International Journal of ChemTech Research
CODEN (USA): IJCRGG ISSN : 0974-4290
Vol.6, No.5, pp 3053-3061, Aug-Sept 2014

Modeling of Irrigation Water Quality using Multilayer Perceptron Back Propagation Neural Network (MLBP-NN)

SaralaThambavani D[1]*, Uma Mageswari T.S.R[2]

[1]Sri Meenakshi Government Arts College for Women (Autonomous), Madurai, Tamilnadu, India.
Research and Development Centre, Bharathiar University, Coimbatore, India.
[2]PSNA College of Engineering & Technology, Dindigul, Tamilnadu, India.

*Corres. Author: umamageswaritsr1980@gmail.com, Contact No: + 91 9487136997

Abstract: Irrigation water quality is one of the main yield factors in the cultivation of agricultural and horticultural crops in arid and semi-arid areas. In past two decades, irrigation water quality and quantity problems are increasing severely because of improper management and industrialization. The main aim of this study is to describe the applicability of artificial neural network that can effectively predict the quality of irrigation water. The study was conducted in Batlagundu, Nilakottai Taluk, Tamil Nadu and about 150 samples were collected. Irrigation samples were analyzed for Physico-chemical properties, various cationic and anionic constituents outlined. Data obtained from chemical analysis were used in the ANN model to predict pH, TDS and SAR. The results of this study proved that MLPBP-NN is effectively predicting irrigation water of the study area.
Key words: pH, SAR,TDS,MLPBP-NN.

Introduction

In arid and semi-arid regions groundwater is the major source for domestic purposes and irrigation. Irrigation water quality has a significant role in crop production and has a profound impact on physical and chemical soil properties. Monitoring of water quality is one of the important tools for sustainable development and provides important information for water management[1]. Due to the correlations and interactions between water quality variables such as anions and cations concentrations, it is interesting to investigate whether a domain-specific mechanism governing observed patterns exists to prove the predictability of these variables. The identification of such forecast models is particularly useful for ecologists and environmentalists, since they will be able to predict water pollution levels and take necessary precaution measures in advance[2]. Variation in groundwater quality is a function of physical and chemical parameters that are greatly influenced by geological formations and anthropogenic activities as well [3].

Management and periodical assessment of ground water quality status is essential to maintain the potential property of land for food production[4]. Generally, quality of irrigation water are governed by factors such as pH, sodium adsorption ratio(SAR), Soluble sodium percentage (SSP),the residual sodium carbonate (RSC), Total dissolved solids (TDS) and Electrical conductivity (EC).Among them pH, TDS and SARare used to determine the desirability of irrigation water.

The artificial neural network (ANN), as its name implies, is a technique for the human brain's problem–solving process. Just as humans apply knowledge gained from experience to new problems or situations, the structure of a neural network can be applied to the powerful computation of complex nonlinear relationships. The ANN is used as an approximation tool rather than a complex mathematical calculation, which results in a ten percent deviation of predicted value from observed data[5]. ANN, is a biological inspired computing methodology that have the ability to learn by imitating the learning method used in the human brain, don't accompany any of the above drawback of conventional methods, physical and statistical. ANNs, especially back propagation network are closely related to statistical methods and are most suitable forpredicting ground water quality applications.

ANN models are usually employed to predict or to optimize the values of qualitative parameters. ANNs are well suited to complex problem as they belong to class of data driven approaches. ANNs are relatively in sensitive to data noise, as they have ability to determine the underlying relationship between input and output resulting in good generalization ability. There are a number of studies in which neural networks are applied to water quality problems. The most common types of ANN used in ecology are supervised multilayer perceptron neural networks with a back propagation learning algorithm[6]. Chau[7] reviewed the development and current progress of the integration of artificial intelligence into water quality modelling. Hatzikos[8] utilized neural networks with active neurons as a modelling tool for the prediction of seawater quality indicators like water temperature, pH, dissolved oxygen (DO) and turbidity. Palani[2] demonstrated the application of ANNs to model the values of selected seawater quality variables, having the dynamic and complex processes hidden in the monitored data itself[9].

