Rainfall Prediction using Neural Net based Frequency Analysis Approach

Seema Mahajan
L. J. Institute of Engineering & Technology, S.G.Road
Ahmedabad-382210, Gujarat, India

Himanshu Mazumdar
Professor E.C and Head R&D Center
D.D.University,
Nadiad-387001, Gujarat, India

ABSTRACT
Rainfall prediction is very complex hydrologic process and is important as it holds the key to any countries' economy. Proposed model presents a new approach for yearly rainfall prediction of 30 Indian subdivisions. Yearly rainfall data of the Indian subdivision is available from IITM, Pune. The combination of Fast Fourier Transform (FFT) and Feed Forward Neural Network (FFNN) is applied for next one year rainfall prediction. Fast Fourier transform with filtering is performed on interpolated rainfall data to separate periodic components. These periodic components and delayed periodic components are given as input and desired output respectively to an FFNN for training. While testing the output of FFNN, inverse FFT gives the predicted rainfall value by amount of training input-output delay. This model is tested with 140 year's Indian subdivisions rainfall data. The experimental results of 30 subdivisions show that next one year rainfall prediction accuracy is above 92%.

Keywords
Rainfall prediction, Fast Fourier Transform, Feed Forward Neural Network

1. INTRODUCTION
India's agriculture production is highly dependent on its precipitation behavior of the monsoon rainfall. Monsoon is the main source of water. Average rainfall prediction is a prime important factor for crop planning. India receives 80% of rainfall during June to September. Monsoon follows cyclic behavior and has seasonality. As cycle period cannot be accurately predicted, this adds complexity of rainfall prediction. Many researchers worked on new strategies for rainfall forecasting. "Guhathakurta (2005)" constructed one year rainfall prediction model based on back propagation neural network for 14 districts of Kerala and also for Kerala as a whole subdivision. His results shows the root mean square error (RMSE) of area weighted values of district forecast is 6.26% and RMSE of Kerala as a whole subdivision is 13.29%. Rainfall prediction at up scaling of regions gives improved result than downscaling. "Somvanshi et al." implemented mean annual rainfall prediction models over Hyderabad region of Andhra Pradesh, India. They compared two models based on ANN and autoregressive integrated moving average (ARIMA). RMSE of ANN and ARIMA model is 145.14 and 262.57 respectively. ANN model found superior to trace the nonlinearity of rainfall data. "Sarah et al." developed a rainfall prediction model with the parameters like southern oscillation index (SOI), North Atlantic oscillation (NAO), sea level pressure (SLP), sea surface temperature (SST) and past month rainfall in Karoon basin. They have used artificial neural network along with fuzzy logic and wavelet functions. RMSE for two years, annual

and six months is 6.22, 7.11 and 13.15 respectively. "Iyengar et al." proposed rainfall forecasting model based on the rainfall data itself which is decomposable into six empirical time series. They used combination of ANN and simple regression technique to handle non linear and linear portion of rainfall respectively. The model efficiently predicts the rainfall with 75% to 80% of the interannual variability. In past years many researchers have introduced several new concepts and ideas in rainfall forecasting, but still rainfall prediction remains a challenge for researchers.

2. METHODOLOGY
There is regional variation in rainfall distribution of India and separate rainfall prediction model for each subdivision is desired. Global climatic factors like, distance from sea, latitude the Himalayan Mountain, distribution of land and water, surface pressure and wind, upper air circulation and western cyclones are affecting Indian sub-divisional rainfall. Rainfall is complex functions of these parameters. Researchers are correlating these independent parameters to predict rainfall. These parameters are embedded in rainfall time series. Rainfall prediction in the proposed model uses effect of these parameters as part of rainfall time series. Monthly rainfall data of all subdivisions (regions) of India shown in Figure. 1 from 1871 to 2010 is available from IITM, Pune. We downloaded these data to train and validate the model. Yearly rainfall values are derived from the corresponding monthly rainfall values of each subdivision. An interactive user friendly application is developed using C # .Net tool to implement the proposed model. Figure 2 shows pictorial representation of the model. The proposed model is performed in three steps of Data Interpolation, Fast Fourier Transform and Feed Forward Artificial Neural Network. Each step is described below.

