Full Length Research Paper

Prediction of water quality parameter in Jajrood River basin: Application of multi layer perceptron (MLP) perceptron and radial basis function networks of artificial neural networks (ANNs)

M. Niroobakhsh¹*, S. H. Musavi-Jahromi², M. Manshouri¹ and H. Sedghi¹

¹Department of Water Sciences and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
²Faculty of Water Sciences Engineering, Shahid Chamran University, Ahwaz, Iran.

Accepted 6 June, 2012

River water quality is a significant concern in many countries, considering agricultural and drinking consumptions. Therefore, prediction of salinity index, as the main water quality condition is a necessary tool for water resources planning and management. This paper describes the application of artificial neural networks (ANNs) models for computing the total dissolved solids (TDS) level in Jajrood River (Iran). Two ANN networks, multi-layer perceptron (MLP) and radial basis function (RBF), were identified, validated and tested for the computation of TDS concentrations. Both networks employed five input water quality variables measured in river water over a period of 40 years. The performance of the ANN models was checked through the coefficient of determination ($R^2$) and root mean square error (RMSE). Jajrood River is one of the most important rivers which is located adjacent to Tehran city and supplies drinking water for people who live in this mega-city and recreational uses. Tehran is the most populous city and largest industrial pole in Iran, which caused the river, to be exposed to various pollutants. Matlab 2007 was selected for modeling goals in this research. Results show that MLP and RBF modeling as two methods of ANN are able to simulate water quality variables of Jajrood River with more than 90% accuracy. After modeling in MLP and RBF formatting and comparing simulation results (output) show that, the RBF result ($R^2$ of validation is 0.9362) are more closely to reality than the MLP ($R^2$ of validation is 0.8968). In other words, because of large number of input data, the RBF modeling performance has a better prediction than MLP modeling.

Key words: Water quality variables, artificial neural network (ANN), multi-layer perceptron (MLP), radial basis function (RBF).

INTRODUCTION

Rivers are the most important water resources on the globe. Knowing the quality of this resource is one of the major actions that must be done before planning for the water consumption. There are four characteristics to define water quality: chemical indicators, physical indicators, biological indicators, and microbial indicators. Each one includes many parameters. Depending on consumptions or needs, one or more of these parameters will be determined; on the other hand, there are standards of water consumptions. There are many methods used worldwide to determine water quality.

In fact, the related standards should dictate the water quality for different consumptions, such as, drinking, agricultural, industrial, and construction. The importance of “water use standard” indicates that more than 95% of water resources are used for drinking and agricultural needs (Hassani and Jalali, 2012).

Electrical conductivity (EC) and total dissolved solids (TDS) are two chemical parameters that show chemical water quality condition. In the necessity of the study using artificial neural network (ANN): there are several known
and unknown elements and mechanism which cause each water quality variable to be changed. However, there is almost no related model or algorithm that can help calculating water quality variables. Therefore, ANN can play the role of non-existent models and mechanisms. In this study, to survey ANN’s ability in water quality prediction using ANNs modeling is determined as the main target. For this, two main methods of ANNs, multi-layer perceptron (MLP) and radial basis function (RBF) were used to estimate EC and TDS concentration in Jajrood River basin.

**MATERIALS AND METHODS**

**Water quality filed data**

EC and TDS are the two most important water quality variables. Water quality condition is determined by measuring them. Both EC and TDS are dependent variables. TDS is depends on summation of anions and cations and a linear relation existed between TDS and EC. This relation is known as:

\[ EC = a \times TDS, \]  

(1)

where a is a constant factor and it is approximately equal to 0.64. Value of the constant depends on time and location.

**Input variables and data processing**

The monthly data of three water quality parameters was measured at the same time for over a period of 40 years at both two sampling sites which were selected for this analysis. The basic statistics of selected variables is presented in Table 1. The TDS and EC are two major parameters in water quality assessment. As mentioned, a linear relationship is established between EC and TDS. Based on existing measured values of different variables and their correlative analysis, five factors including EC, TDS and Q (Discharge) from upstream station and EC and Q from downstream station were selected for the model development.

Nowadays, expanding towns and villages and their populations on one hand, and industries and factories development on the other, cause the environmental pollution to be one of the more important issues. Population growth, rural development and overuse of natural and chemical fertilizers rich in nitrogen and phosphates and other nutrients also, including insignificance to governance and cheap land prices, the people established varied industrial centers in the upstream of Latian basin. Different effects of these actions results in environmental degradation and pollution on water resources of more than ten million dwellers of Tehran city (Tajrishi et al., 2002). The data set used in this study as shown in Table 1, was generated through continuous monitoring of the water quality of Jajrood River basin. Jajrood is one of the most important rivers, which supply drink-water for the city of Tehran, Iran.

