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

Abdel karim M. Baareh
Computer Science Department
Al-Balqa Applied University
Ajlune College, Jordan

Alaa F. Sheta
Information Technology Department
Al-Balqa Applied University
Al-Salt, Jordan

Khaled Al Khnaifes
Mathematics Department
Damascus University
Damascus, Syria

Abstract: This paper provides a solution to the forecasting problem of the river flow for two well known rivers in USA. They are the Black Water River and the Gila River. Feed-forward Neural Network and the Linear Auto-Regressive (AR) models were used to model the flow dynamics. The performance of the two proposed model were compared in both training and testing cases. The model performances were computed in each case. NN model showerd a better modeling capability compared to the AR model.

Key words: River flow forecasting, Feedforwad Neural Networks, Auto-Regression

INTRODUCTION

Water is the life for all creatures. Models for River flow forecasting are a fundamental tool in water resource studies. Forecasting the future in a short term and long term format helps in planning, maintaining, managing, and controlling future events. Forecasting the future events plays an important role in many fields. They include economics, business, military, rivers and many others. Forecasting is an important and animating problem for many researchers in different fields. Rivers flow estimation can protect from water shortage, flood damage, and help in agriculture management. Different models have been proposed for forecasting the daily flow of Rivers [1, 2]. Linear prediction model (LP) [3, 4] such as Auto-Regressive model and Neural Network models were used in a variety of forecasting problems [5]. 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. Neural Networks is an efficient and experimented model widely used in number of applications [6, 7] such as the sales prediction [8], shift failures [9], estimating prices [10] and stock returns [11].

In this paper, two models were used to predict the Black Water and Gila River flow in USA. An Auto-Regressive order seven model for the Black Water and an Auto-Regressive order five model for the Gila River are used as an example of linear models. Auto-Regressive model parameters are estimated using Least-Square Estimation (LSE), multiplying these parameters by the AR matrix, computing the new data flow and finding the Sum Square of Errors (SSE) [12].

Feed-forward Neural Network model of three layers input, hidden and output was constructed and trained using Back-propagation algorithm to forecast the Black Water and Gila River flow. A number of networks were implemented with different number of input delays. Both models were trained and tested on different set of data.

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 [13, 14]. 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 unison to solve specific problem. Neural Network, with their remarkable ability to 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 [15, 16]. They have signals traveling in both directions by introducing loops in network. Errors are propagating backwards from the output nodes to the inner nodes again and again until the weights get adjusted. Feedback networks are very powerful and can get extremely complicated. Feedback networks are dynamic. Their state is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and new equilibrium needs to be found. Back-propagation NN have been successfully applied to a number of differs fields, and proves its efficiency. A 3-layer, Back-propagation NN is shown in Fig. 1. The Neural Network has input layer, hidden layer (middle layer), and output layer. Neurons in any layer are connected to all neurons in the next layer. The neurons in the same layer are not connected among them self.

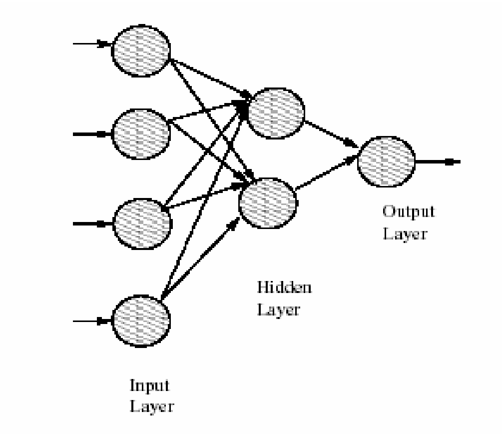

Fig.1: Feedforward Neural Network Structure

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 near Dendron in Virginia [1] and 2nd station No: 0944200 for the Gila River near Clifton in Arizona [1]. The location of these stations is shown in Figure 2.