2.1 Data Interpolation
Though 140 years rainfall data is available, 140 discrete points are not sufficient in time scale to develop a forecast model. Number of samples within the function and sample space are increased through data interpolation to get desired size of data set. Interpolation is performed using C# Drawcurve method. This method draws a cardinal spline passing through each point in the given array. This method draws a cardinal spline passing through each point of the given array with default tension value. 5601 data points are extracted from generated curve. 5601 is the empirical value. Each interpolated sample is equivalent to (140×365)/5600, 9.125 days. Each year is equivalent to 5600/140, 40 samples. Gujarat region interpolated data is shown in Figure 3.
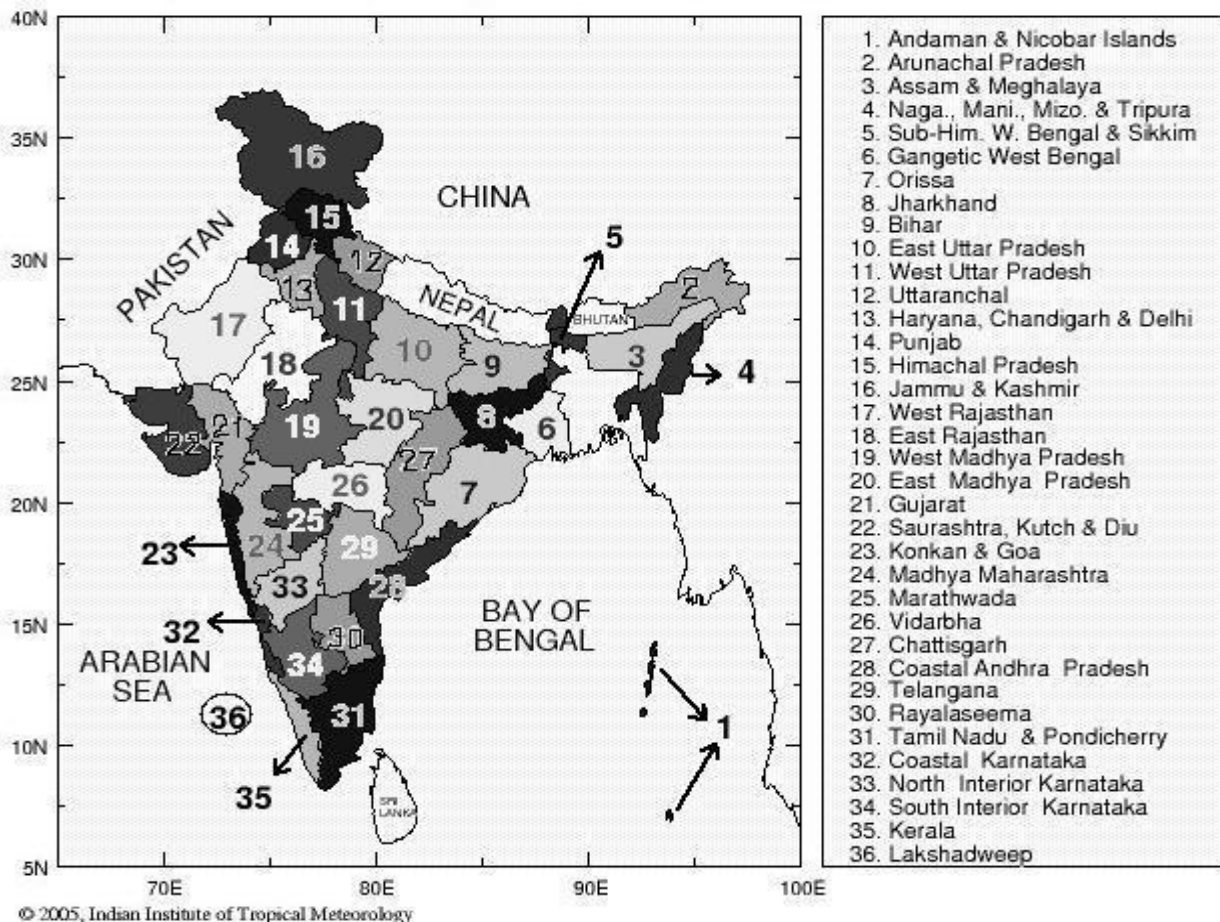
Fig 1: Meteorological subdivisions of India (Figure is made available from IITM)
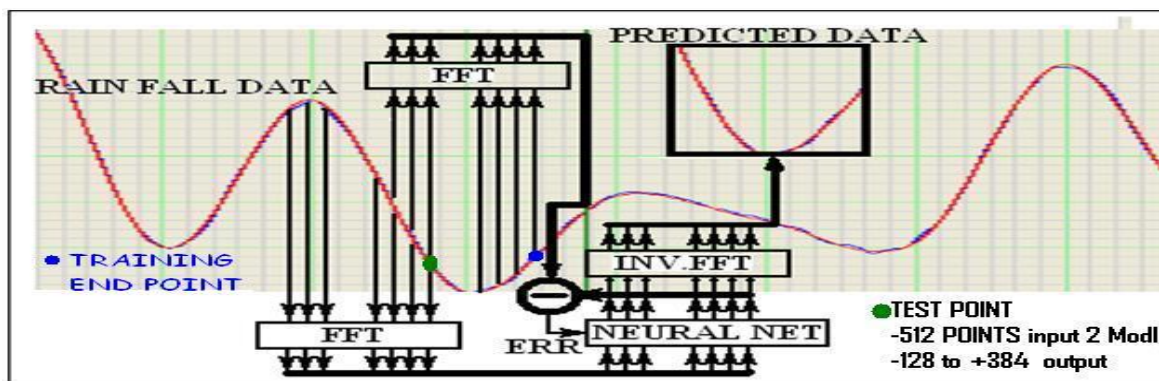