**ANNs modeling**

ANNs have been used increasingly in recent years for the prediction and forecasting of complex hydrological relationships. ANNs have been seen as an attractive alternative to process based modeling approaches, as they are able to extract an underlying relationship from the data when knowledge of the physical process is lacking (Kingston et al., 2005). ANNs are flexible mathematical structures that are capable of identifying complex nonlinear relationships between input and output data sets. The ANNs, as the name implies, employs the model structure of a neural network which is a very powerful computational technique for modeling complex non-linear relationships, particularly, in situations where the explicit form of the relation between the variables involved is unknown (Gallant, 1993; Smith, 1994). It is increasingly being used to simulate and forecast quantitative characteristics of water bodies (Sundarambal et al., 2006).

In the words of Sarle (1994), users of ANNs "...want their networks to be black boxes requiring no human intervention data in, predictions out". More recently, researchers have examined ANN models from a statistical perspective (Cheng and Titterington, 1994; Hill et al., 1994; Ripley, 1994; Sarle, 1994; Warner and Misra, 1996; White, 1989). Such studies indicate that certain models are obtained when ANN geometry, connectivity and parameters are changed to either equivalent, or very close to, the existing statistical models.

The increasingly growing field of computational intelligence techniques has been proposed as an efficient tool in the modeling of dynamic phenomena (Ishmael et al., 2008). The primary objective of this paper was to compare the efficiency of two computational intelligence techniques in water demand forecasting. The techniques under comparison are ANNs and support vector machines (SVMs). In this study, it was observed that ANNs perform significantly better than SVMs. This performance is measured against the generalization ability of the two techniques in water demand prediction.

An artificial neural consists of three components, including weights (W), bias (b), and transfer function (f). These three components are unique for each neural. Figure 1 shows the schematic of artificial neural. In the figure, p and a, are input and output of a neural, respectively. Parameter n is called net input, which is the input of transfer function and it is built according to input p and neural parameters (Kanani et al., 2008). Mentioned artificial neural can be modeled by the following equations.

\[ n = wp + b \]  

(2)
Figure 1. Schematic diagram of artificial neural.

Figure 2. Neural network structures.

\[ a = f(n) = f(wp + b) \]  

In neural instruction process, \( W \) and \( b \) change until the best approximation for an output member corresponding to the input member is obtained. Weight of neural determines the rate of \( p \) effect on "a" and parameter "b" causes neural to be transformed to sub-space of bias input space. There are some types of transfer functions; some of which are as follows:

1) Linear, transfer function;
2) Hard-limit transfer function;
3) Log-Sigmoid transfer function;
4) Tan-Sigmoid transfer function; and
5) Tan-Hyperbolic transfer function.

The signal passing through the neuron is modified by weights and transfer functions. This process is repeated frequently until the output layer to be achieved is achieved (Govindaraju, 2000).

A neural network must be trained to determine the values of the weights that will produce the right outputs. In a training process, a set of input data is used for training and presented to the network many times. The performance of the network is tested after the training step is stopped. The back propagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient). This is the direction in which the performance function is decreasing most rapidly. It turns out, although, the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. Therefore, the basic gradient descent training algorithm is inefficient owing to its slow convergent speed and at times the poor accuracy in model predictions (Huang et al., 2004).

The basic structure of an ANN model usually comprised three distinctive layers; the input layer, where the data are introduced to the model and computation of the weighted sum of the input is performed; the hidden layer or layers, where data are processed; and the output layer, where the results of ANN are produced. Each layer consists of one or more basic element(s) called a neuron or a node. A neuron is a non-linear algebraic function, parameterized with boundary values (Dreyfus et al., 2002).

The number of neurons in the input, hidden, and output layers depends on the problem. If the number of hidden neurons is small, the network may not have sufficient degrees of freedom to learn the process correctly. On the other hand, if the number is too high, the training will take a longer time and the network may over-fit the data (Karunanithi et al., 1994). Neural Networks consist of many patterns as shown in Figure 2.

**MLP network**

Among many neural network architectures, the three-layer-feed forward back propagation network [one kind of MLP] is the most commonly used (Haykin, 1999). This network architecture consists of one hidden layer of neurons with nonlinear transfer functions and an output layer of linear neurons with linear transfer functions.

Considering a network with one hidden layer (Figure 3), the processing of a single neuron is broken into two steps, that is, the weighted sum of the inputs followed by the activation function. For example, consider a neuron in the hidden layer that receives inputs from neurons in the input layer. The net input, \( y_{in} \), to the hidden neuron is the sum of the weighted signals from the input neurons (that is: \( y_{in} = \sum w_i x_i \)). The activation \( y_\) of this hidden neuron is then given by some function of its net input, \( y = f \)
The most common activation function and the one implemented in this study, is a sigmoid function which is described as follows:

$$f(x) = \frac{1}{1 + e^{-x}}$$  \hspace{1cm} (4)

This procedure is repeated for each input vector and at the completion of a pass through the entire data set, all the nodes change their weights based on the accumulated derivatives of the error with respect to each weight and these changes move the weights in the direction in which the error declines most quickly.

