DEVELOPMENT OF STREAMFLOW FORECASTING MODEL USING ARTIFICIAL NEURAL NETWORK IN THE AWASH RIVER BASIN, ETHIOPIA

D.C. EDOSSA AND M.S. BABEL

Abstract

Early indication of possible drought can help in developing suitable drought mitigation strategies and measures in advance. Therefore, drought forecasting plays an important role in the planning and management of water resource in such circumstances. In this study, a non-linear streamflow forecasting model was developed using Artificial Neural Network (ANN) modeling technique at the Melka Sedi stream gauging station, Ethiopia, with adequate lead times. The available data was divided into two independent sets using a split sampling tool of the neural network software. The first data set was used for training and the second data set, which is normally about one fourth of the total available data, was used for testing the model. A one year data was set aside for validating the ANN model. The streamflow predicted using the model on weekly time step compared favorably with the measured streamflow data (R^2 = 75%) during the validation period. Application of the model in assessing appropriate agricultural water management strategies for a large-scale irrigation scheme in the Awash River Basin, Ethiopia, has already been considered for publication in a referred journal.

Keywords: Awash River Basin; ANN; streamflow forecasting; drought early warning system; Ethiopia

1. INTRODUCTION

Knowledge of droughts has been an important aspect in the planning and management of water resource systems. Reservoirs are often planned so that they are able to meet the expected water demands during drought of a certain magnitude and water supply systems are often evaluated to see whether they will be able to withstand a *T*-year drought (Frick *et al*. 1990). In any case, determining drought properties at a time and in space (or region) is an important aspect of water planning and management activities. Drought analysis may be made based on single site data (Yevjevich 1967; Dracup *et al.* 1980) and multi-site data (Tase 1976; Santos *et al.* 1983; Guttman *et al.* 1992; Soule 1992), depending on the specific purpose of the study at hand.

Traditionally, hydrologic variables of interest, such as annual and monthly precipitation and streamflows, have been extensively modeled using linear techniques, such as autoregressive moving average with exogenous inputs (ARMAX) (Salas *et al*. 1985), Box-Jenkins multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) class of models

(McKerchar and Delleur 1974; Panu *et al*. 1978; Cline 1981; and Govindasamy 1991) and also non-linear regression (Chang and Hwang 1999). These methods have been generally accepted by practitioners during the past several decades. However, they are based on the basic assumptions that data is stationary, and has a limited ability to capture non-stationarities and non-linearities in hydrologic data.

It is necessary for hydrologists to consider alternative models when nonlinearity and non-stationarity are important and play a significant role in the forecasting. Artificial Neural Networks (ANN) have shown great ability in modeling and forecasting nonlinear and non-stationary time series data in hydrology and water resources engineering due to their innate nonlinear property and flexibility for modeling (ASCE Task Committee 2000). ANN is a "computational paradigm inspired by the parallelism of the brain". The ANN is particularly valuable in performing classification and pattern recognition functions for processes governed by complex nonlinear interrelationships.

In the recent past years, the use of ANNs in hydrological modeling has been rapid, such as rainfall estimation and forecasting (Hsu *et al*. 1999; Shin and Salas 2000; Luk *et al*. 2001; Kim and Valdes 2003), real-time reservoir operation (Chang and Chang 2001) and streamflow forecasting (Atiya *et al*. 1999; Sajikumar and Thandaveswara 1999; Govindaraju and Rao 2000; Chang and Chen 2001; Chang *et al*. 2001; Kisi 2004; Kisi 2005; Wu *et al*. 2005; Jain *et al*. 2007).

The demonstration of a relatively strong relationship between precipitation and El Nino-Southern Oscillation for many regions (Ropelewski and Halpert 1996) has also aroused considerable interest and encouraged investigations of possibilities of forecasting rainfall and perhaps alleviating some of the socially undesirable effects of sudden and unexpected occurrences of extremes such as floods and droughts. Many studies have reported an approach to predict drought from atmospheric circulation patterns (Pesti *et al*. 1996; Pongracz *et al*. 1999; Pongracz *et al*. 2003).

