ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 6, Issue 10, April 2017

Assessment of Total Dissolved Solid Concentration in Groundwater of Nadia District, West Bengal, India using Artificial Neural Network

Alivia Chowdhury[1]*, Ankita Ghosh[2] and Debaditya Gupta[3]
[1]. Assistant Professor, Department of Soil and Water Engineering, Faculty of Agricultural Engineering, Bidhan Chandra Krishi Viswavidyalaya (BCKV)
[2].Phd Student, Department of Soil and Water Engineering, Faculty of Agricultural Engineering, Bidhan Chandra Krishi Viswavidyalaya (BCKV)
[3]. M.Tech Student, Dept. of Agricultural and Food Engineering, IIT, Kharagpur

Abstract:* An attempt was made to predict the total dissolved solid (TDS) concentration of groundwater for Nadia district, West Bengal using artificial neural network approach. The sole aim of the study was simulating TDS in groundwater as one of the major indicators of groundwater quality using artificial neural network. Five different multilayer perceptron (MLP) neural network structures with different combination of input parameters were investigated in the present study to predict TDS in the study area. A good numbers of trials revealed that the ANN architecture M-3-20-1(3 nodes in input layer as specific conductivity, chloride and TH, 20 nodes in hidden layer and one node in output layer) gave highest R^2 and lowest RMSE. The result also revealed the predominance of total hardness and specific conductivity over chloride and potassium on prediction of TDS.

***Keywords:* Total dissolved solid (TDS), Artificial neural network (ANN), groundwater quality, Nadia district.**

I. INTRODUCTION

A major portion of world's freshwater resources is composed of groundwater. A large portion of the world's population depends on this freshwater resource for their various needs. India which is an agriculture based country is not an exception where groundwater is the major and most important source of agricultural as well as domestic water supply. But the indiscriminate and unplanned use of this natural resource has put a threat on the sustainability of this major freshwater reserve in various parts of our country. The overuse of groundwater is reported from different parts of the country such as Tamil Nadu, Gujarat, Orissa and West Bengal, among several other states. Therefore, sustainable and efficient management of this vulnerable resource has become a challenge to the policy makers. Sustainable management of this resource can be achieved through precise and effective analysis of the existing water resource related data, proper and organized groundwater use planning and social awareness. Apart from the decline of quantity, deterioration of groundwater quality is also a major concern. Contamination of groundwater resources either from anthropogenic activities or from inherent aquifer material composition reduces its supply, posing a threat to development and a challenge to water managers and strategists. A variety of factors contribute to variations in groundwater quality. Their inherent uncertainty carries weight, as more than one variable affect quality of water. The non-homogeneity of the medium has thrown the quality prediction and the approaches adopted by researches into complexity [1]. Therefore, no single parameter is sufficient to express the complete picture of water quality (WQ) of any area. WQ is normally assessed by measuring a broad range of parameters and modeling of water quality requires a complicated non-linear relation among the variables. One basic measure of water quality is the total dissolved solids (TDS), which is the total amount of solids (in milligrams per liter) left when a water sample is evaporated to dryness. Water naturally contains a number of different dissolved inorganic constituents. The major cations are calcium, magnesium, sodium, and potassium; the major anions are chloride, sulfate, carbonate, and bicarbonate. Although not in ionic form, silica can also be a major constituent. These major constituents constitute the bulk of the mineral matter contributing to total dissolved solids (TDS). TDS is directly associated with sodium absorption ratio, salinity of water and drinking water quality ([2]; [3]; [4]).