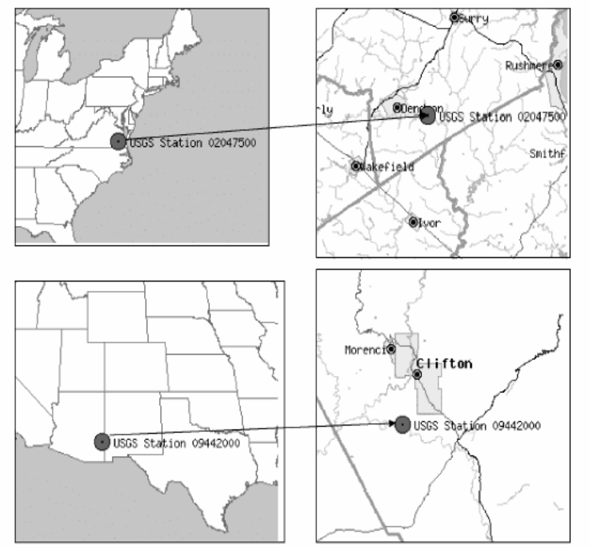

Fig.2: The locations of the stations operated by the USGS. This map was presented in [1].

For the 1st station the training data period was from 01 Oct 1990 to 30 Sept 1996, (6 water years) and tested data were from the period of 01 Oct 1996 to 30 Sept 1997 (1 water year). For the 2nd station the training data period was from 01 Oct 1995 to 30 Sept 1998 (3 water year) and the tested data were from the period of 01 Oct 1998 to 30 Sept 1999 (1 water year).

PROBLEM FORMULATION

In our case two models Linear Auto-regression model and the Back-propagation 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.

A. LINEAR AUTO-REGRESSION MODEL

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.

$$y(k) = a_{\circ} + \sum_{i=1}^{n} a_i y(k-i) \qquad (1)$$

y(k) is the flow at particular day *k*, *y(k-1)* is the previous day flow.

B. PERFORMANCE MEASUREMENTS

We have 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.

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

y(k) is the real flow measuerments and $\hat{y}(k)$ is the estimated flow measuerments.

DEVELOPED AR MODEL FOR FLOW FORECASTING

We developed an AR model for predicting the flow of the Black Water River and Gila River. We expored many model orders in each case. The AR model parameters were estimated using the Least Square Estimation (LSE).

The devloped model for the Black Water river is given by Equation 3.

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

The devloped model for the Gila river is given in Equation 4.

Corresponding Author: Alaa Sheta, Department of Information Technology, Al-Balqa Applied University, Jordan

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

In Figure 3 (a), 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 Figure 3 (b). In Figure 4 (a) and (b) we show the results for both testing and testing cases of the Gila river used AR5 model.