There have been notable droughts in Ethiopia throughout human history. In particular, frequent droughts and floods are the key hazards to life in the Awash River Basin in Ethiopia (Desalegn *et al*. 2009). Droughts entail loss of assets in the form of crops, livestock, and other productive capitals as a result of water shortages and related impacts. The existing coping strategy, in the country in general and in the Awash River Basin in particular, is based on crisis management (reactive approach) which may lead to untimely and costly short-term solution rather than through the formulation and implementation of anticipatory measures commonly referred to as risk management (proactive approach). Edossa *et al.* (2006) reported that there is no drought early warning system in the Awash River Basin.

Moving from crisis to risk management will require the adoption of a new paradigm by land and water managers, governments, international and regional development organizations, and nongovernmental organizations. This approach emphasizes preparedness, mitigation, and improved early warning systems over emergency response and assistance measures. A typical reason mentioned by decision makers for the lack of such drought planning and management is the lack of means to forecast climate conditions with sufficient skills and lead-time due to the randomness of drought events in both time and space dimensions. Therefore, it goes without saying that drought forecasting in the Awash River Basin is necessary in mitigating its impacts on various water sectors and thereby on the livelihood of the basin community.

In this study, a streamflow forecasting model was developed using Artificial Neural Network (ANN) technique at the headwork of a large irrigation scheme, Middle Awash Agricultural Development Enterprise (MAADE) in the Awash River Basin, Ethiopia. Applicability of the developed model in assessing appropriate agricultural water management strategies to be adopted in the irrigation scheme under drought conditions has been published (Edossa and Babel 2011).

2. MATERIALS AND METHODS

A three-layer neural network with back-propagation algorithm was applied to develop a streamflow forecasting model for the Awash River at a point of diversion for MAADE irrigation scheme. WinNN32 neural network shareware (Danon 1997), one of the ANN software families, was used in this study.

Weekly time series of streamflow data (19872001) were used in order to develop a long-term streamflow forecasting model, at least one season. The model was trained using a 10 and half year data (1987 to 1996/97) and tested with the next three and a half year data (1997-1999/2000). Finally, a one year data (2001) were used to validate the model. The optimum weights obtained from the ANN using the historical data were used to establish a non-linear autoregressive model for the streamflow time series.

2.1. Design of ANN Architecture

There are no hard and fast rules governing the correct design of a neural network. It goes without saying that more complex problems will require more complex networks. However, when there are a large number of free parameters, the network will be (a) slower to train and (b) more susceptible to over-fitting. Important factors such as the number of inputs, the number of hidden units, and the arrangement of these units into layers are often determined using 'trial and error' experimental design procedures (Fischer & Gopal 1994) or fixed in advance according to the subjective opinion of each individual designer (Abrahart & Kneale 1997).

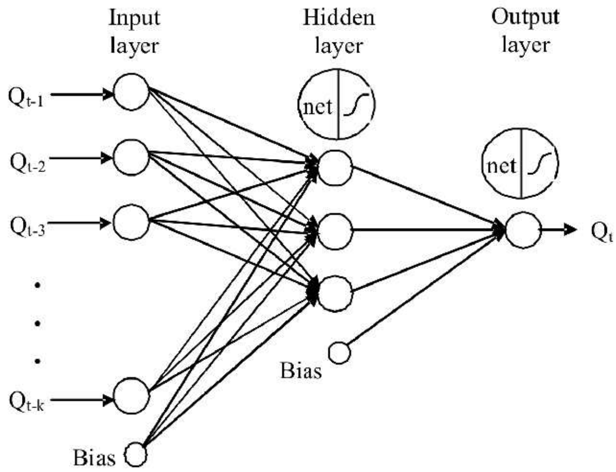

Figure 1. ANN architecture used in the model development

A three-layer ANN, which is typically used in hydrology and water resources engineering, was adopted in this study. It has input, output, and hidden (middle) layers (Figure 1). Each neuron in a layer is connected to all the neurons of the next layer, and the neurons in one layer are not connected among themselves. All the neurons within a layer act synchronously. The number of neurons in the input and output layers were fixed according to the number of input and output variables, respectively. In order to determine the optimal network architecture, the optimum number of neurons in the hidden layer was determined by experimentation. Accordingly, the number of neurons in this layer was varied between 1 and 7 (Ozgur 2004) during the training phase and the configuration that gave the minimum RMSE (residual mean squared error) and the maximum correlation was selected as the best net for modeling the streamflow.