One suitable approach to look into groundwater behavior is applying computerized models. In recent years, ANN models have been successfully employed to the water quality studies ([5]; [6]; [7]). Many statistically based water quality models, assume linear relationship among the variables, however, ANNs can efficiently comprehend and

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 6, Issue 10, April 2017

model the non-linear relationship among the different variables which affect water quality [8]. A great deal of modeling has been done in this regard ([9]; [10]; [11]; [12]). As for the water quality prediction based on neural network, Mehrdadi et al., (2012) [4] made an attempt to predict the TDS parameter with the neural network in Fajr Purification Center in the south of Iran in 2012 with successful results. Other examinations using similar modeling conducted by Abyaneh et al., (2011) [13] to predict nitrate parameter have been successful. Kheradpisheha et al., (2015) [14] Carried a groundwater quality assessment study using ANN in Bahabad plain, Yazd, Iran and inferred that ANN can predict accurately the various groundwater parameters. The present study aimed to model TDS in artificial neural network environment for efficient prediction of this major groundwater quality parameter for Nadia district, West Bengal.

II. MATERIALS AND METHODS

A. Description of the Study Area

In the present study, Nadia district was selected as a study area, which is part of Indo-Gangetic Plain [Fig.1]. Nadia district is located in south eastern part of West Bengal state between 22º53' and 24º11' North latitude and 88º09' and 88º48' East longitude and encompasses a geographical area of 390027 km2. It consists of 17 blocks.

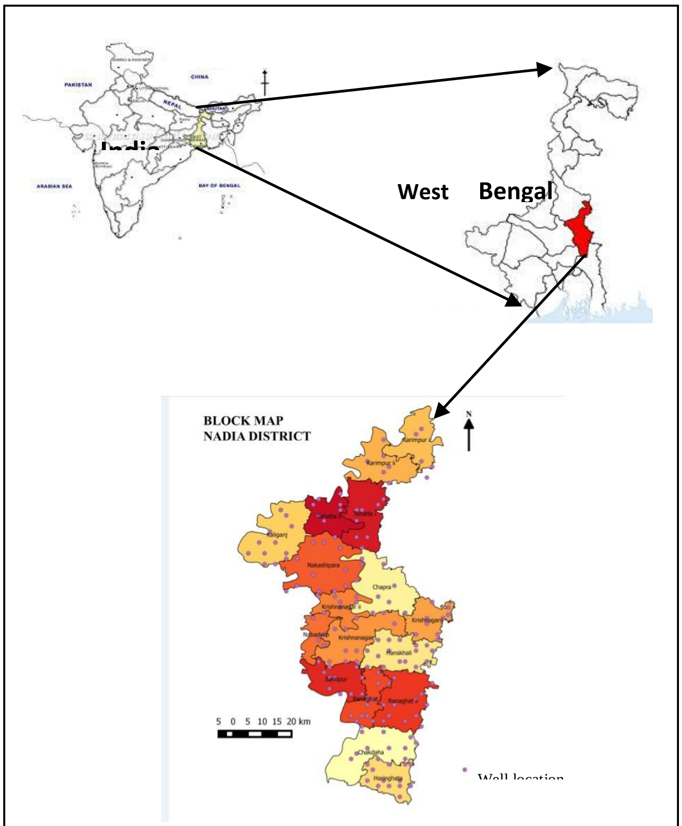


Fig 1. Location map of the study area

The major river systems of the study area are the Jalangi River which flows through the heart of the district and the Hughly River which flows through the western boundary of the district. Other significant river systems are Churni, Mathabhanga and Ichamati. The district is bounded by Bangladesh in the east, Bardhaman and Hugli district on the west, Murshidabad district on the north and north west and North 24 Parganas towards south and south east. The tropic of Cancer divides the district in two parts and is characterized by tropical monsoon with cold and dry winters, and warm and humid summer. January/February are the coldest months with an average temperature of 14 0C and April/May are the hottest months with an average temperature of 35 0C. The mean annual rainfall is 1500 mm, precipitating more than 80% during June through September.

B. Collection of Data

Groundwater quality data on seven water quality parameters namely chloride, specific conductivity, TDS, total hardness, bicarbonate, potassium and sodium has been collected for the year of 2011 at 166 sites (Fig. 1) located in Nadia district from the SWID, Kolkata. Geographical coordinates of each sampling location was also collected from SWID. Artificial neural network modeling was carried out to compute the total dissolve solids (TDS) concentration in the study area using MATLAB R2013 software package.

