International Journal of Scientific & Engineering Research Volume 8, Issue 5, May-2017
ISSN 2229-5518

Application of Neural Networks for River Flow Forecasting

[1]Daniyal Hassan, [2]Kamran Ansari, [3]Rakhshinda Bano, [3]Awais Anwar, [1]Nauman Shaikh, [1]Fatima Zaidi, [1]Mansoor Ali

Abstract— River flow forecasting is pre-requisite for water resources planning and management. Artificial neural networks (ANN) can analyze the processes, simulate the phenomena, perform linear and nonlinear forecasts in conditions where the details of the relationships between the insidious parameters are not available. Being flexible in structure it can recognize the complicated relationships among input and output data. This study describes the use of different Artificial Neural Networks (ANN) for stream-flow forecasts in the Indus river flows. Three distinct ANN models, feed-forward backpropagation networks, radial basis networks, and recurrent neural networks were applied for modeling the stream flows at Guddu, Sukkur and Kotri Barrage gauges' data. Daily Discharge Data was available from 1970-2015, it was divided into three sets, i.e. training data (calibration), test while training data (verification), and forecasting data. The RMSE of the feed-forward backpropagation neural networks, radial basis neural networks, and recurrent neural networks were obtained 0.083, 0.058, and 0.078 respectively. The Feed-forward and recurrent networks produced the superior results of forecasting. The results specify that the ANN has the capability to forecast the stream flows at the rivers where data about the other hydrological parameters are difficult to obtain.

Key Words— Water Resources Planning and Management, River flow forecasting, Indus River, Artificial Neural Networks, Feed-forward backpropagation, Radial Basis Networks, and Recurrent Neural Networks

1 Introduction

Floods are one of the most devastating and frequently occurring natural disasters which strike numerous regions in the world each year. According to World Meteorological Organization (WMO) report (2011), during the last decades, the trend in flood damages has been growing exponentially. The development of hydrological forecasting and warning systems is, therefore, a fundamental and crucial element in provincial and national planning. From 1930, the world has undergone over more than 200 deadliest floods including China's Yellow and Yangtze River flood, Netherland's St. Lucia's flood, Bangladesh's monsoon floods, Iran's flood, Barcelona's flash flood, South and North Korean floods etc. These floods brought almost 8 million deaths across the globe. The statistics are shown by statista.com and worldmapper.org just the flood of China (July 1931) led to 3.7 million deaths and 30.7 billion U.S. dollars economy's loss.

Pakistan is also one of the most affected countries where heavy flood comes after every third or fourth year. According to the Federal Flood Commission (FFC) report, Pakistan has observed 20 major floods from 1950 to 2015. These floods affected 599,459 square kilometers' area, snatched 11,239 precious human lives, and caused losses worth over 39 billion U.S. dollars to the national economy. Due to heavy floods in 2010, 2011 and 2012, Pakistan lost 3,072 precious lives and had a monetary loss of $16 billion. United Nations Secretary-General Ban Ki-Moon interpreted 2010's Flood as the worst disaster he had ever witnessed. An estimated one-fifth of Pakistan's total land area (62,000 square miles) was submerged by the flooding. The flood is approximated to have eventually affected more than 20 million Pakistanis, damaged 1.9 million households, caused $9.7 billion economy's loss and demolished more 39 health care units. According to Pakistan National Disaster Management Agency (NDMA) flood, 2010 caused 1,802 deaths and 2,994 injuries. Therefore, it is essential for Pakistan to develop an effective and efficient stream-flow flood forecasting systems so that huge financial and human losses can be saved.

1.1 Artificial Neural Networks

ANN modeling is a nonlinear arithmetical method; it is capable for simulations that are not possible by using typical statistical and mathematical models [1]. The ANN models could be described by the elements: nodes, weights, and transfer function. Usually, neural networks are trained, as an input indicates to a certain output or target as shown in fig. 1 below. The network is modified using various trial and error iterations, based on a contrast between the output and target, till the network resembles the target.

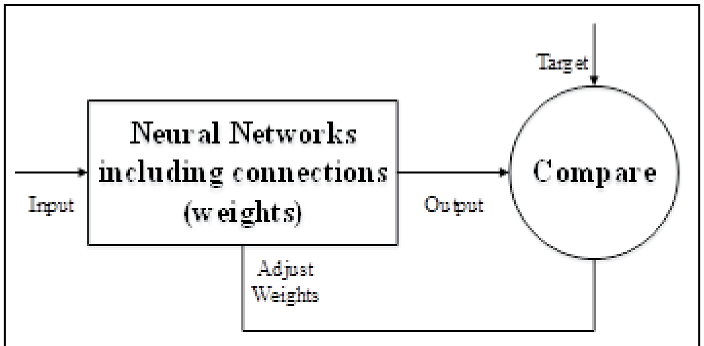


Fig. 1. Schematic Diagram of the neural networks

[1]Research Associate-U. S.-Pakistan Centre for Advanced Studies in Water (USP-CAS-W), Mehran-UET, Pakistan.
[2]Deputy Director USPCAS-W
[3]Assistant Professor USPCAS-W
E-mail: daniyalhassan109@gmail.com

