International Journal of Engineering Research
Volume No.5, Issue No.2, pp : 115- 118
ISSN:2319-6890)(online),2347-5013(print)
1 Feb. 2016

Performance of ANNs for Prediction of TDS of Godavari River, India

Prajot D. Tarke[1], Purushottam R. Sarda[2] and ParagA. Sadgir[3]
Government College of Engineering, Aurangabad, MS, India
[1]ptarke@gmail.com;[2] prsarda04@gmail.com;[3]sadgirp@gmail.com

Abstract: This research investigates the potential of the Artificial Neural Networks (ANNs) on simulating interrelation between water quality parameters for river water quality management. It aims to model Total Dissolved Solids (TDS) values at Pategaon station on Godavari River by application of ANNs. Monthly data from 2001 to 2012 of various water quality parameters is collected. Correlation analysis has been carried out for selecting the most suitable input parameters for the model. The ANN modelling strategy is implemented by Neural Network Toolbox in MATLAB. Several ANN architectures and training possibilities are assessed and the best ANN architecture is selected for finding the best prediction model of TDS. Comparisons between the measured and predicted values show that the ANNs model could be successfully applied and provide high accuracy and reliability for predicting water quality parameters. Coefficient of correlation between observed and predicted TDS values calculated using ANNs and analytical method is found to be similar (R=0.98293), which shows the effectiveness of ANNs in predicting the missing water quality parameters.

Keywords: ANN, Correlation, Regression, TDS, Water Quality Prediction

I. Introduction

All forms of lives from micro-organisms to man are ultimately depend upon water for their survival. Most human activities involve the use of water in one way or other. Water in the Godavari River which is used mainly for agriculture is degraded by diffuse (non-point) sources. Among water quality parameters, Total Dissolved Solid (TDS) or “filterable residue” is defined as the quantity of dissolved material in water; it includes mineral and organic matter (McNeil and Cox 2000). TDS is one of the vital water quality parameters and continuously used to determine the water quality of rivers. TDS has important consideration in determining its suitability for irrigation, drinking and industrial uses. For irrigation water, dissolved solid is a very important criterion due to their gradual accumulation results in salinization of soil, thus, rendering the agriculture land non-productive. For this reason TDS was assessed to evaluate the water quality condition in the basin.

New approaches such as Artificial Intelligence (AI) techniques have proven their ability and applicability for simulating and modelling various physical phenomena in the water engineering field. During last decade one of the most importance advances made in the field of the water resources engineering is the development and adaption of Artificial Neural Networks (ANNs), which is another AI approach, for prediction of missing water quality parameters.McCulloch and Pitts (1943) are recognized as the first designers of ANNs, which are generally inspired by the operation of the brain and nerve systems in biological organisms with capability of self-learning and automatic abstracting.A neural network is an adaptable system that learns relationships from the input and output data sets and then is able to predict a previously unseen data set of similar characteristics to the input set (Haykin, 1999).Extensive literature of the subject reveals that many successful applications of this technique have been made.

Maier and Dandy (1996) were the first to apply the artificial neural networks for assessment of variation of water salinity in Murray River in the south of Australia to find a way for predicting the average water salinity and obtaining the efficient time to pump and operation of the water. Khalil *et al.* (2012) applied the ANN to predict the water quality parameters in Nile delta. Mangeshkumar*et al.* (2012) have developed a neural network model of TDS concentration in Cauvery River water. ANNs through exploring the relationship between the input parameters, is able to predict the parameter TDS (Maedeh *et al*., 2013). ANNs model of TDS in SiminehRiver, Iran was developed by Nemati (*et al.* 2014). Najah *et al.* (2012) have worked on RBFNN and MLPNN models and found that the RBFNN model are the best modelling techniques for predicting the missing water quality parameters. Sarda *et al.* (2015) have provided an easy and rapid method of monitoring water quality parameters.Goyal (*et al.*2013) have proved effectiveness of ANNs tool in predicting water quality for Kalyan-Dombivali Municipal Corporation, they have found that the most efficient network is Generalized Feed Forward Network.