If we let $w_m$ represent the value after iteration $m$ of a weight $w$, which may be either a hidden-node weight $w_{ij}$ or an output node weight $w_{jk}$ then:

$$w_m = w_{m-1} + \Delta w_m$$ \hspace{1cm} (5)

where $\Delta w_m$ is the change in the weight $w$ at the end of iteration $m$ and is calculated as:

$$\Delta w_m = -\varepsilon d_m$$ \hspace{1cm} (6)

where $\varepsilon$ is the parameter controlling the proportion by which the weights are changed. The user sets the value of this parameter and the term $d_m$ is given by:

$$d_m = \sum_{n=1}^{N} \left( \frac{\partial E}{\partial w_m} \right)_n$$ \hspace{1cm} (7)

where $N$ is the total number of examples, and $E$ is the simulation error.

### RBF network

A radial basis function network has a feed-forward structure consisting of a single hidden layer for a given number of locally tuned units which are fully interconnected to an output layer of linear units (Dibike et al., 1999; Mason et al., 1996). The mapping function of a radial basis function network, as schematized in Figure 4 is mostly built up of Gaussians rather than sigmoid as in MLP networks. Learning in RBF network is carried out in two phases: first, for the hidden layer, and then for the output layer. The hidden layer is self-organizing; its parameters depend on the distribution of the inputs, not on the mapping from the input to the output. The output layer, on the other hand, uses supervised learning (gradient decent or linear regression) to set its parameters.

A RBF hidden unit has one parameter associated with each input unit. These parameters $w_i$ are not weights placed on the input; rather they are the co-ordinates in input space of a point, that is, the centre of the hidden units output function. The output from the hidden layer is a function of the radial distance $d_j$ between the datum vector $X = (x_1, x_2, \ldots, x_n)$ and the ‘radial centre’ $W_j = …
(w_1j, w_2j, ..., w_mj) and may be written as:

$$
\delta_j = \sqrt{\sum_{i=1}^{k} x_i - w_{ij}}
$$

(8)

and we can then describe the output for the hidden layer \( y = f(d_j) \). There are various choices of \( f(d) \), and the most popular form is that of Gaussian described as follows (Mason et al., 1996):

$$
f(d_j) = e^{-\lambda d_j^2}
$$

(9)

where \( \lambda \) is a constant.

Objective functions, which are used for learning phase are: root of mean square error (RMSE), mean of absolute error (MAE), and coefficient of determination (R^2). The hidden units thus send to the output units, values that indicate how far the example is from each of them. Each output unit has a parameter for each hidden unit and calculates the output as:

$$
Z_k = \frac{\sum_{j=1}^{l} b_{jk}y_j}{\sum_{j=1}^{l} y_j}
$$

(10)

where \( b_{jk} \) is the parameter on the connection from hidden node \( j \) to output node \( k \) and \( y_j \) is the output of hidden node \( j \). Learning, finding the values of the parameters is carried out in two phases, first for the hidden layer, then for the output layer. In general, RBF networks that employ clustering for locating hidden unit receptive field centers can achieve a performance comparable with back propagation networks, while requiring order of magnitudes less training time. However, the RBF network requires more data to achieve the same accuracy of back propagation networks (Hassoun, 1995).

In the present study, using MATLAB of the aforementioned concepts, a computing code has been developed to modeling process.

**Study area**

Jajrood River basin, at Latiyan Dam site lies within latitudes 51° 22’ N to 51° 51’ N, and longitudes 34° 45’ E to 36° 5’ E (Figure 5). Latin dam is located in Northeast of Tehran. It’s area is about 710 km². Jajrood River is the most important river of this basin, in addition to the role of dam water supply and eventually drinking water for the people living in Tehran, this river has a recreational use for the people who live in this area. Jajrood River is located near the city of Tehran, which is one of the most populous cities in the world and is
Table 2. Performance parameters of the MLP model for computation of the TDS in Jajrood River.

<table>
<thead>
<tr>
<th>Model</th>
<th>ANN-structure</th>
<th>R²</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>0.9070</td>
<td>0.0555</td>
</tr>
<tr>
<td>11-7-1</td>
<td>Test</td>
<td>0.8958</td>
<td>0.0587</td>
</tr>
<tr>
<td>6-7-1</td>
<td>Training</td>
<td>0.8181</td>
<td>0.0753</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>0.7717</td>
<td>0.0525</td>
</tr>
<tr>
<td>10-1</td>
<td>Training</td>
<td>0.8005</td>
<td>0.0780</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>0.8409</td>
<td>0.0828</td>
</tr>
<tr>
<td>3-4-1</td>
<td>Training</td>
<td>0.6790</td>
<td>0.0909</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>0.7127</td>
<td>0.0484</td>
</tr>
</tbody>
</table>

Figure 6. Comparison of the model computed and measured TDS values in the river water training sets, using MLP model.