No study has been conducted to use ANN for predicting pH, TDS and SAR of irrigation water of Batlagundu of Nilakkotai Taluk. In this sense, this study aimed to predict the pH, TDS and SAR of irrigation water using MLP-BP ANN network.

Materials and Methods

Study Area

The study area is Batlagundu town panchayat in Dindigul district in the state of Tamilnadu. It is located 450km south of state capital Chennai and situated at the foot hills of the Kodaikanal mountain range. It has an average elevation of 320 meters (1049 feet). Batlagundu is also known as "BETEL CITY". This area is endowed with 25 medium scale industries and 2000 small scale industries. The economy of the town is mostly dependent on agricultural products like betel leaf, and is a home to banana leaf commission, coconut powder exports and spinning mills.Batlagundu is geographically located at Longitude and Latitude is 77 ^{0} 45' 33.84" E and 10 ^{0} 9' 55.80" N. As per census 2001, Batlagundu had population of 22,007.Average temperature and humidity is 22^{0} C and 86% respectively. In order to achieve the research objective, samples were collected from 18 sample points on a monthly basis from 2012 to 2013.The water samples were analyzed for pH, EC,TDS and various cationic and anionic concentrations as per standard procedure [10,11,12] The map of the study area given in Fig.1.

Fig.1 Map of the study area

Artificial Neural Network

Artificial Neural Network (ANN) is a simulation of the real nervous system, in other words, is a mathematical model based in biological neural networks. It is a system that contains a collection of units “neurons” communicating with each other in a network that works to produce an output stimulus. ANNs are inspired by the activity of human brain. The key is the creation of neural networks (The system structure), which is compose of a large number of highly interconnected basic units (neurons) in layers that work together to solve specific problems. ANNs can be configured for specific applications, such as pattern recognition or data classification using learning process. This learning process as well as the biological system provides adjustments of developed models. The use of Artificial intelligence technologies, specifically Artificial Neural Networks (ANN), is increasing in the drinking water treatment industry as they allow for the development of robust nonlinear models of complex units processes, improving the drinking water quality, and at the same time, reducing operating costs with the advanced process control.

Structure of an ANN

The basic units of a biological neural system are neurons, which are grouped into sets, consisting of millions of them organized in layers and constitute a system with own functionality. A set of these subsystems create a global system. In Fig. 2 it can be seen an ANN as a collection of parallel processors connected in the form of a guided or directed graph, organized as the network structure itself leading us to consider it as a feature to keep in mind when creating an ANN. It can be representedas, each item (unit) that process the information of the network, as a node with connections between units represented by arrows, these arrows also indicate the direction in which information flows.

