Comparative Analysis of Machine Learning Techniques for Rain Prediction

Sudipti Banerjee[1], Amit Kumar Manjhvar[2]

[1, 2]*Department of Computer Science & Engineering, Madhav Institute of Technology & Science, Gwalior (M.P.), India*

Abstract: Heavy rain prediction is a major concern for earth science department because it is closely related to the economy and existence of human. Heavy rain is a cause for natural disasters like flood and drought that are encountered by individuals across the world per annum.

Accuracy of rain prediction has considerable importance for countries like Asian nation whose economy is basically captivated with agriculture. Due to dynamic nature of atmosphere, applied math techniques fail to produce sensible accuracy for rain statement.

Nonlinearity of precipitation information makes Artificial Neural Network a far better technique. Review work and comparison of various approaches and algorithms employed by researches for rain prediction is shown in an exceedingly tabular format. Intension of this paper is to provide non-experts easy accessibility to the techniques and approaches utilized in the sector of rain prediction.

Keywords: Rain Prediction, Artificial Neural Networks, Adaptive, Self-Organising, Robustness.

I. INTRODUCTION

Rain prediction is useful to avoid flood that save lives and properties of humans. Moreover, it helps in managing resources of water. Information of rain in advance helps farmers to manage their crops superior that lead to growth of country's economy. Fluctuation in rain schedule and its amount makes rain prediction a difficult task for earth science scientists. In all the services provided by earth science department, weather foreseeing stands out on prime for all the countries across the world. The task is incredibly complicated because it needs numbers of specialised and conjointly all calls are created with none certainty. Section 2 discusses the various ways used for rain prediction for foretelling with their limitations. Various neural networks algorithmic program that are used for prediction are mentioned with their steps thoroughly. Section 3 categorizes various approaches and algorithms used for rain prediction by various researchers in today's era. Finally, section 4 presents conclusion of paper.

II. BACKGROUND THEORY

Two broadly used strategies for rain foretelling are: applied math strategies and Numerical Weather Prediction (NWP) model [16]. Nature of rainfall data is non-linear. Frequency, intensity and quantity are main characteristics for statistic rain. These values are often varied from one position on earth to different position of earth and from just one occasion to alternative time. Every statistical model has some drawbacks. Combination of AR and MA along forms a general and helpful category of the statistic model called ARMA model. ARMA model is just helpful for stationary time-series information and foretelling of short-term rain. The applied mathematics approaches don't have the flexibility to spot nonlinear patterns and irregular trend within the statistic [16].

A. ARMA Model

Box and Jenkins [29] projected a technique that consists of 4 steps:

Step 1. In the identification stage, determine statement is employed is employed to specify the response of series and identification of candidate. The determine statement considers statistic that may be utilized in later statements, probably for differentiating them, and calculates autocorrelations, cross correlations and partial autocorrelations. Step 2 and 3. In the estimation and diagnostic checking stage, ESTIMATE statement helps ARIMA model to suit to the variable taken within the previous determine statement, and to estimate the parameters of that model. Step 4. In the foretelling stage, FORECAST statement is employed to forecast future values of the statistics. Confidence intervals are generated for these forecasts from the ARIMA model. Autocorrelation operate (ACF) and also the Partial Autocorrelation operate (PACF) are necessary analytical tools used with the statistical analysis and foretelling [30]. Main uses of those models are to meter the applied mathematics relationships between observations in very single information series. ACF has massive advantage of measuring the number of linear dependences between results of a statistic which is able to be separated by a lag k. In order to spot the model (step 1), ACF and PACF have to be compelled to be evaluated. They are used to guess the shape of the model and to get approximate estimates of the parameters in addition [30].
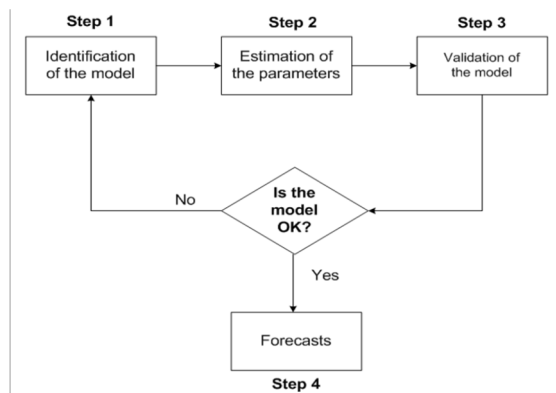
Fig: 1 Outline of Box-Jenkins Methodology.

