

# Short Term Flood Forecasting Using General Recurrent Neural Network Modeling a Comparative Study

Rahul P. Deshmukh

Indian Institute of Technology, Bombay  
Powai, Mumbai  
India

A. A. Ghatol

Former Vice-Chancellor  
Dr. Babasaheb Ambedkar Technological  
University, Lonere, Raigad, India

## ABSTRACT

The artificial neural networks (ANNs) have been applied to various hydrologic problems recently. This research demonstrates dynamic neural approach by applying general recurrent neural network to rainfall-runoff modeling for the upper area of Wardha River in India. The model is developed by processing online data over time using dynamic modeling. Methodologies and techniques by applying different learning rule, activation function and input layer structure are presented in this paper and a comparison for the short term runoff prediction results between them is also conducted. The prediction results of the general recurrent neural network with Momentum learning rule and Tanh activation function with Axon as input layer structure indicates a satisfactory performance in the three hours ahead of time prediction. The conclusions also indicate that general recurrent neural network with Momentum learning rule and Tanh activation function with Axon as input layer structure is more versatile than other combinations for general recurrent neural network and can be considered as an alternate and practical tool for predicting short term flood flow.

**General Terms:-** Prediction, Radial basis function neural network, Time lagged recurrent neural network.

**Keywords-** Artificial neural network, Forecasting, Rainfall, Runoff, Models.

## I. INTRODUCTION

The main focus of this research is development of Artificial Neural Network (ANN) models for short term flood forecasting, determining the characteristics of different neural network models. Comparisons are made between the performances of different parameters for general recurrent artificial neural network models.

The field engineers face the danger of very heavy flow of water through the gates to control the reservoir level by proper operation of gates to achieve the amount of water flowing over the spillway. This can be limited to maximum allowable flood and control flood downstream restricting river channel capacity so as to have safe flood levels in the river within the city limits on the downstream.

By keeping the water level in the dam at the optimum level in the monsoon the post monsoon replenishment can be conveniently stored between the full reservoir level and the permissible maximum water level. Flood estimation is very essential and plays a vital role in planning for flood regulation and protection measures.

The total runoff from catchment area depends upon various unknown parameters like Rainfall intensity, Duration of rainfall, Frequency of intense rainfall, Evaporation, Interception,

Infiltration, Surface storage, Surface detention, Channel detention, Geological characteristics of drainage basin, Meteorological characteristics of basin, Geographical features of basin etc. Thus it is very difficult to predict runoff at the dam due to the nonlinear and unknown parameters.

In this context, the power of ANNs arises from the capability for constructing complicated indicators (non-linear models). Among several artificial intelligence methods artificial neural networks (ANN) holds a vital role and even ASCE Task Committee Reports have accepted ANNs as an efficient forecasting and modeling tool of complex hydrologic systems[22].

Neural networks are widely regarded as a potentially effective approach for handling large amounts of dynamic, non-linear and noisy data, especially in situations where the underlying physical relationships are not fully understood. Neural networks are also particularly well suited to modeling systems on a real-time basis, and this could greatly benefit operational flood forecasting systems which aim to predict the flood hydrograph for purposes of flood warning and control[16].

A subset of historical rainfall data from the Wardha River catchment in India was used to build neural network models for real time prediction. Telematic automatic rain gauging stations are deployed at eight identified strategic locations which transmit the real time rainfall data on hourly basis. At the dam site the ANN model is developed to predict the runoff three hours ahead of time.

In this paper, we demonstrate the use of general recurrent neural network (GRNN) model for real time prediction of runoff at the dam and compare the effectiveness of different learning rules, activation function and input layer structure. General recurrent neural network extend the multilayer perceptron with context units, which are PEs that remember past activity..

At a time when global climatic change would seem to be increasing the risk of historically unprecedented changes in river regimes, it would appear to be appropriate that alternative representations for flood forecasting should be considered.

## II. METHODOLOGY

In this study different parameters like learning rule activation function and input layer structure are employed for rainfall-runoff modeling using general recurrent neural network model of artificial neural network.

General recurrent networks (GRNN) are to temporal data as multi-layer perceptrons (MLP's) are to static data. They are categorized by a layer that feeds back upon itself using adaptable weights.

