International Journal of Pure and Applied Mathematics
Volume 116 No. 8 2017, 353-357
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

IRON TOXICITY OF POLLUTED RIVER WATER BASED ON DATA MINING PREDICTION ANALYSIS

[1]C.Nalini, [2]S.Brintha Rajakumari
[1]Department of CSE, BIST,BIHER,Bharath University, Chennai, India
[2]Department of MCA, BIST,BIHER,Bharath University, Chennai, India
[1]drnalinichidambaram@gmail.com[2]brintha.cse@bharathunive.ac.in

Abstract: In river water resources becomes essential to inspect the information and use the analysed data for optimum use of water resources as there is scarcity of such resources. In this paper multilayer perception of neural networks function and time series forecasting methods were used to predict the values of iron toxic metal in the unfit value of toxic metal of Indian River water data. The experimental result helps to the government to take the necessary precaution step to protect the people living near the rivers.

Keywords: Artificial Neural Networks, Prediction, Toxicity, Water Analysis.

1. Introduction

Due to advanced technology solutions, tons of information will be handled and data analysis become easier and quicker. In water resources it becomes necessary to investigate the information and use the analysed data for optimum use of water resources as there is scarcity of such resources[1-2]. Estimation is a process of assessing the numeric value of an attribute based on given values of other related attributes. When the assessment is for a future time, the problem is called a prediction. The conventional time series modelling methods were also efficient since long time, but they suffer the problem of stationary and linearity and gives only reasonable accuracy[3-4]. The need for giving more accurate forecast for time series has forced the researchers to develop innovative methods to model time series. Artificial Neural Networks (ANNs) were introduced as an efficient tool for modelling and forecasting since two decades. The artificial neural networks can model, map, and as well as can demonstrate the nonlinear relationship of complicated phenomena. The artificial neural networks are widespread and highly flexible function approximates, used in the fields of cognitive science and engineering. The artificial neural networks are widely used and have become increasingly popular with a broad range of fields. The neural networks are less sensitive to the error term assumptions and can bear the noise, hectic components[5-6], and hefty tails better than most other methods. This section presents a study aimed at forecasting river water using neural network approaches. It has been proved that ANN models show better results in river stage-discharge modelling in comparison to traditional models. The development of Artificial Neural Networks began approximately 50 years ago, inspired by a desire to understand the human brain and emulate its functioning.

2. Related works

The use of artificial neural network (ANN) modelling for prediction and forecasting variables in water resources engineering are being increasing rapidly(Mustafa,2012). Infrastructural applications of ANN in terms of selection of inputs, architecture of networks, training algorithms[7-8], and selection of training parameters in different types of neural networks used in water resources engineering have been reported.Therefore, the goal of this study is to examine how effectively ANN has been applied to solve problems in water resources engineering particularly in toxicity of the river. Furthermore[9-10], investigate the kind of infrastructure, such as input selection criterion, selection and division of the data sets, appropriate structure of the network, activation function and algorithms used for training network has been utilized for proper modelling to find the best solution.

This paper (Archana, 2015) presents the concept of artificial neural network technique to measure the dissolved oxygen concentrations at the downstream of Mathura city in India, is located at the bank of River Yamuna in the state of Uttar Pradesh in India. In the data analysis, they used feed forward error back propagation neural network technique. Monthly data sets of Mathura (upstream), Mathura (central) and Mathura (downstream) have been used for the analysis about flow discharge, temperature, pH, biochemical oxygen demand (BOD) and dissolved oxygen (DO). Shilpi et al presents an ANN approach for forecasting of reservoir water level using ten daily data of inflow, water level and release (Shilpi, 2014).

The paper (Mozejkoet al.,2008) describes the application of ANNs for the time series modelling of

total phosphorous concentrations in the Odra River. Two models were proposed to prove the forecast of phosphorus concentrations such as a simpler with a single input and more complex with 14 inputs.A neural network model was developed (marina et al., 1999) to analyse and forecast the behaviour of the river Tagliamento in Italy during heavy rain periods[13-14]. The model makes use of distributed rainfall information coming from several rain gauges in the mountain district and predicts the water level of the river at the section closing the mountain district.

The paper (Zaheer,2003) presented an artificial neural network based on a decision-making approach to water quality management to control environmental pollution. Theeffects of pollution sources were evaluated for strategic planning of water quality management using a backpropagation algorithm of a feed forward neural network. Hanjiang River data of china were used to demonstrate the procedure.

The aim of this study(Bowers, 2000) was to predict water quality in small streams using an Artificial Neural Network. The selected input variables were local precipitation[11-12], stream flow rates and turbidity for the initial prediction of suspended solids within the stream.

3. Material and methods

3.1 Neural Network

The human brain always stores the information as a pattern. Any capability of the brain may be viewed as a pattern. Development of ANN and its application in field of pattern recognition. ANN is a computing model that tries to mimic the human brain and the nervous system in a very primitive way to emulate the capabilities of the human being in a very limited sense.