The literature review revealed the fact that in recent years attempts have been made by researchers to develop water quality models of variety. Results of their research have been found positive and their work can be referred for further development of new models.The present work describes the development of the ANNs model for the purpose of predicting TDS. From 2001 to 2012 a set of recorded data ofvarious water quality parameters at Pategaon stationon the Godavari Riveris used to predict TDS.

Theory of Artificial neural network

Considering natural neural and its components, scientists have developed artificial neural, which is the smallest unit of an ANNs. An artificial neural consists of three components: weighting (W), bias (b), and transfer function (f). These three components are unique for each neural. Fig. 1 shows schematic structure of an artificial neural. In Fig. 1 “p” and “a” are the input and output to a neural, respectively. The weight and bias associated with the neural is indicated by “w” and “b” and “f” is the symbol of the transfer function. Parameter “n” is called the net input, which is subjected to the “f” to give an output. Mathematical representation of the artificial neural can be given as follows (Menhaj 1998).

$$n = wp + b \quad (1.1)$$

$$a = f(n) = f(wp + b) \quad (1.2)$$

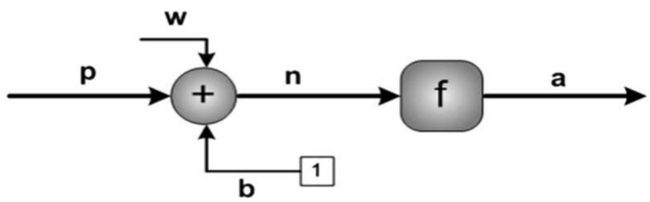


Fig. 1 Schematic of Artificial Neural

In neural instruction process, “w” and “b” neural parameters change until the best approximation for an output member corresponding to the input member is obtained (Menhaj 1998). ANN as a “black-box” technique is an appropriate solution for modelling nonlinear and complex phenomena. Among various types of ANN models, the feed-forward neural network (FFNN) with back-propagation learning algorithm has been known as a universal predictor (Hornik 1989).

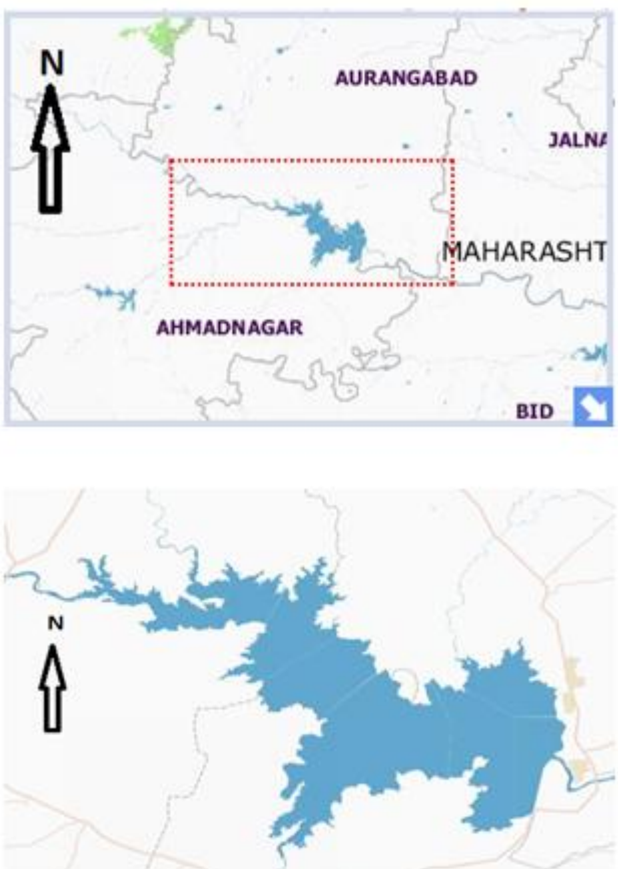


Fig. 2 Location map of Jayakwadi Reservoir

II. Study Area

The study area is situatedat downstream of Jayakwadi reservoir across Godavari River in Maharashtra State, India. Jayakwadi reservoir is a multipurpose and multiobjective project mainly used for drinking, irrigation, industrial, and hydro-electricity. It is located between latitude 1927’55N and longitude 7524’27 E with catchment area of 21,750 sq.km, length 10.20 km and gross storage capacity 2909 Mm^3. River Godavari is the lifeline of millions of people in the state of Maharashtra, Telangana and Andhra Pradesh. Godavari is polluted by industries, chemicals and fertilizers used for agriculture purpose, sewage from cities and by open defecation all along the river banks by the people who are daily using the water for drinking and bathing. The fertilizer also is discharged into the river by rainfall runoff. So, it is necessary to model the critical water quality parameters and predict them. TDS is exceeding the limits for the purpose of domestic water supply and agriculture use and it is very useful, if this parameter has been modelled and predicted.