Recurrent networks are fundamentally different from feed-forward networks. Even with a constant input, they do not necessarily settle to a constant output. They can exhibit limit cycles and even chaotic behavior. Nevertheless, Li [1992] showed that general recurrent neural network (GRN's) are a universal approximator of a differentiable trajectory. This means that given a sufficient number of neurons, general recurrent neural network (GRN's) can capture the dynamics imbedded within input data.

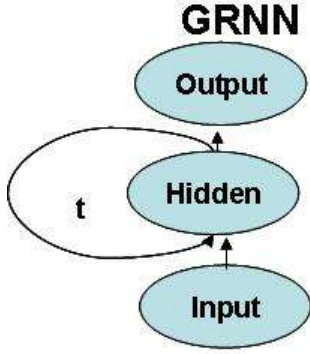


Figure 1. The General Recurrent Neural Network

In this study we applied different input structure to the general recurrent neural network and studied the optimum performance for different learning rules and activation function. We applied input structures like Axon, TDNN Axon, Gamma Axon and Laguarre Axon to the network. The learning rule studied are Momentum, Deltabar Delta, step, Conjugate Gradient and Quick prop with activation function Tanh, Linear Tanh, Sigmoid and Linear Sigmoid.

#### Performance Measures:

The learning and generalization ability of the estimated NN model is assessed on the basis of important performance measures such as MSE (Mean Square Error), NMSE (Normalized Mean Square Error) and r (Correlation coefficient)

##### A. MSE (Mean Square Error):

The formula for the mean square error is:

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N d_{ij} - y_{ij}^2}{NP} \quad \dots (1)$$

Where

P = number of output PEs,

N = number of exemplars in the data set,

$y_{ij}$  = network output for exemplar i at PE j,

$d_{ij}$  = desired output for exemplar i at PE j.

##### B. NMSE (Normalized Mean Square Error):

The normalized mean squared error is defined by the following formula:

$$NMSE = \frac{PNMSE}{\sum_{j=0}^P \frac{N \sum_{i=0}^N d_{ij}^2 - \left( \sum_{i=0}^N d_{ij} \right)^2}{N}} \quad \dots (2)$$

Where

P = number of output processing elements,

N = number of exemplars in the data set,

MSE = mean square error,

$d_{ij}$  = desired output for exemplar i at processing element j.

##### C. r (correlation coefficient):

The size of the mean square error (MSE) can be used to determine how well the network output fits the desired output, but it doesn't necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, the MSE can be changed without changing the directionality of the data. The correlation coefficient (r) solves this problem. By definition, the correlation coefficient between a network output x and a desired output d is:

$$r = \frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{\sqrt{\sum_i (d_i - \bar{d})^2} \sqrt{\sum_i (x_i - \bar{x})^2}} \quad \dots (3)$$

$$\text{where } \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{and} \quad \bar{d} = \frac{1}{N} \sum_{i=1}^N d_i$$

The correlation coefficient is confined to the range [-1, 1]. When r = 1 there is a perfect positive linear correlation between x and d, that is, they co-vary, which means that they vary by the same amount.

### III. STUDY AREA AND DATA SET

The Upper Wardha catchment area lies directly in the path of depression movements which originates in the Bay of Bengal. When the low pressure area is formed in the Bay of Bengal and cyclone moves in North West directions, many times this catchment receives very heavy intense cyclonic precipitation for a day or two. Occurrence of such events have been observed in the months of August and September. Rainfall is so intense that immediately flash runoff, causing heavy flood has been very common feature in this catchment.

For such flashy type of catchment and wide variety in topography, runoff at dam is still complicated to predict. The conventional methods also display chaotic result. Thus ANN based model is built to predict the total runoff from rainfall in Upper Wardha catchment area for controlling water level of the dam.

In the initial reaches, near its origin catchment area is hilly and covered with forest. The latter portion of the river lies almost in plain with wide valleys.

The catchment area up to dam site is 4302 sq. km. At dam site the river has wide fan shaped catchment area which has large variation with respect to slope, soil and vegetation cover.



Figure 3- Location of Upper Wardha dam on Indian map



Figure 4- The Wardha river catchment

**Data:** Rainfall runoff data for this study is taken from the Wardha river catchment area which contains a mix of urban and rural land. The catchments is evenly distributed in eight zones based on the amount of rainfall and geographical survey. The model is developed using historical rainfall runoff data , provided by Upper Wardha Dam Division Amravati, department of irrigation Govt. of Maharashtra. Network is trained by rainfall information gathered from eight telemetric rain-gauge stations distributed evenly throughout the catchment area and runoff at the dam site [3].