1.2 Back Propagation networks

The backpropagation indicates that the gradient is figured for nonlinear multilayer networks. Backpropagation integrates the Widrow-Hoff learning rule, multiple-layer networks, and nonlinear differentiable transfer functions. Input vectors and the subsequently targeted vectors are employed to direct the networks till the approximation of the function, secondary input vectors including output vectors, or organize input vectors in a suitable way as stated by you. Mostly, the linear transfer function purelin is utilized in backpropagation.
A basic neuron having R inputs is illustrated in fig.2. Each input is biased with a suitable w. The total weighted inputs and the bias forms are put into the transfer function f. Neurons might engage any differentiable transfer function f to produce their output.

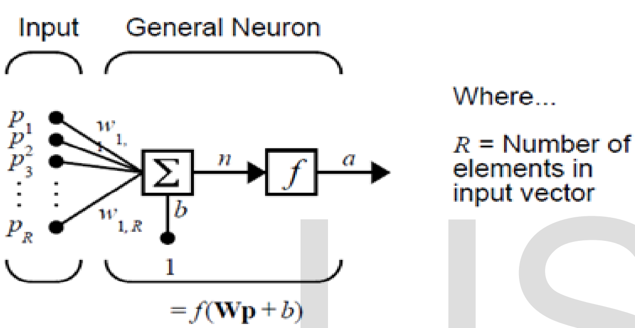


Fig. 2. Back Propagation Network

1.3 Feed-Forward Backpropagation Neural Networks (FFNN)

Generally, the feed-forward ANN is used for the linear forecast. The network's architecture consists of layers and nodes within all layers. Every single node in input and inner layers obtains input values, and the weights administer and forwards to the subsequent layer. The total number of neurons in the layers at input and output are decided by the numbers of parameters. The architecture of the feed forward model is shown in Fig. 3.

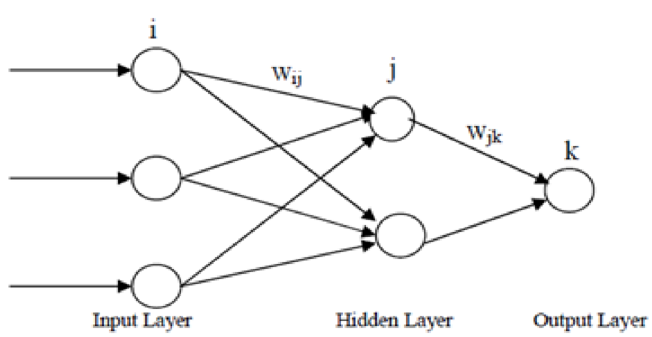


Fig. 3. Feedforward Artificial Neural Network

Here, the i, the input, j, the hidden, and k represents the output layers. The 'w' shows the weight of the nodes. Subscripts specify the connections between the nodes.
Feedforward networks might have more than one hidden layers of transfer functions followed by linear neurons output layer. Several layers of neurons along with nonlinear transfer functions let the network to learn linear and nonlinear correlations linking input and output vectors.

1.4 Radial basis networks

Radial basis functions (RBF) are robust techniques in multidimensional forecasts. It works on the radius or the distance from the center. These radial functions could be utilized in the interpolation and smoothing of data. Radial basis functions are introduced as the replacement of the sigmoidal transfer function. The RBF mostly have 3 layers, input, the hidden (i.e. non-linear radial), and an output layer. Radial basis functions usually require extra neurons than any backpropagation networks, and they are capable of forecasting in a time fraction to train standard feed-forward networks.

1.5 Recurrent neural networks (RNN)

Recurrent networks having two layers and work on the principle of the backpropagation, they might have an additional path for feedback from the hidden layers output to its input. This feedback connection enables Recurrent networks to recognize, to train, and to produce the spatial and temporal patterns. Hydrological forecasts are made in the Recurrent networks are based on the prior series values dependent on the persistence components. A primary step in designing Recurrent networks is the hidden neurons selection. In RNN, the target function is unspecified, therefore it is tricky to predict the optimal network size. The appropriate network should be balanced neither overfit nor underfit the data.