In this study, ANNs were identified to predict TDS concentration, as one of the water quality variables. For ANN identification, the complete river water quality data set was divided into two sub-sets. The calibration (or training) and validation (or testing) of data subsets comprised many division such as (60 and 40%), (65 and 35%), (70 and 30%), etc., on each sample, respectively. Result of TDS modeling is as shown in Table 2 and Figures 6 to 9).

In this study, ANNs were identified to predict TDS concentration, as one of the water quality variables. For ANN identification, the complete river water quality data set was divided into two sub-sets. The calibration (or training) and validation (or testing) of data subsets comprised many division such as (60 and 40%), (65 and 35%), (70 and 30%), etc., on each sample, respectively. Result of TDS modeling is as shown in Table 2 and Figures 6 to 9).

In this study, ANNs were identified to predict TDS concentration, as one of the water quality variables. For ANN identification, the complete river water quality data set was divided into two sub-sets. The calibration (or training) and validation (or testing) of data subsets comprised many division such as (60 and 40%), (65 and 35%), (70 and 30%), etc., on each sample, respectively. Result of TDS modeling is as shown in Table 2 and Figures 6 to 9).
was run. For each structure, two popular indexes for validate result of models were used; coefficient of correlation ($R^2$) and RMSE. After more run models away, the best structure with the highest in $R^2$ and lowest in RMSE was found. In this case, a three layer perceptron, one kind of MLP, with the more than 0.90 and 0.89 $R^2$ and RMSE were 0.055 and 0.058, respectively train and test.
Table 3. Performance parameters of the RBF model for computation of the TDS in Jajrood River.

<table>
<thead>
<tr>
<th>Model</th>
<th>( R^2 )</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Test</td>
</tr>
<tr>
<td>RBF</td>
<td>0.8248</td>
<td>0.0572</td>
</tr>
<tr>
<td></td>
<td>0.9032</td>
<td>0.0518</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.0408</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.0402</td>
</tr>
</tbody>
</table>

Figure 10. Comparison of the RBF and MLP models, computed TDS values in the river water.

The appropriate program of the best MLP and RBF models for the prediction of TDS in the river water was developed and is as shown in Figures 6 to 9. The selected ANN for the TDS model composed one input layer with five input variables, two hidden layers with eighteen neurons and one output layer with one output variable. The constructed ANN models (TDS) were trained using the scaled conjugate gradient (SCG).

According to RBF’s structure and its calculation algorithm shows that the radial of clusters is important and affective. So, as shown in Table 3, the best performance of RBF modeling is achieved by changing this parameter. According to RBF principles, there is no structure like as MLP network. As told in this case, only one hidden layer exists. This property helps network to models nature more closely and fast. Value of \( R^2 \) coefficient and RMSE, confirm it. Although, the achieved result of RBF modeling is greater adaptive with reality, but ability of MLP modeling in river water quality should not be ignored in Jajrood River basin. The ability of both MLP and RBF modeling to predict TDS in Jajrood River is as shown in Figures 6 to 9.

As shown in Figure 6, the MLP is one of the good adaptable networks to nature. In this figure, coordination and fit proper of the model simulation with nature is very clear, but Figure 7 shows the model results, and there is no good coordination between nature and model result values. But the situation in Figures 8 and 9 is in different way. In the beginning of modeling process, a brief confusing can be seen (Figure 8), that is, finding a good harmony in the end of process. This is related to the absence of multi layer and neurons, basic principle of RBF. In the last figure, good ability of RBF network to simulate the TDS concentration can be seen. However, the RBF at the beginning was poorer than MLP, but at the end of modeling process and output results, which is the most important part of modeling, it worked best.

To clarify the difference between MLP and RBF methods in water quality predictions, the results were made by comparing them with themselves as shown in Figure 10.

Conclusion

In this paper, two models of ANNs were identified for
computation of the TDS concentrations in the water of Jajrood River (Iran). The identified models were trained and tested on monthly data sets of TDS measured over a period of 40 years.

The feed-forward network with back propagation learning algorithm was employed. The present study shows that the optimal networks are capable to capture long-term trends observed for the river water quality variable in time and space. We propose the neural networks as effective tool for the computation of river water quality and it could also be used in other areas to improve the understanding of river pollution trends. The ANN can be seen to be a powerful predictive alternative to traditional modeling techniques. Studying Figures 5 to 8 and their relative result table modeling in MLP and RBF indicate that RBF modeling was closer to natural system.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Hrid Abghari at Urmeyer University, Iran for his valuable comments on the article.

REFERENCES