C. ANN Modeling

An Artificial Neural Network (ANN) is a flexible mathematical structure, which is capable of identifying complex nonlinear relationships between input and output datasets. The basic structure of a network that is common in hydrological applications consists of three layers namely: input layer (where data are introduced to the network), hidden layers (where data are processed) and output layer (where the results for the given inputs are produced). The optimum ANN architecture (i.e. the numbers of hidden layers and neurons in each layer) which can effectively capture the relationship between the input and output data is usually determined by trial and error.

In the present study, multilayer perceptron (MLP) models with one input layer, hidden layer and output layer were used to simulate the TDS for the study area. Six post monsoon groundwater quality parameters namely; specific conductivity, bicarbonate, chloride, total hardness, sodium and potassium were used in different combinations as

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 6, Issue 10, April 2017

inputs to the different ANN models tested for the study. In order to understand the relative importance of different input variables on TDS, a correlation matrix between different input parameters and output parameter (total dissolved solids, TDS) was constructed. A correlation analysis is a simplified statistical tool to show the degree of dependency of one variable to the other. From matrix plot, strong (r = 0.8 to 1), moderate (r = 0.6 to 0.8) and low (r = 0.5 to 0.6) correlation between selected variables was found out and accordingly the different input parameters for different ANN models were selected

Only one output was selected for all the ANN models which are the measured TDS values for the study area. The selection of an appropriate neural network structure (i.e. the selection of number of nodes in hidden layer) is a very important task as this affects the capability of the model to simulate the actual process. A larger network may tend to over-fit the training samples whereas too-small networks may have problems in learning the training data. In the current study, the numbers of the nodes in the hidden layer were selected by trial and error. For each new ANN architecture with same number of input -output nodes and different numbers of nodes in hidden layer was trained to minimize the mean sure error at the output layer.

Model	ANN Architecture	Input parameters	R^2	RMSE
M-1	M-3-20-1	Sp. Conductivity, Chloride,TH	0.97	14.63
M-2	M-6-35-1	Sp.Conductivity,Chloride,TH,Sodium, Bicarbonate,Potassium	0.96	40.53
M-3	M-4-20-1	Sp.Conductivity,Chloride,TH,Sodium	0.94	22.73
M-4	M-5-20-1	Sp.Conductivity,Chloride,TH,Sodium, Bicarbonate	0.89	35.92
M-5	M-4-25-1	Chloride,Sodium,Bicarbonate,Potassium	0.57	57.54

Table 2. Summary of model statistics

In the current study, the three- layered feed-forward models were trained using Levenberg-Marquardt (LM) back propagation algorithm. During training, the ANNs provide information on mean square error with each epoch (iteration) and the actual output value was compared with the desired output and the error was calculated. The error values were then propagated back into the network to update connection weights between the different layers. These processes were repeated until the network has been trained to the lowest RMSE. In this study 'logsig' activation function was used in the hidden layer whereas, 'purlin' activation function was set as default in the tool box as output layer activation function. The number of the nodes in the hidden layer was selected by trial and error method i.e., the method which was giving least error values. The appropriate architecture of the neural network models was determined through training, testing and validating of the models. 70% of the total data set was used in ANN model training. After optimizing the number of nodes in the hidden layer and the network structure, the models were tested with rest 30% data. The network was designed in neural network tool box of MATLAB-2013.

D. Performance Evaluation of the ANN models

The prediction performance of selected ANN model was investigated by RMSE and coefficient of determination (R^2) between network output and network target outputs in training and validation groups. The equations for calculating RMSE is given below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \qquad \ldots 1$$

Where, y_i and x_i are actual TDS and ANN predicted TDS respectively. The best method is the one with the smallest RMSE and the highest R^2 ([2]; [12]; [11]).

