Flood Forecasting using Adaptive Neuro-Fuzzy Inference System (ANFIS)

Dushyant Patel[1] , Dr. Falguni Parekh[2]

PG Student, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410,Gujarat, India[1]

Offg. Director & Associate Professor, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410, Gujarat, India[2]

***Abstract*— The aim of the present study is to explore applicability of artificial intelligence techniques such as ANFIS (Adaptive Neuro-Fuzzy Inference System) in forecasting flood for the case study, Dharoi Dam on the Sabarmati river near village Dharoi in Kheralu Taluka of Mehsana District in Gujarat State, India. The proposed technique combines the learning ability of neural network with the transparent linguistic representation of fuzzy system. ANFIS models with various input structures and membership functions are constructed, trained and tested to evaluate the models. Statistical indices such as Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of Determination (R^2) and Discrepancy Ratio (D) are used to evaluate performance of the ANFIS models in forecasting flood. This objective is accomplished by evaluating the model by comparing ANFIS model to Statistical method like Log Pearson type-III method to forecasting flood. This comparison shows that ANFIS model can accurately and reliably be used to forecast flood in this study.**

Keywords— ***Adaptive Neuro-Fuzzy Inference System, Flood forecasting, Statistical method.***

I. INTRODUCTION

Flood forecasting play very important role in water resources and flood management. Flood forecasting is the study of extreme values of the inflow of a river. The basic formulae of statistical methods are computation of mean flow, standard deviation and probability through normal distribution curve. With the advent of computers, mathematical models are increasingly being used for flood forecasting. A mathematical model is a set of mathematical expressions, relationships and logic rules that mimic the behavior of a physical process. Modern artificial intelligence methods such as neuro-fuzzy systems can be used for forecasting. These methods are very fast and reliable. Another advantage of these methods is that they can handle dynamic, non-linear and noisy data, especially when the relations are very complex and not able to understand easily. The purposes of this study are to explore the applicability of ANFIS in predicting peak discharge in the Sabarmati River and to identify the most fitted model to the study area.

To forecast flood accurately, inclusion of all past discharge and rainfall data is essential in this ANFIS model. In the present study to develop ANFIS model, past discharge and rainfall data of study area are used to forecast flood in a river system. ANFIS models with various input structures and membership functions are constructed, trained and tested to evaluate the models. Statistical parameters such as Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of Determination (R^2) and Discrepancy Ratio (D) are used to evaluate performance of the ANFIS models in forecasting flood. To show the accuracy and reliability of the ANFIS model; the model is compared with the statistical method i.e. Log Pearson type-III.

II. STUDY AREA

The Dharoi dam is the most important structure of the Sabarmati basin. It is located on Sabarmati River near village Dharoi in Kheralu taluka of district Mehsana in Gujarat State, India, 103 km from the source of the river. The latitude and longitude of the dam are 24^o 00' N and 72^o 52'E, respectively. Its purposes are water supply to the cities of Ahmedabad and Gandhinagar, irrigation, flood control and power generation of power.

The catchment area of the river at the dam site is 5,540 sq. km. The dam was completed in the year 1976. The FRL and HFL of the dam are at a level of 189.59 m and 175.87 m respectively. The dead storage and live storage capacity of the reservoir (as per revised capacity plan after 50 years) are 131.99 and 775.89 MCM, respectively.

For this study, daily rainfall data and peak discharge data are collected for Flood Forecasting of Sabarmati River.

III. METHODOLOGY

Adaptive Network-Based-Fuzzy Inferences System (ANFIS) approach was employed in this study.

The ANFIS architecture consists of 5 layers such as input layer, fuzzification layer, inferences process, defuzzification layer, and final output layer. Typical architecture of ANFIS is shown by Figure 1.

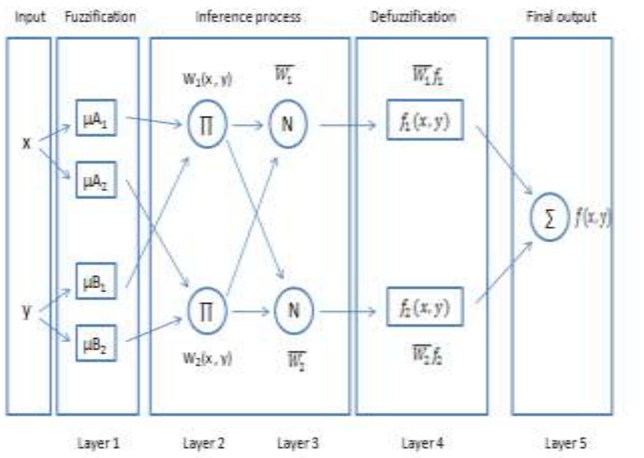

Figure 1: Typical Architecture of ANFIS

In this process there are five layers. In layer 1, input is to be *fuzzyfied.* Inference process and rules are applied in layer 2 and 3. Calculation of output for each corresponding rules are carried out in layer 4 and then in layer 5 all outputs from layer 4 are combined to get one final output.