1.6 STUDY AREA

Sindh being the lower riparian of Indus river and having a subtropical region with relatively flat topography is affected in both situations of rain and no rains. It remains hot in summer and bit cold in winter. Temperatures often rise above 40 °C (between May-August), and the least average temperature of 1 °C occurs in December and January. This research was conducted on Guddu, Sukkur and Kotri Barrages of Sindh. 1) The Guddu Barrage is in Kashmore District of Sindh it controls water for irrigation and flood control purposes. 2) The Sukkur barrage is located near Sukkur District and having the discharge capacity of 1.2 cusecs. 3) Kotri Barrage is located between Hyderabad and Jamshoro (last barrage on Indus River and Pakistan at downstream). During this study, the streamflow is forecasted at above mentioned three Barrages. The Data of the 46 years (1970-2015) discharges of Guddu, Sukkur and Kotri Barrages was used for forecasting.

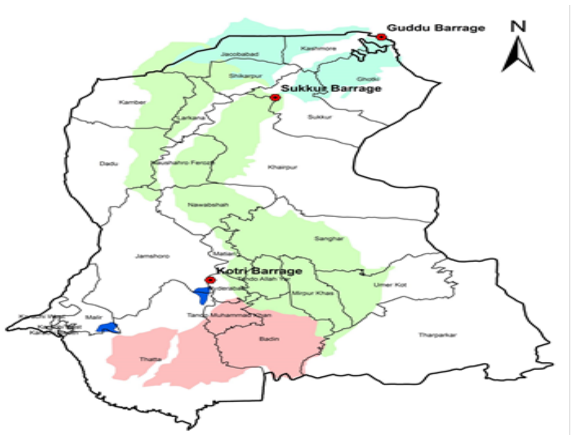


Fig. 4. Study Area, Guddu, Sukkur and Kotri Barrage

2 LITERATURE REVIEW

Hydrologists and engineers used the skills of distributed, lumped, stochastic and time series models for predictions [1, 2, 3, & 4]. Conceptual and physically based models normally consider all the physical processes involved in the Rainfall and Runoff process and time series stochastic models is complicated due to non-stationary behavior and nonlinearity in the data. These models usually need a big expertise of the modeler [5]. Recent development in the field of the hydrological modeling is the Data-Driven Modelling that uses the machine learning mechanism. The DDM is based on the data analysis by finding the connections between input, internal and output variables [6]. Approximately, from the last two decades, Artificial Neural Networks (ANNs) has emerged as a powerful computing tool in DDM for highly complex and nonlinear systems. ANNs consist of many simple processing elements called neurons or nodes. Each node is then connected to other nodes by means of direct links. Each link is associated with a weight that represents the strength of the outgoing signal. The processing of each node is carried out in two steps, that is, the weighted sum of the inputs is taken, and is followed by the application of the activation fiction computer. They are characterized by (1) their patterns of connections between the neurons (called its architecture), (2) their methods of determining the weights on the connections (called their training).
ANN was first developed in the 1940s [7], and the development has experienced a renaissance with Hopfield's effort [8] in 5 iterative auto-associable neural networks. Many researchers [for example 9, 10, 11, 12, 13, & 14] have shown that an Artificial Neural Network is capable of simulating runoff from rainfall. Once the ANN is properly trained, the computational time is small. One difficulty is that the ANN is a "prisoner of its training data" because the predicted values will not be greater than the maximum trained value put into the ANN [15]. It is generally accepted that ANN cannot extrapolate beyond the range of the training data [16 & 17.)]. Approximations in determining stage–discharge relationships or changes in the cross section of the stream after a flood, for example, may affect the accuracy of measurement [5]. Abraham used an artificial neural network with scaled conjugate gradient algorithm (ANN-SCGA) and evolving fuzzy neural network (EfuNN) for predicting the rainfall time series [18]. In the study, monthly rainfall was used as input data for training model. Another application was described by Koizumi [19], who employed an ANN model using radar, satellite, and weather-station data together with numerical products generated by the Japan Meteorological Agency (JMA) Asian Spectral Model for 1-year training data. Koizumi found that the ANN skills were better than persistence forecast (after 3 h), the linear regression forecasts, and numerical model precipitation prediction [20].
Past research presents the successful applications of ANN models in the simulation of future runoffs with a high degree of accuracy compared ANN and Box & Jenkins techniques and concluded that ANN is an improvement on Box & Jenkins model. [25, & 26].
In Pakistan, many of the prediction has been made by using ANN for the inflation rate, veterinary (prediction of the live weight), for stock exchange [27, & 28]. The ANN was also used to measure the Rainfall-Runoff in Hub River catchment located in Karachi [29]. In the downstream of the Indus River where the problem is most severe (which is affected in both flood and drought cases) any forecasting model including the ANN was not used before. In this study, the Feed forward ANN is used to predict the Indus river flows at Sindh.