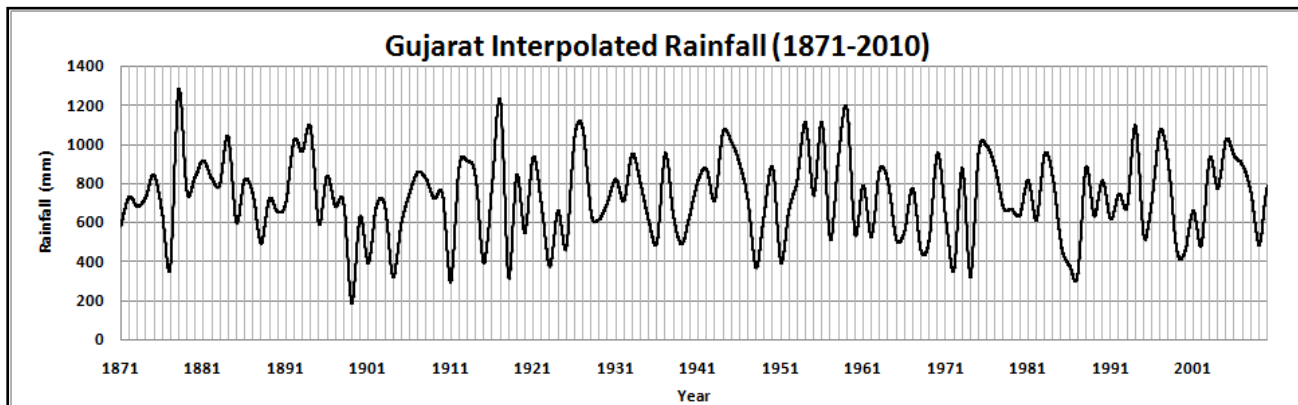**Fig 2: Pictorial representation of a proposed model**