III. RESULTS AND DISCUSSIONS

A. Development of ANN models for estimating TDS

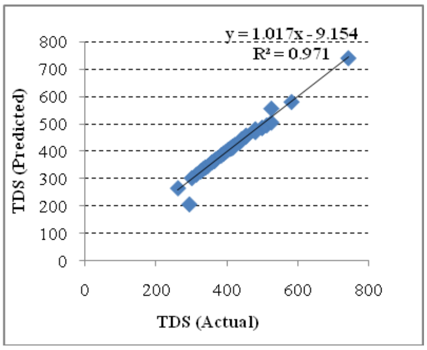
The present section discusses the selection of different architecture of ANN models which suits most to estimate the total dissolved solids (TDS) for the study area. Six groundwater quality parameters namely specific conductivity, bicarbonate, chloride, total hardness, sodium, and potassium for the 166 well locations for the post-monsoon period were used in different combinations as inputs to the different ANN models tested for the study. Only one output was selected for all the ANN models which are the measured post-monsoon TDS values for the

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 6, Issue 10, April 2017

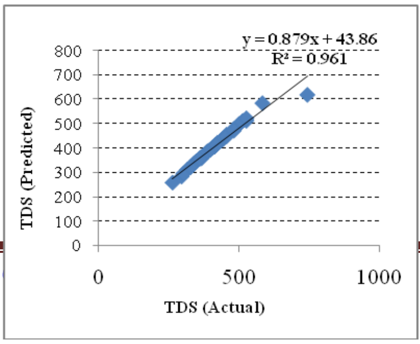
study area. In order to understand the relative importance of different input variables on TDS, a correlation matrix between different input parameters and output parameter (TDS) was constructed and the correlation coefficient values were estimated. Correlation of a particular input parameter with the output parameter was considered as the key factor in selecting the various combinations of input nodes in different model. Based on the correlation between selected variables and the output variable, five different ANN architectures were selected with different combinations of input nodes in the input layer. In this study, one hidden layer was used for all the five ANN architectures with different nodes in the hidden layer. The numbers of nodes in the hidden layer was determined by trial and error method during training. Each model was trained using the training data set which used 70% of the patterns (110) and tested with remaining 30% (46) of the patterns. The architectures of the selected models along with R^2 value and RMSE values are presented in the Table 2.

B. Comparison of Different Models

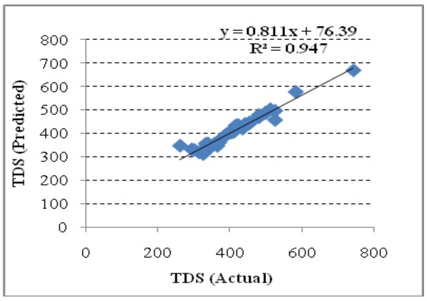
The ANN architecture M-3-20-1(3 nodes in input layer, 20 nodes in hidden layer and one node in output layer) (Table 2) gives highest R^2 and lowest RMSE. The least correlation and highest RMSE was obtained for model M-4-25-1. The results reveal that sp conductivity plays an important role in predicting TDS in the study area. The study also inferred that TDS can be successfully modeled using ANN for the study area with proper selection of input parameters. Fig. 2 a-e presents the scatter plot of ANN predicted TDS values and measured TDS values for M-1, M-2, M-3, M-4 and M-5, respectively. The simulated TDS values and observed TDS values for each model was showed graphically in Fig. 3 a-e.



