

Forecasting River Flow in the USA: A Comparison between Auto-Regression and Neural Network Non-Parametric Models

¹Abdel karim M. Baareh, ²Alaa F. Sheta and ³Khaled Al Khnaifes

¹Computer Science Department Al-Balqa Applied University Ajlune College, Jordan

²Information Technology Department Al-Balqa Applied University Al-Salt, Jordan

³Mathematics Department Damascus University Damascus, Syria

Abstract: Forecasting a time series became one of the most challenging tasks to variety of data sets. The existence of large number of parameters to be estimated and the effect of uncertainty and outliers in the measurements makes the time series modeling too complicated. Recently, Artificial Neural Network (ANN) became quite successful tool to handle time series modeling problem. This paper provides a solution to the forecasting problem of the river flow for two well known Rivers in the USA. They are the Black Water River and the Gila River. The selected ANN models were used to train and forecast the daily flows of the first station no: 02047500, for the Black Water River near Dendron in Virginia and the second station no: 0944200 for the Gila River near Clifton in Arizona. The feed forward network is trained using the conventional back propagation learning algorithm with many variations in the NN inputs. We explored models built using various historical data. The selection process of various architectures and training data sets for the proposed NN models are presented. A comparative study of both ANN and the conventional Auto-Regression (AR) model networks indicates that the artificial neural networks performed better than the AR model. Hence, we recommend ANN as a useful tool for river flow forecasting.

Key words: River flow forecasting, Feed forward Neural Networks, Auto-Regression model.

INTRODUCTION

Water is the sources of life for all creatures. Rivers flow forecasting can protect from water shortage, flood damage and in agriculture management. Different models have been proposed for forecasting the daily flow of Rivers^[1,2,3]. Linear prediction model (LP)^[4,5] such as Auto-Regressive and Neural Network models were used in variety of forecasting problems^[6]. Selecting a suitable model for forecasting is very complicated and difficult process. These difficulties include the data availability, the size of the basins of interest and the different sensing and measuring instruments being used.

Recently, artificial neural networks have been introduced as a useful tool which can be used for modeling hydrologic processes. ANN showed a strong capability in handling diversity of problems including rainfall-runoff, water quality, sedimentation and rainfall forecasting. It has been also an efficient and experimented model widely used in number of applications^[7,8] such as the sales prediction^[9], shift failures^[10], estimating prices^[11] and stock returns^[12].

In this paper, we are presenting yet another neural network forecasting application, namely the river flow forecasting of the Black Water River and the Gila River

in the USA. The proposed NN models have been developed and evaluated for its performance for forecasting the river flow of two rivers in the USA. Many authors were able to develop a variety of NN models to solve river flow forecasting problem^[13,14,15]. We are investigating the use of the AR and NN based regression model in solving such a problem. The proposed NN was also trained using the Levenberg Marquardt technique to provide better forecasting capabilities^[16].

Artificial neural network: Artificial Neural Network (ANN) is an information processing paradigm that is inspired from biological nervous systems, such as the brain process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in union to solve specific problem. Neural Network can derive meaning from complicated or imprecise data. There are many types of Neural Network but Back-propagation Neural Networks are the most famous neural type^[17,18].

There are mainly three different types of layers presented in most ANNs. The first layer is called the input layer. Its main task is to receive input from the outside world. This layer has number of neuron equal to the number of model input. The layer next to the input

layer is called the hidden layer. This layer is receiving input from the immediately preceding layers. The final layer of the network is called the output layer. The neuron present in this layer presents the output of the network. Neurons in any layer are fully connected to all neurons in the next layer. The neurons in the same layer are not connected among each other. A weighted sum of the neuron inputs specifies the activation (i.e. sigmoid) function argument. This activation function is assumed to be nonlinear.

An example of a three layer feed forward ANN is shown in Fig. 1. The shown network structure can simulate the behaviors of a model which has four inputs and one output. The number of neurons in the hidden layer depends on the problem complexity.

