	0.64	283.66

Table 2 | Assigned weight to the parameters according to the Fuzzy Topsis method

Parameter	Weight
BOD	0.076
COD	0.076
TSS	0.088
FC (fecal coliform)	0.17
NH_4	0.167
PO_4	0.157
pH	0.134
TDS	0.135

Table 2 shows the results achieved by the Fuzzy Topsis method. In step 2 the sub-indices should be determined. The associated best-fit formulas to each parameter rating curve were used to calculate aggregated index. Each of eight parameter sub-indexes (SI) used to calculate the overall effluent quality index are listed in Table 3.

Thus, by substituting the value of sub-indices, the EQI formula is calculated by the following relationship:

$$\text{EQI} = \sum_{n=1}^{n=8} w_n I_n = 0.076 I_{BOD} + 0.076 I_{COD} \\ + 0.088 I_{TSS} + 0.17 I_{Fc} + 0.167 I_{NH4} \\ + 0.157 I_{PO4} + 0.134 I_{PH} + 0.135 I_{TDS} \tag{4}$$

where I is the normalized sub-indices of the parameters. Treated wastewater can be reused for agricultural, industrial, or recreational purposes if their quality meets specific legal requirements. In Iran, the limitations set by the Iranian environmental agency focused on agricultural purposes, since other purposes are not widely used in Iran. Therefore, in the case where the Iran Environmental standard is incomplete, EPA and Jordanian limitations have been used (WHO 2005; EPA 1992). Major problems usually arise due to high concentrations of the parameters selected to develop the EQI. Their legal limitations for wastewater reuse and disposal are reported in Table 4.

As can be seen in Table 4, the thresholds for different reuse purposes are calculated through the EQI formula and it can be concluded that the EQI formula can suggest the best reuse option for each effluent sample.

Step 2: developing ANN model

The ANNs represent an innovative and attractive solution to the problem of relating output to input variables in complex systems (Dawson & Wilby 2001) and prediction is a common reason for employment of the neural network technology. The major steps for the development of ANN models include defining the suitable model inputs, specifying network type, pre-processing and partitioning of the available data; determining network architecture; defining model performance criteria; training (optimization of connection weights); and validating the model (Govindaraju 2000; Maier & Dandy 2000; Dawson & Wilby 2001). These steps are outlined below with some detail.

In this study, ANN modeling is done by ANN toolbox in MATLAB 2013. The first step aimed at determining the numbers of input, hidden, and output neurons; and the optimal data splitting plan while the second phase involved specifying the training algorithm, learning rate, number of iterations, and training stopping criteria. Thereupon, this study employed a neural network with one input layer, one hidden layer, and the EQI as the output layer, thus producing a three-layer perceptron (TLP) network.

The hidden layers provide the network with its ability to generalize. In theory, a network with a hidden layer and adequate number of hidden neurons can simulate any continuous function and represents a rich and flexible class of universal approximates (Dawson & Wilby 2001; Fischer 2006; Palani *et al.* 2008).

Determining the number of hidden neurons is the significant part and is usually achieved through a trial and error task in ANN modeling (Özesmi *et al.* 2006; Palani *et al.* 2008; Singh *et al.* 2009). The Alyud Research Company (2003) suggested that the N should fall in I/2 to 4I where the I is the number of inputs. In this study, a different number of neurons from 4 to 15 with different train algorithms has been assessed.