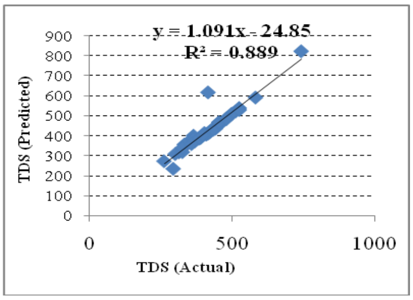
a



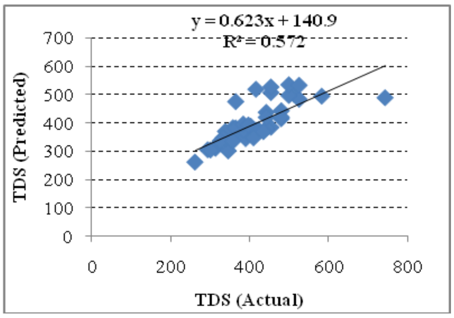
b



c

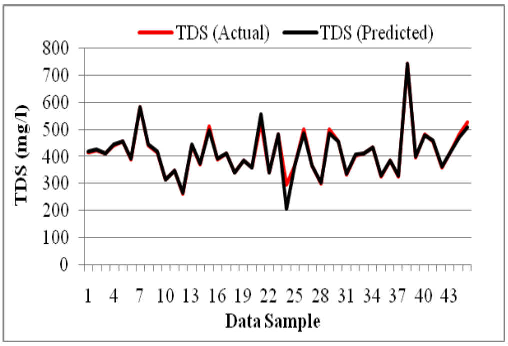


d

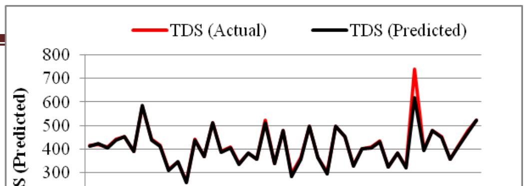


e

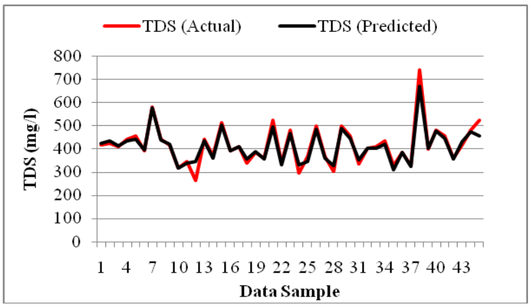
Fig 2 a-e. The distribution diagrams of the predicted and observed TDS for models M-1, M-2, M-3, M-4 and M-5, respectively.



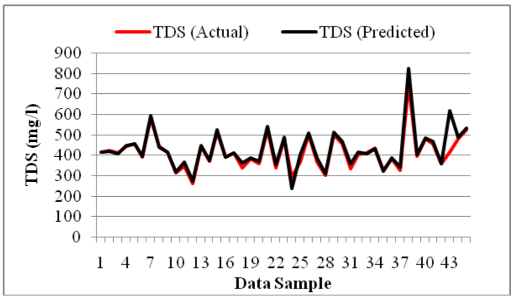
a



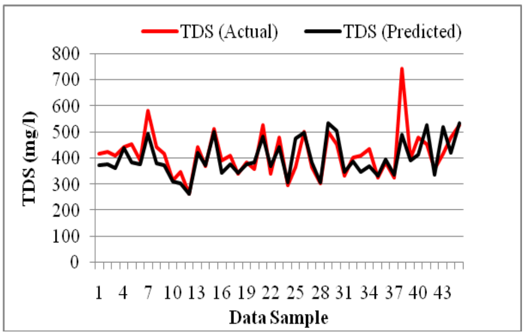
ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 6, Issue 10, April 2017



c



d



e

Fig 3 a-e. Comparison of ANN-simulated TDS concentration and observed TDS concentration for models M-1, M-2, M-3, M-4 and M-5, respectively.

IV. CONCLUSIONS

The TDS (total dissolved solids) parameter constitutes one of the fundamental parameters as regards to drinking and agricultural water. In the present study an attempt was made to understand the efficacy of Artificiel Neural network (ANN) in modelling the TDS concentration in groundwater for Nadia district, West Bengal, India. Artificial neural network modeling was carried out to compute the total dissolve solids (TDS) concentration in the study area using MATLAB R2013 software package. Six different groundwater quality parameters were used. The results of the study reveals that the ANN architecture M-3-20-1(3 nodes in input layer, 20 nodes in hidden layer and one node in output layer) gives highest R^2 and lowest RMSE. The least correlation and highest RMSE was obtained for model M-4-25-1. The results reveal that specific conductivity plays an important role in predicting TDS in the study area. The study also inferred that TDS can be successfully modeled using ANN for the study area. However, the current study has considered only few numbers of groundwater quality parameters. The study would have been more precise if more groundwater quality parameters could have been considered. The groundwater of district has a prominent arsenic contamination problem. Similar studies on this aspect also help in efficient management groundwater in the study area.