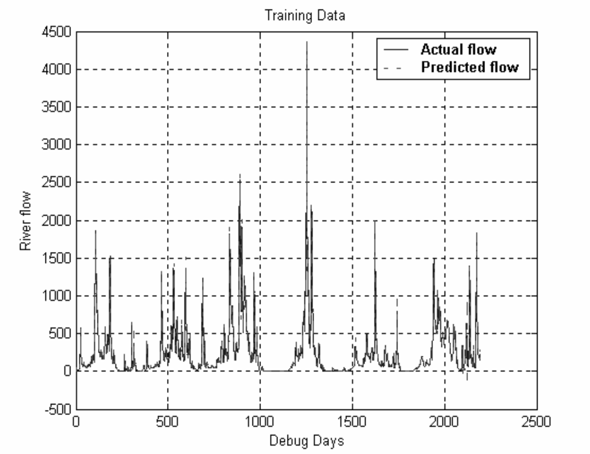

Figure 3: (a) Actual and predicted flow AR7 Training case: Black Water River

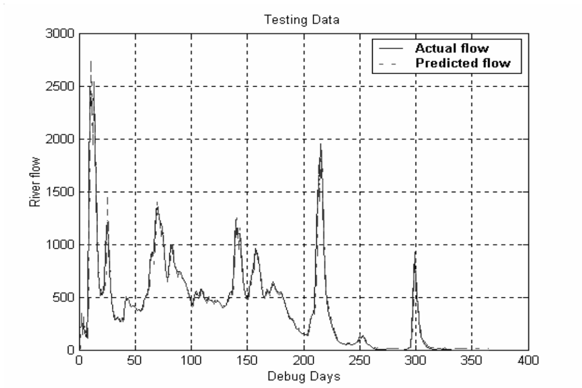

Figure 3: (b) Actual and predicted flow AR7 Testing case: Black Water River

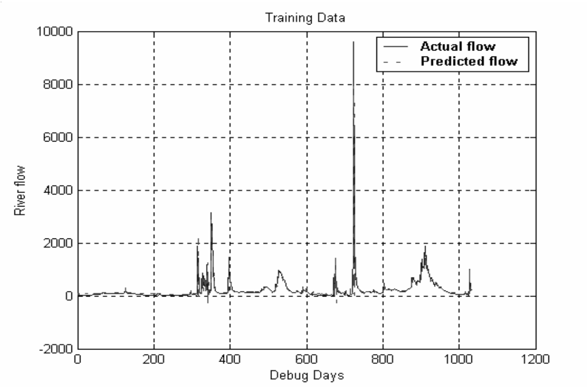

Figure 4: (a) Actual and predicted flow AR5 Training case: Gila River

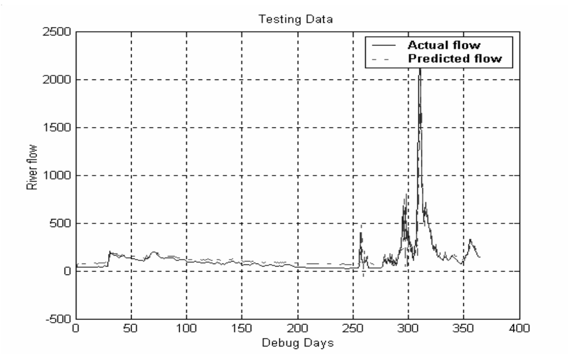

Figure 4: (b) Actual and predicted flow AR5 Testing case: Gila River

DEVELOPED NEURAL NETWORK MODEL FOR FLOW FORECASTING

Our developed Neural Network model was used to estimate the flow for the two Rivers. The proposed NN consists of three layers. They are the input layer, the hidden layer and the output layer. 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 one.

Weights from the Input-to-hidden layer and hidden-to-output layer were adjusted using backpropagation learning algorithms. The output layer consists of one output neuron to produce the prediction of the flowThe Network was trained using the BP algorithm. Table 1 and Table 2, show the values of the SSE in the two cases under study.

Corresponding Author: Alaa Sheta, Department of Information Technology, Al-Balqa Applied University, Jordan

J. Computer Sci., 2 (2): xx-xx, 2006

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

No. Of Inputs	Training Data	Validation Data	Training / Testing Sum
	SSE	SSE	
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

It can be seen, that in the case of the Black Water River, order seven provided the best result since the error in both training and testing cases was minimum.

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

No. Of Inputs	Training Data	Validation Data	Training / Testing Sum
	SSE	SSE	
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

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

The minimum sum of the training and testing results was the evaluation criteria for pointing and choosing the best model order in both the NN and AR models. The results obtained shows that Neural Network model performance is better than the Auto-Regression model performance.

In Figure 5 (a), we show the actual flow with 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 Figure 5 (b).

In Figure 6 (a) and (b) we show the results for both testing and testing cases of the Gila river used NN5 model.