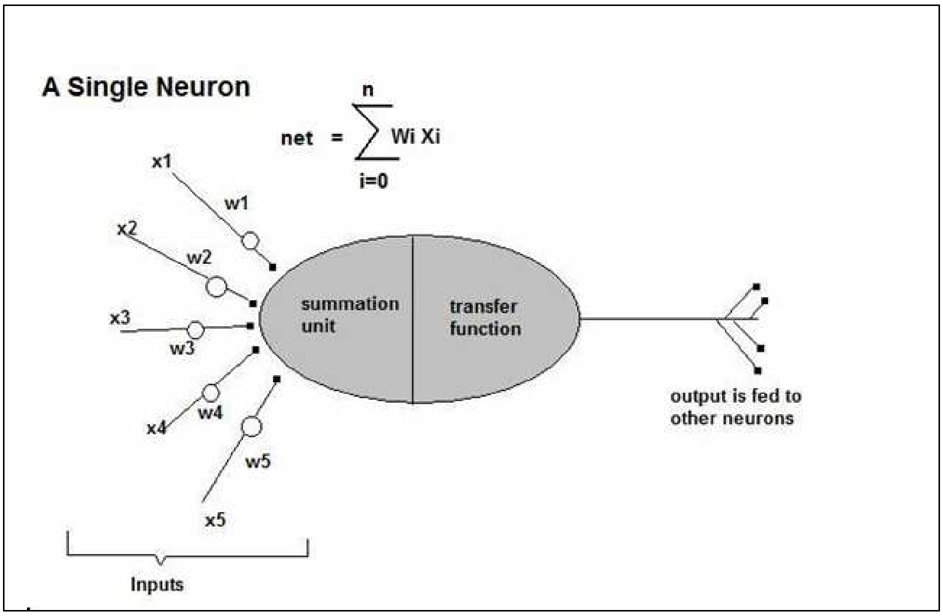


Fig. 2 Basic elements of an artificial neuron

Components of ANN

In this section, most of the components in a neural networkis described. These components are valid even if the neuron is used like input, output or hidden layer.

Inputs

X_j are inputs to the neuron and the appropriate selection of these variables or inputs in a group of potential measurement, to the system under investigation, is a vital step in model development. This is particularly important in data driven techniques, such as artificial neural networks and fuzzy systems, as the performance of the final model is heavily dependent on the input variables used to develop the model.

Weight

Typically a neuron receives many simultaneous and multiple inputs. Each input has its own relative weight which gives the importance of the input within the activation function of the neuron. These weights do the same role performed by the biological neurons in synaptic. In both cases, some inputs are more important than others so they have more effect on the processing of the neuron combined to produce a neuronal answer. The weights are coefficients that can be adapted within the network to determine the intensity of the input signal, received by the artificial neuron. They are the measure of the strength of an input connection. These forces can be modified in response to the training examples according to the specific topology or because of the training rules.

Summing Part

This rule provides from the inputs and weights the potential postsynaptic value h_i of the neuron.

$$\mathbf{h_i(t) = \sigma_i\left(w_{ij} * x_i\right)} \text{ - - - - - - - - - - - - - - } \mathbf{(1)}$$

The most common function is the sum of all weights and inputs, by grouping the inputs and weights in two vectors (x_1, x_2, ...,x_n) and (w_{1J}, w_{2j}, ..., w_{nj}) and them calculate this amount making the scalar product of two vectors.

$$\mathbf{h_i(t) = \sum_j w_{ij} * x_i} \text{ - - - - - - - - - - - - - - } \mathbf{(2)}$$

The role of this summing part may be more complex than a simply sum of products. The inputs and weights can be combined in different ways before passing the value to the activation function. The specific algorithm for the propagation of neural inputs is determined by the choice of architecture.

Activation Or Transfer Function

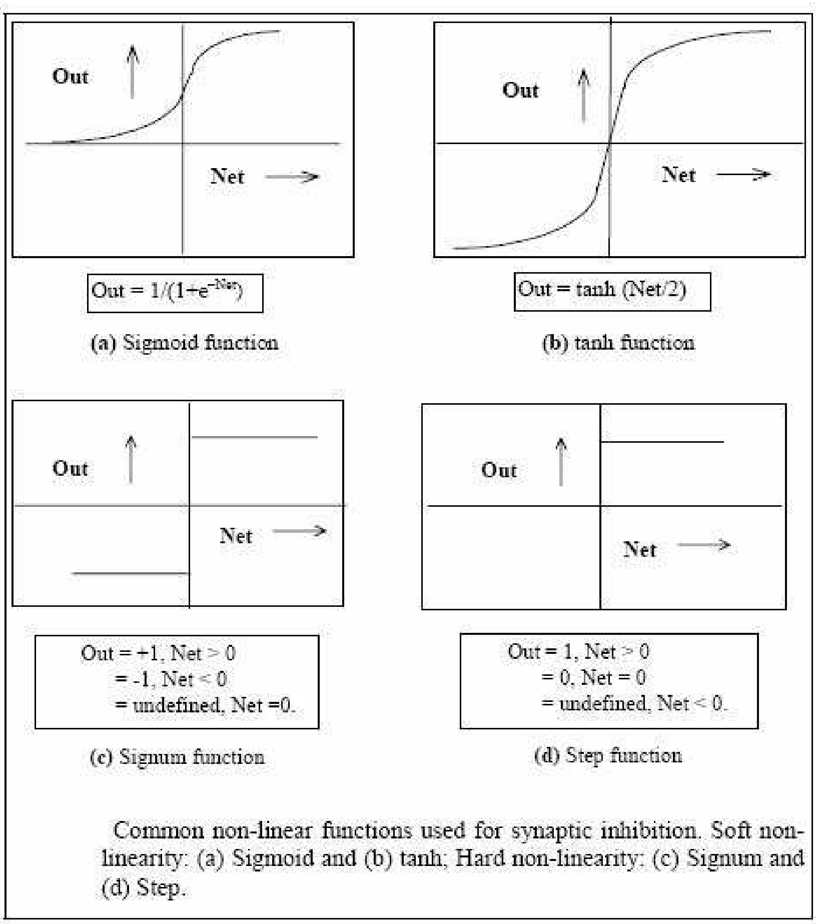
The result of the summing part in most cases is a weighed sum, which is transformed into the actual output of the neuron through an algorithmic process known as activation function.