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \qquad (3.1)$$

III. Methodology

Correlation Analysis

Correlation analysis is carried out using equation 3.1 to find the correlation among various water quality parameters. It also helps in deciding input parameters for ANN model.

Where X_i is observed value and $Y_i$is predicted value in series of data.$\overline{X}$and $\overline{Y}$ are the mean of observed and predicted values respectively. While ‘r’reaches 1, it shows strong correlation between real data and predicted ones, and when ‘r’ approaches zero, there is a weak correlation between two types of the data.

Table 1. Pearson’s Correlation Coefficient

	TURB	TH	pH	Na	Ca	SS	SO_4	Mg	EC	TDS
TURB	1.000									
TH	0.470	1.000								
pH	-0.059	-0.057	1.000							
Na	-0.057	0.184	0.090	1.000						
Ca	0.470	0.980	-0.098	0.197	1.000					
SS	0.053	0.047	-0.098	-0.096	0.075	1.000				
SO_4	0.631	0.223	-0.174	0.369	0.223	-0.071	1.000			
Mg	0.127	0.009	-0.137	0.167	0.035	0.134	0.462	1.000		
EC	0.095	0.020	0.184	0.248	0.017	-0.177	0.243	0.403	1.000	
TDS	0.107	0.134	0.202	0.326	0.131	-0.151	0.290	0.443	0.946	1.000

From table 1, it is observed that out of 55 correlation coefficients, one correlation i.e., TDS-EC is found to be highly significant (r=0.946) which shows the direct relationship between TDS and EC. Suspended solids shows negative correlation with TDS (r=-0.151). Turbidity, Hardness, pH, Na, Ca, SO_4 shows poor correlation with TDS. Mg has moderate correlation with TDS (r=0.443).

Development of Model

ANN model architecture refers to the layout of neurons and the number of hidden layers. The feed forward back-propagation training algorithm is a supervised training mechanism and is normally adopted in most of the engineering application.

ISSN:2319-6890)(online),2347-5013(print)
1 Feb. 2016

Table 2. Variation in R value with change in number of neurons.

Output_Input	Architecture	Train R	Valid R	Test R	All R	Remark
T_tHpNCSSO$_4$ME	9-1-1	0.92048	0.94732	0.94134	0.92511	Architecture with 6 neurons is selected as the optimum model for prediction of TDS.
T_tHpNCSSO$_4$ME	9-2-1	0.92036	0.96393	0.92033	0.92467	
T_tHpNCSSO$_4$ME	9-3-1	0.97032	0.92505	0.90818	0.95202	
T_tHpNCSSO$_4$ME	9-4-1	0.96096	0.90016	0.89989	0.9468	
T_tHpNCSSO$_4$ME	9-5-1	0.97072	0.92275	0.84849	0.95383	
T_tHpNCSSO$_4$ME	9-6-1	0.97625	0.92857	0.98293	0.97002	
T_tHpNCSSO$_4$ME	9-7-1	0.9652	0.94574	0.9606	0.96204	
T_tHpNCSSO$_4$ME	9-8-1	0.98393	0.96686	0.95428	0.97438	

**Note:T=Total Dissolved Solids, t=turbidity, H=hardness, p=pH, N=Na, C=Ca, S=suspended solids, SO_4 =Sulphate, M= Magnesium, E= electrical conductivity.*

The primary goal is to minimize the error at the output layer by searching for a set of connection strengths that cause the ANNs to produce outputs that are equal to or closer to the targets.

A typical ANN model with a back propagation algorithm is constructed to predict TDS values. Table 2 shows the variation in R value with varying number of neurons. TANSIG transfer function with 6 numbers of neurons shows the best model performance. The network was trained in 1000 epochs using the Levenberg–Marquardt learning algorithm.