Eventually, this search process depicted that the optimal network architecture was 8–7–1 and that the optimum associated partitioning scheme was 70–15–15% (i.e. the proportions of the samples allocated to the training, cross-validation, and testing sets were 70%, 15%, and 15%, respectively). This network had a testing error (RMSE) of 0.32 and the EQI predictions it produced had a significant,

Table 3 | Sub-indices formula achieved from rating curves

Parameter		Sub-index
BOD	If $0 \leq$ BOD ≤ 12	$I_{BOD} = -0.3299x^2 + 11.731x - 22.184$ $R^2 = 0.9975$
	If $12 <$ BOD ≤ 107	$I_{BOD} = -12.991\ln(x) + 38.244$ $R^2 = 0.9935$
COD	If COD ≤ 11	$I_{COD} = 0$
	If $11 <$ COD ≤ 53	$I_{COD} = -0.0759x^2 + 6.9783x - 67.439$ $R^2 = 0.9957$
	If $53 <$ COD ≤ 165	$I_{COD} = 8.3501\ln(x) + 57.746$ $R^2 = 0.9465$
	If $165 <$ COD	$I_{COD} = 100$
TSS	If TSS ≤ 1	$I_{TSS} = 0$
	If $1 <$ TSS ≤ 98	$I_{TSS} = 22.908\ln(x) - 0.7474$ $R^2 = 0.9841$
	If $98 <$ TSS	$I_{TSS} = 100$
Fecal coliform	If FC ≤ 130	$I_{FC} = 0$
	If $130 <$ FC ≤ 580	$I_{FC} = 0.1617x - 19.801$ $R^2 = 0.9863$
	If $580 <$ FC $\leq 3{,}700$	$I_{FC} = 9E\text{-}09x^3 - 6E\text{-}05x^2 + 0.132x + 6.6894$ $R^2 = 0.9892$
	If $3{,}700 <$ Fc	$I_{FC} = 100$
NH$_4$	If NH4 ≤ 0.1	$I_{NH4} = 0$
	If $0.1 <$ NH4 ≤ 53	$I_{NH4} = 16.259\ln(x) + 40.013$ $R^2 = 0.9775$
	If $53 <$ NH4	$I_{NH4} = 100$
PO$_4$	If PO4 ≤ 2.6	$I_{PO4} = 0$
	If $2.6 <$ PO4 ≤ 5	$I_{PO4} = -12.083x^3 + 136.29x^2 - 465.36x + 501.72$ $R^2 = 0.9938$
	If $5 <$ PO4 ≤ 17	$I_{PO42} = 6{,}069x + 57.506$ $R^2 = 0.9815$
	If $16 <$ PO4	$I_{PO4} = 100$
TDS	If TDS ≤ 296	$I_{TDS} = 0$
	If $296 <$ TDS $\leq 1{,}600$	$I_{TDS} = 1E\text{-}07x^3 - 0.0004x^2 + 0.4839x - 110.99$ $R^2 = 0.9783$
	If $1{,}600 <$ TDS	$I_{TDS} = 100$
pH	If pH ≤ 6	$I_{pH} = 100$
	If $6 <$ pH ≤ 9.2	$I_{pH} = 36.394x^2 - 553.95x + 2{,}100.5$ $R^2 = 0.9814$
	If $9.2 <$ pH	$I_{pH} = 100$

Table 4 | Reuse and disposal limitations of treated effluent

	BOD (mg/L)	COD (mg/L)	TSS (mg/L)	Fecal coliform	NH$_4$ (mg/L)	PO$_4$ (mg/L)	pH	TDS (mg/L)	EQI
Recreational limitations	5	10	30	400	0.02	1	6–9	750	26
Industrial reuse	30	75	30	200	2	4	6–9	1,000	48
Surface water disposal	30	60	40	400	2.5	6	6.5–8.5	1,500	56
Ground water disposal	30	60	40	400	1	6	5–9	1,500	53
Agriculture limitations	100	200	100	400	50	15	5–9	1,500	75

very high, positive correlation with the measured EQI values ($R = 0.96$).

The Levenberg–Marquardt (LM) training algorithm and tansig activated function achieved the best performance among the various training algorithms. Figure 1 illustrates the performance of different training algorithms. If the validation error was increased six times sequentially, the training would stop. Training was restarted using network weights, which had been obtained from the previous run until acceptable results on the testing block were reached.