$$\mathbf{a_i(t) = f_i(a_i(t-1),\ h_i(t))} \text{-------------------------} \mathbf{(3)}$$

In this case the activation function depends on the postsynaptic potential h_i(t) and its previous state of activation. However, in many models of ANN is considered that the current state of the neuron does not depend on its previous state a_i(t-1), but only the current.

$$\mathbf{a_i(t) = f_i(h_i(t))} \text{---} \mathbf{(4)}$$

In the Activation function, the value of the output combination can be compared with a threshold value for determining the output of the neuron. If the sum is greater than the threshold value, a neuron signal is generated. If the sum is less than the threshold, no signal is generated. Usually the threshold value, or transfer function value is typically nonlinear. The use of linear functions is limited since the value of the output is proportional to the input; in fact this was one of the problems in the early models of artificial neural networks in Perceptrons. The activation function could be something as simple as it only depend on whether the result of the combination function is positive or negative. Some transfer or activation function can be seen in Fig.3.



Common non-linear functions used for synaptic inhibition. Soft non-linearity: (a) Sigmoid and (b) tanh; Hard non-linearity: (c) Signum and (d) Step.

Fig. 3 Examples of activation functions

From the functions presented in Fig.3, stand out the sigmoid function. From a mathematical point of view, the usefulness of these functions is that the function and its derivative are continuous. These functions work quite well and are usually elected. There are other activation functions that are specific to some architectures. Before applying the activation function, we can add some noise to the inputs. The source and amount of this noise is determined by the training of a particular network. This noise is commonly known as temperature of the neuron. In fact by adding different noise levels to the result of the combination or summing, leads to create a model more similar to the brain. The use of the noise by temperature is still under investigation and is not usually applied in the praxis.

Neural Network Architectures

Architecture is called to the topology, structure or connection pattern of a neural network. In an ANN, nodes are connected by synapses, this structure of synaptic connections determines the behavior of the network. In general, neurons are usually grouped into structural units that are called layers and finally, the set of one or more layers is the neural network. There are three types of layers:

• Input: An input layer or sensory layer consists of neurons that receive data or signals from the environment.

• Hidden: It is the one that has no direct connection with the contour, i.e. is not directly connected, no body sensors nor effectors.

• Output: It is the layer in which the neurons provide the response of the neural network.

The connections between neurons can be excitatory or inhibitory: a synaptic weight defines a negative inhibitory connection, while a positive determines excitatory connection.Intra-layer connections, also called side connections, take place between neurons in one single layer, while the inter-layer connections occur between neurons in different layers. There are also feedback connections that have an opposite way input - output. Based on these concepts, different neural architectures can set:

- Single- layer Networks are those composed by just one layer of neurons.
- Multi-layer Networks (layered networks) are those whose neurons are organized in several layers.

  In response to the data flow in a neural network,
- Feed-forward Networks which circulates the information in one direction from the input neurons to the output.
- Feedback Networks circulates the information between the layers in any direction.