In this modeling, the use of daily streamflow data was constrained by the limitation imposed by the WinNN32 software with regard to the maximum number of weights to be determined (180 in this case) vis-à-vis the period of forecast. The total number of weights depends on the number of inputs to the network. In this study, the neural network was made to model a streamflow process of the following form:

$$Q_t = f\left(Q_{t-n}\right)$$

Where: Qt is current streamflow, Qt-n is antecedent streamflow (at t-1, t-2, … t-n time steps). Therefore, antecedent streamflows (Qt-1, Qt-2, Qt-3… etc) at different lag times (t-1, t-2, t-3,…etc) were used as inputs to the network to predict current streamflow, Qt, at time, t, on weekly time step. As a result, the output layer had one neuron to estimate the current streamflow, Qt.

2.2. Model development

A streamflow time series data at the point of diversion was used as input data to the ANN to develop a non-linear streamflow forecasting model. After preprocessing, the historical data were divided into three sets, one for training and one for testing of the ANN model and the remaining set was set aside for validating the model.

2.3. Pre- and post-processing of data

The output of the Logistic Activation Function (LAF), which was used in this study, lies in the interval [0,1]; for this reason the original data need to be transformed to the interval [0.05,0.95] before being presented to the network. Each input and output values were normalized with their own specific normalization factors as follows: suppose a and A are the minimum and maximum values of the data series, respectively, then an actual flow value of Qt was transformed to the interval [0.05,0.95] using the formula:

$$Q_t^{'} = \frac{0.90(Q_t - a)}{A - a} + 0.05$$

where, Qt = actual value; a = minimum value of Qt; A = maximum value of Qt; and = transformed value.

After the best network was found, all the transformed data were retransformed back to their original range by the equation:

$$Q_t = \frac{(A - a)(Q_t^{'} - 0.05)}{0.9} + a$$

2.4. Training, testing and validation

The practice adopted in the training and testing was to divide the available data into two independent sets using split sampling tool in the WinNN32 software. The first data set was used for training and the second data set, which is normally about one fourth of the total available data, was used for testing. A one year data was set aside for validating the ANN model.

During the training, the learning parameters such as Eta (η), Alpha (α), weight noise, and temperature were set to their default values of 0.2, 0.5, 0, and 1, respectively. Quickprop (Fahlman 1988), a modification of the backpropagation, was used as training algorithm. It uses a second order weight-update function, based on measurement of the error gradient at two successive points to accelerate the convergence over simple first-order gradient descent.

Quickprop is one of the earliest modifications designed to speedup backpropagation.

The backpropagation (BP) method uses a set of input and output patterns. An input pattern was used by the system to produce an output, O, which then was compared with the target output, Q. The data passing through the connections from one neuron to another were multiplied by weights that control the strength of a passing signal, the product summed and then passed through a transfer function to produce results. This summation of product is termed the net, N, and must be calculated for each neuron in the network. Artificial neurons or nodes are simple processing units which produce outputs as nonlinear functions of weighted sums of the inputs to that node. Inputs were applied to the ANN from a set of streamflow data created from the same time series at various lag times. After net was calculated, an activation function was applied to modify it, thereby producing the output signal, O. One of the most popular activation function used in neural network studies (Blum 1992), which was also used in the present study for the neurons in the hidden and output layers is the Logistic Activation Function (LAF) or simply sigmoid function of the form:

$$Q_{out} = f(Q_{net}) = \frac{1}{1+e^{-Q_{net}}}$$

where Qnet and Qout are neuron net input and output, respectively. The logistic function has an S shape and its output ranges between 0 and 1. The use of such LAF introduces non-linearity in the operation of the neural network thereby underpinning and enhancing their ability to model non-linear processes.

If there is no difference between the net and target outputs, then no learning takes place and the signal emanating from the output node is the network's solution to the input problem. Otherwise, the weights are changed while moving backward through the network to reduce the difference. When these weights are modified, the data transferred through the network changes; consequently, the network output also changes. The objective is to minimize the overall network error E for all input patterns in the training set. The error of pattern p for a network having a single output variable was compared as follows:

$$E_p = \frac{1}{2}\sum (Q_t - O_t)^2$$

where the summation is for all patterns. The backpropagation method tries to minimize this error by adjusting the weights in an iterative process.