Figure 2 illustrates the validation of ANN model by comparing with target EQI values. We used scatterplot to illustrate the relationship between the output EQI and target EQI and to see if the model fits the data (Figure 3). This line is a graphical representation of the mathematical regression equation. It is plotted using the least squares method which minimizes the sum of the squared distances between the points and the fitted line.

Sensitivity analysis

A sensitivity analysis was carried out to evaluate the relative importance of each of the eight variables in prediction of the EQI. Sensitivity analysis refers to assessment of the importance of predictors in the fitted models. This analysis ranks the predictor variables according to the deterioration in model performance that occurs if a variable is removed from the model. Ultimately, it identifies the variables that can be ignored safely in subsequent analyses as well as the essential variables which must be retained (Olszewski et al. 2008; Gazzaz et al. 2012). In this research each input parameter has been removed separately and the simulation has been conducted without that parameter. The ratio is obtained from the network error after individual variables were removed to the network error from using all variables in the network, i.e. the error of the reduced model against that of the full model. If the ratio is below 1, the parameter can be removed from the model since it is not so important in ANN modeling (Olszewski et al. 2008). Table 5 demonstrates the ratio of each parameter. As can be concluded from Table 5, pH and COD ratio was below 1 and can be omitted from modeling.

The ANN method for EQI calculation and forecasting offers some advantages over the traditional method. The traditional method uses manual calculations whereby the raw data of eight quality variables (BOD, COD, TSS, NH_3-N, fecal coliform, PO_4, TDS and pH) have to be converted into sub-indices (Table 2) before the EQI can be calculated. Since the calculations are not performed on the parameters and their sub-indices which consist of a set of equations should be used for calculations, EQI estimation becomes a time consuming task. On the contrary, the ANN approach utilized archived data to establish a model that can be used for direct calculation of the EQI from raw data without
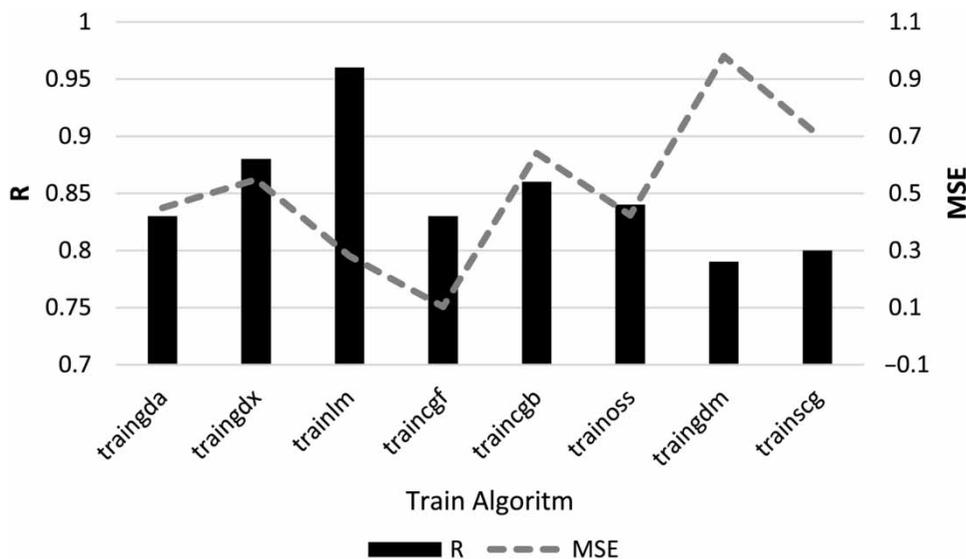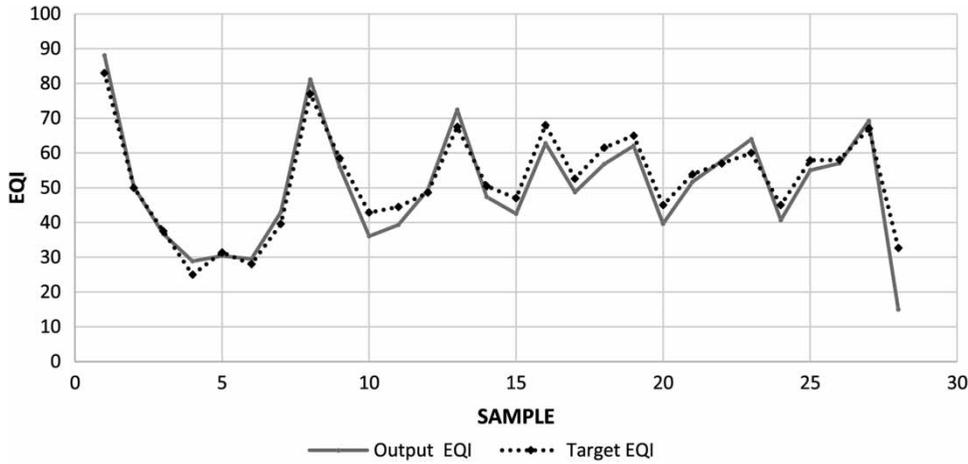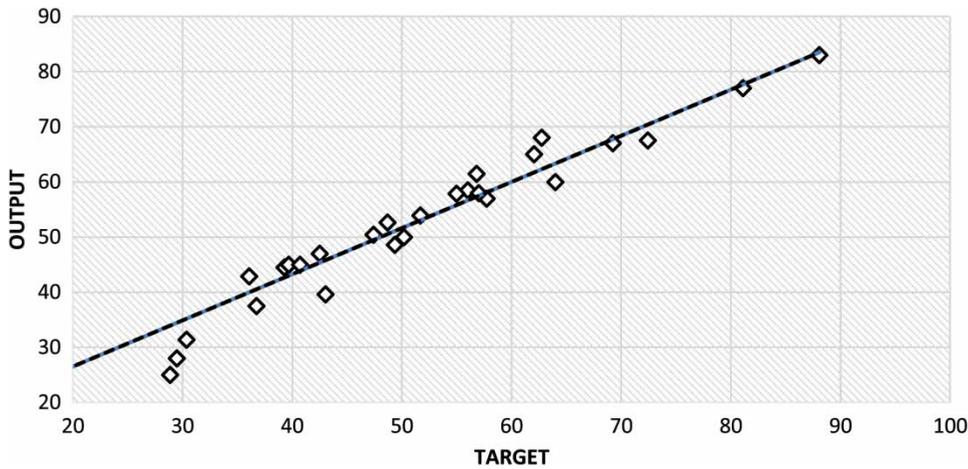