Fig 3: Interpolated rainfall of Gujarat (1871-2010)

2.2 Fast Fourier Transform (FFT)

Interpolate rainfall time series data is a continuous and periodic in nature. Frequency domain conversion using FFT is used to achieve temporal association within rainfall time series. FFT separates frequency components of given segment of time series data. There are claims that a sunspot cycle (10-12 years) is associated with change in rainfall17. 512 rainfall samples of 12.8 years are used for FFT, where each sample is of 9.125 days. Training segment is divided into overlapping sub segments of 512 points each for input and output frequencies. Offset between input and output depends on the prediction period. An offset of 40 samples is kept one year prediction. A neural network is trained using these two FFT outputs after filtering out the high frequency components. 25% low frequencies are selected as input to neural network. High frequency component are having very low amplitude and do not contribute much to prediction error.

2.3 Feed Forward Artificial Neural Network (FFNN)

Neural Network is a mathematical model motivated by the structure of the brain. In a proposed study FFNN is used with 128 input neurons, 128 output neurons and 300 hidden neurons. Figure 4 shows block diagram of neural network. Y_i is the input layer, y_j is the hidden layer and y_k is the output layer. NN is calculated as

$$x_j = \sum_{i=0} y_i \times w_{ji} \quad \text{---------- (1)}$$

$$x_k = \sum_{j=0} y_j \times w_{kj} \quad \text{---------- (2)}$$

Activation function of each neuron is sigmoid function of the form

$$y = \frac{1}{1 + e^{-x}} \quad \text{---------- (3)}$$

Where y is the output and x is the input to neuron. In proposed method initially weights are random values. Weights are updated as

$$w_{kj(new)} = w_{kj(old)} + \eta(y_k - d_k) \times (1 - y_k) \times y_k \times y_{jk} \quad \text{---------- (4)}$$

Weights of hidden layer connection are updated as

$$w_{ji(new)} = w_{ji(old)} + \eta \left\{ \sum_{k=0} ((y_k - d_k)(1 - y_k) y_k) \right\} y_j (1 - y_j) y_i \quad \text{---------- (5)}$$

Experiments are carried out to optimize the model performance using different number of neurons in hidden layer from 50 to 350 neurons. Figure 5 show that prediction accuracy increases with increased number of hidden neurons. Experimentally 300 neurons are optimized for model. FFNN is trained up to tolerance level of training error. Though the numbers of weights used during training process are high, pruning shows that only 3-4 % out of them is effectively taking part in the prediction. Inverse FFT of FFNN output is carried out to get

the rainfall data in time domain. 5 sets of experiments are conducted for each year rainfall prediction. Each experiment is a new set of initial random weights. These weights are used to train the network
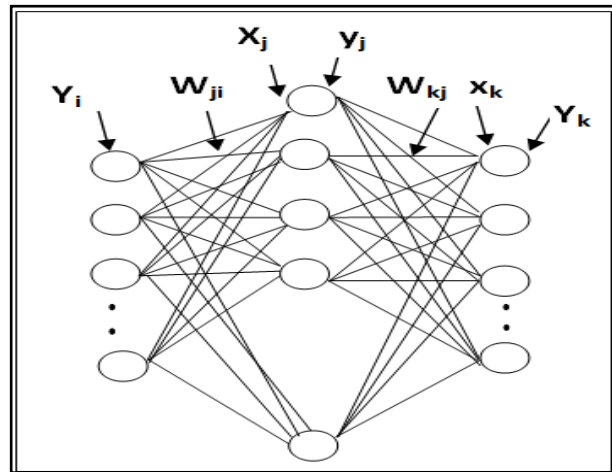