ACKNOWLEDGEMENT

Authors gratefully acknowledge the State Water Investigation Directorate (SWID), Kolkata for providing the necessary data for the successful completion of the study.

REFERENCES

[1] Esmaeili Varaki, M., Khayat Khalaghi, M., Shafiei, M. (2004), ″Provide a model for intelligent water level fluctuations estimated alluvial groundwater aquifer using an artificial neural network″, Articles first annual confer-ence of Iran water resources management, pp. 1 – 11.

[2] Asghari Moghaddam. A., Nadiri, A., Fijani, E. (2006), ″Ability to study different Models of Artificial Neural Networks to Evaluate Groundwater Water Level in the Hard Formation, Tenth″, Conference of Geological Society, Tehran.

[3] Jamshidzadeh, Z. and Mirbagheri, S.A. (2011), ″Evaluation of groundwater quantity and quality in the kashan Basin, Central Iran″, Desalination, Vol. 270, pp. 23 – 30.

[4] Mehrdadi, N., Hasanlou, H., Jafarzadeh, M.T., Hasanlou, H., Abodolabadi, H. (2012), ″Simulation of low TDS and biological units of Fajr industrial wastewater Treatment plant using artificial neural network and principal component analysis hybrid method″, Journal of water resource and protection, Vol. 4, pp. 370 – 376.

[5] Bowers J, Shedrow C. (2000). Predicting stream water quality using artificial neural networks. Miscellaneous series. Westinghouse Savannah River Co. (112).

[6] Kuo J.T., Hsieh M.H., Lung W..S, She, N. (2007). Using artificial neural network for reservoir

eutrophication prediction. Ecological modeling 200:171-177.

[7] Kuo Y.M., Liu C.W., Lin, K.H. (2004). Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of Blackfoot disease in Taiwan. Water Research 38:148-158.

[8] Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S. (1996). Application of neural networks to modeling nonlinear relationships in ecology. Ecological modeling 90: 39-52.

[9] Kumar, M, Raghuwanshi, N., Singh, R., Wallender, W., and Pruitt, W. (2002), ″Estimating evaporate transpiration using artificial neural networks. Journal of Irrigation and drainage engineering″, ASCE 128, Vol. 4, pp. 224 – 233.

[10] Coppola, E., Szidarovszky, F., Poulton, M., Charles, E. (2003), ″Artificial Neural Network Approach for Predicting Transient Water Levels in a Multi Layered Groundwater system under variable state″, Pumping and climate conditions, Hydrologic engineering, Vol. 8 No. 6, pp. 348 – 360.

[11] Hosaini, M.T., Siosemarde, A., Fathi, P., Siosemarde, M. (2007), ″Application of artificial neural networks (ANN) and multiple regressions for estimating assessing the performance of dry farming wheat yield in Ghorveh Region, Kurdistan province″, Agricultural research: Water and soil and plant, Vol. 7 No. 1, pp. 41 – 54.

[12] Khalili, S.R., Davari, k., and Mousavi Baygi, M. (2008), ″Monthly Precipitation Forecasting Using Artificial Neural Networks: A case Study for Synoptic Station of Mashad″, Journal Water and Soil, Agricultural Science & Technology Ferdowsi University of Mashhad, Vol. 22 No. 1, pp. 39 – 99.

[13] Zare Abyaneh, H., Bayat Varkeshi, M., and Daneshkare Arasteh, P. (2011), ″Forecasting nitrate concentration in groundwater using artificial neural network and linear regression models″, International agrophysics, Vol. 25 No. 2, pp. 187 – 192.

[14] Kheradpisheha, I.Z., Talebib, A., Rafatia, L., Ghaneeiana, M.T. and Ehrampousha, M.H. (2015), Groundwater quality assessment using artificial neural network: A case study of Bahabad plain, Yazd, Iran Z. Vol20 (1): 65-71.