3 METHODOLOGY

Different artificial neural networks (feed forward, back propagation, and radial based neural network) are used to forecast the stream flow of River Indus at Guddu, Sukkur, and Kotri Barrages. They have been widely used for hydrological applications because they are simple, accurate and allow high processing speeds. The Discharge Data for the 46 years (1970-2015) of the above-mentioned Barrages was acquired from WAPDA. It was divided into three sets. I.e. training data which is used to minimize the error of the model (calibration), while the testing set was used to check generalization capability as calibration proceeds (verification), and forecasting data [23, & 30]. According to previous work and recommendations by various researchers, the 75% of the data must be dedicated to training (calibration and verification) and 25% for forecasting purpose. The training data was further divided into a 2/3rd part of the training set and 1/3rd for the testing set [16, & 31].

3.1 EVALUATION CRITERIA

ROOT MEAN SQUARE ERROR (RMSE)

The scaled RMSE (R) is the ratio between Root Mean Square Error and the standard deviation of observations σ. A predictive model having very satisfactory performance, if the scaled RMSE's result is less than 0.5 and satisfactory if it is less than 0.7.

ISSN 2229-5518

$$R = RMSE/\sigma$$

$$RMSE = \sqrt{\sum_{i=0}^{Nv} (yi - oi)^2 / Nv}$$

Here y, o and Nv are representing the predicted, observed values and the number of target data used for testing.

4 Results

4.1 Feed-Forward Backpropagation Networks

To acquire the optimal number of hidden nodes, 2, 3, 4, 5, and 7 were tried. In this array, the model, having 5 inputs and three hidden layers with 2 nodes and 1000 iteration numbers, showed the best performance. During the performance evaluation of the forecasts made by any Artificial Neural Networks model, to calculate the average prediction error and the distribution of these errors are equally important. The statistical performance evaluation criteria used in this study are R^2 and RMSE. Table 1 below shows the statistical tests results for feedforward network during the simulation at Guddu barrage in the test phase.

Table 1 Feed Forward Statistical Results for Test Period

Statistical Test	2004	2005	2006	2007
RMSE	0.0778	0.0224	0.0919	0.0380
R	0.8389	0.9993	0.9888	0.8971
R^2	0.7932	0.9271	0.9354	0.8396

4.2 Radial Basis Networks

Various numbers of hidden layer neurons are tested to obtain the best radial basis network. After many trial and error practices, RBN model, having 6 inputs and 8 spread constants, and 800 iteration number, gave the best results. Table 2 below shows the statistical tests results for radial basis networks during the simulation at Sukkur barrage in the test phase.

Table 2 Radial Basis Network Statistical Results for Test Period

Statistical Test	2004	2005	2006	2007
RMSE	0.1725	0.0596	0.0481	0.0502
R	0.96474	0.9797	0.9854	0.9686
R^2	0.9337	0.9475	0.9493	0.9360

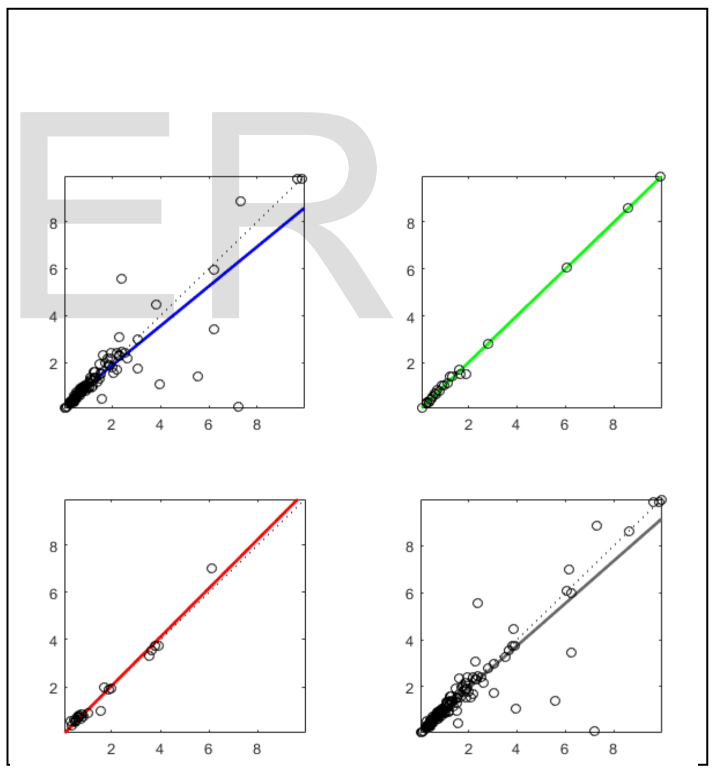
4.3 Recurrent Neural Networks

The 6 inputs and one hidden layer having iteration numbers 15000 were obtained after the trial and error practices for recurrent networks. The Artificial Neural Networks are very sensitive for the hidden layers nodes. A very few can under-predict and too many nodes might produce the over-predictions.

