Journal of Water Sciences Research
ISSN: 2251-7405 eISSN: 2251-7413
Vol.4, No.1, Fall 2012, 71-77, JWSR

Technical note

Stream Flow Prediction in Flood Plain by Using Artificial Neural Network
(Case Study: Sepidroud Watershed)

Alireza Mardookhpour [1*]

[1]Assistant Professor, Department of civil engineering, Islamic Azad University, Lahijan, Iran

Received: 29 Sep. 2011 Accepted: 30 Jun. 2012

ABSTRACT

In order to determine hydrological behavior and water management of Sepidroud River (North of Iran-Guilan) the present study has focused on stream flow prediction by using artificial neural network. Ten years observed inflow data (2000-2009) of Sepidroud River were selected; then these data have been forecasted by using neural network. Finally, predicted results are compared to the observed data. Results showed that neural network could predict stream flow with high precision and the maximum error percentage in data prediction was about 3.

Keywords

Stream flow; Neural network; Water management; Sepidroud watershed

1. Introduction

Short term forecasts are generally applicable in real time operation of water management and flood warning (Deepika and Sharma 2011), whereas long term forecasts are applicable for water supply system management (French et al. 1992). Proper planning of hydropower generation, water storage reservoir, planning of dams, flood mitigation, potable water distribution, drainage system, and river transportation planning are dependent upon suitable prediction of watersheds input data. Non-linear relationship between input and output flow makes it difficult to predict stream flow process and a large uncertainty may exist in stream flow data (Deshmukh et al. 2008). Artificial neural network (ANN) has become one of the most common used techniques in predicting time series data which has a non-linear mathematical structure leaning on interactions between input and output data (Raman et al. 1995). It consists of data processing units (neurons) that are connected via adjustable connections (weights) (Sivanandam et al. 2006). The function is based on the connections between network elements. The network learns by applying a back propagation algorithm that compares simulated neural network outputs to the actual values and calculates a prediction error. The error is propagated back through the network and weights are adjusted as the network attempts to decrease the prediction error by optimizing the weights (Jason et al. 2009). Application of ANN for the problems

[*] Corresponding Author Email: (A.mardookh@yahoo.com)

involving rainfall-runoff modeling and river flow prediction has been investigated by several researchers (Cheng and Noguchi 1996; Hsu et al. 1995; Maier and Dandy 2000; Smith and Eli 1995). Present study is focused on stream flow prediction of Sepidroud River by using neural network.

2. Materials and methods

Many of the activities associated with planning and operation of water resources systems require future events prediction. Water resources management requires inputs from hydrological events. This is mainly in the form of estimation or prediction of the hydrological parameters. In order to predict hydrological components which are the input of water resources systems, both short term and long term stream flow events predictions are needed. In this study, ten years historical inflow data of Sepidroud River were selected for the model development. Two years inflow data of Sepidroud River has been predicted by using ANN and validated by observed data. Sepidroud watershed characteristics (North of Iran- Guilan) are shown in Table 1.

Observed inflow data at Manjil station of Sepidroud River are also shown in fig. 1. Historical time series data for stream flow prediction were taken during 10 years. Stream flow prediction was done by using ANN based on 2 years average inflows.

Table 1. Basic characteristics of Sepidroud catchment

Catchment area (km2)	Sp	10.80
catchment forested area (km2)	SL	9.84
Forestation (%)	l	90.14
river length (km)	L	6.438
Length of inflows (km)	ΣLpi	9.263
Catchment perimeter (km)	O	14.905
talweg length (km)	Lu	6.834
Max. catchment altitude (a.s.l.)	H max	1458
Min. catchment altitude (a.s.l.)	H min	569
Average catchment altitude (a.s.l.)	H ave	909.86
Average catchment width (km)	Bp	1.580
Average river slope (%)	It	15.75
Average talweg slope (%)	Iú	12.34
Average catchment slope (%)	Is	31.15