Figure 1 | The performance parameters of training algorithms.

Figure 2 | A graphical representation of the results of ANN model validation. This figure compares between the ANN-output EQI and the target EQI values for each individual observation.

Figure 3 | Scatter plot of the output and target EQI.

Table 5 | Sensitivity analysis result

Parameter	Ratio	Rank
Fecal coliform	1.26	1
NH_4	1.19	2
TDS	1.18	3
TSS	1.17	4
PO_4	1.13	5
BOD	1	6
PH	0.95	7
COD	0.94	8

the need for sub-indexing. Of course, ANN approach is a more direct, rapid, and convenient means of calculation of the EQI than the traditional method.

Moreover, there is a chance by using ANNs to assess more than eight quality variables, which the EQI formula does not cover, since ANNs can be trained and calibrates the weights in order to reach satisfactory results.

The EQI obtained from the traditional method were considered as reference values for the ANN method. Gazzaz *et al.* (2012) used both the traditional method and the ANN approach to estimate EQI.

The performance of the ANN model can be evaluated by comparing its EQI outputs with those of the traditional method through regression analyses. The results demonstrate that a reasonable approximation was made by the ANN model across the calculated EQI values. The overall agreement between the calculated and simulated EQI value was very satisfactory ($r = 0.96$, $P < 0.01$; $R^2 = 92\%$; $Mse = 0.12$)

CONCLUSIONS

In this research two different nonlinear methods have been used in order to estimate effluent quality index for treated municipal wastewater samples. In the first part, a nonlinear formula was developed using Fuzzy Topsis weights and sub-indicing based on the observed data. In the second part, the applicability of feed-forward back propagation neural networks effluent quality index prediction was assessed. The EQI for 200 samples of treated wastewater is calculated through the statistical formula and sub-indicing was carried out. The network parameters to be tuned in designing the feed forward back propagation network were the number of hidden neurons and the connection weights. It could be concluded from the obtained results that a BPNN network with LM training algorithm, Tansig activation function and the architecture of 8:7:1 was the best for EQI prediction. The observed determination coefficient (R^2) for the BPNN was 0.92, and the mean square error was 0.1.

Findings from this study emphasize that the ANN enables easy modeling of the EQI and allows for identification of the comparative importance and contribution of input parameters to the model predictions.

REFERENCES

Abdoli, M., Falahnezhad, M., Salehi, R. & Behboudian, S. 2011 Long-term forecasting of solid waste generation by the artificial neural networks. *Environ. Prog. Sustain. Energy* **31** (4), 483–649.

Alfarra, A., Kemp-Benedict, E., Hötzl, H., Sader, N. & Sonneveld, B. 2011 A framework for wastewater reuse in Jordan: utilizing a modified wastewater reuse index. *Water Resour. Manage.* **25**, 1153–1167.

Aydiner, C., Demir, I. & Yildiz, E. 2005 Modeling of flux decline in crossflow microfiltration using neural networks: the case of phosphate removal. *J. Membr. Sci.* **248**, 53–62.

Banerjee, P., Singh, V. S., Chatttopadhyay, K., Chandra, P. C. & Singh, B. 2011 Artificial neural network model as a potential alternative for groundwater salinity forecasting. *J. Hydrol.* **398**, 212–220.

Behboudian, S., Tabesh, M., Falahnezhad, M. & Alavian, F. 2013 A long-term prediction of domestic water demand using preprocessing in artificial neural network. *J. Water Supply Res. Technol.* **63** (1), 31–42.

Boyacioglu, H. 2007 Development of a water quality index based on a European classification scheme. *Water SA* **33** (1), 101–106.

Dawson, C. W. & Wilby, R. L. 2001 Hydrological modelling using artificial neural networks. *Prog. Phys. Geog.* **25** (1), 80–108.

EPA 1992 *Manual guidelines for water reuse*. EPA/ 625/R-92/004. US Agency for International Development, Washington, DC.

Fischer, M. M. 2006 Neural networks: a general framework for non-linear function approximation. *Trans. GIS* **10** (4), 521–533.

Gazzaz, N. M., Yusoff, M., Aris, A., Juahir, H. & Ramli, M. 2012 Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar. Pollut. Bull.* **64**, 2409–2420.

Govindaraju, R. S. 2000 Artificial neural networks in hydrology. I, preliminary concepts. *J. Hydrol. Eng.* **5** (2), 115–123.

Güler, C., Thyne, G. D., McCray, J. E. & Turner, A. K. 2002 Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol. J.* **10** (4), 455–474.

Hwang, C.-L. & Yoon, K. 1981 *Multiple Attribute Decision Making: Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag, Berlin, New York.

Kashaninejad, M., Dehghani, A. A. & Kashiri, M. 2009 Modeling of wheat soaking using two artificial neural networks (MLP and RBF). *J. Food Eng.* **91**, 602–607.

Kaurish, F. W. & Younos, T. 2007 Development of a standardized water quality index for evaluating surface water quality. *J. Am. Water Res. Assoc.* **43**, 533–545.

Khalil, B., Ouarda, T. B. M. J. & St-Hilaire, A. 2011 Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis. *J. Hydrol.* **405** (3–4), 277–287.

