Rainfall-Runoff Modeling Using Artificial Neural Network (A Literature Review)

Jignesh Joshi  
PG Student, Civil Engineering Dept.,  
Shantilal Shah Engineering College  
Bhavnagar, India  
joshijignesh2010@gmail.com

Vinod M. Patel  
Asso. Prof., Civil Engineering Dept.,  
Shantilal Shah Engineering College  
Bhavnagar, India  
ympatel33@yahoo.com

Abstract—This is a review paper in which three neural network methods, Feed Forward Back Propagation (FFBP), Radial Basis Function (RBF) and Generalized Regression Neural Network (GRNN) were employed for rainfall-runoff modelling of Maleshri hydrometeorologic data. It was seen that all three different ANN algorithms compared well with conventional Multi Linear Regression (MLR) technique. It was seen that only GRNN technique did not provide negative flow estimations for some observations. The rainfall-runoff correlograms was successfully used in determination of the input layer node number.

Keywords—neural network, rainfall-runoff modelling, Correlograms, multi linear regression.

I. INTRODUCTION

The rainfall-runoff relationship is one of the most complex hydrologic phenomena to comprehend due to the tremendous spatial and temporal variability of watershed characteristics and precipitation patterns, and the number of variables involved in the modelling of the physical processes. Conceptual models provide daily, monthly, or seasonal estimates of stream flow for long-term forecasting on a continuous basis. The entire physical process in the hydrologic cycle is mathematically formulated in conceptual models that are composed of a large number of parameters. The accuracy of model predictions is very subjective and highly dependent on the user’s ability, knowledge, and understanding of the model and of the watershed characteristics.

The artificial neural network (ANN) approach is extensively used in the water resources literature. In these studies either the conventional feed forward error back propagation method (FFBP) or the radial basis function was employed to train the neural networks. As well known the FFBP algorithm has some drawbacks. It is very sensitive to the selected initial weight values and may provide performances differing from each other significantly. Another problem faced during the application of FFBP is the local minima issue. The performance of another ANN algorithm, Generalized Regression Neural Network (GRNN) is not extensively analyzed in water resources problems. The white meteorologic data employed in the whole study belonged to a Maleshri river basin.

II. NEURAL NETWORK METHODS IN THE STUDY

A. The Feed Forward Back Propagation (FFBP)

A FFBP distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons of hidden units. In this study the FFBP were trained using Levenberg-Marquardt optimization technique.

B. The Radial basis function- neural networks (RBF)

RBF networks were introduced into the neural network. The RBF network model is motivated by the locally tuned response observed in biological neurons. The theoretical basis of the RBF approach lies in the field of interpolation of multivariate functions. The solution of exact interpolating RBF mapping passes through every data point. Different number of hidden layer neurons & spread constants were tried in study.

C. The Generalized Regression Neural Networks (GRNN)

The GRNN consists of four layers: input layer, pattern layer, summation layer and output layer. The number of input units in the first layer is equal to the total number of parameters, including from one to six previous daily flows. The first layer is fully connected to the second, pattern layer, where each unit represents a training pattern and its output is a measure of the distance of the input from the stored patterns. Each pattern layer unit is connected to the two neurons in the summation layer: S-summation neuron and D-summation neuron. The optimal value of spread parameter (s) is often determined experimentally. The larger that spread is the smoother the function approximation will be. In this study the optimum spreads were found between 0.0015 and 0.019 for this investigation and available CW is not only beneficial for raising crops, but also helps in lowering the groundwater table.

III. DATA ANALYSIS

According to the data collection, the rainfall and river flow data belongs to Maleshri River in Gujarat region. The daily
mean flow data of a measurement station are provided by state water data centre (swdc). The data covers a period of 10496 days. Rainfall data for this region was provided by State Meteorological Service (DM). Several rainfall measurement stations were subjected to correlation analysis. The rainfall stations having highest correlations with the

Figure 1. Flow estimations for station by FFBP (a) based only rainfall values and (b) based both rainfall and flow values for the testing period.

flow station was selected. Final selection was made by considering the location coordinates of the rainfall and runoff stations with respect to each other.

IV. METHOD APPLICATIONS

3 MATLAB codes were written for three different ANN methods (FFBP, RBF and GRNN) as explained before. The data was divided into two parts for training and testing. The flow estimation simulations were carried out in two steps. First, only rainfall data was employed for the input layer. Then previous daily flow value was also incorporated into the input data group. The improvement with adding flow value was discussed for all ANN method applications. Mean square error (MSE) value for the testing period was considered for performance evaluation and all testing stage estimates were plotted in the form of hydrograph. The ANN training data covered the first 9000 daily total rainfall and daily mean flow values of the record. However because of the computer memory limitation only first 5000 and 4500 daily values could be used for RBF and GRNN training stages, respectively. The following 1490 daily values were kept for testing purpose (Table 1). The initial training simulations have shown that one hidden layer was sufficient. The number of hidden layer nodes was found by trial and error. The numbers of nodes in the input layer were determined using the rainfall-runoff correlograms as suggested by [6]. The correlograms for the station pair 71

Figure 2. Flow estimations for station by RBF (a) based only rainfall values and (b) based both rainfall and flow values for the testing period.