Model efficiency

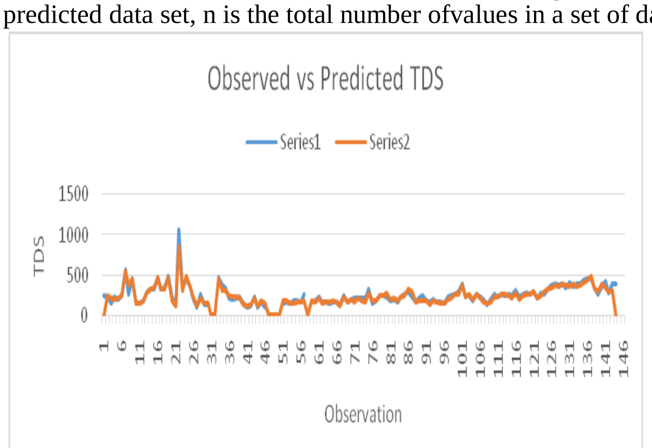
To assess the applicability of the models, two statistical criteria were adopted, i.e., Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).These criteria can be defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{iobserved} - Y_{ipredicted})^2} \quad (3.2)$$

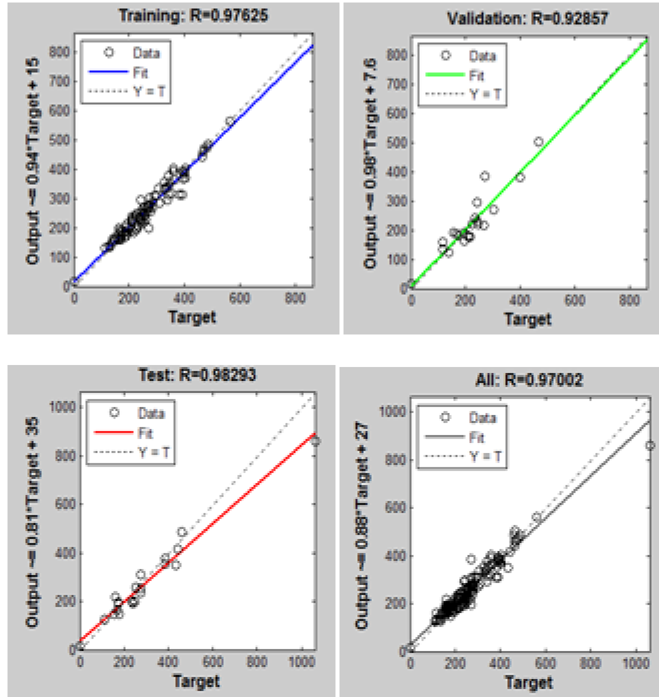
$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_{iobserved} - Y_{ipredicted}\right| \quad (3.3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(X_{iobserved} - Y_{ipredicted})^2}{\sum_{i=1}^{n}(X_{iobserved} - \overline{X_{iobserved}})^2} \quad (3.4)$$

Where, $X_{i\ observed}$ is the available data set, $Y_{i\ predicted}$ is the predicted data set, n is the total number ofvalues in a set of data.



(a)



(b)

Fig.3 Comparison of predicted ANNs time series with observed values.

(a) Sequence plots (b) Scatter plots

V. Results and Discussion

Sequence plot and scatter plot of the results obtained from the optimum ANNs model are shown in fig.3.The model gave close approximations of the actual observations, suggesting that these approaches are applicable for modelling the TDS dataset.Table 3 shows the analytical results of observed and predicted values. The value of R found using ANNs and using standard formula are similar, which shows the accuracy of predicted dataset. RMSE and MAE value is found to be 31.76 and 21.39 respectively which indicates the minimum error between observed and predicted dataset.