Fig 4: Block Diagram of Neural Net

3. RESULTS

The model efficiency is evaluated using many parameters like a) Number of Neurons – Initially model is trained with 50 neurons. Experimentally it is found that lower numbers (10-100) of hidden neurons are faster to learn but produces smooth curve with more error. More (100 – 500) number of neurons gives more accurate results. Finally 300 neurons are used for training. Figure 5 shows the Gujarat (21) rainfall prediction error of 2011 with varying number of neurons. b) Test cases with varying training and testing period- Rainfall data is available from 1871 to 2010. Random segment from this rainfall data is used during the training for prediction of immediate next year rainfall. The rainfall data of only 30 sub divisions of India is available out of 36 sub divisions shown in figure 1. Here 2 results are shown in detail

3.1 Gujarat (21) rainfall prediction of 1996 to 2010-

Model is trained for rainfall prediction of each year of Gujarat (21) from 1996 to 2010. Model input for each year prediction is rainfall data from all the years prior to the prediction year. A separate neural net is trained for each year rainfall prediction. 1996 rainfall data is predicted using all the previous years data from 1871 to 1995.Similarly all the other year's rainfall is predicted. Figure 6 shows actual and predicted rainfall data of Gujarat (21). Root Mean Square Error (RMSE) of each year is also shown in figure. RMSE is calculated as % of Long Period Average (LPA). Here LPA is considered for the period of 1871-2010. Model error for one year rainfall prediction is below 6%. Similarly other subdivision rainfall can be predicted but only Gujarat (21) result is shown here.

3.2 All 30 subdivision rainfall prediction of 2011 –

Each subdivision's 2011 rainfall is predicted. The model trained with the entire data excluding year 2011. 5 experiments are tried for each subdivision. Table 1 shows the rainfall prediction RMSE for all the 30 subdivisions. The RMSE is calculated as percentage error with respect to LPA. Table 1 also shows average error of 5 experiments. This average error is considered as final error for 2011 rainfall prediction of corresponding sub division. In all the cases the error is below 8 %. Table 2 shows prediction error of the

models proposed by other researchers. Graphical data is not available from other researchers.
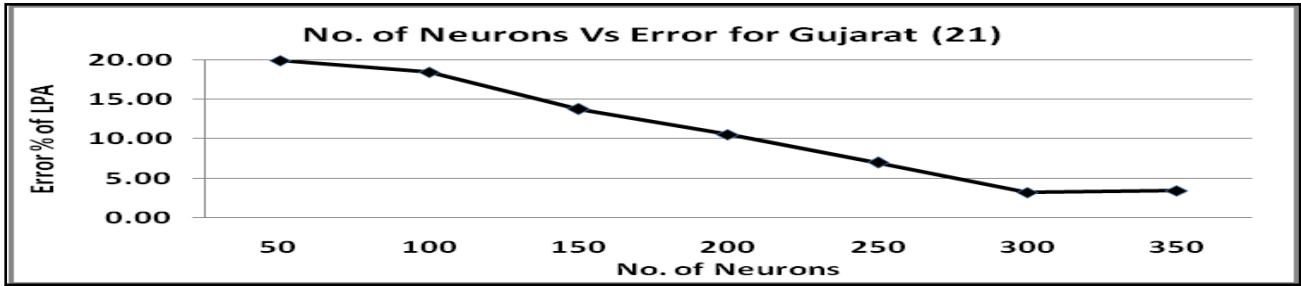