The data is received at the central control room online through this system on hourly basis. The Upper Wardha dam reservoir operations are also fully automated. The amount of inflow, amount of discharge is also recorded on hourly basis. From the inflow and discharge data the cumulative inflow is calculated. The following features are identified for the modeling the neural network .

Table 1- The parameters used for training the network

Month	RG1	RG2	RG3	RG4	RG5	RG6	RG7	RG8	CIF
-------	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Month – The month of rainfall
- Rain1 to Rain8 – Eight rain gauging stations.
- Cum Inflow – Cumulative inflow in dam

Seven years of data on hourly basis from 2001 to 2007 is used. It has been found that major rain fall (90%) occurs in the month of June to October Mostly all other months are dry hence data from five months. June to October is used to train the network

#### IV. RESULT

The different structures of neural network are employed to learn the unknown characterization of the system from the dataset presented to it. The dataset is partitioned into three categories, namely training, cross validation and test. The idea behind this is that the estimated NN model should be tested against the dataset that was never presented to it before. This is necessary to ensure the generalization. An experiment is performed at least twenty five times with different random initializations of the connection weights in order to improve generalization.

The data set is divided in to training , testing and cross validation data and the network is trained for all combinations of general recurrent neural network for 5000 epochs.

The performance results obtain on parameters by applying learning rules Momentum, Deltabar Delta, step, Conjugate Gradient, Quick prop with activation function Tanh, Linear Tanh, Sigmoid, Linear Sigmoid with different input structures like Axon, TDNN Axon, Gamma Axon, Laguarre Axon are listed in Table II through Table V.

Table II - Comparison of performance parameters for Axon structure

No.	Axon	Tanh		Linear Tanh		Sigmoid		Linear Sigmoid	
		MSE	r	MSE	r	MSE	r	MSE	r
1	Step	0.0263	0.875	0.0314	0.712	0.0351	0.764	0.0341	0.601
2	Momentum	0.0106	0.914	0.0214	0.878	0.0239	0.831	0.0241	0.624
3	Conjugate Gradient	0.0349	0.752	0.0412	0.769	0.0384	0.654	0.0421	0.625
4	Quickprop	0.0432	0.624	0.0413	0.612	0.0391	0.751	0.0370	0.714
5	Deltabar Delta	0.0394	0.714	0.0419	0.659	0.0412	0.667	0.0421	0.683

Table III - Comparison of performance parameters for TDNN Axon structure

No.	TDNN Axon	Tanh		Linear Tanh		Sigmoid		Linear Sigmoid	
		MSE	r	MSE	r	MSE	r	MSE	r
1	Step	0.0316	0.852	0.0334	0.884	0.0361	0.876	0.0370	0.834
2	Momentum	0.0376	0.824	0.0384	0.889	0.0364	0.871	0.0391	0.815
3	Conjugate Gradient	0.0431	0.763	0.0426	0.726	0.0391	0.729	0.0380	0.762
4	Quickprop	0.0361	0.669	0.0380	0.692	0.0389	0.712	0.0357	0.684
5	Deltabar Delta	0.0328	0.701	0.0331	0.628	0.0341	0.604	0.0352	0.629

Table IV - Comparison of performance parameters for Gamma Axon structure

No.	Gamma Axon	Tanh		Linear Tanh		Sigmoid		Linear Sigmoid	
		MSE	r	MSE	r	MSE	r	MSE	r
1	Step	0.0324	0.657	0.0331	0.663	0.0341	0.672	0.0351	0.668
2	Momentum	0.0376	0.704	0.0369	0.694	0.0354	0.680	0.0341	0.710
3	Conjugate Gradient	0.0384	0.712	0.0374	0.704	0.0389	0.726	0.0352	0.713
4	Quickprop	0.0376	0.669	0.0346	0.676	0.0361	0.722	0.0380	0.741
5	Deltabar Delta	0.0412	0.712	0.0392	0.759	0.0436	0.741	0.0425	0.732

