Predictive Potentiality of Artificial Neural Networks for predicting the Electrical Conductivity (EC) of drinking water of Hyderabad city

NIAZ A. MEMON  M.A. UNAR  A.K. ANSARI
Dept: of civil Eng:  Dept: of Computer Systems Eng:  Dept: of Chemical Eng:
Quaid-e-Awam University of Eng  Mehran University of Eng:  Mehran University of Eng:
Nawabshah  Jamshoro  Jamshoro
PAKISTAN  PAKISTAN  PAKISTAN

G.B.KHASKHELI  BASHIR AHMED MEMON
Department of Civil Engineering  Department of Civil Engineering
Mehran University of Engineering & Technology  Quaid-e-Awam University of Eng: Sciences
JAMSHORO, PAKISTAN  NAWABSHAH, PAKISTAN

Abstract: - Electrical Conductivity (EC) is an important parameter of drinking water. The determination of electrical conductivity provides a prompt and expedient way to measure the accessibility of electrolytes in the water. There are convinced health effects on human life through these electrolytes, like disorder of salt and water balance in infants, heart patients, individuals with high blood pressure, and renal diseases. Salty taste is one of the aesthetic effects of EC if it exceeds 150 mS/m and if greater than 300 mS/m it does not slake the thirst. The drinking water supplied to Hyderabad city is taken from River Indus and the EC of this river remains questionable. The values of EC in drinking water of Hyderabad at selected locations were recorded. From 49 samples, the average values ranged from 658 to 762. In order to determine the optimal value of EC with in the distribution system, where it deteriorates, it is necessary to predict it at different locations. The use of conventional methods to predict parametric values in the distribution systems is suffered from certain precincts. To get better drinking water quality by tumbling operational costs, Advance process control and automation technologies are the tools to be used normally. The application of Artificial Neural Networks in Water Supply Engineering is enticing and more accepted because of its high predictive accuracy. In this paper Radial Basis Neural Network has been demonstrated. The data sets are prepared for training the model. It was observed that the model has high predictive potentiality to predict the values of Electrical conductivity at 07 locations of distribution system of water supply in Hyderabad city.

Key-Words: - Electrical Conductivity, Drinking water, Distribution System, ANNs, RBF, modeling

1 Introduction

Hyderabad lies in the latitude 25° 22’ N and longitude 65°-41’East. Geologically, the city is low flat-topped and typically of arid topography. The climate is subtropical, semi-desert type. It is characterized by low and highly erratic rainfall, low relative humidity, and high rate of temperature. The mean annual rain fall is 12.92 cm mainly concentrated in the months of July and August, which together accounts for 12.2 cm. [1]. River Indus is a major source of drinking water supplied to Hyderabad city having population of 1.8 million. Domestic water supply of Hyderabad comes from canals emanating from the river Indus, which receives an unexpected and unannounced discharge of contaminated water from Manchhar Lake, one of the biggest natural fresh reservoirs in Asia. [2] From this river at Jamshoro, water is supplied to the lagoons named North and South lagoons of 400 MG for pre-settlement and then brought to the “New treatment Plant” (NTP) with its capacity of 30 MGD. This plant and the pipeline distribution system came in being in the early 80s, in comparison with Old treatment Plant (OTP) with capacity of 10 MGD, commissioned in the early 60s, which is now out of work. It is observed that the most of waste water drainage pipe lines are also laid in parallel and are about 152-244 cm away from the drinking water pipeline, hence, often causing mixing of wastewater with drinking water [3] Under these conditions parametric pollution is evident from the distribution system of the city where all the samples were bacteriological positive at all seven locations concluded in a weeks study carried out by Pakistan Council for Research on water Resources
(PCRWR) [4]. The parameters include pH, Turbidity, Electrical conductivity, HCO₃, Cl, SO₄, Ca, Mg, Na, K, Hardness and Bacteriology. In this study the Electrical Conductivity is undertaken, being an important parameter of drinking water quality. The adverse health effect include disturbance of salt and water balance in infants, heart patients, persons with high blood pressure, and renal disease. Aesthetic effects include a salty taste to the water (if conductivity > 150 mS/m) while water with conductivity > 300 mS/m does not slake thirst. [5] The dissolved or soluble fraction of the water’s total solids load is referred to as total dissolved solids, known as TDS, normally the weight of this material. The Electrical conductivity EC provides a simple measure of TDS and is the measure of the water’s ability to conduct an electric conductivity. A relationship between TDS and EC may be considered with respect to the following relationship. This relationship gives an idea about the approximate calculation of Total Dissolved solids (TDS) when Electrical conductivity EC is taken with respect to a specific temperature of 25°C. The relationship is given as under:

\[
\text{TDS(mg/l)} = 0.67 \times \text{EC} \quad 25 (1)
\]

It is observed that the Electrical conductivity EC is temperature sensitive. It increases with increasing temperature. This is maintained automatically when measured with latest equipment having probes. These modern probes automatically standardize all the readings to 25°C temperature, and this EC is called as EC₂₅ [6]. Recent research indicates towards the use of new technologies and computer based applications for predicting the Drinking Water Pollution Level (DWPL) for all its measurable parameters. Over the last two decades, artificial neural networks have been a primary focus of interest for research in computer based civil and environmental engineering by providing convenient and often highly accurate solutions to problems from all branches including drinking water related issues. At first momentary look, artificial neural networks appear to be one of the great accomplishments in the history of computing in civil engineering. Replying to the question: “why there has been such a high and sustained level of interest in applying artificial neural networks to civil/environmental engineering”, the answer is that ANNs, despite their presently elementary form, are very good at solving direct mapping problems that are non-linear and comprise several independent variables, a common class of problems in civil engineering. In this context they often provide more accurate solutions than the alternative modeling techniques and place a little demand on the modeller in terms of understanding the basic form of the function being represented [7]. Artificial neural network models are capable of modeling data whose functional relationship are not known in advance. By choosing appropriate architecture and activation, neural networks can be trained to capture knowledge from the data available with acceptable performance. In general Multilayer Perceptron MLP Artificial Neural Networks are commonly used in almost every filed of engineering. This type of ANN is supposed to perform well in a number of hydrologic and water resources applications [8]. The most commonly used algorithm for training multilayer feed-forward networks is the ‘error back-propagation algorithm’. Often it is referred as back-propagation training algorithm. This algorithm involves calculating the derivatives of the network training error with respect to the weights by the application of chain-rule and gradient decent optimization to adjust the weights to minimize the errors [9]. Despite its popularity, the back propagation algorithm suffers from the following disadvantages.

- It is slow.
- It may stuck to global minima of the error surface rather than the global minimum during the training.
- While using this algorithm, the number of the hidden layer neurons is to be selected manually; this may be far from the optimal.

In this study an RBF network has been used which does not suffer from any of the above mentioned disadvantages and is therefore a powerful alternative to MLP network. The RBF networks were originally applied to the multivariable interpolation problems [10]. An RBF network is essentially a three layers network. That is, it has input layer, a hidden layer and an output layer, as shown in figure 1.

**Fig.1**

**Structure of an RBF Network**

The output layer is always linear because each neuron of this layer has a linear activation function. The hidden layer is always non-linear. The neurons in this layer use a radial basis function as an activation function. A most widely used radial basis function is the Gaussian function which is defined as follows:

\[
g_k(\bar{x}) = \exp\left( -\frac{||\bar{x} - \bar{c}_k||^2}{\sigma_k^2} \right)
\]  (2)
In the above equation, $\vec{x}$ is the input vector, $c_k$ is the centre of the kth Gaussian Function and $\sigma_k^2$ is the corresponding variance. The overall output of the network is computed as

$$f(\vec{x}) = \sum_{k=1}^{M} w_k \cdot g_k(\vec{x})$$

Where $M$ is the number of RBFs and $W_k$ is the kth weight.