2.5 Model performance

Accuracy of models can be evaluated by plotting line graphs that show the actual data versus the values predicted by the models. However, the five more formal quantitative measures of accuracy of time series modeling techniques include: the mean absolute deviation (MAD), the mean absolute percent error (MAPE), the mean square error (MSE), and the root mean square error (RMSE) (Rogsdale 2001) and the efficiency index (EI). These indices measure the differences between the actual values in the time series and the predicted, or fitted, values generated by the model. In this study, two methods were used to check performance of the developed model: plots of line graphs and RMSE.

3. RESULTS AND DISCUSSION

3.1 ANN Architecture

The selection of number of inputs is one of the critical issues in ANN-based streamflow forecasting model, since it provides the important information about the complex autocorrelation structure in the data. Too many or too few inputs can affect both the network training and forecast capability of the model. Therefore, it is important to monitor the performance measures of both the training and testing sets during the modeling process to avoid overtraining of the network. A network is considered to be over-trained when the training error keeps reducing while the error over the testing set is getting increased.

Figure 2 shows the change in the performance of the model (RMSE) with change in the number of inputs. Different combinations of number of nodes in the input and hidden layers were evaluated and an ANN architecture with 40 and 4 nodes in the input and output layers, respectively, was found to give reasonably minimum RMSE. Therefore, in order to forecast streamflow into the future with a reasonable lead time (one season), the 40-4-1 ANN architecture was selected in this modeling where 40, 4 and 1 are indicating the number of nodes (neurons) in the input, hidden and output layers, respectively. It can be noted that RMSE of the training set decreases whereas that of the testing set shows irregular patterns (neither increases nor decreases) with an increase in the number of inputs.

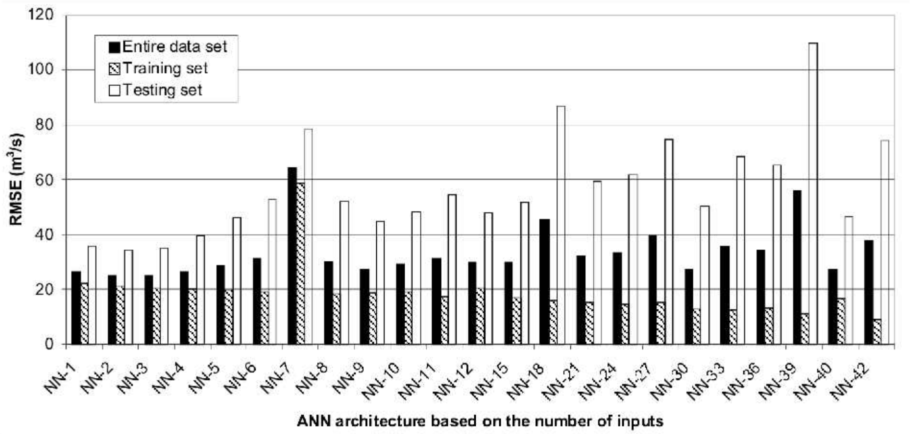

Figure 2. Performance of neural network (NN) models in relation to the number of inputs

3.2 Streamflow Forecasting Model

Figures 3, 4, and 5 show plots of measured and predicted time series data of streamflow for the training, testing and validation sets, respectively at the Melka Sedi stream gauging station. These results show that the proposed model tends to fit the data in the low flows range fairly well in all the three data sets. However, it tends to underestimate the peak flows in all of the three data sets.

