

Research Article

ANN-Based Approach for Predicting Rating Curve of an Indian River

Arun Goel

Department of Civil Engineering, National Institute of Technology, Kurukshetra, Haryana 136119, India

Correspondence should be addressed to Arun Goel, drarun.goel@yahoo.co.in

Received 5 April 2011; Accepted 17 May 2011

Academic Editor: I. Raftoyiannis

Copyright © 2011 Arun Goel. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prediction of stage-discharge relation or a rating curve is of immense importance for reliable planning, design, and management of most of the water resources projects. Measurement of discharge in a river is a time-consuming, expensive, and difficult process, and the conventional approach of regression analysis of stage-discharge relation does not provide encouraging results especially during the floods. Therefore, the present study is aimed at the application of soft computing techniques such as back propagation feed forward neural network-based algorithm for modelling stage-discharge relation. A data set of discharge-measuring station located on an Indian river has been used for analysis in the present study. A multilinear regression model was also employed on the same data in order to compare the performance of the results. The performance of each model has been compared by calculating correlation coefficient and root mean square error for the used data set. The outcome of the study suggests that the back propagation feed forward ANN works quite well for the data sets and produced promising results in comparison to the linear regression technique.

1. Introduction

It is important to predict a rating curve in most of the water resources projects. A common practice is to measure the river stages at regular intervals and use them for discharge calculations, which can be used for future hydrological analysis. The rating curve is determined by assuming that there exists a unique relation between stage and discharge of the river at the given site. However, the stage-discharge relationship is time dependent and very often exhibits a random phenomenon with fluctuations. Mostly a power equation is used to establish a relation between stage and discharge whose variables can be determined by the polynomial regression analysis. Recently, there has been a growing interest in the analysis of the complex hydrological processes by using modelling techniques like artificial neural networks [1–6]. The present study explores the potential of back propagation feed forward ANN in predicting the rating curve and discharge prediction using data of a gauging sites Tikrarpara from an Indian river Mahanadi. The performance of back propagation feed forward ANN was compared with a multilinear regression modelling approach as well.

2. Artificial Neural Networks

A neural network is an artificial intelligence technique that mimics a function of the human brain. Neural networks are general-purpose computing tools that can solve complex nonlinear problems in the field of pattern recognition, classification, speech, vision, and control systems. The network comprises a large number of simple processing elements linked to each other by weighted connections according to a specified architecture. A neuron consists of multiple inputs and a single output. The number of neurons in the input and output layers is fixed by the problem being modeled as the number of input variables equals the number of input neurons and the number of output variables equals number of output neurons. The determination of optimal number of hidden layers and hidden neurons is usually cumbersome, as no general methodology is available for their determination. These networks learn from the training data by adjusting the connection weights. There is a range of artificial neural network architectures designed and used in various fields of hydrology and hydraulics. Most of the studies employing neural networks for water resource problems have used back

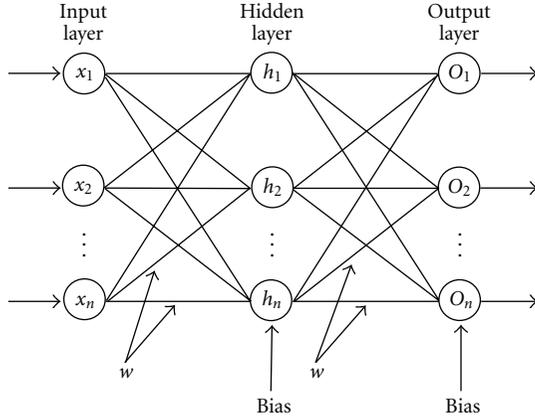


FIGURE 1: Three-layer feed forward neural network.

propagation and radial basis function types of neural networks. In this study, a feed forward neural network with back propagation learning algorithm is applied. The basic element of a back propagation neural network is processing node and structure of commonly used back propagation neural network (Figure 1). A three-layer feed forward ANN has been shown in Figure 1, which consists of three layers known as input, hidden, and output layers. Input layer neurons are called x_1, x_2, x_3 ; hidden layers neurons are h_1, h_2, h_3 ; and output layers neurons are O_1, O_2, O_3 . A neuron consists of multiple inputs and a single output. The sum of inputs and their weights leads to a summation function. The output of a neuron is decided by an activation function, which can be step, sigmoid, threshold, linear, and so forth.