Results and Discussion

The type of ANN model used was the well known Multilayer Perceptron (MLP). The MLP is a feed forward ANN model that maps the sets of input data onto a set of appropriate output. A MLP consist of multiple layers of nodes in a directed graph, which is fully connected from one layer to the next. Each node in the hidden layer and the output layer uses a nonlinear activation function in MLP and utilizes the supervised learning technique "back propagation" to train the network. For the model calibration, the data set was treated using the following analysis: the data set was divided into three subsets, The first subset was used to train the network (Learning phase), the second part was used to test the ANN models in order to determine when to stop the training stage (Testing phase) and the last part was used to validate the model data not involved in the training process (Validation phase).

The software SPSS 15 is used to analyze and examine the relation among the parameters, forming a correlation between them. The correlation coefficients thus appearing brings out the relative susceptibility of each parameter. Finally, we will turn to an examination of the recommended models with the designed neural networks to predict the pH, TDS and SAR. To predict pH of irrigation water, dissolved oxygen, electrical conductivity, total dissolved solids, total alkalinity, total hardness, calcium and magnesium are chosen as the input parameters. Out of 150 samples, 70% of the samples were used to train ANN model and 15% of the samples were used to validate the ANN model. Remaining 15% of the samples were used to test the ANN model. To find out the best ANN architecture, the number of nodes in the input layer and output layer was fixed at 7 and 1 respectively. A total of 3 networks were selected by trial and error method. Out of this network 7-3-1 network architecture is given high correlation coefficient (r=0.93) and low mean squared error (0.0038). A scatter diagram of actual and predicted value of pH using multilayer perceptron artificial neural network is given in Fig.4.

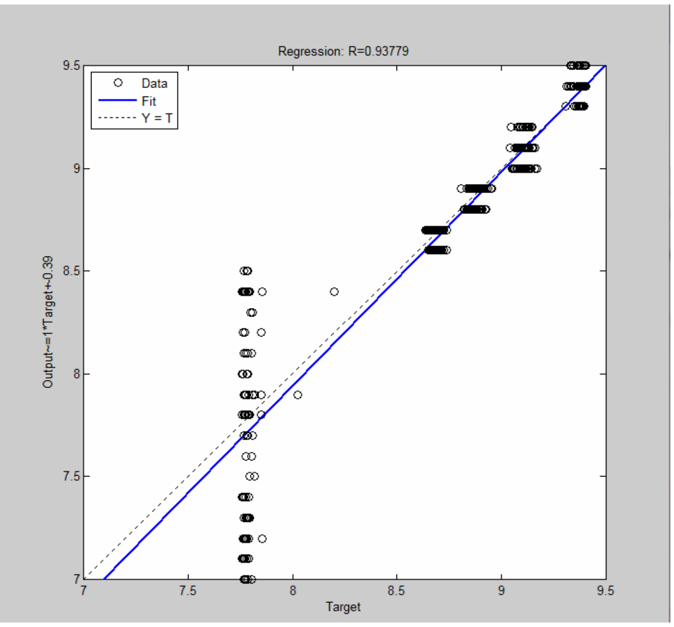


Fig. 4. A scatter diagram of actual and predicted value of pH using MLP-NN

As mentioned before, 7 parameters were used to predict ground water TDS concentration using different ANN models. The parameters used for the model is pH, dissolved oxygen, electrical conductivity, total alkalinity, total hardness, calcium and magnesium. The best ANN architecture was chosen and evaluated among various architectures with different number of neurons in the hidden layer. Table 1 shows that best architecture to predict TDS is 7-3-1.(r=0.97 and MSE=0.1352). Fig.5 represents the distribution diagram of the TDS predicted and measured values using Feed Forward Back Propagation Neural Network (FFBPNN).Since the comparison of ANN results with measured values of TDS, artificial neural network is showing high accuracy. Therefore, it can be stated that the performance of an ANN is suitable for predicting output parameter.