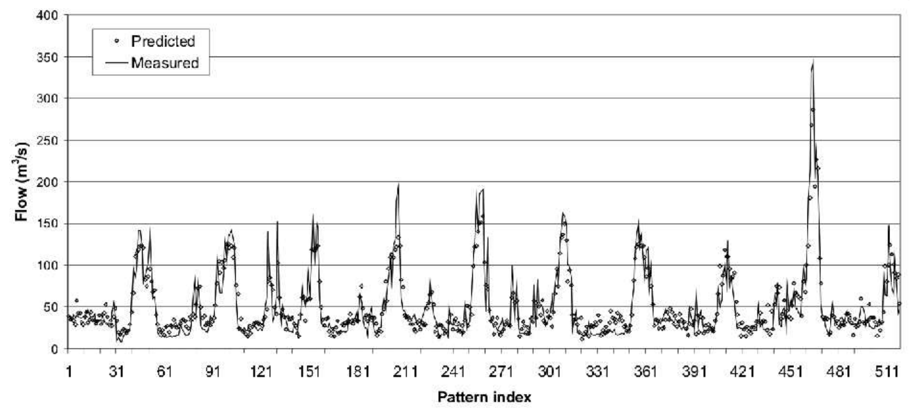

Figure 3. Forecasted and observed streamflow at the Melka Sedi stream gauging station during training stage

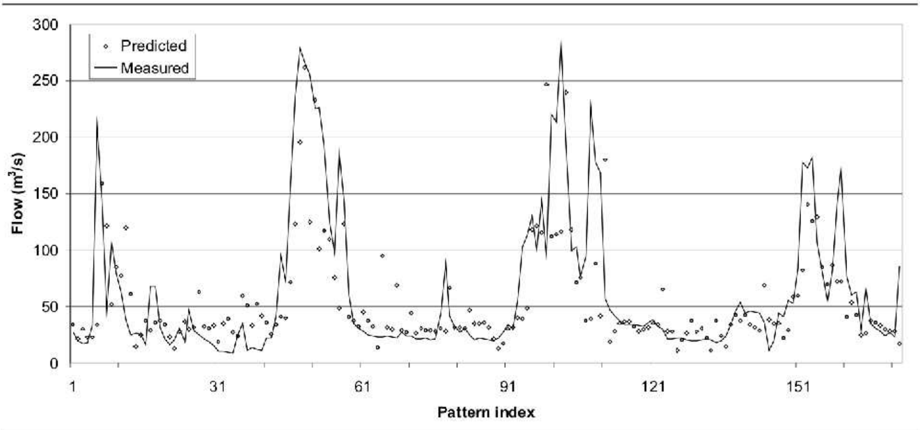

Figure 4. Forecasted and observed streamflow at the Melka Sedi stream gauging station during testing stage

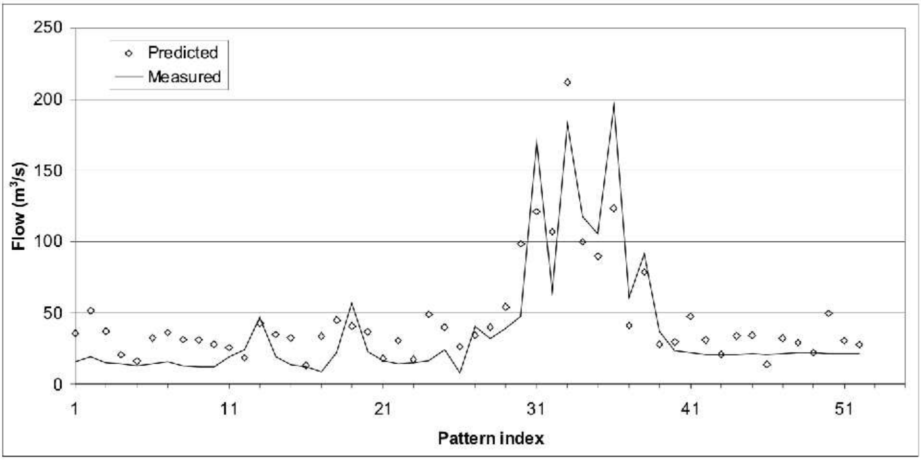

Figure 5. Forecasted and observed streamflow at the Melka Sedi stream gauging station during validation stage

Table 1 shows performance measures of the proposed model. The model produced the smallest RMSE during the training set and the largest RMSE during the testing set while the validation set produced intermediate RMSE value. The correlation value is worse during the testing than during training, as is expected. Again the model produced an intermediate correlation value during validation period.

The model efficiency index shows that there is fairly good agreement between the measured and predicted values over the entire range of recorded data.