Each processing node behaves like a biological neuron and performs two functions. First, it sums the values of its inputs. This sum is then passed through an activation function to generate an output. Any differentiable function can be used as activation function. All the processing nodes are arranged into layers, each fully interconnected to the following layer. There is no interconnection between the nodes of the same layer. In a back propagation neural network, generally, there is an input layer that acts as a distribution structure for the data being presented to the network. This layer is not used for any type of processing. After this layer, one or more processing layers follow, called the hidden layers. The final processing layer is called the output layer in a network. This process is repeated until the error rate is minimized or reaches an acceptable level, or until a specified number of iterations have been accomplished. All the interconnections between each node have an associated weight. The values of the interconnecting weights are not set by the analyst but are determined by the network during the training process, starting with randomly assigned initial weights. There are a number of algorithms say gradient descent method that can be used to adjust the interconnecting weights to achieve minimal overall training error in multilayer networks. The generalized delta rule, or back propagation is one of the most commonly used methods as suggested by [7]; the first

derivative of the total error with respect to a weight (1) determines the extent to which that weight is adjusted.

$$\Delta w = - \epsilon \frac{\partial E}{\partial w}, \quad (1)$$

where ϵ is the learning constant, $\partial E/\partial w$ is the first derivative of the total error with respect to weight, and Δw is weight change. A neural network-based modelling approach requires setting up several user-defined parameters like learning rate, momentum, optimal number of nodes in the hidden layer, and the number of hidden layers, so as to have a less complex network with a better generalization capability.

3. Error Measure Criteria

(1) *Correlation Coefficient (r).*

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}, \quad (2)$$

where $x = X - X'$, $y = Y - Y'$ where X = observed values, X' = mean of X , Y = predicted values, and Y' = mean of Y .

(2) *Root Mean Square Error.*

$$\text{RMSE} = \left[\frac{\sum (X - Y)^2}{n} \right]^{0.5}. \quad (3)$$

4. Application of ANN for Discharge Prediction/Rating Curves

To assess the usefulness of back propagation feed forward ANN-based modelling approach in predicting the rating curve and discharge, the data sets collected on a gauging site (Tikrapara) located on river Mahanadi in Orrissa (India) were used. The back propagation feed forward ANN is used in calculating correlation coefficients and root mean square errors (RMSEs) to generate the model on the input data set and predicting the discharge for the data sets with input as stage. With 70% data, training was done and the remaining 30% data testing and validation of the results was done. A most common neural network architecture namely, back propagation feed forward multilayer perceptron neural network was employed as well for this data set. A neural network-based modelling approach requires setting up several user-defined parameters like learning rate, momentum, optimal number of nodes in the hidden layer, and the number of hidden layers, so as to have a less complex network with a relatively better generalization capability.

The suitable value of user-defined parameters in case of back propagation feed forward ANN (Table 1) was obtained by comparing the correlation coefficients and root mean square error (RMSE) values after a number of trials for the data sets used in the present study. Table 2 shows correlation and RMSE for back propagation feed forward ANN and linear regression techniques.

TABLE 1: User-defined parameters of back propagation feed forward neural network modeling.

S No.	Type of Input parameter	Type of parameter			
		Momentum	Learning rate	No. of hidden nodes	No. of iterations
1	Stage-discharge relation	0.2	0.3	0.1	500
2	Discharge prediction	0.1	0.05	0.1	500

TABLE 2: Values of correlation coefficient and RMSE for input parameter using back propagation feed forward ANN and linear regression.