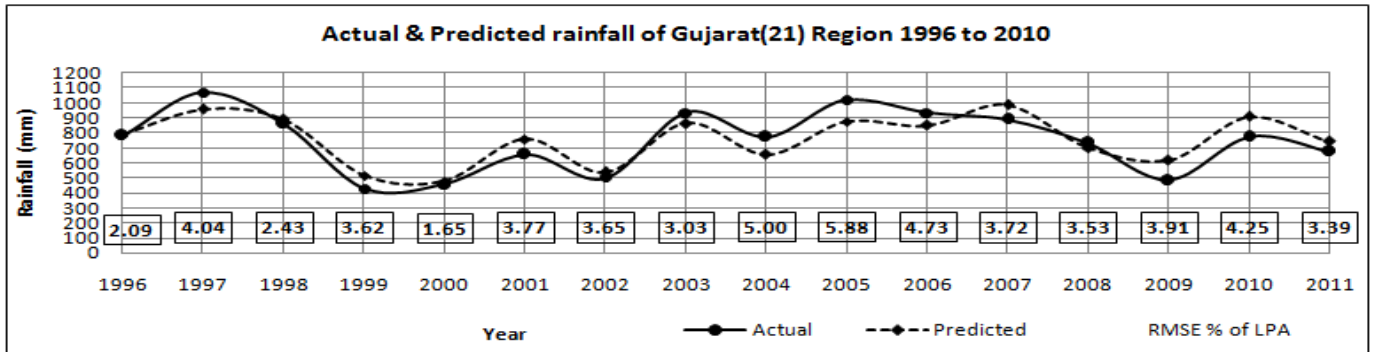**Fig 5: Error of rainfall prediction of 2011 with varying neurons.**

Fig 6: Shows actual and predicted rainfall data of Gujarat (21)

4. CONCLUSION

Artificial Neural Network model based on frequency analysis approach using Fast Fourier Transform is used for rainfall prediction. This model is tested as forecasting tool for one year rainfall prediction on regional (subdivision) scale of India. Model error is below 8%. It predicts the quantity of rainfall 1 year ahead of time which is adequate time for crop planning in agriculture country like India.

5. REFERENCES

[1] A. K. Sahai, M.K. Soman and V. satyan, All India summer monsoon rainfall prediction using an artificial neural network, Climate Dynamics (2000)

[2] A. El-shafie, M.Mukhlisin, Ali A. Najah and M.R. Taha, Performance of artificial neural network and regression techniques for rainfall-runoff prediction, International Journal of the Physical Sciences vol. 6(8)

[3] 8P. Guhathakurta, Long range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network, Scientific Correspondance

[4] V. K. Somvanshi, O. P. Pandey, P. K. Agrawal, N. V. kalanker, M. Ravi Prakash and Ramesh Chand, Modelling and prediction of rainfall using artificial neural network and ARIMA techniques, Indian Geophysical Journal (April 2006).

[5] Sarah Afshin, Hedayat Fahmi, Amin Alizadeh, Hussein Sedghi and Fereidoon Kaveh, Long term rainfall forecasting by integrated artificial neural network-fuzzy logic-wavelet model in Karoon basin, Scientific Research and Essays Vol. 6(6)

[6] Turgay Partal and Özgür Kişi , Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, Journal of Hydrology, Volume 342, Issues 1–2

[7] Lingzhi wang and Jiansheg Wu, Application of Hybrid RBF Neural Network Ensemble Model Based on Wavelet Support Vector Machine Regression in Rainfall Time Series Forecasting, Computational Sciences and Optimization (CSO), 2012 Fifth International Joint Conference

[8] Khaing Win Mar, Thinn Thu Naing, Optimum Neural Network Architecture for Precipitation Prediction of Myanmar, World Academy of Science, Engineering & Technology, 48 2008

[9] D. Nagesh Kumar, M. Janga Reddy and Rajib Maity, Regional Rainfall Forecasting using Large Scale Climate Teleconnections and Artificial Intelligence Techniques, Journal of Intelligent Systems, vol 16, No 4,2007

[10] Iyengar, R. N. and Raghukanth, S. T. G., Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. Meteorol. Atmos. Phys., 2004

[11] S. Karmakar, M. K. Kowar, P. Guhathakurta, Long-Range Monsoon Rainfall Pattern Recognition and Prediction for the Subdivision 'EPMB' Chhattisgarh Using Deterministic and Probabilistic Neural Network, 7th International conference on Advance in Pattern Recognition

[12] Sulochana Gadgil*, M. Rajeevan and Ravi Nanjundiah, Monsoon prediction – Why yet another failure?, Current Science, Vol. 88, NO. 9, 10 May 2005.