ANN is network of parallel, distributed information processing system that relates an input and output vectors. It has a quantity of information processing elements such as cells or nodes and that are grouped into layers. The input layer processing elements receive the input vector and transmit the values to the next layer of processing elements across connections where this process is continued. This type of network, where data flow one way (forward)[15-16], is known as a back propagation network. A feed forward back propagation ANN has mainly input layer, an output layer, and one or more hidden layers between the input and output layers. Each of the neurons in a layer is connected to all the neurons of the next layer, and the neurons in one layer are connected only to the neurons of the immediate next layer. The strength of the signal passing from one neuron to the other depends on the weight received from the interconnections[17-18]. It is initiated that ANNs are robust tools for modelling many of the nonlinear hydrologic processes such as rainfall-runoff modelling, ground-water administration, water quality recreation, stream flow, and reservoir water levels. The hidden layers enhance the network's ability to model advanced functions. Performance of BPANN (Back Propagation Artificial Neural Network) models is compared to the developed linear transfer perform (LTF) model and was found superior. ANNs will effectively model the stage-discharge relationship.

Comparison to a conventional statistical stage-discharge model shows the superiority of an approach using ANN. Enlargement of ANN is based on the following rules:

- Information processing occurs at nodes that are single elements and are also denoted as units, neurons or cells.
- Signals are passed between nodes through connection links.
- Each connection link has an associated weight that represents its connection strength.
- Each node applies an activation function to its net input to determine its output signal.

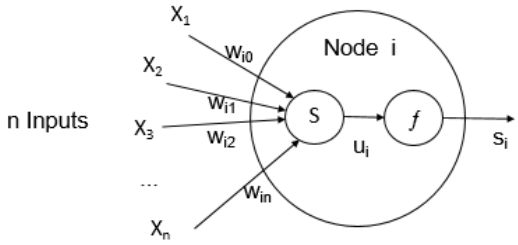

Figure 1. Function of Single Node in Neural Networks

Single neurons are not able to solve complex tasks and creating networks by hand is too expensive. Each node outputs an activation function applied over the weighted sum of its inputs: $s_i = f(w_{i,0} + \sum_{j \in I} w_{i,j} \times s_j)$ and the sigmoid activation function is $y = \frac{1}{1+e^{-x}}$Figure 1 shows the single node activation function of feed forward neural networks.

3.2 Multilayer Perceptron

Multilayer Perceptron is a feed forward neural networks are the most popular and widely used neural network models in many real world applications. It is a finite acyclic graph, consists of input layers, hidden layers and output layer. In feed forward networks, the input is introduced at the nodes in the input layer and then propagated through the network in one direction, without any loop[19-20], to the output layer.

3.3 Time Series Forecasting

Time series have been extensively studied for decades by many researchers because so many variables important to our well-being evolve with time, such as daily stock price, weekly sales and monthly

flows in rivers. A time series typically consists of an observation of one variable taken at equally spaced intervals of time[21]. The usefulness of time series analysis lies in the ability to identify the evolving pattern of the series and then project that pattern into the future to get estimates of the future values.

4. Result and discussion

Figure 2 illustrates the comparison of actual verses predicted toxicity of iron metals using multilayer perceptron for the training data samples. To illustrate that the Indian River water unfit value of toxic metal data from 2011to 2013 was collected from central

The iron toxic metal data were used for analysis. The multilayer perceptron gives best result with Mean Absolute Error (MAE) was 0.08 and Root Mean Square Error (RMSE) was 0.10. In the last data sample, Actual value is 0.769 and the predicted output value is 0.708. So the MLP method gives the best approximate value compared to the observed value. From that result, we can easily predict the future value and identify the polluted area.

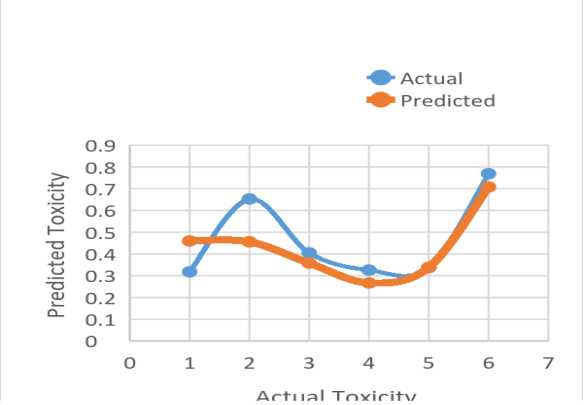

Figure 2. Comparison of Actual and Predicted Toxicity of Iron Metal using Multilayer Perceptron

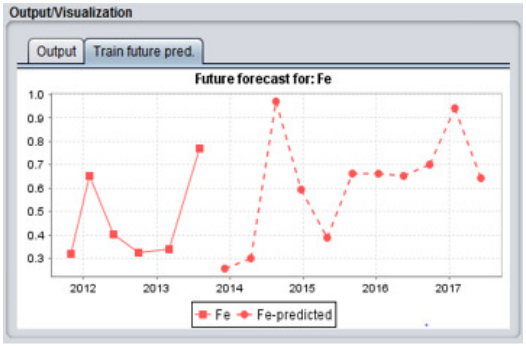

Figure 3. Prediction of Iron Toxic Metal based on Time Series Forecasting

The research analysed the river water data and predict the future values using time series forecasting method. The predicted values of iron toxic metal in the river Ghaghara of siteid 37 is in the figure 3. The result obtained by using Forecast data package of WEKA 3.9.1 tool and used the multilayer perceptron method with two input layers and 5 hidden layers. As a result the multilayer perceptron provides less Mean Square Errors and Root Mean Square Error to predict the values.