Table V - Comparison of performance parameters for Laguarre Axon structure

No.	Laguarre Axon	Tanh		Linear Tanh		Sigmoid		Linear Sigmoid	
		MSE	r	MSE	r	MSE	r	MSE	r
1	Step	0.0384	0.718	0.0389	0.746	0.0394	0.734	0.0409	0.724
2	Momentum	0.0412	0.694	0.0380	0.680	0.0387	0.712	0.0396	0.670
3	Conjugate Gradient	0.0378	0.650	0.0342	0.654	0.0335	0.662	0.0360	0.640
4	Quickprop	0.0364	0.630	0.0369	0.621	0.0374	0.634	0.0349	0.640
5	Deltabar Delta	0.0365	0.720	0.0374	0.712	0.0392	0.695	0.0384	0.704

The parameters and performance for general recurrent model with different learning rule and activation function are compared on the performance scale and are listed in the Table II to V for Axon, TDNN Axon, Gamma Axon and Laguarre Axon respectively. The comparative analysis of the MSE and r (the correlation coefficient) is done.

After training the network the optimum performance is studied and it is found that Momentum learning rule and Tanh activation function with Axon input layer structure produce optimal result. In the Table-VI the parameters and the best performances for general recurrent neural network are listed.

Table VI- GRN network parameters

Parameter	Performance
MSE	0.0106
NMSE	0.0126
Min Abs Error	0.0043
Max Abs Error	0.6238
r	0.9140

Fig 4 shows the plot of actual Vs predicted runoff for optimum values for general recurrent neural network found with Momentum learning rule and Tanh activation function for Axon input structure.

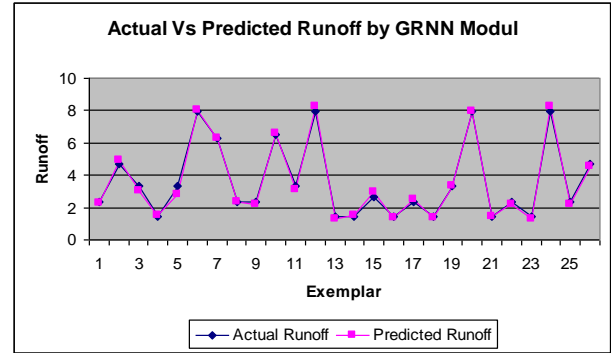


Figure 4.– Actual Vs. Predicted runoff by GRNN for Axon with momentum and Tanh

The error found in the actual and predicted runoff at the dam site is plotted for GRNN model is as shown in the Figure 5.

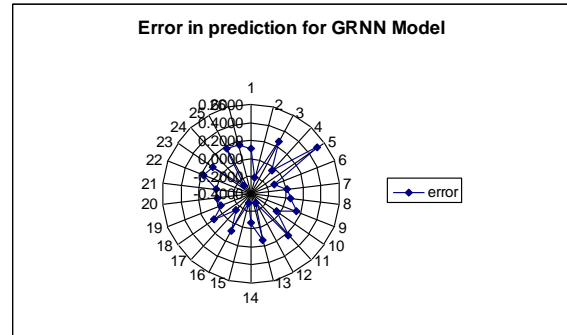


Fig 5 – Error graph of GRNN

The main advantage of general recurrent neural network is that they have a potentially unlimited memory depth and thus, can actually capture the dynamics of the system that produced a temporal signal

## V. CONCLUSION

An ANN-based short-term runoff forecasting system is developed in this work. A comparison between five different learning rules with four activation function for four different input structures is made for optimal performance for general recurrent neural network model. We find that general recurrent neural network with momentum learning rule and Tanh activation function for Axon input structure is more versatile than other approaches studied. General recurrent neural network with Momentum learning rule and Tanh activation function with Axon as input structure is performing better as compare to other approaches studied as far as the overall performance is concerned for forecasting runoff for 3 hours lead time. Other approaches studied are also performing optimally. Which means that dynamic model of general recurrent neural network with momentum learning rule and Tanh activation function with Axon input structure is powerful tool for short term runoff forecasting for Wardha River basin.

## ACKNOWLEDGMENT

This study is supported by Upper Wardha Dam Division Amravati, department of irrigation Govt. of Maharashtra, India

REFERENCES.