Table 3 Recurrent Based Statistical Results for Test Period

Statistical Test	2004	2005	2006	2007
RMSE	0.1894	0.0047	0.0587	0.0581
R	0.9961	0.9956	0.9774	0.9936
R^2	0.9452	0.9447	0.9472	0.9415

In this research, different neural networks feed forward back propagation, radial, and recurrent basis networks were employed to forecast the Indus River stream flows and the models were applied to daily stream flows of Guddu, Sukkur and Kotri Barrages located at Indus river. The RMSE of the feed-forward backpropagation neural networks, radial basis neural networks, and recurrent neural networks were obtained 0.083, 0.058, and 0.078. While, the mean R^2 were obtained 0.94, 0.87, and 0.944. The Feed forward and recurrent networks produced the reliable forecasting results.



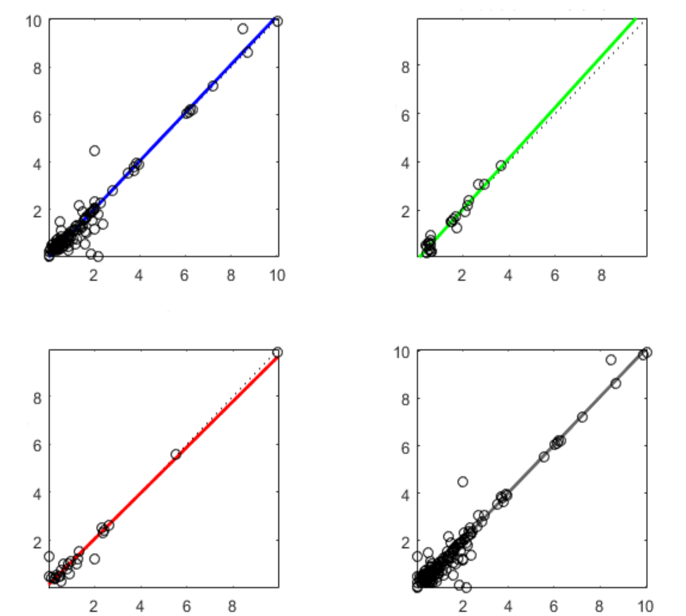


Fig.6. R- Radial Basis Networks

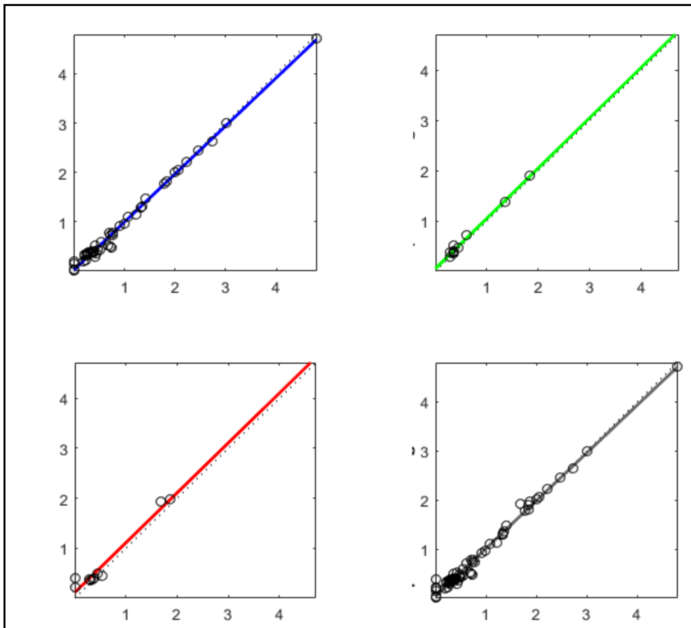


Fig.7. R- Recurrent Neural Networks

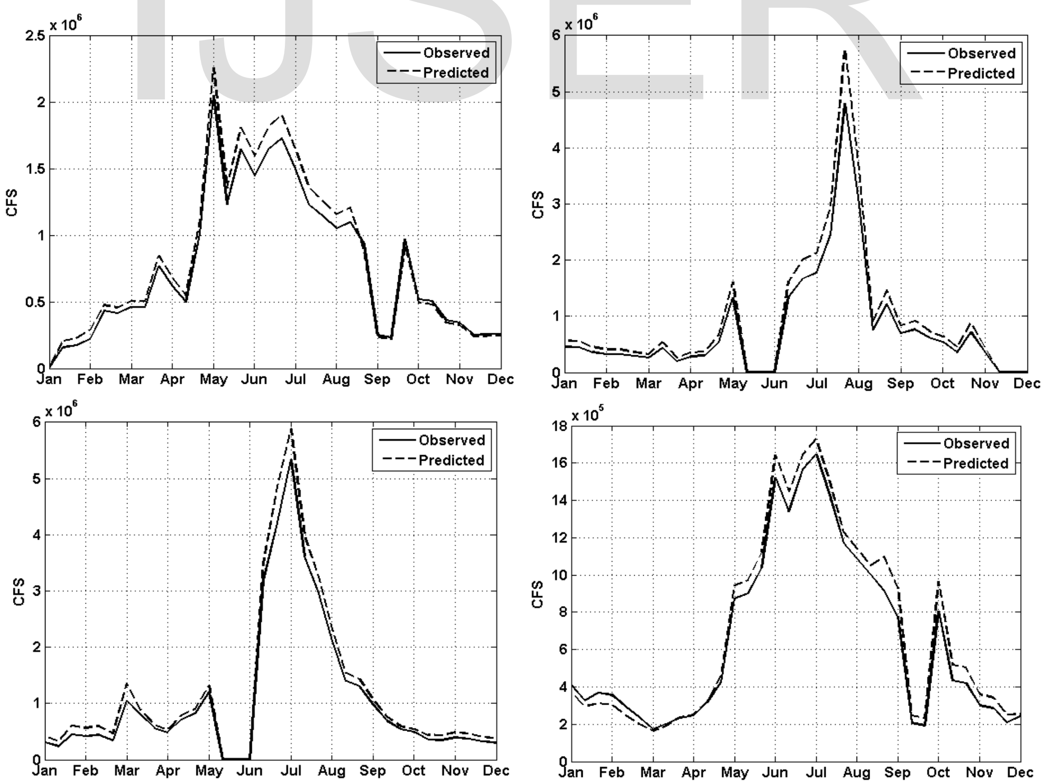


Fig.8. Feed forward ANN results at Guddu Barrage (2004-2007 test period)

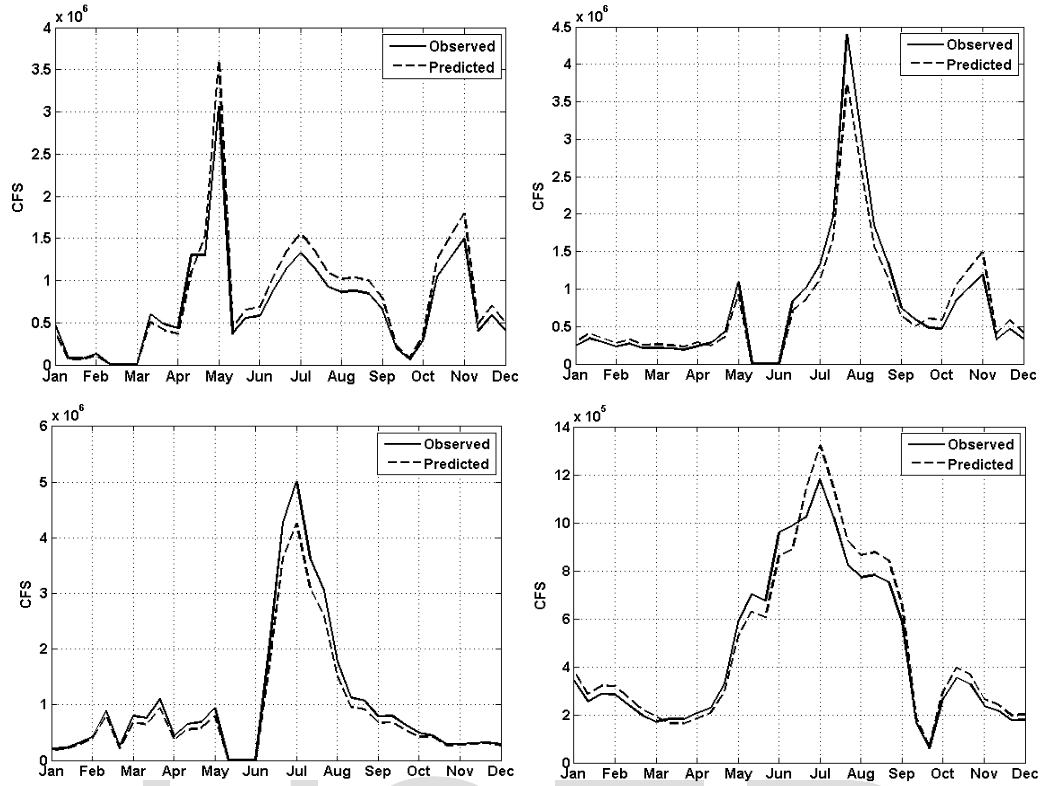


Fig. 9. Radial basis network results at Sukkur Barrage (2004-2007 test period)