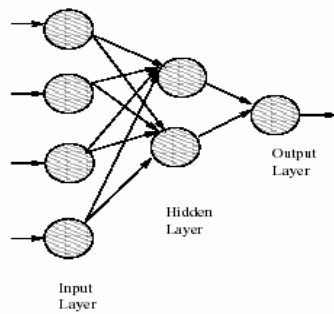


Fig. 1: An example of a Feedforward NN structure

The study area: The data flow were recorded and collected from two stations operated by the U.S. Geological Survey (USGS). The 1st station No: 02047500, for the Black Water River isolated near Dendron, Virginia^[1] and 2nd station No: 09442000 for the Gila River is located near Clifton, Arizona^[1]. The location of these stations is shown in Fig. 2.

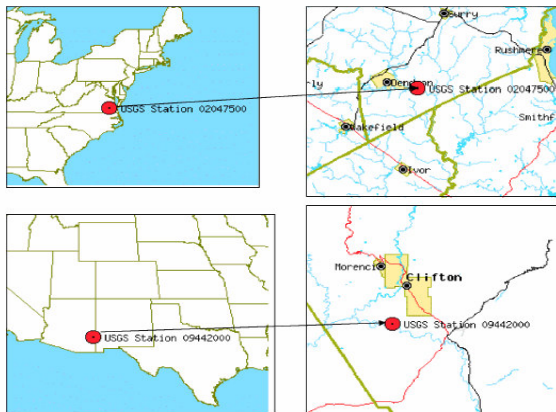


Fig. 2: The locations of the stations operated by the USGS. This map was presented in^[1]

For the 1st station the training data period was from 01 Oct 1990 to 30 Sept 1996, (six water years) and tested data were from the period of 01 Oct 1996 to

30 Sept 1997 (one water year). For the 2nd station the training data period was from 01 Oct 1995 to 30 Sept 1998 (three water years) and the tested data were from the period of 01 Oct 1998 to 30 Sept 1999 (one water year).

Problem formulation: In our case two models, the linear Auto-Regression model and the Backpropagation model, were used to predict the future flow for both the Black Water and Gila Rivers. Both models were trained and tested on different set of data.

We used the Sum Square of Errors (SSE) as the evaluation criterion for the developed models. The SSE was computed for both training and testing cases.

$$SSE = \sum_{k=i}^m \left| \hat{y}(k) - y(k) \right|^2 \quad (1)$$

$y(k)$ is the real flow measurements and $\hat{y}(k)$ is the estimated river flow. It is very important to minimize the error difference between the real and estimated flow such that we have an accurate model.

Developed AR model for flow forecasting: Auto-Regressive model is one of the most well known models among traditional linear models. It was used for variety of modeling applications. The model is described by Equation 1.

$$y(k) = a_0 + \sum_{i=1}^n a_i y(k-i) \quad (1)$$

$y(k)$ is the flow at particular day k , $y(k-1)$ is the previous day flow. Auto-Regressive model parameters are estimated using Least-Square Estimation (LSE). Minimizing the Sum Square of Errors (SSE) helps in proving the best set of parameters for the AR model.

We developed an AR model for predicting the flow of the Black Water River and Gila River. We explored many model orders in each case. The AR model parameters were estimated using the Least Square Estimation (LSE) to minimize the SSE between the actual flow and the predicted flow based the developed models.

The developed AR model for the Black Water River is given by Equation 3. The model parameters were estimated using Least Square Estimation (LSE).

$$\begin{aligned} \hat{y}(k) = & 10.6851 + 1.7055 * y(k-1) \\ & - 1.2359 * y(k-2) + 0.8573 * y(k-3) \\ & - 0.6173 * y(k-4) + 0.3220 * y(k-5) \\ & - 0.1468 * y(k-6) + 0.0714 * y(k-7) \end{aligned} \quad (3)$$