S No.	Type	ANN		Linear	
		Correlation coefficient	Root mean square error	Correlation coefficient	Root mean square error
1	Stage-discharge relation	0.9349	1800.035	0.9201	1952.5818
2	Discharge prediction	0.9630	1234.8113	0.9327	1646.6751

5. Model Development and Performance Criteria

One data set consisting of daily stage and discharge records of the site Tikarapara (1440 pairs) on Mahanadi river in Orrisa (India) with a drainage area equal to 124450 sq km at site number 44, having 1440 number of data set pairs, for a period from June to Oct (1981 to 1986) has been used in the present study. For modelling rating curves, different combinations of stage and discharges were considered together and three techniques already mentioned were applied. To model the discharge curves, all stage and discharge values were considered and divided in two parts for the Tikarapara site. The 70% data was used for training and the remaining 30% was used for testing the model for predicting of rating curves. For the discharge prediction, a tenfold cross validation was used. The whole of the data was considered as input data for training. Normalisation of the data was carried out so as to bring all input variables with a range of 0 to 1. Correlation coefficient and root mean square error (RMSE) values for different data sets obtained are as given in the Table 2 and are used for the performance evaluation of the models and comparison of the results for establishing the stage-discharge curve and discharge prediction using back propagation feed forward ANN and linear regression. A higher value of correlation coefficient with a smaller value of RMSE for given set of input parameters is considered to be a better performing model. To study the scatter, a line of perfect agreement (i.e., a line at 45 degrees) was plotted in the resulting graphical output between the actual and predicted discharge values.

6. Analysis and Discussion of Results

The optimal values of user-defined parameters learning rate, momentum, optimal number of nodes in the hidden layer, and the number of hidden layers for this data set are also given in Table 1. A value of correlation coefficient of 0.9349 and RMSE = 1800.035 was achieved with input as stage using back propagation feed forward ANN, and correlation coefficient of 0.9201 and RMSE = 1952.5818 with linear regression for the rating curves are obtained as given in Table 2. The rating curves as shown in Figure 2 have been plotted for

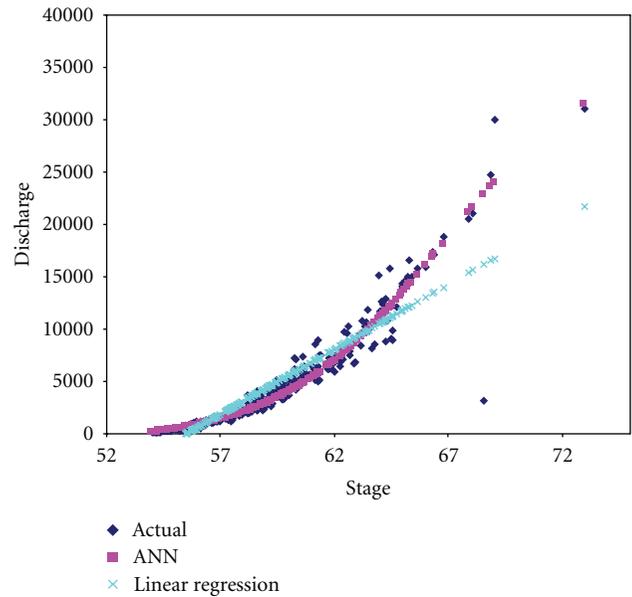


FIGURE 2: Variation of rating curve with actual versus predicted by ANN and linear regression.

observed and predicted results by back propagation feed forward ANN, and results are compared with the linear regression modelling. The examination of these figures indicates that the linear regression points are not following the actual rating curve as compared to predicted rating curve by the back propagation feed forward ANN modelling. Similarly by comparing Figure 3, it is clear that the rating curve predicted by back propagation feed forward ANN modelling is conforming to actual rating curve more closely.

A value of correlation coefficient of 0.9630 and RMSE = 1234.8113 was achieved with input as stage using ANN, and correlation coefficient of 0.9327 and RMSE = 1646.6751 with linear regression for the discharge prediction are obtained as given in Table 2. Figure 3 provides the plot for actual and predicted discharges using back propagation feed forward ANN against the linear regression modelling. The perusal of these two figures tells that the predicted discharges by ANN and M5 model tree are showing superior results as compared to linear regression. The perusal of these figures shows that