[13] Surajit Chattopadhyay, Goutami Chattopadhyay, Comparative study among different neural net learning algorithms applied to rainfall time series, Meteorological Applications Volume 15, Issue 2, pages 273–280, 2008

[14] Venkatraman Prasanna, Juvy Subere, Dwijendra K. Das, Srinivasan Govindarajan and Tetsuzo Yasunari,

Development of daily gridded rainfall dataset over the Ganga, Brahmaputra and Meghna river basins, Meteorological Applications, Royal Meteorological Society.

[15] http://www.tropmet.res.in/ - Indian Institute of Tropical Meteorology, Pune

Table 1. Indian Subdivision rainfall prediction error for year 2011

Subdivison no_ Subdivison Name	Experiment-1	Experiment-2	Experiment-3	Experiment-4	Experiment-5	Avg. of all 5 Experiment
3_Assam_Meghalaya	3.36	3.17	4.01	3.63	3.53	3.54
4_Naga_Mizo_Mani	3.09	3.73	3.54	3.07	3.18	3.32
5_Him_WBengal	4.61	4.01	3.64	3.41	4.19	3.98
6_GangticWestBengal	2.57	3.02	3.03	2.68	2.25	2.71
7_Orissa	2.58	3.12	3.22	2.34	2.97	2.85
8_Jharkhand	1.36	3.64	2.62	2.64	2.57	2.56
9_Bihar	2.71	3.98	3.14	2.83	3.38	3.21
10_EastUttarPradesh	2.47	2.57	2.82	2.43	2.73	2.60
11_WestUttarPradesh	2.85	3.57	2.72	3.02	4.02	3.24
13_Haryana_Chadigarth	4.28	4.35	3.72	4.33	4.92	4.32
14_Punjab	3.11	3.69	3.37	4.22	4.22	3.72
17_West_Rajasthan	4.54	5.61	5.37	5.65	5.55	5.34
18_East_Rajastan	4.78	6.52	5.57	5.45	6.89	5.85
19_West_MadhyaPradesh	3.25	3.10	2.70	3.69	3.41	3.23
20_East_Madhya_Pradesh	1.55	2.67	2.92	2.45	3.45	2.61
21_Gujarat	3.93	4.07	3.93	4.19	3.26	3.88
22_Saurashtra	3.22	6.65	7.96	5.51	7.59	6.19
23_Konkan_Goa	3.89	4.78	4.93	5.65	4.32	4.71
24_MadhyaMaharashtra	3.67	4.91	2.63	3.57	4.26	3.81
25_Marathavada	5.59	6.86	5.81	6.16	5.29	5.94
26_Vidarbha	4.14	3.59	4.03	5.15	2.54	3.89
27_Chattisgarth	1.65	3.36	3.34	3.25	3.14	2.95
28_Coastal Andhra Pradesh	5.81	4.86	3.95	3.85	3.96	4.49
29_Telangana	2.91	3.51	3.63	3.76	3.69	3.50
30_Rayalaseema	3.70	4.63	4.22	4.62	5.16	4.47
31_TamilNadu_Pandicherry	3.24	3.64	3.14	3.84	3.20	3.41
32_CoastalKarnataka	2.88	4.18	3.64	2.56	4.39	3.53
33_NorthInteriorkarnataka	3.71	4.61	3.57	3.33	4.63	3.97
34_SouthInteriorKarnataka	1.98	3.79	3.07	4.68	5.02	3.71
35_Kerala	3.58	2.70	3.59	1.69	2.32	2.78

Table 2. Rainfall prediction model error comparison

Model	RMSE % of LPA (Long Period Average)
Long range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network	13.29
Long term rainfall forecasting by integrated artificial neural network-fuzzy logic-wavelet model in Karoon basin	7.11
Intrinsic mode functions and a strategy for forecasting Indian monsoon	20
Proposed Model	8