Table 1 Configuration and performance measures of the ANN model

Item/parameter	**Value**
Model architecture	40-4-1
Efficiency Index	0.70
RMSE (m^3/s)	
⇒Entire data set	27.38
⇒Training set	16.66
⇒Testing set	46.63
⇒Validation	24.75
Correlation	
⇒Entire data set	0.84
⇒Training set	0.93
⇒Testing set	0.70
⇒Validation	0.88
Target error	0.05
Input noise	0.03

Figure 6 shows plot of measured versus predicted flows at the stream gauging station over the entire data set. To get an overview of the deviation between the measured and predicted flows, a 1:1 line was superimposed on the plot. Minimum deviation is observed in the region below 50 m^3/s and increases with the flow magnitude suggesting that the model performs better in the low flow regions. Therefore, the model performance is justified with respect to the task for which it was developed drought planning and management.

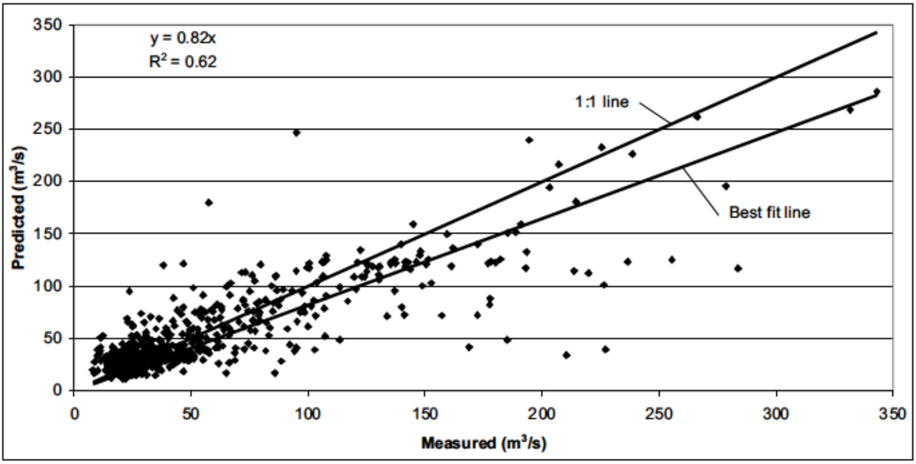

Figure 6. Measured versus ANN predicted flows for long-term forecasting at the Melka Sedi stream gauging station

3.3 Forecasting results

The proposed model was used to forecast the streamflow 40-weeks ahead into the future at the Melka Sedi stream gauging station which might be used for analyzing and assessing appropriate water management practices to be implemented in the area. Figure 7 shows plot of measured and forecasted streamflows during the forecasting period. The figure shows that the forecasted streamflows follow similar general patterns with the measured streamflow data during the forecasting period with a correlation value of 0.78.