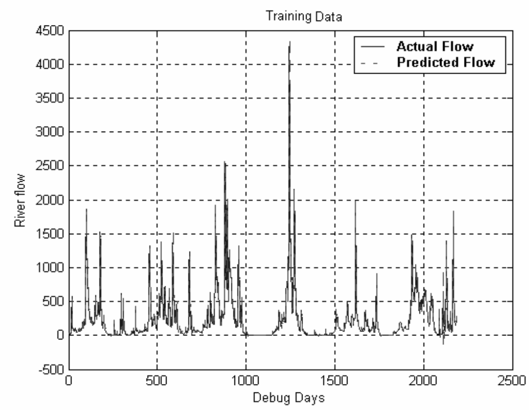

Figure 5: (a) Actual and predicted flow NN7 Training case: Black Water River

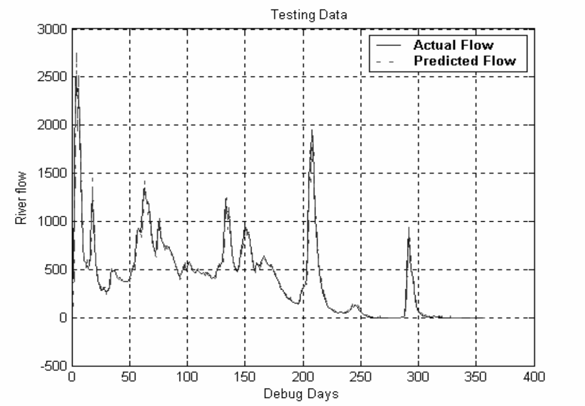

Figure 5: (b) Actual and predicted flow NN7 Testing case: Black Water River

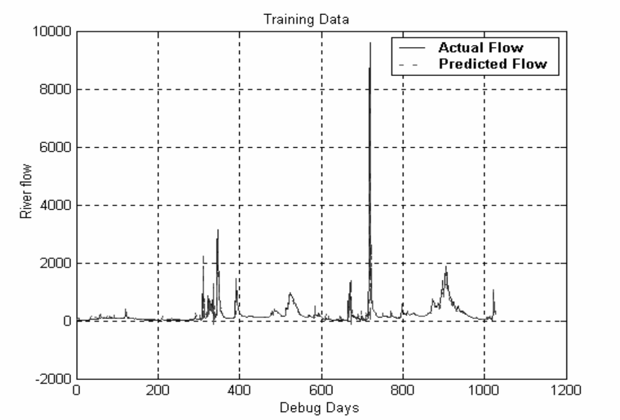

Figure 6: (a) Actual and predicted flow NN5 Training case: Gila River

Corresponding Author: Alaa Sheta, Department of Information Technology, Al-Balqa Applied University, Jordan

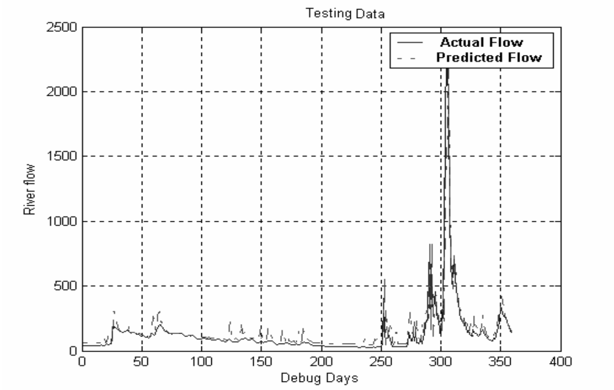

Figure 6: (a) Actual and predicted flow NN5 Testing case: Gila River

NEURAL NETWORK CONVERGANCE

The Error convergence of the backprogatation learning algorithsm in the cases studeies in this article is presented in Figure 7 and Figure 8. It can be seen that NN convergence successfully to the minimum error difference between the actual and estimated responses.

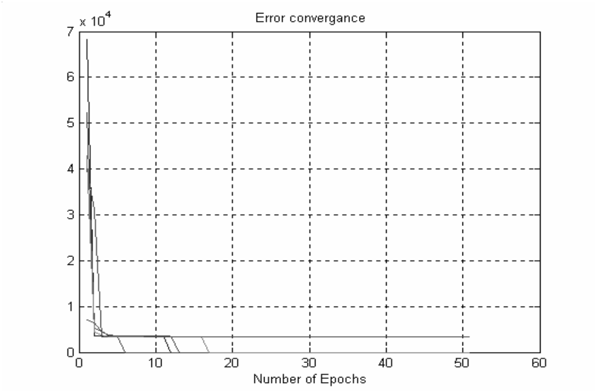

Figure 7: Back propagation error convergence for models of order 3 to 7: Black Water River

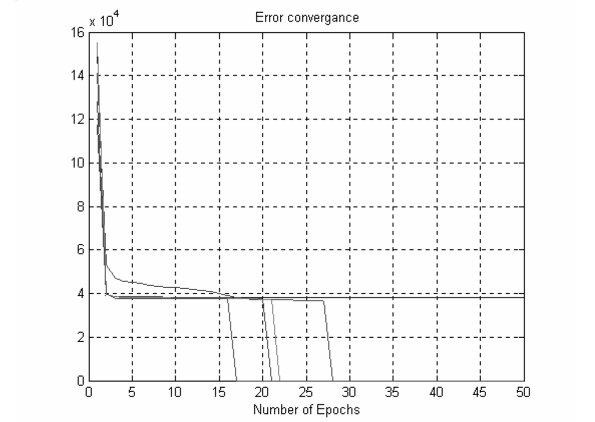

Figure 7: Back propagation error convergence for models of order 3 to 7: Gila River

From the developed results, we can see that the Neural Network model has more advantage than the traditional linear Auto-Regressive model in predicting the flow of the Black Water and Gila Rivers in USA.