The main objective of the ANFIS is to determine the optimum values of the equivalent fuzzy inference system parameters by applying a learning algorithm. The parameter optimization is done in such a way during training session that the error between the target and the actual output is minimized. Hybrid algorithm is used for optimization, which is combination of least square estimate and gradient descent method. The parameters to be optimized in ANFIS are the premise parameters. These parameters define the shape of the membership functions. The following parameters designate the overall output of the system. The optimized parameters are used in testing session to calculate the prediction.

The objectives for the study are,

i. Development of river stage-discharge ANFIS model.
ii. Validation of the model.
iii. Performance evaluation of the model for the Sabarmati river system.
iv. Comparison of ANFIS model and Statistical method.

In this study, basic model is constructed by 2 inputs and 1 output. The inputs are rainfall data and peak discharge data while for the output is predicted discharge. In this ANFIS model 70% data is used for Training and 30% data is used for Validation. After load the data in ANFIS Editor the Fuzzy Inference System (FIS) is generated with taking 5 number of linear triangular membership function.

The statistical method i.e. Log Pearson Type III method was applied to compare with ANFIS model. In this statistical method 70% data is used for Training and 30% data is used for Validation.

In Log Pearson Type III method, the variate (i.e., Q in this case) is first transformed into logarithmic form before analyzing the data. If Q is the variate of a random hydrologic series, then another series of y variates is thus obtained such that ,

$$y = log_{10} Q$$

The value of the variate Y_T for any reccurence interval T_r is given by

$$Y_T = \bar{y} + K.\sigma$$

Where, K= The frequency factor which depends upon the recurrence interval T_r and the coefficient of skew (g). The value of g is given by,

$$g = \frac{n\sum(y-\bar{y})^3}{(n-1)(n-2)\sigma^3}$$

Where, $\bar{y}$ = mean of y- values = $\frac{\sum y}{n}$

N = sample size (= no. of years of record)

And σ = standard deviation of y = $\sqrt{\frac{\sum(y-\bar{y})^2}{n-1}}$

The value of Q_T can be obtained as

$$Q_T = antilog(Y_T) = (10)^{Y_T}$$

The model is validated on the remaining 30% of the data by evaluating the following statistic performance indicators: Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of determination (R^2) and Discrepancy Ratio (D) which is given below:

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q(i)-\hat{Q}\ (i))^2}{n}}$$

Correlation coefficient:

$$R = \frac{\sum_{i=1}^{n}(Q\ (i)-\bar{Q})(\hat{Q}(i)-\tilde{Q})}{\sqrt{\sum_{i=1}^{n}(Q\ (i)-\bar{Q})^2\sum_{i=1}^{n}(\hat{Q}(i)-\tilde{Q})^2}}$$

Coefficient of determination:

$$R^2 = \frac{\left(\sum_{i=1}^{n}(Q(i)-\bar{Q})(\hat{Q}(i)-\tilde{Q})\right)^2}{\sum_{i=1}^{n}(Q(i)-\bar{Q})^2\sum_{i=1}^{n}(\hat{Q}(i)-\tilde{Q})^2}$$

Discrepancy Ratio (D):

$$D = \frac{\sum_{i=1}^{n} Q(i)}{\sum_{i=1}^{n} \hat{Q}(i)}$$

Where $\hat{Q}(i)$ is the n estimated discharge value, Q(i) is the n observes discharge value, $\bar{Q}$ is the mean of the observed discharge values, and $\tilde{Q}$ is the mean of the estimated discharge values.

IV. RESULT AND ANALYSIS

70% data is used for Training and 30% data is used for Validation. The results obtained from ANFIS are shown below:

ANFIS

The Training data results for ANFIS is shown below in fig.2 and fig.3.