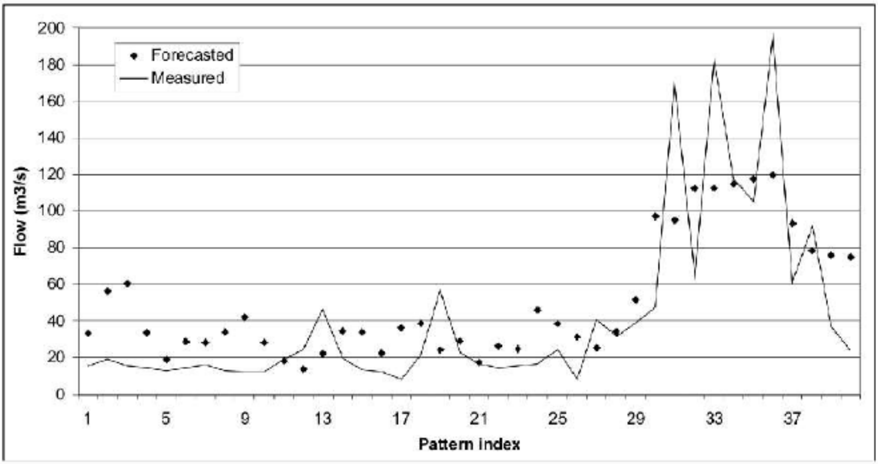

Figure 7. Measured and forecasted flows at the Melka Sedi stream gauging station

4. SUMMARY AND CONCLUSION

A non-linear long-term streamflow forecasting model was developed for the Melka Sedi stream gauging station using three-layer back-propagation neural network algorithm. The model was validated using a one year streamflow data which were not used in the model development. Finally, streamflow time series is forecasted with sufficient lead time using the developed model for planning drought mitigation strategies. This study found that on the basis of the developed long-term streamflow forecasting model, there is an opportunity of developing a drought watch system for mitigating impacts of droughts in a large-scale irrigation scheme supplied by water from Awash River through a diversion canal.

5. ACKNOWLEDGMENTS

This study was funded by the International Water Management Institute (IWMI) under the PhD research scholarship program for doctoral studies of the first author.

6. REFERENCES

Abrahart, R. J. and Kneale, P. E. 1997 Exploring Neural Network Rainfall-Runoff Modeling. In Proceedings of the 6th British Hydrological Society Symposium, Salford University, 15-18 September 1997, 9.35-9.44.

ASCE Task Committee 2000 Artificial neural networks in hydrology. II: Hydrologic applications. Journal of Hydrologic Engineering, ASCE, 5(2), 124-137.

Atiya, A. F., El-Shoura, S. M., Shaheen, S. I. and El-Sherif, M. S. 1999 A comparison between neural network forecasting techniques-case study: river flow forecasting. IEEE Transactions on Neural Networks, 10(2): 402-409.

Fahlman, S. E. 1988 Faster-learning variations on back-propagation: An empirical study. 1988 Connectionist Models Summer School.

Blum, A. 1992 Neural networks in C++: An object-oriented framework for building connectionist systems. John Wiley and Sons, Inc., USA.

Chang, F. J. and Chen, Y. C. 2001 Couterpropagation fuzzy-neural network modeling approach real-time streamflow prediction. Journal of hydrology, 245: 153-164.

Chang, F. J. and Hwang, Y. Y. 1999 A self-organization algorithm for real-time flood forecast. Hydrological processes, 13(2): 123-138.

Chang, F. J., Hu, H. F. and Chen, Y. C. 2001 Counterpropagation fuzzy-neural network for streamflow reconstructing. Hydrological processes, 15(2): 219-232.

Chang, L. C. and Chang, F. J. 2001 Intelligent control of modeling of real-time reservoir operation. Hydrological processes, 15(9): 1621-1634.

Cline, T. B. 1981 Selecting seasonal streamflow models. Water Resources Research, 17(4), 975-984.

Danon, Y. 1997 Scientific software. Available at http://www.nuceng.com/WinNN.htm, access date: 15/01/2004

Dracup, J. A., Lee, K. S., and Paulson, E. N. Jr. 1980 On the statistical characteristics of drought events. Water Resource Research, 16(2), 289296.

Edossa, D. C., and Babel, M. S. 2011 Application of ANN-Based Streamflow Forecasting Model for Agricultural Water Management in the Awash River Basin, Ethiopia. Water Resources Management, 25:17591773, DOI 10.1007/s11269-010-9773-y.

Edossa, D. C., Babel M. S., and Gupta A. D. 2009 Analysis of drought severity in the Awash River Basin, Ethiopia. Journal of Water Resource Management, DOI: 10.1007/s11269-009-9508-0.

Edossa, D. C., Babel, M. S., Gupta, A. D., Seleshi, B. A. and Merrey, D. 2006 Farmers' perception of water management under drought conditions in the upper Awash Basin. Journal of International Water Resource Development, 22(4), 589-602.

Fischer, M. M. and Gopal, S. 1994 Artificial neural networks: a new approach to modelling interregional telecommunication flows. Journal of Regional Science, 34, 503-527.

Frick, D. M., Bode, D., and Salas, J. D. 1990 Effect of drought on urban water supplies. I: Drought analysis. Journal of Hydrologic Engineering, ASCE, 116(6), 733753.

Govindaraju, R. S. and Rao, A. R. 2000 Artificial neural networks in hydrology. Kluwer, Rotterdam.