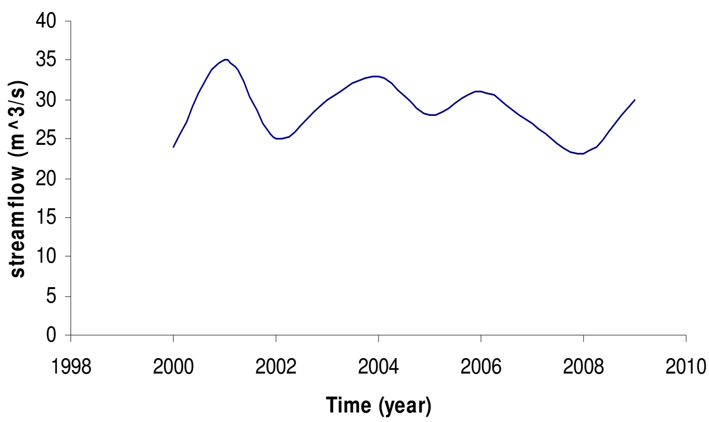

Fig. 1. Mean observed stream flow at Manjil station–Sepidroud River

2.1. Neural network overview

Neural networks are inspired by nervous systems as biological organisms. It is composed of data processing units (neurons) connected via adjustable connection weights. Neurons are arranged in layers including input layers, hidden layers, and output layers. There is no rule governing the hidden layers. A neural network is characterized by its architecture that represents the pattern of connection between nodes. It is a method for determination of the connection weights and the activation function. The function is established based on the connections between network elements. The most popular ANN architecture in hydrologic modeling is the feed forward neural network trained with a back propagation algorithm.

This network architecture and the corresponding learning algorithm can be viewed as generalization of the least –mean square algorithm. In the input layer each neuron is connected to one of the input parameters. The network learns by applying a back propagation algorithm which compares the neural network simulated output values to the actual values and calculates a prediction error. The error is then back propagated through the network.

Then weights are adjusted as the next network attempt to decrease the prediction error. The training or learning of the network occurs through cycles of data patterns to the network. The training is stopped as soon as validation error rate starts stabilizing or increasing. It is common practice to divide the data set into a learning data set to avoid the tendency of the network to memorize the training data after an extended learning phase. Moreover, if the network over learns the training data, it is more difficult for the network to produce a data set that was not seen by the network during training. Here in this paper, neural network fitting tool of MATLAB is used (Beale et al. 1996). This research employed supervised learning where the target values for the output are presented to the network to update networks weights. After training, network verification is applied in which only the input values are presented to the network so that the success of the training can organize an algorithm that trains artificial neural network 10 to 100 times faster than the usual back propagation algorithm. Levenberg-Marquardt algorithm, which is an approximation to Newton method, is selected for this research (Hagan and Menhaj 1994). In order to investigate artificial neural network suitability, three ratios between training, validation and testing sets were considered as 60:20:20, 80:10:10 and 90:05:05. fig. 2 shows the prediction accuracy by employing different numbers of hidden nodes for 90:05:05 ratio.

Hyperbolic function was applied for hidden layer according to Equation 1 and the linear transfer function was used in the output layer.

$$F(x)=\frac{1}{(1+exp(x))} \tag{1}$$

where X is stream flow parameter.
After normalization process input data was applied between -1 and +1. To evaluate neural network performance, initialization of connection weights, training, validation and testing, five independent trials were performed. Mean Square Error (MSE) that indicates the average squared difference between outputs and targets and is used to assess network performance is calculated by using equation 2.

$$MSE=\frac{(\sum_{m=1}^{M}(Ym-Dm)\hat{}2)}{M} \tag{2}$$

Where *Ym* and *Dm* are the network output and the desired output at any "m" sample, respectively, and M is the length of the investigated data set. Correlation coefficient, R also provides how well the model is close to the actual values. In other words, it provides a measure of how well future outcomes are likely to be predicted by the model. It is desired for R values to be very close to 1.

3. Results and discussions

In order to choose the best prediction model three different data ratios (60:20:20, 80:10:10 and 90:05:05) were selected. Five different random trails with two, three, four, and five neurons in the hidden layer have been conducted by utilizing Levenberg-Marquardt algorithm and MATLAB tool box software.

In order to predict stream flow data in Sepidroud watershed by using neural network, the use of Levenberg –Marquardt algorithm has been investigated. The network was trained by MATLAB software and the best three ratios between training, validation and testing sets were considered as 90:05:05 by training. To check sensitivity of the neural network, 5 independent random training, testing and validation trials have been performed to achieve the best prediction accuracy. fig. 2 shows prediction accuracy of the model and table 2 shows MSE and R values.