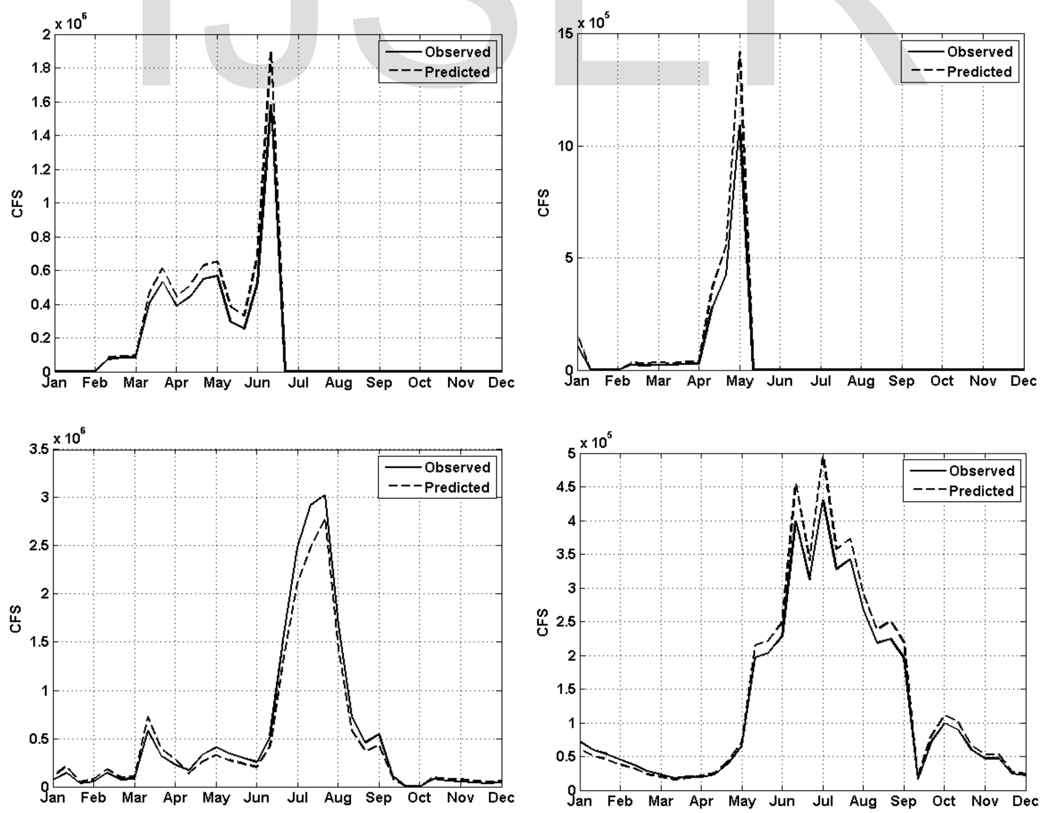


Fig. 10. Recurrent neural network results at Kotri Barrage (2004-2007 test period)

5 Conclusion

Artificial Neural Networks are the powerful computing tool and have been successfully used all over the world for linear forecasts. In this study, 46 years' flow data of the Guddu, Sukkur, and Kotri gauges located at Indus river were used to forecast the streamflow. The three different ANN models feed-forward, radial based, and recurrent based networks were employed for streamflow forecasts. The 75% of the data was dedicated for training of the networks while remaining 25% was utilized for forecasts. The Feedforward and Recurrent networks were found best in the production of the superior forecasting results and could be employed for flood and streamflow forecasts in the streams where the meteorological and hydrological data is rarely available for advanced models.

References

[1] ASCE Task Committee, "Artificial Neural Networks in Hydrology. I: Preliminary concepts." J. of Hydrologic Engineering, ASCE, 5(2), pp. 115–123, 2000.

[2] Abbott, M. B., and A. W. Minns, "Computational Hydraulics", Ashgate Hampshire, England, 1998.

[3] Feng, S., S. Kang, et al., "Neural networks to simulate regional ground water levels affected by human activities." Ground Water 46(1): pp. 80-90, 2008

[4] Govindaraju, R. S. "Artificial neural networks in hydrology. I: Preliminary concepts." Journal of Hydrologic Engineering 5(2): pp. 115-123, 2000.

[5] Gaume, E. and R. Gosset. "Over-parameterization, a major obstacle to the use of artificial neural networks in hydrology?" Hydrology and Earth System Sciences Discussions 7(5): pp. 693-706, 2003.

[6] Solomatine, D., L. See, et al., "Data-driven modeling: concepts, approaches, and experiences." Practical hydro-informatics, Springer: pp. 17-30, 2009.