Govindasamy, R. 1991 Univariate Box-Jenkins forecasts of water discharge in the Missouri River. International Journal of Water Resources Development, 3, 168-177.

Guttman, N. B., Wallis, J. R. and Hosking, J. R. M. 1992 Spatial comparability of the Palmer drought severity index. Water Resource Bulletin, 28(6), 11111119.

Hsu, K. L., Gupta, H. V., Gao, X. G. and Sorooshian, S. 1999 Estimation of physical variables from multichannel remotely sensed imagery using a neural network: application to rainfall estimation. Water resources research, 35(5): 1605-1618.

Jain, A. and Kumar, M. A. 2007 Hybrid neural network models for hydrologic time series forecasting. Applied Soft Computing, 7(2): 585-592.

Kim, T. W. and Valdes, J.B. 2003 A Nonlinear Model for Drought Forecasting Based on Conjunction of Wavelet Transforms and Neural Networks. Journal of Hydrologic Engineering, ASCE, 8(6), 319-328.

Kisi, O. 2004 River flow modeling using artificial neural networks. Journal of Hydrologic Engineering, Technical notes, Vol. 9, No. 1.

2005 Daily river flow forecasting using artificial neural networks and auto-regressive models. Turkish Journal of Environmental Science, Vol 29.

Luk, K. C., Ball, J. E. and Sharma, A. 2001 An application of artificial neural networks for rainfall forecasting. Mathematical and computer modeling, 33: 683-693.

McKerchar, A. I. and Delleur, J. W. 1974 Application of seasonal parametric linearstochastic models to monthly flow data. Water Resources Research, 10(2): 246-255.

Ozgur, K. 2004 River flow modeling using artificial neural networks. Journal of hydrologic engineering, 9(1), 60-63.

Panu, U. S., Unny, T. E. and Ragade, R. K. 1978 A feature prediction model in synthetic hydrology based on concepts of pattern recognition. Water Resources Research, 14(2), 335-344.

Pesti, G., Shrestha, B. P., Duckstein, L. and Bogardi, I. 1996 A fuzzy rule-based approach to drought assessment. Water Resources Research, 32(6), 1741-1747.

Pongracz, R., Bogardi, I. and Duckstein, L. 1999 Application of fuzzy rule-based modeling technique to regional drought. Journal of Hydrology, 224, 100-114.

Pongracz, R., Bogardi, I. and Duckstein, L. 2003 Climatic forcing of drought: a Central European example. Hydrologic Sciences Journal, 48(1), 39-50.

Rogsdale, C. T. 2001 Spreadsheet modeling and decision analysis. South-Western College publishing, USA.

Ropelewski, C. F. and Halpert, M. S. 1996 Quantifying southern oscillation-precipitation relationships. Journal of Climate, 9, 1043-1059.

Sajikumar, N. and Thandaveswara, B. S. 1999 A non-linear rainfall-runoff model using an artificial neural network. Journal of Hydrology, 216, pp 32-55.

Salas, J. D., Delleur, J. W., Yevjevich, V. and Lane, W. L. 1985 Applied modeling of hydrologic time series. Water resources publications: Littleton, Colorado.

Santos, M. A. B., Correia, F. N. and DaCunha, L. V. 1983 Drought characterization and drought impacts in Portugal. Memoria 591, Laboratorio Nacional de Engenharia Civil, Lisbon, Portugal.

Shin, H.-S. and Salas, J. D. 2000 Regional drought analysis based on Neural Network. Journal of Hydrologic Engineering, 5(2), 145-155.

Soule, P. T. 1992 Spatial patterns of drought frequency and duration in the contiguous USA based on multiple drought event definitions. International Journal of Climatology, 12, 1124.

Tase, N. 1976 Area-deficit-intensity characteristics of drought. PhD dissertation, Colorado State University, Fort Collins, Colorado, USA.

Wu, J. S., Han, J., Annambhotla, S., and Bryant, S. 2005 Artificial Neural Networks for Forecasting Watershed Runoff and Stream Flows. Journal of Hydrologic Engineering, 10(3).

Yevjevich, V. 1967 An objective approach to definitions and investigations of continental hydrologic droughts. Hydro. Paper 23, Colorado State University, Fort Collins, Colo.