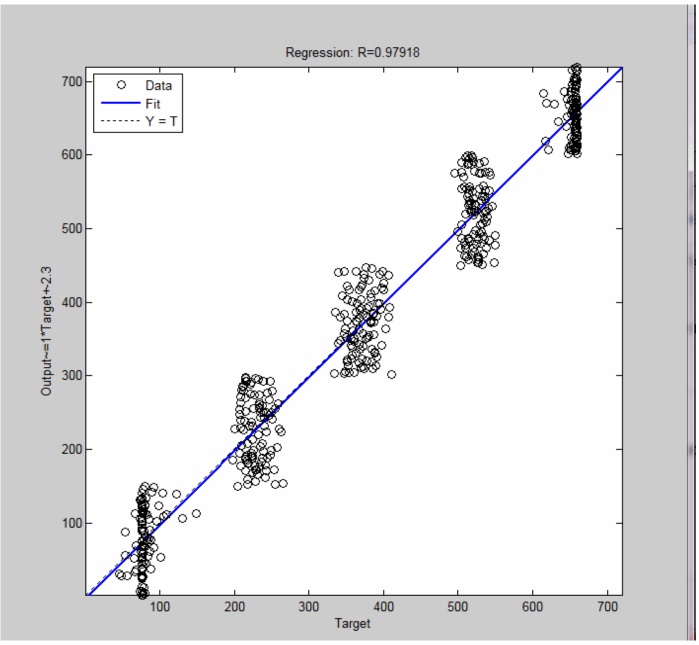


Fig. 5 A scatter diagram of actual and predicted value of TDS using MLP-NN

To predict SAR (Sodium Adsorption Ratio) using artificial neural network, input parameters introduced to the network, including pH, electrical conductivity, total alkalinity, total hardness, calcium and magnesium. In this study, three different MLP-ANN architectures were used to perform the best performance. Levenberg Marquadt algorithm is used for the training procedure. The best network architecture was found to be 7-2-1. The correlation coefficient and MSE was found to be 0.99 and 0.00281 respectively. Fig.6 shows the graph of measured and predicted values of SAR by ANN. The training, testing and validation results for each of these models are given in Table 6.

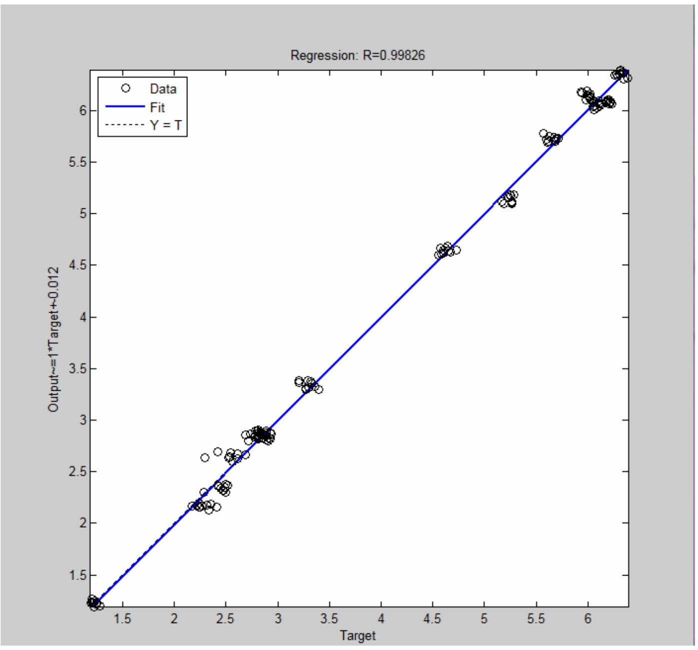


Fig. 6 A scatter diagram of actual and predicted value of SAR using MLP-NN

Table 1. Correlation coefficient and MSE for Datasets with different architecture