Training, validation and testing operations were performed several times through the analysis and the best fitted model with minimum MSE was chosen. Each training session was carried out with different initial weights. The best pre-dictions accuracies were obtained by 5 neurons in the hidden layer on fifth trial. The best chosen value of R for the model was 0.999 that showed a perfect fit in which outputs and targets were approximately equal to each other. Stream flow prediction results for 10 years with data ratio of 90:05:05 are shown in Table 3 and fig. 3.

Table 2. Performance evaluation of training, validation and testing the data

Number of hidden layer neurons	operation	MSE	R
2	Training	124	0.741
2	Validation	106	0.750
2	Testing	53	0.800
3	Training	82	0.880
3	Validation	43	0.890
3	Testing	22	0.890
4	Training	29	0.975
4	Validation	16	0.978
4	Testing	13	0.980
5	Training	28	0.999
5	Validation	22	0.989
5	Training	18	0.991

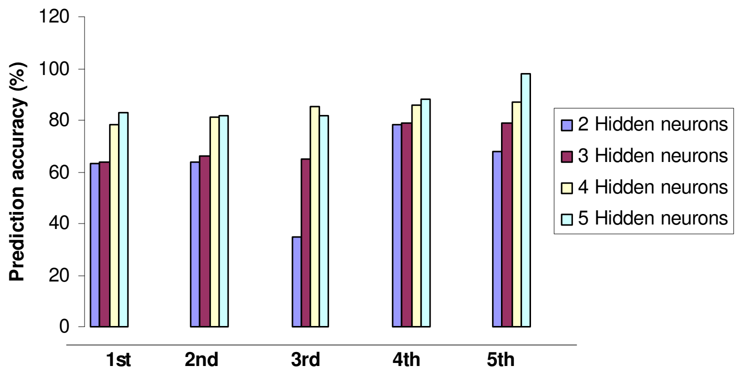

Fig. 2. Prediction accuracy of various trials with different hidden neuron numbers

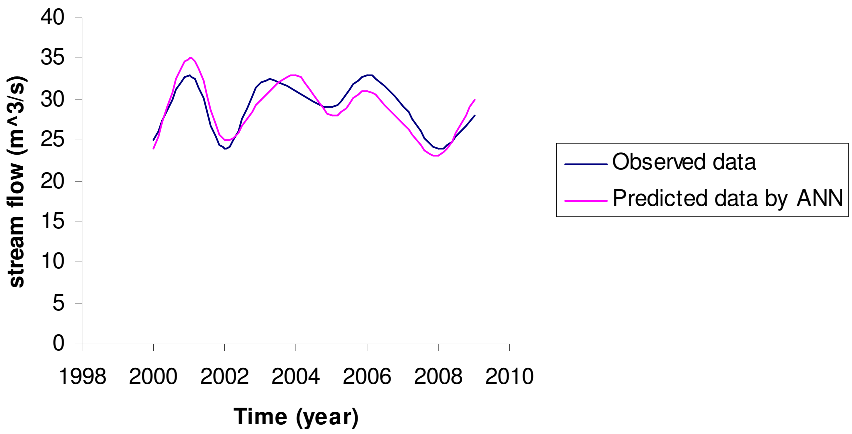

Fig. 3. Comparison of the best predicted results by using ANN model to the observed data

Table 3. Prediction error percentages in ANN model with data ratio of 90:05:05

Ratio	PME	Accuracy (%)
90:05:05	0.0015	97

If Zt and Zi are predicted data and actual observed data, respectively, prediction error would be as:

$$Ei = Zi - Zt \quad (3)$$

Therefore, mean error percentage is obtained from Equation 4 as follows:

$$PME = (\frac{1}{n}) \times (\sum_{i=1}^{n} (\frac{Ei}{Zi}) \quad (4)$$

According to table 3 and fig. 3, the accuracy of stream flow prediction is about 97 percent. Thus, the ANN prediction model would be useful in evaluating the performance of inflow and various operating policies.