- [1] P. Srivastava, J. N. McVair, and T. E. Johnson, "Comparison of process-based and artificial neural network approaches for streamflow modeling in an agricultural watershed," *Journal of the American Water Resources Association*, vol. 42, pp. 545-563, Jun 2006.
- [2] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, pp. 359-366, 1989.
- [3] M. C. Demirel, A. Venancio, and E. Kahya, "Flow forecast by SWAT model and ANN in Pracana basin, Portugal," *Advances in Engineering Software*, vol. 40, pp. 467-473, Jul 2009.
- [4] A. S. Tokar and M. Markus, "Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models," *Journal of Hydrologic Engineering*, vol. 5, pp. 156-161, 2000.
- [5] S. Q. Zhou, X. Liang, J. Chen, and P. Gong, "An assessment of the VIC-3L hydrological model for the Yangtze River basin based on remote sensing: a case study of the Baohe River basin," *Canadian Journal of Remote Sensing*, vol. 30, pp. 840-853, Oct 2004.
- [6] R. J. Zhao, "The Xinanjiang Model," in *Hydrological Forecasting Proceedings Oxford Symposium*, IASH, Oxford, 1980 pp. 351-356.
- [7] R. J. Zhao, "The Xinanjiang Model Applied in China," *Journal of Hydrology*, vol. 135, pp. 371-381, Jul 1992.
- [8] D. Zhang and Z. Wanchang, "Distributed hydrological modeling study with the dynamic water yielding mechanism and RS/GIS techniques," in *Proc. of SPIE*, 2006, pp. 63591M1-12.
- [9] J. E. Nash and I. V. Sutcliffe, "River flow forecasting through conceptual models," *Journal of Hydrology*, vol. 273, pp. 282-290, 1970.
- [10] D. Zhang, "Study of Distributed Hydrological Model with the Dynamic Integration of Infiltration Excess and Saturated Excess Water Yielding Mechanism." vol. Doctor Nanjing: Nanjing University, 2006, p. 190-529
- [11] E. Kahya and J. A. Dracup, "U.S. Streamflow Patterns in Relation to the El Niño/Southern Oscillation," *Water Resour. Res.*, vol. 29, pp. 2491-2503, 1993.
- [12] K. J. Beven and M. J. Kirkby, "A physically based variable contributing area model of basin hydrology," *Hydrological Science Bulletin*, vol. 43, pp. 43-69, 1979.
- [13] N. J. de Vos, T. H. M. Rientjes, "Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation", *Hydrology and Earth System Sciences, European Geosciences Union*, 2005, 9, pp. 111-126.
- [14] Holger R. Maier, Graeme C. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications", *Environmental Modelling & Software*, ELSEVIER, 2000, 15, pp. 101-124.
- [15] T. Hu, P. Yuan, etc. "Applications of artificial neural network to hydrology and water resources", *Advances in Water Science*, NHRI, 1995, 1, pp. 76-82.
- [16] Q. Ju, Z. Hao, etc. "Hydrologic simulations with artificial neural networks", *Proceedings-Third International Conference on Natural Computation, ICNC*, 2007, pp. 22-27.
- [17] G. WANG, M. ZHOU, etc. "Improved version of BTOPMC model and its application in event-based hydrologic simulations", *Journal of Geographical Sciences*, Springer, 2007, 2, pp. 73-84.
- [18] K. Beven, M. Kirkby, "A physically based, variable contributing area model of basin hydrology", *Hydrological Sciences Bulletin*, Springer, 1979, 1, pp.43-69.
- [19] K. Thirumalaiah, and C.D. Makarand, *Hydrological Forecasting Using Neural Networks Journal of Hydrologic Engineering*. Vol. 5, pp. 180-189, 2000.
- [20] G. WANG, M. ZHOU, etc. "Improved version of BTOPMC model and its application in event-based hydrologic simulations", *Journal of Geographical Sciences*, Springer, 2007, 2, pp. 73-84.
- [21] H. Goto, Y. Hasegawa, and M. Tanaka, "Efficient Scheduling Focusing on the Duality of MPL Representatives," *Proc. IEEE Symp. Computational Intelligence in Scheduling (SCIS 07)*, IEEE Press, Dec. 2007, pp. 57-64.
- [22] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, "Artificial neural networks in hydrology I: preliminary concepts", *Journal of Hydrologic Engineering*, 5(2), pp.115-123, 2000