Figure 3. Flow estimations for station 71 by GRNN (a) based only rainfall values and (b) based both rainfall and flow values for the testing period

Figure 4. Flow estimations for station 71 by MLR (a) based only rainfall values and (b) based both rainfall and flow values for the testing period (flow)-17886 (rainfall) shows that the correlation values do not decrease noticeably after lag 4. The highest correlation corresponds for lag 0 meaning that the concentration time of the basin is shorter than one day. The simulation results for three ANN methods are presented in Figs. 1a-3a. For the purpose of comparison, multi linear regression (MLR) was employed. Accordingly the independent variables of MLR were the input nodes of ANNs and the calibration period was the ANN training duration (Table 1). The MLR estimations are depicted in Fig.4a. The most convenient ANN configuration for each method is given in Table 1. An ANN structure FFBP (5, 4, 1) consists of 5, 4 and 1 nodes in input, hidden and output layers, respectively. RBF (5, x, 1) with s=0.5 has 5 and 1 nodes in input and output layers, respectively, and a spread parameter equal to 0.5. Similarly GRNN (5,x,1) with s=0.10 represents a structure with 5 input- and 1 output nodes and a spread equal to 0.10. The spread parameters are found simply by trial and error. Generally the ANN and MLR estimates showed deviations from the observations. For station 701, FFBP, RBF and MLR provided MSE values close to each other for the testing period. The MSE obtained by GRNN is slightly higher compared with other methods. Then the flow at the precedent day ($Q_{t-1}$) was also added to the input layer in order to increase the estimation performance. The estimation results are presented in Figs.1b-3b and Table 1. It is clear that a noticeable improvement in estimation performance was obtained with the incorporation of flow value into the input layer.

It was seen that RBF and MLR simulations resulted in negative flow estimations for some low flows. This problem, however, was not faced in GRNN simulations and this approach does not require an iterative training procedure as in back propagation method as explained by [10]. The forecast of GRNN is bounded by the minimum and maximum of the observed series preventing the network to provide forecasts physically not possible. The GRNN forecast can not converge
to poor solutions corresponding to local minima of the error criterion which is one of the drawbacks of FFBP algorithm. It was shown that FFBP method performance is very sensitive to the randomly assigned initial weight values. Several simulations are needed in order to obtain the best FFBP performances and only these superior FFBP results are presented in the study. RBF and GRNN, however, provide estimations with a unique simulation.

V. CONCLUSIONS

TABLE I

<table>
<thead>
<tr>
<th>Variables in the input layer</th>
<th>ANN Structure</th>
<th>MSE ($m^3/s^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFBP</td>
<td>FFBP(5,4,1)</td>
<td>96.13</td>
</tr>
<tr>
<td>RBF</td>
<td>RBF(5,x,1)</td>
<td>95.06</td>
</tr>
<tr>
<td>GRNN</td>
<td>GRNN(5,x,1)</td>
<td>105.45</td>
</tr>
<tr>
<td>MLR</td>
<td>--</td>
<td>94.92</td>
</tr>
<tr>
<td></td>
<td>--</td>
<td>53.22</td>
</tr>
</tbody>
</table>

In this study three different ANN methods, FFBP, RBF and GRNN, were successfully applied to rainfall-runoff modelling study. It was seen that GRNN flow estimation performances were close to those of the FFBP, RBF and MLR. During the simulations GRNN did not provide physically non-plausible estimations (negative values). The training time of a single FFBP application is relatively shorter than GRNN.

This handicap can be easily overcome with GRNN algorithms by including clustering, which was not employed in this study [11]. Besides, it was seen that multiple FFBP simulations were required until obtaining a satisfactory performance criteria. This total duration for FFBP simulations was longer than the unique GRNN application. Another important result of the study is the positive contribution of the utilization of initial statistical analysis results in determining the ANN input layer node number. The rainfall-runoff correlation diagrams point to the number of rainfall values significant in flow estimations. This can be a time saving feature since the input layer nodes are found by trial and error in general. Further suggestion given by the author of this paper, namely Jignesh Joshi, reveals that like this method, the most accurate method of FFBP can also be used for other watershed analysis, too and as it gives good and perfect results, the author also finds feasibility of recommending this method for important projects, too.

VI. ACKNOWLEDGMENT

This work has been supported by the state water data centre (SWDC). The data is provided by General Directorate of water Resources Survey and State Meteorological Service (DMI).

REFERENCES

Sciences Journal 40: 146–163