B. Artificial Neural Network

ANN is a process model that's impressed by the human brain [31]. ANN contains an enormous range of interconnected neurons, that principally operate in parallel, and are well structured. Categories of neural networks are either single layer or multi-layer. Layer between input layer and output layer is named as hidden layer. A single-layer feed forward (SLFF) neural network consists one input layer whose nodes have weights assigned and one output layer. A multi-layer feedforward (MLFF)

neural network architecture can be developed by adding hidden layers in SLFF neural network.

1) *Back-Propagation Neural Network:* BPNN is formed of MLFF neural network that contains one input layer, hidden layers and one output layer. BPNN design with one hidden layer is shown in figure. The ultimate goal of BPNN is to decrease the calculated error obtained from the difference between the calculated output and desired output of the neural network by adjusting the weights (self-organising) after each iteration. So in BPNN, each information is propagated in backward direction until the calculated error is very small or zero. There are three phases of BPNN training: (a) using FFNN for training process of input. Adjustable of weights and nodes are created during this section, (b) to calculate the error, and (c) modification of weights. ANN model has nice ability to be told by doing correct adjustment of those parameters for achieving the required output. During the training process, this output may fit to the data very well, but it may provide poor results during the testing process. This suggests that the neural network may not generalize well. This may well be attributable to overfitting or overtraining of Information [32], which can be controlled by analysing the error during training process and stopping the process when the error reaches a minimum threshold with reference to the testing set [33]. Alternate choice to create the neural network generalise enough is by doing little changes within the range of layers and neurons within the inputs, while not dynamical the output parts. However, best neural spec choice may be a heuristic approach. Solution is to stay the design of neural-network comparatively straight forward and tiny [34], because complex architectures are much more prone to overfitting [35].
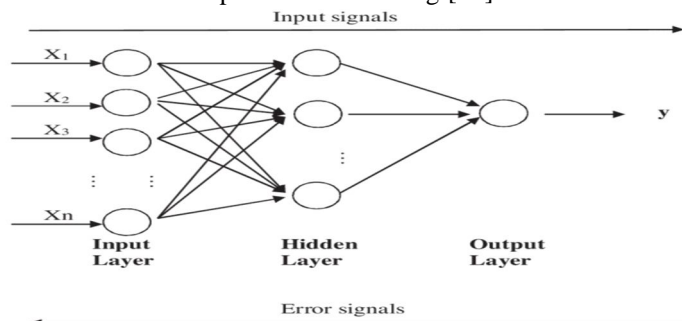

Fig: 2 A BPNN architecture with one hidden layer [39].