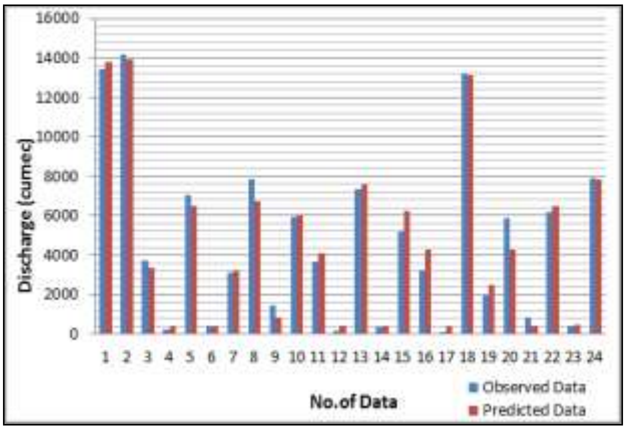

Fig.2 Comparison of observed & predicted peak discharge for Training using ANFIS

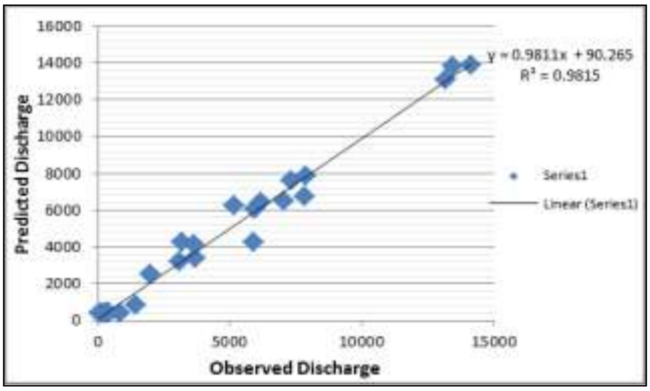

Fig.3 Scatter plot for Observed vs Predicted Peak Discharge for Training using ANFIS

The Validation data results for ANFIS is shown below in fig.4 and fig.5.

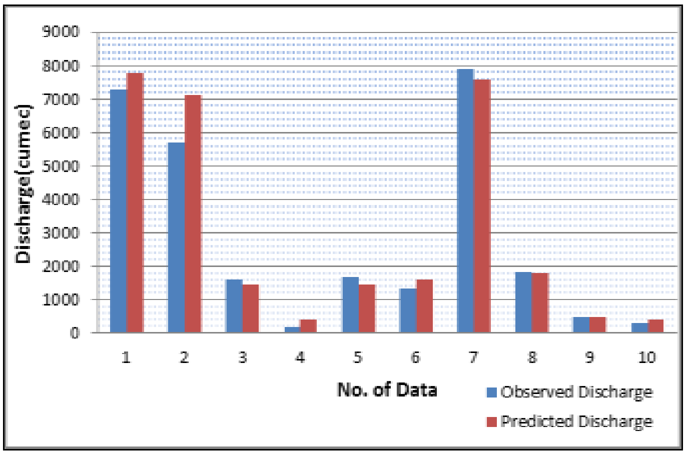

Fig.4 Comparison of observed & predicted peak discharge for Validation using ANFIS

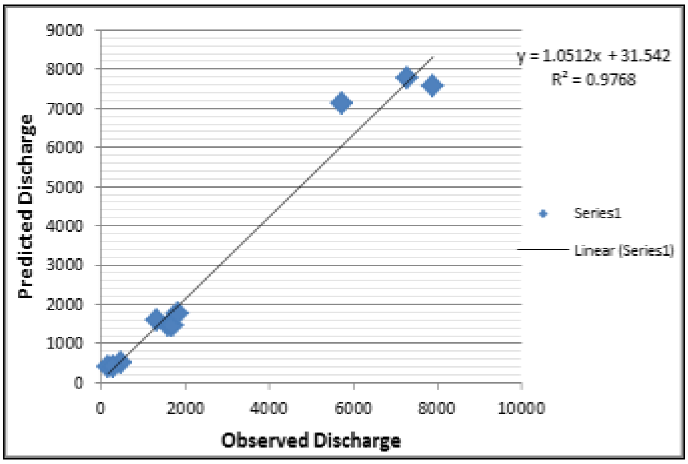

Fig.5 Scatter plot for Observed vs Predicted Peak Discharge for Validation using ANFIS

The forecast accuracy of model using ANFIS Tool was evaluated by calculating the following statistic performance indicators: Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of determination (R^2) and Discrepancy Ratio (D) described in Table I.