Like other feed forward neural networks, the RBF network is trained in such a way that the error between the desired and actual output(s) of the network is minimum. In other words, the weights are optimized during training so that the actual output is sufficiently close enough to the desired output. Before optimization of the weights, the network must select an appropriate number of RBFs. Moreover, the centres and width of the RBFs must also be optimized to get good approximation.

In this paper, the orthogonal least Mean Squares Algorithm (LMS) proposes of Chen et al [11] has been used to train the RBF network. This algorithm automatically selects the number of hidden layer neurons by optimizing the centers of the RBFs of a given width.

2 Case Study

The case study considers the important parameter of EC of drinking water of Hyderabad, the 4th biggest city of Pakistan with 2.28 million souls. The drinking water is treated in the filter plant bearing the capacity of 30 MGD. Analytical/experimental observations on the drinking water of Hyderabad city were available with the concerned quarters like Water and Sewerage Authority (WASA), Environmental Protection Agency (EPA), and Pakistan Council for Research in Water Resources (PCRWR) Lahore. Successful application of artificial neural network model requires proper input data preparation [12]. In this study complete data sets were prepared according to the specifications normally used for input data preparation. The data were based on the results of experimental analysis carried out by PCRWR and WASA, during last 5 years. The data sets include the observations of the samples taken at the locations K.B.Federal Police Station Khurshid Colony, City water supply tank Kotri, Hussainabad Pacca Tank HDA-8, Board of Intermediate & Secondary Education, Latifabad Number 7, Tayab Masjid Unit 12 Latifabad, New Wahdat Colony, Qasimabad, and Mustafa Town, Qasimabad.

Normalization (scaling) of the data is a process to get it usable for training a neural network model trained with back propagation algorithms [13] In this paper two parameters, i) Electro conductivity and ii) Turbidity were normalized in order to optimize the weights and train the network for desired outputs.

3. Results and Discussions

A single hidden layer Radial Basis Function (RBF) model with Gaussian function as an activation function in hidden layer neurons, and linear function in output neurons was selected and trained. The data sets were prepared from the available results of the PCRWR and WASA authorities taken during the last 5 years. One complete set was reserved for training and the other data set was prepared to testify and verify the validity of the model. The Input parameters were EC, pH, Turbidity, Alkalinity, HCO$_3$ and Cl. and the targeted value to be predicted out was Electrical conductivity. These results showed that this RBF model doesn’t require any scaling of the data. The Sum of the squared errors SSE was 0.0397799 during the training. It was found that 2 neurons in hidden layer are sufficient for successful training. As there is one output so the number of the neurons in output layer is one. The output (predicted EC values) of the trained network is shown in the Figure 2.

The experimental results are also plotted in the same graph for comparison purpose. The input weights, output weights, input bias, output bias, the time elapsed during training and the SSE is given in the Table 1.
The model was tested for using different testing sets and each time the results were satisfactory because the error in each sample for each testing set was negligible.

Conclusions
The findings of this research suggest that RBF Neural Network Model has the powerful capability of predicting the accurate EC values in combination of the parameters like Turbidity, pH, Cl at various locations in the water supply distribution system of Hyderabad. This has specified the potential to be implemented as an online tool to aid the determination of EC values in drinking water distribution network of Hyderabad city. This Model does not require any robust training termination even in small data sets because of its fast training efficiency which was recorded as low as 0.281000 seconds. This learning proves the advantages, competence, capability and knack of applying the Artificial Neural Network Modeling for monitoring the quality of Drinking water of Hyderabad city of Pakistan.

References:

[2] Zainab dar (Water-for life or death @ www.sandeonline.com/newsletter/downloadable/pdf) (last viewed 03-03-2008)