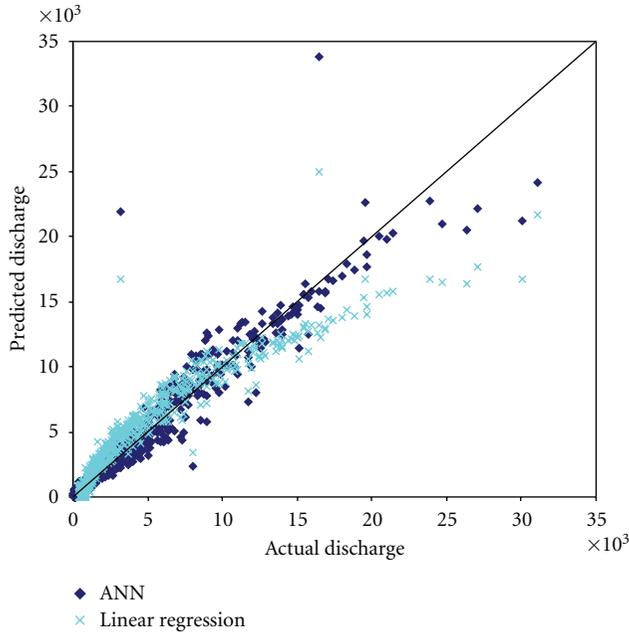


FIGURE 3: Variation of actual discharge with predicted discharge with ANN and linear regression.

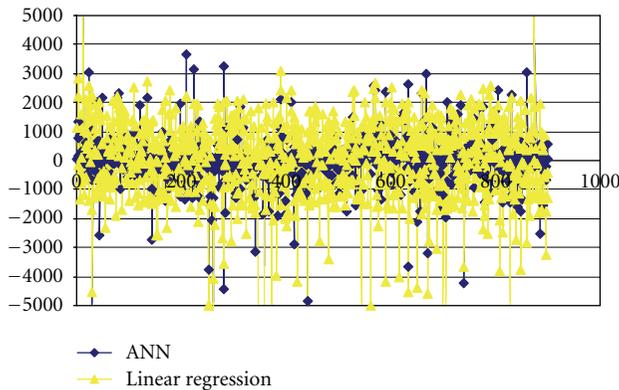


FIGURE 4: Variation of error in actual discharge with predicted discharge with ANN and linear regression.

there is a perfect relation between observed and predicted discharge back propagation feed forward ANN as compared to the linear regression technique.

The error between actual and predicted discharges for the data set is also plotted for discharge prediction in Figure 4 for back propagation feed forward ANN modelling. It can be observed that the error is lesser in case of ANN as compared to linear regression.

7. Conclusions

The back propagation feed forward ANN modelling approach is applied to establish the rating curves and discharge prediction for a data set at Tikarapara gauging site on Indian river. Results from this study suggest and encourage performance by the feed forward back propagation-based back

propagation feed forward ANN approach for rating curve establishment and prediction of discharges as compared to linear regression approach. In view of the above results, this research indicates that ANN approach can be applied as an effective and potential tool for establishing stage-discharge relationship and discharge prediction on a river successfully.

Acknowledgment

The author is thankful to the Chief Engineer, Eastern gauging division, Bhuvneshwar, Central Water Commission, MOR, GOI for providing stage-discharge data, without which it would not have been possible to complete this study.

References

- [1] R. S. Govindaraju, "Artificial neural networks in hydrology. I: preliminary concepts," *Journal of Hydrologic Engineering*, vol. 5, no. 2, pp. 115–123, 2000.
- [2] R. S. Govindaraju, "Artificial neural networks in hydrology. II: hydrologic applications," *Journal of Hydrologic Engineering*, vol. 5, no. 2, pp. 124–137, 2000.
- [3] B. Bhattacharya and D. P. Solomatine, "Application of artificial neural network in stage discharge relationship," in *Proceedings of 4th International Conference on Hydro informatics*, Iowa, USA, 2000.
- [4] S. K. Jain and D. Chalisgaonkar, "Setting up stage-discharge relations using ANN," *Journal of Hydrologic Engineering*, vol. 5, no. 4, pp. 428–433, 2000.
- [5] S. K. Jain, "Development of integrated sediment rating curves using ANNS," *Journal of Hydraulic Engineering*, vol. 127, no. 1, pp. 30–37, 2001.
- [6] K. P. Sudheer and S. K. Jain, "Radial basis function neural network for modeling rating curves," *Journal of Hydrologic Engineering*, vol. 8, no. 3, pp. 161–164, 2003.
- [7] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representation by error propagation," in *Parallel Distributed Processing: Explorations in Microstructures of Cognition*, pp. 318–362, MIT Press, Cambridge, UK, 1996.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