2) *Cascade Forward Back Propagation Network:* The CFBP network shown in Figure a pair of 3 is one of the artificial neural network types, which is used for the prediction of new output data. All the layers in network aren't solely connected with its previous layer however conjointly connected with input. Inputs are provided to each layer in network. CFBP network is adaptive & robust in nature.
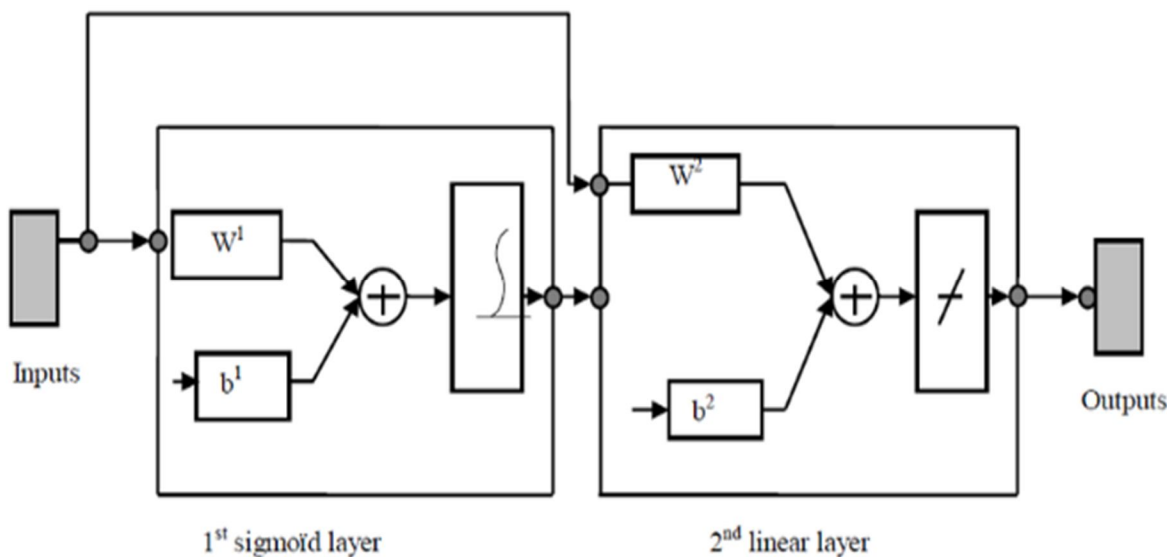
Fig: 3 Cascade Forward Back Propagation Network.

3) *Layer Recurrent Network:* In this variety of neural network, connections between units create a directed cycle. Unlike different feedforward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs [36]. RNN are neural networks with a feedback loop. The previous processes of hidden layer and practical outputs are feedback to the network as a part of the input to consequent hidden layer processes.
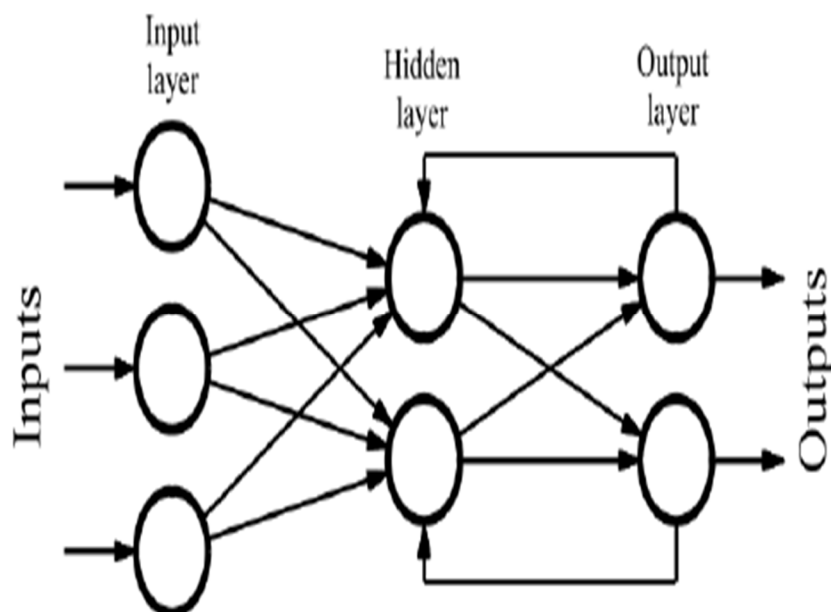

Fig: 4 Layer Recurrent Neural Network.

C. *Support Vector Machine (SVM)*

SVM is additionally one variety of feed forward network. Support vector machines are applicable for tasks like pattern classification, nonlinear regression etc. SVM were created by Vapnik and his co-workers that has been used for supervised learning because of- (i) higher generalization ability than different NN models. (ii) SVM answer is identical, optimal and absent from local minima. (iii) Applicable to non-vectoral data (Strings and Graphs) and (iv) Very few parameters are needed for tuning the learning m/c. Very few scientists have applied this system for rain prediction and results were acceptable.

III.   LITERATURE REVIEW

P. Goswami and Srividya [1] have combined RNN and TDNN characteristic and conclusion of their work was that composite models provides higher accuracy than the single model.

C. Venkatesan et al. [2] used Multilayer Feed Forward Neural Networks (MLFNN) for predicting Indian summer monsoon rainfall. Error Back Propagation (EBP) algorithmic rule is trained and applied to predict the rain. Three network models with two, three and ten input parameters have analysed. They conjointly compared the output result with the applied math models.