Kowalkowski, T., Zbytniewski, R., Szpejna, J. & Buszewski, B. 2006 Application of chemometrics in river water classification. *Water Res.* **40** (4), 744–752.

Liou, S. M., Lo, S. L. & Wang, S. H. 2004 A generalized water quality index for Taiwan. *Environ. Monit. Assess.* **96** (1–3), 35–52.

Maier, H. R. & Dandy, G. C. 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environ. Model. Softw.* **15** (1), 101–124.

Najufpoor, Z., Vojodur, M. H., Dehgani, H. & Alidadi, H. 2007 Quality assessment of Kashif river in North-east of Iran in 1996–2005. *J. Appl. Sci.* **7** (2), 253–257.

Obaid, S. & Al-Shujairi, H. 2013 Develop and apply water quality index to evaluate water quality of Tigris and Euphrates rivers in Iraq. *Int. J. Mod. Eng. Res.* **3** (4), 2119–2126.

Olszewski, T., Ryniecki, A. & Boniecki, P. 2008 Neural network development for automatic identification of the endpoint of drying barley in bulk. *J. Res. Appl. Agric. Eng.* **53** (1), 26–31.

Özesmi, S. L., Tan, C. O. & Özesmi, U. 2006 Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecol. Model.* **195** (1–2), 83–93.

Palani, S., Liong, S. & Tkalich, P. 2008 An ANN application for water quality forecasting. *Mar. Pollut. Bull.* **56** (9), 1586–1597.

Palese, A., Pasquale, V., Celano, G., Figliuolo, G., Masi, S. & Xiloyannis, C. 2009 Irrigation of olive groves in Southern Italy with treated municipal wastewater: effects on microbiological quality of soil and fruits. *Agric. Ecosyst. Environ.* **129**, 43–51.

Riad, S., Mania, J., Bouchaou, L. & Najjar, Y. 2004 Rainfall-runoff model using an artificial neural network approach. *Math. Comp. Model.* **40** (7–8), 839–846.

Sapkal, R. S. & Valunjkar, S. S. 2013 Development and sensitivity analysis of water quality index for evaluation of surface water for drinking purpose. *Int. J. Civil Eng. Technol.* **4** (4), 119–134.

Sarkar, B., Sengupta, A., De, S. & DasGupta, S. 2009 Prediction of permeate flux during electric field enhanced cross-flow ultrafiltration – a neural network approach. *Separ. Purif. Technol.* **65**, 260–268.

Shamseldin, A. Y., Nasr, A. E. & O'Connor, K. M. 2002 Comparison of different forms of the multi-layer feed-forward neural network method used for river flow forecast combination. *Hydrol. Earth Syst. Sci.* **6** (4), 671–684.

Singh, K. P., Basant, A., Malik, A. & Jain, G. 2009 Artificial neural network modeling of the river water quality. A case study. *Ecol. Model.* **220** (6), 888–895.

Teschl, R. & Randeu, W. L. 2006 A neural network model for short term river flow prediction. *Nat. Hazards Earth Syst. Sci.* **6** (4), 629–635.

Verlicchi, P., Masotti, L. & Gallett, A. 2011 Wastewater polishing index: a tool for a rapid quality assessment of reclaimed wastewater. *Environ. Monit. Assess.* **173**, 267–277.

WHO 2005 *A regional overview of wastewater management and reuse in the Eastern Mediterranean Region*. World Health Organization (WHO), Regional Office for the Eastern Mediterranean, Cairo, Egypt.

Wu, C. L. & Chau, K. W. 2011 Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *J. Hydrol.* **399** (3–4), 394–409.

Wu, Z. & Chen, Y. H. 2007 The maximizing deviation method for group multiple attribute decision making under linguistic environment. *J. Fuzzy Sets Syst.* **158**, 1608–1617.

Zandbergen, P. A. & Hall, K. G. 1998 Analysis of the British Columbia water quality index for watershed managers: a case study of two small watersheds. *Water Qual. Res. J. Can.* **33** (4), 519–549.