Table 3. Analytical results based on observed and predicted values

Architecture	Performance Indicators			
	RMSE	**R**	**R^2**	**MAE**
9-6-1	31.76	0.9683	0.9377	21.39

V. Conclusion

A study of TDS time series is reported in this paper using local water quality parameters for a set of recorded missing data in Godavari River at Pategaon station during 2001 to 2012. The general objective of this study is to investigate the performance of ANNs for the estimation of the TDS without applying the significant knowledge of physics of process. Correlation analysis shows the highest correlation between TDS and EC (r=0.946), which indicates that EC is directly proportional to the TDS.Missing data has been computed by ANN model and compared with observed data. The validation of the neural network model by performance indicators showed good agreement for predictions of the TDS.Coefficient of correlation between observed and predicted TDS values calculated using ANNs andanalytical method is found to be similar (R=0.98293), which shows the effectiveness of ANNs in predicting the missing water quality parameters.The modelling results indicated that reasonable prediction accuracy was achieved for the ANNs model.The outcomes and model may be applicable for water quality management for planning of water usage.

References

i. *Bureau of Indian Standards (2012), "IS: 10500:2012, Drinking Water Specification (2nd Revision)", Drinking Water Sectional Committee, FAD25, India, pp. 1-11.*

ii. *Goyal R., Patel H. and Mane S. (2013), "Artificial Neural Network: An effective tool for predicting water quality for Kalyan-Dombivali Municipal Corporation", International Journal of Science and Research, Vol. 4(6), pp. 2863-2866.*

iii. *Haykin S. (1999). "Neural networks: A comprehensive foundation." New York: Macmillan Publishing.*

iv. *Hornik K. and White H (1989) "Multilayer feed forward networks are universal approximators", Neural NetworksVol.2, pp.359–366*

v. *Khalil B. M., Awadallah A. G., Karaman H. and El-Sayed A. (2012), "Application of Artificial Neural Networks for the Prediction of Water Quality Variables in the Nile Delta", Journal of Water Resource and Protection, Vol. (4), pp.388-394.*

vi. *Maier H.R. and Dandy G.C. (2000), "Application of Artificial Neural Networks to forecasting of surface water quality variables: issues, application and challenges", Journal of Water Resources Research, Vol. (15), pp. 101-124.*

vii. *Maedeh P.A., Mehrdadi N., Bidhemdi G.R. and Abyaneh H.Z. (2013), "Application of Artificial Neural Networks to predict TDS variations in groundwater of Tehran plain, Iran", International Journal of Environment and Sustainability, Vol. (2), pp. 10-20.*

viii. *Mangeshkumar P., Pradip T., Stalin J.M.R., Amal Raj and Anandkumar S. (2012), "Neural Network Modelling of TDS concentrations in Cauvery River Water, Tamilnadu, India", International Journal of Applied Sciences and Engineering Research,Vol. 1 (6), pp. 739-746.*

ix. *McCulloch W.S. and Pitts W. (1943), "A logicalcalculus of the ideas immanent in neurons activity", Bulletin of Mathematical Biophysics, Vol. 5pp.115-133.*

x. *McNeil V.H., and Cox M.E. (2000), "Relationship between conductivity and analysed composition in a large set of natural surface-water samples, Queensland, Australia." Journal of Environment Geology, Vol. 39(12), pp.1325–1333.*

xi. *Menhaj M.B. (1998), "Artificial neural networks basis" Vol. (1), Amir Kabir Univ., Tehran, Iran.*

xii. *NajahA., El-Shafie A., Karim O.A. and El-Shafie H. (2013), "Application of Artificial Neural Networks for water quality prediction",Journal ofNeural Computation and Application, Springer, Vol. 22(1), pp. 187-201.*

xiii. *Nemati S., Naghipour L. and Farad M.H.F. (2014), "Artificial Neural Network Modelling of Total Dissolved Solid in the Simineh River, Iran", Journal of Civil Engineering and Urbanism, Vol. 4 (1), pp. 08-14.*

xiv. *Sarda P. and Sadgir P. (2015), "Computation of water quality parameters and prediction tool ANN for modelling of water quality of reservoir", Innovative Journal of Innovative Research in Science and Technology, Vol. 4 (9), pp. 8906-8911.*

xv. *Sarda P. and Sadgir P. (2015), "Statistical Analysis for Predicting Missing Data and Developing Parameter Inter-relationship for Jayakwadi Reservoir in Maharashtra State, India", KSCE Journal of Civil Engineering, Springer,DOI: 10.1007/s12205-015-0050-8.*