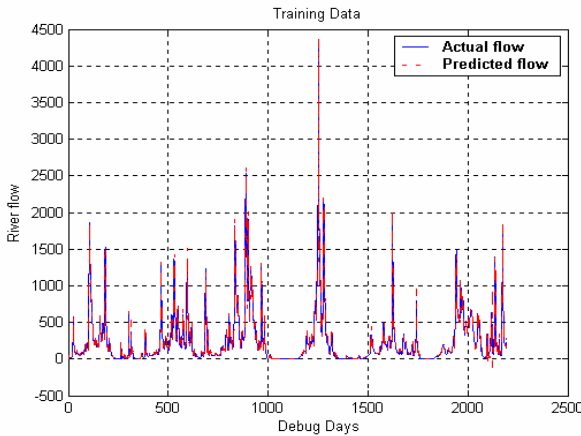


Fig. 3a: Actual and predicted flow AR7 training case: Black Water River

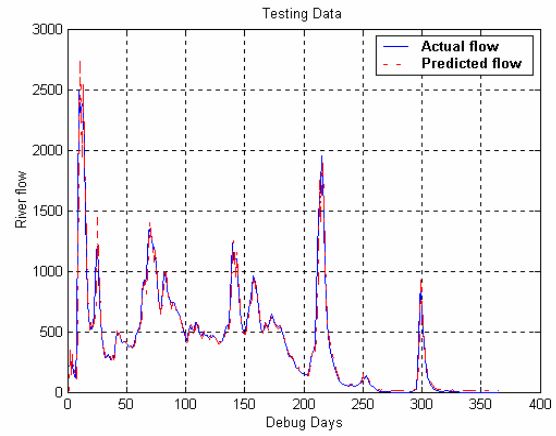


Fig. 3b: Actual and predicted flow AR7 testing case: Black Water River

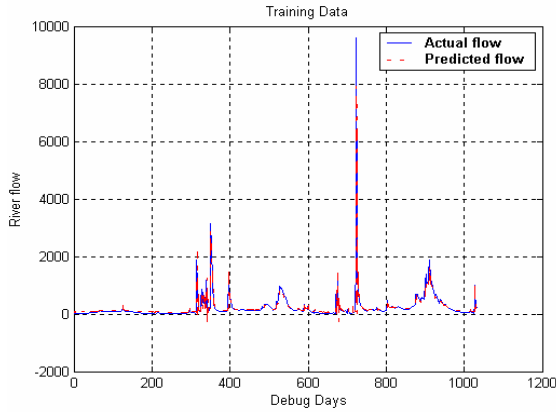


Fig. 4a: Actual and predicted flow AR5 training case: Gila river

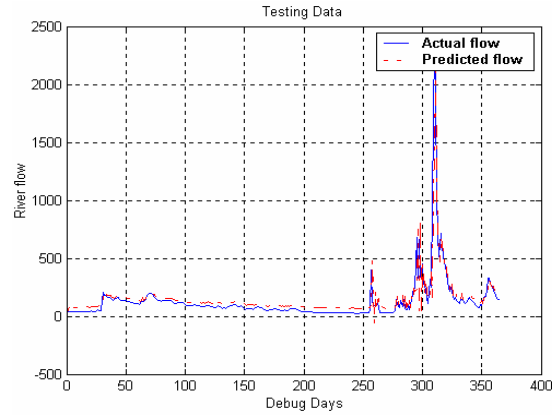


Fig. 4b: Actual and predicted flow AR5 testing case: Gila River

In Fig. 3a, we show the actual flow with solid line and the estimated flow with the dotted line, based on AR7 model for the training period of Black Water River. The developed model was validated as shown in Fig. 3b.

The developed model for the Gila River is given in Equation 4.

$$\begin{aligned} \hat{y}(k) = & 47.2256 + 1.1571*y(k-1) \\ & -0.6574*y(k-2) + 0.3652*y(k-3) \\ & -0.1370*y(k-4) + 0.0692*y(k-5) \end{aligned} \quad (4)$$

In Fig. 4a and b we show the results for both training and testing (i.e. validation) cases of the Gila River used AR5 model. The actual flow is shown with the solid line and the estimated flow with the dotted line.

Developed neural network model for flow forecasting: Our developed Neural Network model was used to estimate the flow for the two River flows. A number of networks were implemented with different number of input delays were used.