A. Sahai et al. [3] used error back propagation algorithmic rule for Summer Monsoon Rain Prediction of India on monthly and seasonal statistic. They used knowledge of previous five years of monthly and periodic mean rain values for rain prediction.

N. Philip and K. Josheph [5] used ABF neural network for yearly rain prediction Kerala region. The work suggests that ABFNN performs better than the Fourier analysis.

V. Somvanshi et al. [7] predicted rainfall of Hyderabad, INDIA region using ANN model. They also compared ANN with ARIMA technique. They used past four months rainfall data as inputs to neural network model.

S. Chattopadhyay and M. Chattopadhyay [9] have used two parameters minimum temperature and most temperature for rain foretelling.

S. Chattopadhyaya and G. Chattopadhyaya [10] used Conjugate Gradient Decent (CGD) and Levenberg-Marquardt (LM) learning algorithmic rule for tutoring.

Performances of each algorithms were same in prediction task.

Wu et al. [12] predicted the rainfall of India and China. They applied Modular

Artificial Neural Network (MANN).

MANN's performance was compared with LR, K-NN and ANN.

K. Htike and O. Khalifa [13] used yearly, biannually, quarterly and monthly rainfall data for rainfall prediction.

They trained four totally different targeted Time Delay Neural Networks (FTDNN) for rain foretelling. Highest prediction accuracy was provided by the FTDNN model when yearly rainfall data is taken for training.

S. Kannan and S. Ghosh [14] contributed towards developing K-mean cluster technique combined with decision tree algorithm, CART, is used for rainfall states generation from large scale atmospheric variables in a river basin. Rainfall state on daily basis is derived from the historical daily multi-site rainfall data using K-mean clustering.

M. Kannan et al. [15] predicted short term rainfall. Empirical method technique is used for prediction task. Data of three specific months for five years is analysed for specific region. Clustering is used for grouping the elements.

G. Geetha and R. Selvaraj [16] used ANN model for predicting monthly rain of Chennai region.

M. Sharma and J. Singh [17] considered parameters like rain, maximum and minimum temperature, and relative humidity. They predicted weekly rainfall over Pant nagar region. ANN obtained higher prediction accuracy than multiple linear regression model.

J. Abbot and J. Marohasy [18] used Time Delay Recurrent Neural Network (TDRNN) for monthly rain prediction over Australia region.

A. Kumar et al. [19] predicted average rainfall over Udipi district of Karnataka. They used ANN models for prediction task of rain. They complete that Back-Propagation Algorithm (BPA) was higher than the layer recurrent and cascaded back propagation.

Soo-Yeon Ji et al. [21] predicted the hourly rainfall. CART and C4.5 are used for prediction, which may provide hidden important patterns with their reasons. There were 18 variables used from weather station. 10-fold cross validation technique is performed for validation purpose. CART performed better than C4.5.

S. Nanda et al. [24] expected rain using a advanced applied math model ARIMA and three ANNs models which square measure MLP, LPE (Legendre Polynomial Equation) and FLANN (Functional-Link Artificial Neural Network). In comparison, FLANN gives better prediction accuracy compared to the ARIMA model.

A. Naik and S. Pathan [25] used the ANN model for rain prediction. They modified back propagation algorithm which was more robust than the simple back propagation algorithm

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue II, Feb 2019- Available at www.ijraset.com