[7] McCulloch, W. S., and W. Pitts. "A logical calculus of the ideas immanent in nervous activity." The bulletin of mathematical biophysics 5(4): pp. 115-133, 1943.

[8] Hopfield, J. J. "Neural networks and physical systems with emergent collective computational abilities." Proceedings of the national academy of sciences 79(8): pp. 2554-2558, 1982.

[9] Zhu, M.-L., M. Fujita, et al. "Application of neural networks to runoff prediction. Stochastic and statistical methods in hydrology and environmental engineering." Springer: pp. 205-216, 1994.

[10] Minns, A. and M. Hall. "Artificial neural networks as rainfall-runoff models." Hydrological sciences journal 41(3): pp. 399-417, 1996.

[11] Hall, T., H. E. Brooks, et al. "Precipitation forecasting using a neural network." Weather and forecasting 14(3): pp. 338-345, 1999.

[12] Shamseldin, A. Y. "Application of a neural network technique to rainfall-runoff modelling." Journal of Hydrology 199(3): pp. 272-294, 1997.

[13] Campolo, M., A. Soldati, et al. "Forecasting river flow rate during lowâ€□flow periods using neural networks." Water resources research 35(11): pp. 3547-3552, 1999.

[14] Dawson, C. W., and R. Wilby. "An artificial neural network approach to rainfall-runoff modelling." Hydrological sciences journal 43(1): pp. 47-66, 1998.

[15] Flood, I., and N. Kartam. "Neural networks in civil engineering. I: Principles and understanding." Journal of computing in civil engineering 8(2): pp. 131-148, 1994.

[16] Maier, H. R., and G. C. Dandy. "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." Environmental modelling & software 15(1): pp. 101-124, 2000.

[17] Bowden, Gavin J., Holger R. Maier, and Graeme C. Dandy. "Real-time deployment of artificial neural network forecasting models: Understanding the range of applicability." Water Resources Research 48, no. 10, 2012.

[18] Abraham, A., D. Steinberg, et al. "Rainfall forecasting using soft computing models and multivariate adaptive regression splines." IEEE SMC Transactions, Special Issue on Fusion of Soft Computing, and Hard Computing in Industrial Applications 1: pp. 1-6, 2001.

[19] Koizumi, K. "An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network." Weather and forecasting 14(1): pp. 109-118, 1999.

[20] Hung, N. Q., M. S. Babel, et al. "An artificial neural network model for rainfall forecasting in Bangkok, Thailand." Hydrology and Earth System Sciences 13(8): pp. 1413-1425, 2009.

[21] Ahmad, S. and S. P. Simonovic. "An artificial neural network model for generating hydrograph from hydro-meteorological parameters." Journal of Hydrology 315(1): pp. 236-251, 2005.

[22] Ali, M., E. Eyduran, et al. "Comparison of Artificial Neural Network and Decision Tree Algorithms used for Predicting Live Weight at Post Weaning Period from Some Biometrical Characteristics in Harnai Sheep." Pakistan Journal of Zoology 47(6): pp. 1579-1585, 2015.

[23] Bishop, C. M. "Neural networks for pattern recognition.", Oxford university press, 1995.

[24] Campolo, M., A. Soldati, et al. "Artificial neural network approach to flood forecasting in the River Arno." Hydrological sciences journal 48(3): pp. 381-398, 2003.

[25] Castellano-Méndez, Marı́a, Wenceslao González-Manteiga, Manuel Febrero-Bande, José Manuel Prada-Sánchez, and Román Lozano-Calderón. "Modelling of the monthly and daily behaviour of the runoff of the Xallas river using Box–Jenkins and neural networks methods." Journal of Hydrology 296, no. 1 pp. 38-58, 2004.

[26] Sohail, A., K. Watanabe, et al. "Runoff analysis for a small watershed of Tono Area Japan by back propagation artificial neural network with seasonal data." Water resources management 22(1): 1-22, 2008.

[27] Haider, A. and M. N. Hanif. "Inflation forecasting in Pakistan using artificial neural networks." Pakistan economic and social review: pp. 123-138, 2009, 2009.

[28] Noor-Ul-Amin, M. "Forecasting with Neural Networks: A comparative study using the data of emergency service." arXiv preprint arXiv:1010.3501, 2010.

[29] Ghumman, A., Y. M. Ghazaw, et al. "Runoff forecasting by artificial neural network and conventional model." Alexandria Engineering Journal 50(4): pp. 345-350, 2011.

[30] Haykin, S. "Neural networks: MacMillan College Publ." Co., New York, 1994.

[31] Chau, K., C. Wu, et al. "Comparison of several flood forecasting models in Yangtze River." Journal of Hydrologic Engineering 10(6): pp. 485-491, 2005.