Standard deviation is a measure that summarized the amount by which every value in a data set varies from the mean. It indicates how tightly the values in data set are bunched around the mean value. The more tightly is the bunch, the less is the standard deviation. MSE, the mean squared errors and R, correlation coefficient of data derived from the standard deviation are shown in table 2. Comparison of the Mean Square Error (MSE) values indicates that the difference between outputs (targets) and inputs, which is used to evaluate the model, is very low. Table 3 shows how well the model results are close to the actual values. In other words, it provides a measure of how well future outcomes are likely to be predicted by the model. In this study, value of R for neuron layer with 5 hidden layers and MSE of 28 is achieved as 0.999 for the best chosen model on the basis of Levenberg –Marquardt algorithm. Validation set errors is monitored during training. In this study input-output pair's reliability for the training test is up to 5% of error. When the model started to over-fit training stage, resulted error was 3%. This indicates that the network is operating very well. Predicted stream flow data by using neural network are compared to stream flow observed data in fig. 3. According to this fig., there is a good convergence between these two data sets so that the error due to ANN reached 3% approximately. Therefore, ANN model is able to significantly predict inflow data. Prediction error is less than 3% that indicates the chosen model for prediction of inflow in Sepidroud River has magnificent convergence to observed data.

This long-term inflow prediction would help in evaluating the performance of flood plain management especially in large watersheds which has missing data in seasonal flows, whereas there is no need for frequent updating of the parameters. In order to check the precision of ANN model, 90:05:05 ratios were analyzed for training, validation and testing which predicted stream flow quite near to the observed data flows. In summary, ANN model would be helpful in evaluating real time management polices and provides a probabilistic frame-work for the decision support systems.

4. Conclusions

In this study, stream flow prediction in long-term time series is investigated by using ANN model. Results indicate that ANN predicting model is an applicable tool in predicting inflows, especially in long-term condition. In order to check the sensitivity of the ANN prediction model, three different ratios were analyzed including 60:20:20, 80:10:10 and 90:05:05 for training, validation and testing. Results indicated that lower ratios give better results with higher accuracy and predicted flows were quite near to the observed flows. In order to evaluate the accuracy of ANN model, observed and predicted data were compared and analyzed. The best accuracy of the model was about 97% in Sepidroud River. It indicates that ANN prediction model is an applicable tool in predicting flood plain inflows especially in long-term conditions.

Acknowledgement,

The writer wishes to thank the Islamic Azad University, Lahijan Branch, Iran, for preparation of the facilities and used software.

References

Beale M., Demuth B., Hagan, M. (1996) Neural network design. PWS publishing company, Boston Massachusetts, USA

Cheng X., Noguchi M. (1996). Rainfall-runoff modeling by neural network approaches", Proc. Of conf. on water resources and environment research towards the 21st century, 29-31 October Kyoto, Japan, 2:143-150.

Deepika Y.K., Sharma V. (2011) Stream flow forecasting using Levenberg- Marquart algorithm approach. International journal of water resources and environment engineering, 3(1): 30-40.

Deshmukh S.R., Doke D.J., Nerkar Y.P. (2008) Optimal generation scheduling under ABT using forecasted load and frequency. Proceedings of POWERCON 2008 and 2008 IEEE Power India conferences (joint Int.Conf.), 1-6.

French M.N., Krajewski W.F. Cuykendall R. (1992) Rainfall forecasting in space and time using a neural network", J. Hydro., 137:1-4.

Hagan M., Menhaj M. (1994) Training feed forward networks with Levenberg-Marquardt algorithm. IEEE Trans Neural Networks, 5(6): 989-993

Hsu K., Gupta HV, Sorooshian S. (1995) Artificial neural network modeling of the rainfall-runoff process. Water Resour Res., 31(10):2517-2530

Jason H., Orlando D.j, Martin T. (2009) Spurious Valleys in the error surface of recurrent networks-analysis and avoidance ",IEEE Trans Neural Networks, 20(4):686-700.

Maier H., Dandy G. (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications, environmental modeling software, 15:101-124

Raman H., Suni kumar N. (1995) Multivariate modeling of water resources time series using artificial neural networks, Journal of

Hydrological Sci, 40(2):145-163.

Sivanandam S.N., Sumathi S. Deepa S. (2006) Introduction to neural networks using MATLAB 6.0. Mc Graw-hill.

Smith J., Eli R. (1995) Neural network models of rainfall-runoff process, Journal of Water Resour.Plann.Mannage, 121(6):499-508.