TABLE I

Authors	Region	Dataset Time Period	Techniques	Accuracy Measure	Rainfall Predicting Attribute
P. Goswami, Srividya (1996) [1]	Global (all over India)	Yearly (135 years)	Artificial Neural Network (EBP)	Relative percentage error	Mean rainfall
C. Venkatesan et al. (1997) [2]	Global (all over India)	Monthly (1939-1994)	Artificial Neural Network (EBP)	RMSE	Min-Max temperature
A. Sahai et al. (2000) [3]	Global (all over India)	Monthly (1876-1994)	Artificial Neural Network (EBP)	RMSE, correlation coefficient	Min-Max temperature
N. Philip et al. (2001) [4]	Local (Kerala)	Monthly (1893-1933)	ABFNN	RMSE	Wind, temperature, precipitation, latitude-longitude, sea surface pressure
N. Philip, K. Joseph (2002) [5]	Local (Kerala)	Yearly (1893-1933)	ABFNN	RMSE	Wind, temperature, precipitation, latitude-longitude, sea surface pressure
N. Chantasut et al. (2004) [6]	Local (Chao Pharya River)	Monthly (1941-1999)	Artificial Neural Network (EBP)	MSE	Temperature
V. Somvanshi et al. (2006) [7]	Local (Hyderabad)	Yearly (103 years)	ANN, ARIMA	RMSE, MAE	Humidity, Min-Max temperature
S. Chattopadhyay (2007) [8]	Global (all over India)	Monthly	Artificial Neural Network (EBP)	MSE	Temperature, Rainfall
S. Chattopadhyay, M. Chattopadhyay (2007) [9]	Global (all over India)	Monthly	Multilayer perceptron	MSE	Min-Max temperature
S. Chattopadhyay, G. Chattopadhyay (2008) [10]	Global (all over India)	Monthly	Artificial Neural Network (EBP)	MSE	Min-Max temperature
P. Guhathakurta (2008) [11]	Global (all over India)	Yearly (1941-2005)	Artificial Neural Network (EBP)	RMSE	Min-Max temperature
C. Wu et al. (2010) [12]	Global (India, China)	Daily, Monthly	Modular Artificial Neural Network	RMSE	Min-Max temperature
K. Htike, O. Khalifa (2010) [13]	Global (India)	Yearly, Monthly, Bi-annually, Quarterly	Focused Time Delay Neural Network	MAPE	Temperature, solar radiation. evaporation
S. Kannan, S. Ghosh	Local (River)	Daily (50	Decision tree,	MSE	Temperature, pressure, wind

		years)	CART, K-mean		speed, rainfall
M. Kannan at el. (2010) [15]	Global (India)	Quarterly (5 years)	Regression	MSE	Min-Max temperature, wind direction, humidity, rainfall
G. Geeta, R. Selvaraj (2011) [16]	Local (Chennai)	Monthly (1978-2009)	Multilayer Back Propagation Neural Network	RMSE	Wind speed, mean temperature, relative humidity, aerosol values
M. Sharma, J. Singh (2011) [17]	Local (Pantnagar, India)	Weakly (39 Years)	Multiple Regression Model, ANN(EBP)	Absolute mean difference	Min-Max temperature, relative humidity, pan evaporation
J. Abbot, J. Marohasy (2012) [18]	Local (Australia)	Monthly (1900-2009)	TDRNN	RMSE, Pearson correlation co-efficient	Rainfall, climatic indices, atmospheric temperature, solar data
A. Kumar et al. (2012) [19]	Local (Udipi)	Monthly (1960-2010)	EBPNN, CBPNN, LRN	MSE	Average humidity, Average wind speed
R. Deshpande (2012) [20]	Local	Monthly	Elman neural network	MSE	Rainfall
G. Shrivastava et al. (2013) [22]	Local (Ambikapur)	Yearly (1951-2011)	Artificial Neural Network (EBP)	MSE	Humidity, dew point, pressure
C. Wu, K. Chau (2013) [23]	Global (India, China)	Daily, Monthly	Moving Average, ANN	RMSE	Min-Max. temperature
Priya et al. (2014) [26]	Global (India)	Monthly (1871-2010)	Artificial Neural Network (EBP)	RMSE	Min-Max temperature
V. Dabhi, S. Chaudhary (2014) [27]	Local (Anand)	Daily (1991-2002)	Wavelet-postfix-GP model, Wavelet ANN	MAE, MSE, Adjusted fitness	Min-Max. temperature, evaporation, relative humidity, rainfall

Table Presents Categorization of various Approaches of rain Prediction.

IV. CONCLUSION

The estimation of rain is of significant importance in terms of water resources management, human life and their surroundings. It is met with the inaccurate or incomplete estimation issues because rain estimation is affected from the geographical and regional changes and properties. This paper bestowed review completely different (of various) ways used for rain prediction and issues one would possibly encounter while applying different approaches for rain foretelling.