5. Conclusion

Water pollution becomes series issue in the world. Identifying the toxicity of river water leads to take necessary steps to product the people those who are all living nearby rivers. Also suggest the government to take necessary action against the factory and mining industries which is located near the river. And the research also identify the polluted river and helps to conduct medical camp to the polluted area.

References

[1]Archana Sarkar, Prashant Pandey,2015, River Water Quality Modelling using Artificial Neural Network Technique, International Conference On Water Resources, Coastal And Ocean Engineering, Aquatic Procedia4, 1070 – 1077.

[2]Bowers A. &Shedrow C B., 2000, Predicting stream water quality using Artificial Neural Networks (ANN), Computer Techniques in Environmental Studies, WIT Press,pp89-96

[3] Marina Campolo, Paolo Andreussi, and Alfredo Soldati, 1999, River flood forecasting with a neural network model, water Resources Research 35(4), 1191-1197.

[4] Mustafa R M, Isa M.H., Rezaur R.B, 2012, Artificial Neural Networks Modeling in Water Resources Engineering: Infrastructure and Applications, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering, 6(2),pp8-136.

[5] Możejko J, R. Gniot R, 2008, Application of Neural Networks for the Prediction of Total Phosphorus Concentrations in Surface Waters, Polish J. of Environ. Stud. 17(3), 363-368.

[6]Udayakumar R., Kaliyamurthie K.P., Khanaa, Thooyamani K.P., Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, v-29, i-14, pp-86-90, 2014.

[7]Kaliyamurthie K.P., Parameswari D., Udayakumar R., QOS aware privacy preserving location monitoring in wireless sensor network, Indian Journal of Science and Technology, v-6, i-SUPPL5, pp-4648-4652, 2013.

[8]BrinthaRajakumari S., Nalini C., An efficient cost model for data storage with horizontal layout in the

cloud, Indian Journal of Science and Technology, v-7, i-, pp-45-46, 2014.
[9]BrinthaRajakumari S., Nalini C., An efficient data mining dataset preparation using aggregation in relational database, Indian Journal of Science and Technology, v-7, i-, pp-44-46, 2014.
[10]Khanna V., Mohanta K., Saravanan T., Recovery of link quality degradation in wireless mesh networks, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4837-4843, 2013.
[11]Khanaa V., Thooyamani K.P., Udayakumar R., A secure and efficient authentication system for distributed wireless sensor network, World Applied Sciences Journal, v-29, i-14, pp-304-308, 2014.
[12]Udayakumar R., Khanaa V., Saravanan T., Saritha G., Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction, Middle - East Journal of Scientific Research, v-16, i-12, pp-1781-1785, 2013.
[13]Khanaa V., Mohanta K., Saravanan. T., Performance analysis of FTTH using GEPON in direct and external modulation, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4848-4852, 2013.
[14]Kaliyamurthie K.P., Udayakumar R., Parameswari D., Mugunthan S.N., Highly secured online voting system over network, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4831-4836, 2013.
[15]Thooyamani K.P., Khanaa V., Udayakumar R., Efficiently measuring denial of service attacks using appropriate metrics, Middle - East Journal of Scientific Research, v-20, i-12, pp-2464-2470, 2014.
[16]R.Kalaiprasath, R.Elankavi, Dr.R.Udayakumar, Cloud Information Accountability (Cia) Framework Ensuring Accountability Of Data In Cloud And Security In End To End Process In Cloud Terminology, International Journal Of Civil Engineering And Technology (Ijciet)
Volume 8, Issue 4, Pp. 376–385, April 2017.
[17]R.Elankavi, R.Kalaiprasath, Dr.R.Udayakumar, A fast clustering algorithm for high-dimensional data, International Journal Of Civil Engineering And Technology (Ijciet), Volume 8, Issue 5, Pp. 1220–1227, May 2017.
[18]R. Kalaiprasath, R. Elankavi and Dr. R. Udayakumar. Cloud. Security and Compliance - A Semantic Approach in End to End Security, International Journal Of Mechanical Engineering And Technology (Ijmet), Volume 8, Issue 5, pp-987-994, May 2017.
[19]Thooyamani K.P., Khanaa V., Udayakumar R., Virtual instrumentation based process of agriculture by automation, Middle - East Journal of Scientific Research, v-20, i-12, pp-2604-2612, 2014.
[20]Udayakumar R., Thooyamani K.P., Khanaa, Random projection based data perturbation using geometric transformation, World Applied Sciences Journal, v-29, i-14, pp-19-24, 2014.
[21]Udayakumar R., Thooyamani K.P., Khanaa, Deploying site-to-site VPN connectivity: MPLS VsIPSec, World Applied Sciences Journal, v-29, i-14, pp-6-10, 2014.