S.NO	Parameters	ANN architecture	Data set	Correlation Coefficient(r)	Mean square error(MSE)
1	pH	7-1-1	Training	0.88688	0.086
			Validation	0.89804	0.070
			Testing	0.89332	0.077
			All	0.88734	0.092
		7-2-1	Training	0.92387	0.0045
			Validation	0.92853	0.0050
			Testing	0.92236	0.0049
			All	0.92102	0.0051
		7-3-1	**Training**	**0.9411**	**0.0026**
			Validation	**0.93697**	**0.0033**
			Testing	**0.94666**	**0.0029**
			All	**0.93779**	**0.0038**
2.	TDS	7-1-1	Training	0.97758	0.1953
			Validation	0.97383	0.1978
			Testing	0.9806	0.1867
			All	0.97755	0.1958
		7-2-1	Training	0.97692	0.1897
			Validation	0.97429	0.1857
			Testing	0.98312	0.1765
			All	0.97771	0.1873
		7-3-1	**Training**	**0.9785**	**0.1345**
			Validation	**0.98283**	**0.1287**
			Testing	**0.97854**	**0.1376**
			All	**0.97918**	**0.1352**
3.	SAR	7-1-1	Training	0.98067	0.0111
			Validation	0.98199	0.0112
			Testing	0.97726	0.0176
			All	0.97971	0.0113
		7-2-1	**Training**	**0.99837**	**0.00284**
			Validation	**0.99877**	**0.00234**
			Testing	**0.99733**	**0.00267**
			All	**0.99826**	**0.00281**
		7-3-1	Training	0.97644	0.0132
			Validation	0.99143	0.0129
			Testing	0.96771	0.0141
			All	0.9760	0.0137

Conclusion

The irrigation water quality of Batlagundu was evaluated by predicting pH,TDS and SAR by Artificial Neural Network. Out of 150 samples, 70% of the samples were used to train ANN model and 15% of the samples were used to validate the ANN model. Remaining 15% of the samples were used to test the ANN model. Performance of ANN models were tested by using correlation coefficient and MSE. pH shows high correlation coefficient (r=0.93) and low mean squared error (0.0038) for the 7-3-1 network architecture. The best architecture to predict TDS is 7-3-1.(r=0.97 and MSE=0.1352). The correlation coefficient and MSE was found to be 0.99 and 0.00281 respectively for SAR and suitable architecture is 7-2-1. Hence, with the proposed model applications, it ispossible to manage irrigation water resources invery effective manner.

References

1. Jalali M.,Groundwater geochemistry in the Alisadr, Hamadan, Western Iran. Environ. Monit. Assess.2009,166, 359-369.
2. Palani S. Liong S.Y. and Tkalich P., An ANN applicationfor water quality forecasting. Marine Poll. Bull., 2008, 56, 1586-1597.
3. Yesilnacar M.I. Sahinkaya E. Naz M. and OzkayaB.,.Neural network prediction of nitrate in groundwater of Harran Plain, Turkey. Environ Geol., 2008, 56, 19-25.
4. Durdu, O.F. A hybrid neural network and ARIMA model for water quality time series prediction, Engineering Applications of Artificial Intelligence.2010, 24: 586-594.
5. Lingireddy S. and Ormsbee L.E., Neural Networks in Optimal Calibration of Water Distribution Systems, Artificial Neural Networks for Civil Engineering: Advanced Features and Applications (Eds I. Flood, R. Kartam). ASCE Press, New York, USA.1998.
6. Maier H.R. and Dandy G.C. Neural networks for the prediction and forecasting of water resource variables: a review of modelling issues and applications. Environmental Modelling & Software, 2000, 15,101– 124.
7. Chau K.W.,Areview on integration of artificial intelligence into water quality modeling. Marine Poll. Bull., 2006, 52, 726-733.
8. HatzikosE.,AnastasakisL.,BassiliadesN., andVlahavas I., Simultaneous prediction of multiple chemical parameters of riverwater qualitywith tide. Proc. 2nd Int. Sci.Conf. Computer Sci., IEEE Computer Society,May11-13, Varna, Bulgaria.2005
9. Faruk D., A hybrid neural network and ARIMA model for water quality time series prediction. Eng. Appl. Artificial Intell., 2010, 23(4), 586-594.
10. APHA, AWWA and WPCF, *Standard Methods for the Examination of Water and Wastewater,* American Public Health Association, Washington D.C., 14 ,1995.
11. Trivedy D.K. and Goel P.K., Chemical and biological methods for water pollution studies. Environment publication. Karad , India. 1984
12. NEERI, Manual on water and waste water Analysis, National Environment Engineering Research Institute Nagpur, 3402, 1986.

*****