Due to nonlinear relationships in rain information and skill of learning from the past makes Artificial Neural Network a desirable approach from all offered approaches.

REFERENCES

[1] P. Goswami and Srividya, "A novel Neural Network design for long range prediction of rainfall pattern," Current Sci.(Bangalore), vol. 70, no. 6, pp. 447-457, 1996.

[2] C. Venkatesanet, S. D. Raskar , S. S. Tambe , B. D. Kulkarni , and R. N. Keshavamurty , "Prediction of all India summer monsoon rainfall using Error-Back-Propagation Neural Networks," Meteorology and Atmospheric Physics, pp. 225-240, 1997.

[3] A. K. Sahai, M. K. Soman, and V. Satyan, "All India summer monsoon rainfall prediction using an Artificial Neural Network," Climate dynamics, vol. 16, no. 4, pp. 291-302, 2000.

[4] N. S. Philip and K. B. Joseph, "On the predictability of rainfall in Kerala-An application of ABF neural network," Computational Science- ICCS, Springer

Berlin Heidelberg, pp. 1-12, 2001.

[5]   N. S. Philip and K. B. Joseph, "A Neural Network tool for analyzing trends in rainfall," Comput. & Geosci.,vol. 29, no. 2, pp. 215-223,2003.

[6]   N. Chantasut, C. Charoenjit, and C. Tanprasert, "Predictive mining of rainfall predictions using artificial neural networks for Chao Phraya River," 4th Int Conf. of the Asian Federation of Inform. Technology in Agriculture and the 2nd World Congr. on Comput. in Agriculture and Natural Resources, Bangkok, Thailand, pp. 117-122, 2004.

[7]   V. K. Somvanshi, O. P. Pandey, P. K. Agrawal, N.V.Kalanker1, M.Ravi Prakash, and Ramesh Chand, "Modeling and prediction of rainfall using Artificial Neural Network and ARIMA techniques," J. Ind. Geophys. Union, vol. 10, no. 2, pp. 141-151, 2006.

[8]   S. Chattopadhyay, "Anticipation of summer monsoon rainfall over India by Artificial Neural Network with Conjugate Gradient Descent Learning," arXiv preprint nlin/0611010, pp. 2-14, 2006.

[9]   S. Chattopadhyay and M. Chattopadhyay, "A Soft Computing technique in rainfall forecasting," Int. Conf. on IT, HIT, pp. 19-21, 2007.

[10] S. Chattopadhyay and G. Chattopadhyay, "Comparative study among different neural net learning algorithms applied to rainfall time series, "Meteorological applicat., vol. 15, no. 2, pp. 273-280, 2008.

[11] P.Guhathakurta, "Long lead monsoon rainfall prediction using deterministic Artificial Neural Network model," Meteorology and Atmospheric Physics 101, pp. 93-108, 2008.

[12] C. L. Wu, K. W. Chau, and C. Fan, "Prediction of rainfall time series using Modular Artificial Neural Networks coupled with datapreprocessing techniques," J. of hydrology, vol. 389, no. 1, pp. 146- 167, 2010.

[13] K. K. Htike and O. O. Khalifa, "Rainfall forecasting models using Focused Time-Delay Neural Networks," Comput. and Commun. Eng. (ICCCE), Int. Conf. on IEEE, 2010.

[14] S. Kannan , Subimal Ghosh, "Prediction of daily rainfall state in a river basin using statistical downscaling from GCM output", Springer-Verlag, July- 2010.

[15] M.Kannan, S.Prabhakaran, P.Ramachandran, "Rainfall Forecasting Using Data Mining Technique", International Journal of Engineering and Technology Vol.2 (6), 397-401, 2010.

[16] G. Geetha and R. S. Selvaraj, "Prediction of monthly rainfall in Chennai using Back Propagation Neural Network model," Int. J. of Eng. Sci. and Technology, vol. 3, no. 1, pp. 211 213, 2011.

[17] M. A. Sharma and J. B. Singh, "Comparative Study of rainfall forecasting models," New York Sci. J., pp. 115-120, 2011.