TABLE I

PERFORMANCE EVOLUTION OF MODEL ON TRAINING AND VALIDATION PERIOD FOR ANFIS MODEL

	RATIO 70-30 %				
ANFIS	PHASE	RMSE	COEFFICIENT OF CORRELATION (R)	R^2	DISCREPANCY RATIO (D)
	Training	577.00	0.99	0.98	0.99
	Validation	508.15	0.98	0.97	0.94

LOG PEARSON TYPE III METHOD

The Training data results for LOG PEARSON TYPE III method is shown below in fig.6 and fig.7.

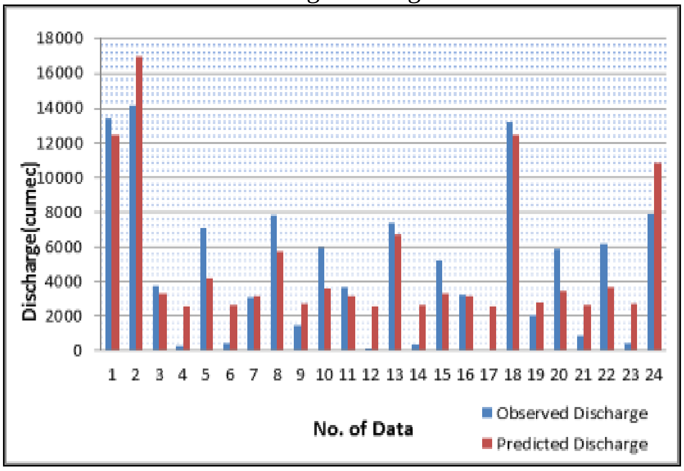

Fig.6 Comparison of observed & predicted peak discharge for Training using Log Pearson Type III.

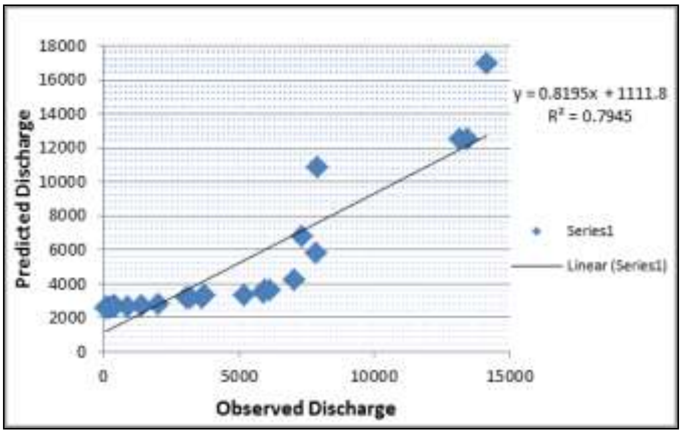

Fig.7 Scatter plot for Observed vs Predicted Peak Discharge for Training using Log Pearson Type III

The Validation data results for LOG PEARSON TYPE III method is shown below in fig.8 and fig.9.

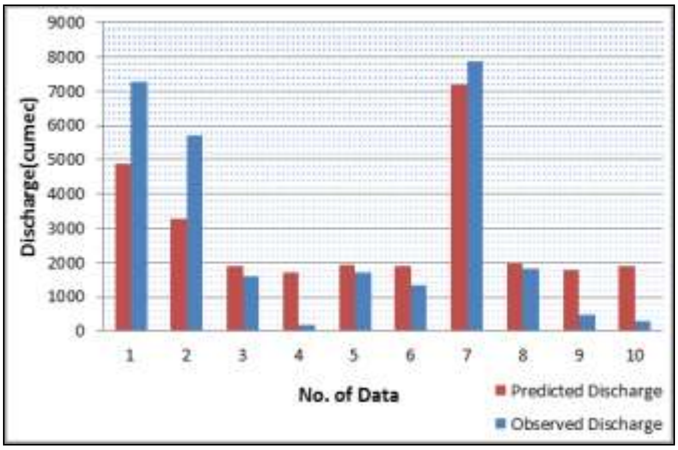

Fig.8 Comparison of observed & predicted peak discharge for Validation using Log Pearson Type III