The proposed NN consists of three layers. The input layer contains number of neurons varied from 3 to 7 based on the developed model order. The hidden layer has seven hidden nodes. This number was an arbitrary chosen. Weights from the Input-to-hidden layer and hidden-to-output layer were adjusted using back-propagation learning algorithms. The output layer consists of one output neuron to produce the prediction of the flow.

The Network was trained using the BP algorithm. To develop our results we used the NNSYSID Matlab toolbox^[17]. The adopted NN was the Neural Network Auto-Regression Matrix model (NNARX). We used the Levenberg Marquardt technique to train the neural network^[16]. In Fig. 5a, we show the actual flow in solid line and the estimated flow with the dotted line, based on NN7 model for the training period of Black Water River. The developed model was validated as shown in Fig. 5b. In Fig. 6a and Fig. 6b, we show the results for both training and testing cases of the Gila River used NN5 model.

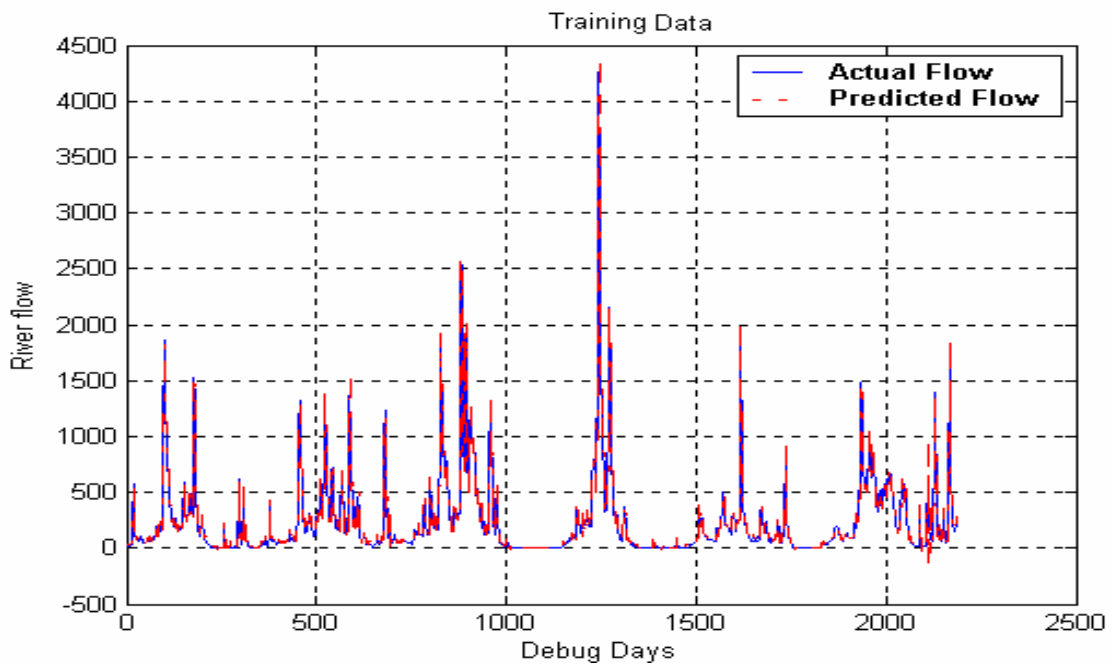


Fig. 5a: Actual and predicted flow NN7 training case: Black Water River

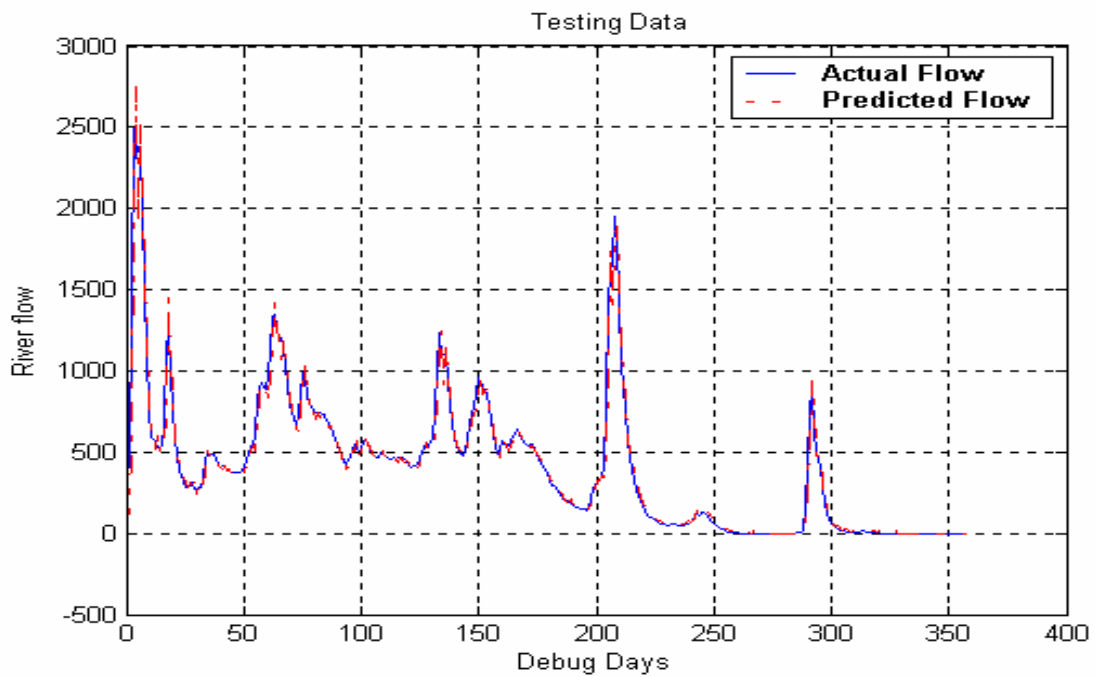


Fig. 5b: Actual and predicted flow NN7 testing case: Black Water River

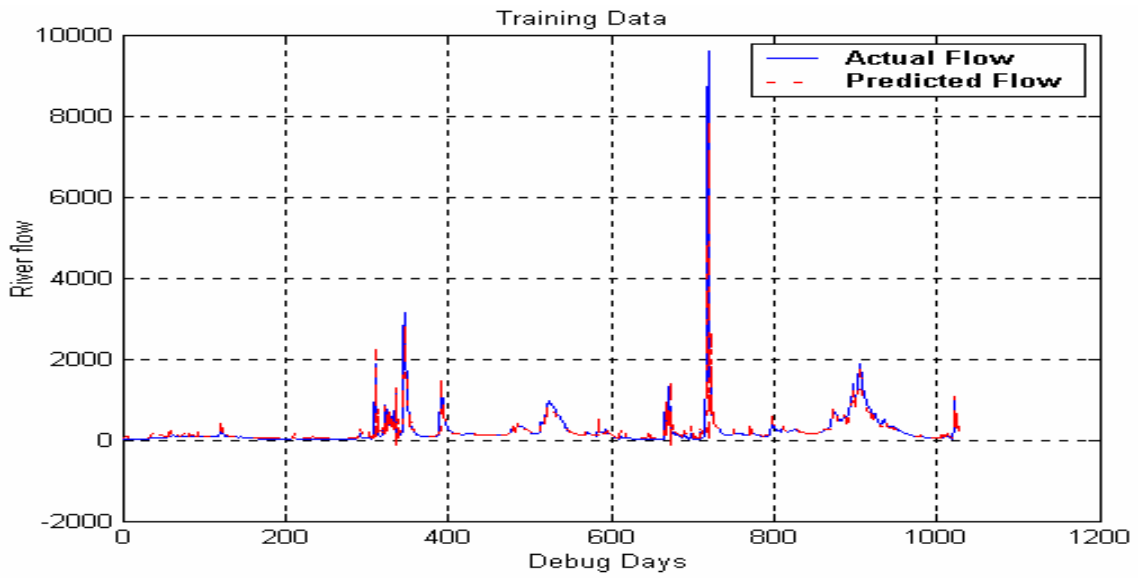


Fig. 6a: Actual and predicted flow NN5 training case: Gila River

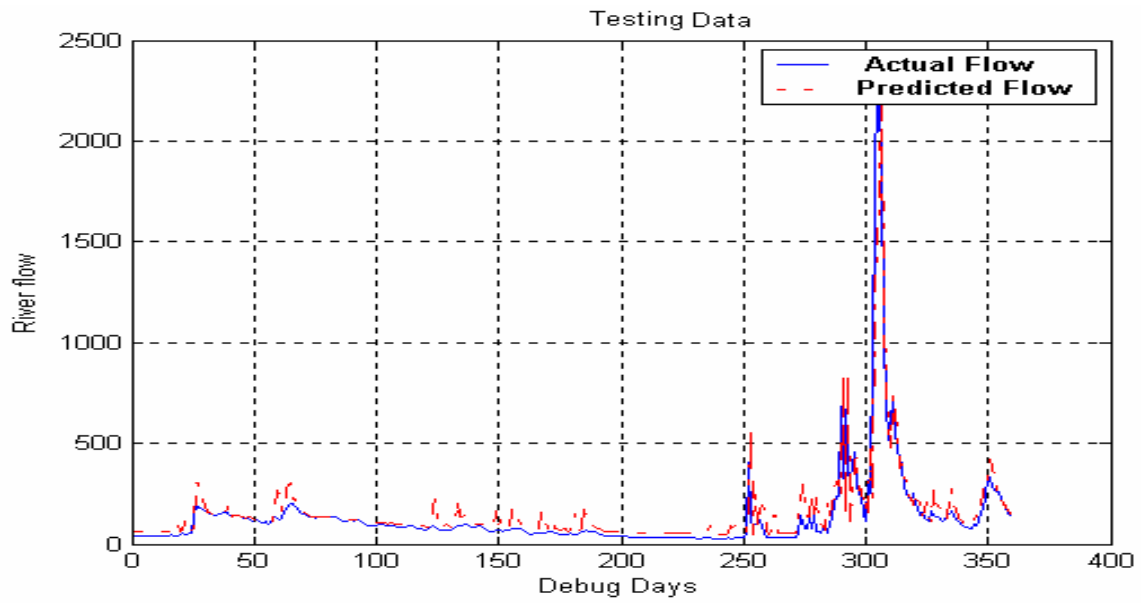


Fig. 6b: Actual and predicted flow NN5 testing case: Gila River

In Table 1 and 2 we show the values of the SSE in the two cases under study. In the case of the Black Water River, Table 1, model with order seven provided the best result since the error in both training and testing cases were minimum and less than that in the case of the AR model.

No. of Inputs	Training Data SSE	Validation Data SSE	Both Sum
NN(3)	1.57E+07	3.12E+06	1.89E+07
NN(4)	1.52E+07	3.02E+06	1.83E+07
NN(5)	1.45E+07	2.88E+06	1.74E+07
NN(6)	1.44E+07	2.83E+06	1.72E+07
NN(7)	1.45E+07	2.81E+06	1.73E+07
AR(3)	1.59E+07	3.21E+06	1.91E+07
AR(4)	1.52E+07	3.07E+06	1.83E+07
AR(5)	1.47E+07	2.94E+06	1.76E+07
AR(6)	1.46E+07	2.95E+06	1.76E+07
AR(7)	1.46E+07	2.92E+06	1.75E+07

Table 1: SSE for ANN and AR models-training and testing data of the Black Water River

In the case of the Gila River, order five provided the best result since the error in both training and testing cases was minimum and less than that in AR model.