[18] J. Abbot and J. Marohasy, "Application of Artificial Neural Networks to rainfall forecasting in Queensland, Australia," Advances in Atmospheric Sci.,vol. 29, no. 4, pp. 717-730, 2012.

[19] A. Kumar, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," Control and Syst. Graduate Research Colloq. (ICSGRC), pp. 82-87, 2012.

[20] R. R. Deshpande, "On the rainfall time series prediction using Multilayer Perceptron Artificial Neural Network," Int. J. of Emerging Technology and Advanced Eng., vol. 2, no. 1, pp. 148-153, 2012.

[21] Soo-Yeon Ji, Sharad Sharma, Byunggu Yu, Dong Hyun Jeong, "Designing a Rule-Based Hourly Rainfall Prediction Model", IEEE IRI 2012, August – 2012.

[22] G. Shrivastava, S. Karmakar, and M. K. Kowar, "BPN model for long range forecast of monsoon rainfall over a very small geographical region and its verification for 2012," Geofizika, vol. 30, no. 2, pp. 155-172, 2013.

[23] C. L Wu and K. W. Chau, "Prediction of rainfall time series using modular soft computing methods," Eng. Applicat. of Artificial Intell., vol. 26, no. 3, pp. 997-1007, 2013.

[24] S. K. Nanda, D. P. Tripathy, S. K. Nayak, and S. Mohapatra, "Prediction of rainfall in India using Artificial Neural Network (ANN) models," Int. J. of Intell. Syst. and Applicat., vol. 5, no. 12, pp. 1-22, 2013.

[25] A. R. Naik and S. K. Pathan, "Indian monsoon rainfall Classification and Prediction using Robust Back Propagation Artificial Neural Network," Int. J. of Emerging Technology and Advanced Eng., vol. 3, no. 11, pp. 99-101, 2013.

[26] Priya, Shilpi, Vashistha and V. Singh, "Time Series Analysis of Forecasting Indian Rainfall," Int. J. of Innovations & Advancement in Comput. Sci., vol. 3, no. 1, pp. 66-69, 2014.

[27] V. K. Dabhi and S. Chaudhary, "Hybrid Wavelet-Postfix-GP model for rainfall prediction of Anand region of India," Advances in Artificial Intell., pp. 1-11, 2014.

[28] Pinky Saikia Dutta, Hitesh Tahbilder, "Prediction Of Rainfall Using Data mining Technique Over Assam", Indian Journal of Computer Science and Engineering (IJCSE), Vol. 5 No.2 Apr May 2014.

[29] Pankratz, A., Forecasting With Univariate Box-Jenkins Models Concepts and Cases.John Wiley Sons, Inc. New York, pp: 414.

[30] M. T.Hagan, H.B.Demuth and M.Beale, Neural Network Design Thomson Asia Pte. Ltd , Singapore, 2002.

[31] Bose NK, Liang P (1998) Neural network fundamentals with graphs, algorithms, and applications. Tata McGraw-Hill, New Delhi.

[32] Weigend A (1994) An overfitting and the effective number of hidden units. Lawrence Erlbaum Associates, Hillsdale, p 335342.

[33] Sivanandam SN, Deepa SN (2007) Principles of soft computing. Wiley India (P) Ltd., New Delhi.

[34] Beale MH, Hagan MT, Demuth HB (2010) Neural network toolbox 7. The Math- Works, Inc., Natick.

[35] Haykin S (1999) Neural networks, a comprehensive foundation. Macmillan College Publishing Co., New York

[36] https://en.wikipedia.org/wiki/Recurrent-neural-network

[37] S. Haykin, 1999: Neural Networks- A Comprehensive Foundation. Addison Wesley Longman

[38] T. Kohonen, (1990), The self organizing map. Proceedings of IEEE, 78(9), 1464-1478.

[39] Singh, P. and Borah, B., 2013. Indian summer monsoon rainfall prediction using artificial neural network. 27(7), pp.1585-1599.

[40] http://www.mdpi.com/2073-4433/5/2/454/htm

[41] https://en.wikibooks.org/wiki/Artificial-Neural-Networks/Recurrent- Network

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)