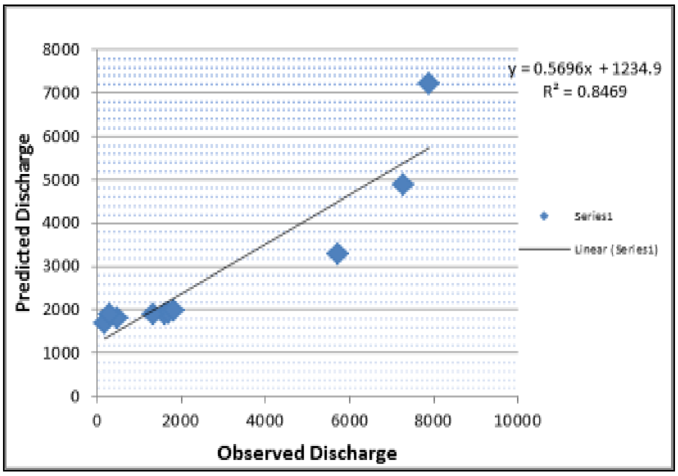

Fig.9 Scatter plot for Observed vs Predicted Peak Discharge for Validation using Log Pearson Type III

The accuracy of model using LOG PEARSON TYPE-III method was evaluated by calculating the following statistic performance indicators: Root Mean Square Error (RMSE), Correlation Coefficient (R), Coefficient of determination (R^2) and Discrepancy Ratio (D) described in Table II.

TABLE II

PERFORMANCE EVOLUTION OF MODEL ON TRAINING AND VALIDATION PERIOD FOR LOG PEARSON TYPE III METHOD

	RATIO 70-30 %				
LOG PEARSON TYPE III	PHASE	RMSE	COEFFICIENT OF CORRELATION (R)	R^2	DISCREPANCY RATIO (D)
	Training	1946.55	0.89	0.79	0.94
	Validation	1383.38	0.92	0.84	0.99

V. CONCLUSION

In this study, an Adaptive Neuro-Fuzzy Inference System (ANFIS) model has been developed to run real time flood forecasting at Dharoi Dam. Statistical method i.e. LOG PEARSON TYPE III method was compared with ANFIS model to identify the best result.

ANFIS:

It is observed that using ANFIS, the coefficient of correlation of observed peak discharge and predict peak discharge for Training is 0.99 and for Validation is 0.98 as shown in Table: I and coefficient of determination of observed peak discharge and predict peak discharge for Training is 0.98 and Validation is 0.97 which is very good i.e. ANFIS model accurately predict the flood.

Log Pearson Type III Method:
It is concluded from the observed peak discharge and predicted peak discharge that the coefficient of correlation for Training is 0.89 and for Validation is 0.92 as shown in Table: II and coefficient of determination of observed peak discharge and predict peak discharge for Training is 0.79 and Validation is 0.84.
The value of R, R^2 and D are very near to one for training and validation of ANFIS Model compared to that of Log Pearson Type III method.

From the overall study, it is concluded that it is possible to forecast the peak discharge very precisely by soft computing technique Adaptive Neuro-Fuzzy Inference System (ANFIS) compared to statistical method, Log Pearson Type III at Dharoi dam.

REFERENCES

[1] A.K. Karl and A.K Lohani,. "Development of flood forecasting system using Statistical and ANN Techniques in the downstream catchment of Mahanadi Basin, India". *Journal of Water Resource and Protection, 2*(10),880-887, 2010.

[2] Dinesh C. S. Bisht and A. Jangid, "Discharge Modelling using Adaptive Neuro - Fuzzy Inference System" International Journal of Advanced Science and Technology Vol. 31, June, 2011

[3] H. Galavi, and T. S. Lee,"Neuro-fuzzy modelling and forecasting in water resources", Scientific Research and Essays, Vol. 7(24), pp. 2112-2121, 2012.

[4] Jang, J.-S.R.. ANFIS: Adaptive network based fuzzy inference system, *IEEE Transactions on Systems, Man and Cybernetics*, 23(3), 665-685, 1993.

[5] Jayawardena, A. W. & Lai Feizhou "Analysis and Prediction of Chaos in Rainfall and Stream Flow Time Series". J. Hydrol.153. 23-52,1994.

[6] K.W. Chau, C.L. Wu and Y.S. Li, "Comparison of several flood forecasting models in Yangtze River", *Journal of Hydrologic Engineering, ASCE*, Vol. 10, No. 6, 2005, pp. 485-491, 2005.

[7] N.Ullah and P.Choudhury"Flood Forecasting in River System Using ANFIS", American Institute of Physics, pp 694-699, 2010

[8] Nazrin Ullah ,"Flood Flow Modeling in a River System Using Adaptive Neuro-Fuzzy Inference System", Environmental Management and Sustainable Development, Vol. 2,No. 2, 2013.

[9] R. D. Singh "Real time Flood Forecasting –India Experiences" National Institute of Hydrology, Roorkee, 1987.