No. of Inputs	Training Data SSE	Validation Data SSE	Both Sum
NN(3)	7.82E+07	3.99E+06	8.22E+07
NN(4)	7.80E+07	4.10E+06	8.21E+07
NN(5)	7.55E+07	4.18E+06	7.97E+07
NN(6)	7.69E+07	4.14E+06	8.10E+07
NN(7)	7.70E+07	4.21E+06	8.12E+07
AR(3)	7.82E+07	3.99E+06	8.22E+07
AR(4)	7.80E+07	4.12E+06	8.21E+07
AR(5)	7.76E+07	4.15E+06	8.18E+07
AR(6)	7.76E+07	4.14E+06	8.18E+07
AR(7)	7.75E+07	4.12E+06	8.16E+07

Table 2: SSE for ANN and AR models-training and testing data of the Gila River

CONCLUSION

In this study, we presented a detailed comparison between Artificial Neural Networks and the Auto-Regression models in solving the River flow forecasting problem. We concluded that neural networks can offer several advantages over conventional modeling approaches. The most important among them is their ability to develop a generalized solution to the forecasting problem from a given set of examples. We showed that ANN models can be used to train and forecast the daily flows of the Black Water River near Dendron in Virginia and the Gila River near Clifton in Arizona. ANN model were found to perform better for forecasting daily river flow than the conventional AR model.

REFERENCES

1. Kisi, O., 2005. Daily river flow forecasting using artificial neural networks and auto-regressive models. *Turkish J. Eng. Env. Sci.*, pp: 9-20.

2. Sheta, A.F. and M.S. El-Sherif, 1999. Optimal prediction of the Nile river flow using neural networks. *International Joint Conference on Neural Networks, IJCNN '99*, Volume: 5, pp. 3438-3441, Washington D.C.
3. Cigizoglu, H.K. and Kisi, O., 2005, Flow Prediction by Three Back Propagation Techniques Using k-fold Partitioning of Neural Network Training Data", *Nordic Hydrology*, vol. 36.
4. Kothiyari, U.C., V. Aravamuthan and V.P. Singh, 1993. Monthly run-off generation using linear perturbation model. *J. Hydrol.*, 144 : 371-379.
5. Said, R., 1993. *The Nile River*. Pergamon Press, Oxford.
6. Hsu, K., Gupta, H.V. and Sorooshian, S., 1995, Artificial Neural Network Modeling of the Rainfall, Runoff Process, *Water. Resour. Res.*, 31, 2517-2530.
7. Karunanithi, N., W. Grenney, D. Whitley and K. Bovee, 1993. Neural networks for river flow prediction. *J. Computing in Civil Engg.*, 8: 371-379.
8. Bulando, P., R. Rosso and J. Salas, 1993. Forecasting of short-term rainfall using ARMA models. *J. Hydrol.*, 144: 193-211.
9. Hruschka, H., 1993. Determining market response functions by neural networks modeling: A comparison to econometric techniques. *Eur. J. Oper. Res.*, 66: 867-888.
10. Li, E.Y., 1994. Artificial neural networks and their business applications. *Information and Managements*, 27: 303-313.
11. Chakraborty, K., 1992. Forecasting the behavior of multivariable time series using neural networks. *Neural Networks*, 5: 962-970.
12. Swales, G. and Y. Yoon, 1992. Applying artificial neural networks to investment analysis. *Financial Analyst J.*, pp: 78-82.
13. Tan Danh, N., H. Ngoc Phien and A. Das Gupta, 1999. Neural network models for river flow forecasting. 25: 1.
14. Adya, M. and F. Collopy, 1998. How effective are neural networks at forecasting and prediction. *J. Forecasting*, pp: 481- 495.
15. Baratti, R. and B. Cannas, 2003. River flow forecast for reservoir management through neural networks. *Neurocomputing*, pp: 421-437.
16. M. Norgaard, *Neural Network Based System Identification Toolbox (NNSYSID)*. Technical University of Denmark, 2000
17. Kisi, O., 2004. River flow modeling using artificial networks. *ASCE J. Hydrol. Engg.*, 9: 60-63.
18. Cigizoglu, H.K. and O. Kisi, 2005. Flow prediction by three back-propagation techniques using k-fold partitioning of neural network training data. *Nordic Hydrol.*, vol. 36.