CONCLUSION

In this paper, we presented a detailed comparison between AR and NN models in solving the river flow forecasting problem. They are the Black Water River and the Gila River. Feed-forward Neural Network and the Linear Auto-Regressive (AR) models were used to model the flow dynamics. The performance of the two proposed model were compared in both training and testing cases. The model performance was computed in each case. NN model showed better modeling capability compared to the AR modelIt was found that NNs provides small error reduction result compared to the AR model.

REFERENCES

[1] O. Kisi, "Daily River Flow Forecasting Using Artificial Neural Networks and Auto-Regressive Models" Turkish J. Eng. Env. Sci, pp. 9-20, 2005.

[2] A. F. Sheta, M. S. El-Sherif, "Optimal Predection of the Nile River Flow Using Neural Networks", In the Proceedings of the Joint Conference on Neural Network, Washington, D.C., July, 1999.

[3] U. C. Kothyari, V. Aravamuthan, and V. P. Singh, "Monthly run-off generation using linear perturbation model," Journal of Hydrology, Vol. 144, pp. 371-379, 1993.

[4] R. Said, The Nile River. Pergamon Press, Oxford, 1993.

[5] B. Cannas, A. Montisci, A. Fanni, L. See, G. M. Sechi, "Comparing Artificial Network Vector Machines For Modelling RainFall-Runoff", In the Proceeding of the 6th International Conference on Hydroinformatics-Liong, Phoon & Babovic (eds), ISBN. 981- 238-787-0, 2004.

[6] N. Karunanithi, W. Grenney, D. Whitley, and K. Bovee, "Neural networks for river flow prediction" Journal of Computing in Civil Engineering, vol. 8, no. 2, pp. 371-379, 1993.

[7] P. Bulando, R. Rosso, J. Salas, "Forecasting of short-term rainfall using ARMA models" Journal of Hydrology, vol. 144, pp. 193-211, 1993.

Corresponding Author: Alaa Sheta, Department of Information Technology, Al-Balqa Applied University, Jordan

J. Computer Sci., 2 (2): xx-xx, 2006

[8] H. Hruschka, "Determining market response functions by neural networks modeling : a comparison to econometric techniques", European Journal of Operation Research, vol. 66, pp. 867-888, 1993.

[9] E. Y. Li, "Artificial neural networks and their business applications," Information and managements, vol. 27, pp. 303-313, 1994.

[10] K. Chakraborty, " Forecasting the behavior of multivariable time series using neural networks," Neural Networks, vol. 5, pp. 962-970, 1992.

[11] G. Swales and Y. yoon, "Applying artificial neural networks to investment analysis", Financial Analyst Journal, pp. 78-82, 1992.

[12] N. Tan Danh, H. Ngoc Phien, A. Das Gupta, "Neural network models for river flow forecasting", Vol. 25, No. 1, ISSN 0378-4738, 1999.

[13] M. Adya, F. Collopy "How effective are neural networks at forecasting and prediction" Journal of Forecasting, pp. 481- 495, 1998.

[14] R. Baratti,, B. Cannas, "River flow forecast for reservoir management through neural networks." Neurocomputing, pp, 421-437, 2003.

[15] O. Kisi "River flow modeling using artificial networks", ASCE J. of Hydrologic Engineering, vol 9, pp, 60-63, 2004.

[16] H.K. Cigizoglu and O. Kisi "Flow Prediction by three Back-propagation Techniques Using k-fold Partitioning of Neural Network Training Data", Nordic Hydrology, vol 36, 2005.

Corresponding Author: Alaa Sheta, Department of Information Technology, Al-Balqa Applied University, Jordan