International Journal of Pure and Applied Mathematics
Volume 117 No. 20 2017, 301-310
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

A Non Linear Based Hybrid Model for Rainfall Forecasting

M. Nirmala[1] R. Rajeswari[2]
[1]Professor, Sathyabama University, Chennai, India
mnirmalamaths@gmail.com
[2]Professor, Sathyabama University, Chennai, India
rajeswarivel1998@gmail.com

Abstract. Forecasting Rainfall is a major scientific issue in the field of meteorology. Many researchers have attempted to establish linear models to predict the rainfall, but with the discovery of nonlinearity in the nature of weather data, the focus has shifted towards the nonlinear prediction of rainfall. A non-linear model based on statistical time series modeling and hybrid based Artificial Neural Network algorithm were applied to forecast the annual rainfall in the state Tamilnadu, India. A dataset containing the annual rainfall in Tamilnadu for a period of 144 years (1871-2014) were collected for this purpose and analyzed. The research findings of this study were compared using the mean average percentage error measure and it was found that the hybrid based Neural Network Algorithm predicted better when compared to the classical non-linear statistical model.

Keywords: Non Linear, BDS test, Hougen Watson Model, Auto Regressive Integrated Moving Average, Artificial Neural Network, Hybrid Model.

1 Introduction

Weather forecasting is one of the most imperative and demanding operational responsibilities carried out by meteorological services all over the world. Rainfall information is an important input in the hydrological modeling. As a country whose economy is heavily dependent on low- productivity rainfed agriculture, rainfall trends are often cited as one of the most important factors in explaining various socio-economic problems such as food insecurity. Rainfall prediction has great significance in planning and understanding the rainfall variability that helps the agricultural management in decision-making process. Several stochastic models have been attempted to forecast the occurrence of rainfall, to investigate its seasonal variability, to forecast yearly rainfall over Tamilnadu. The objective of this study is to evaluate the effectiveness of the Artificial Neural Network algorithms over the traditional non linear statistical techniques. The different models discussed in this research work for prediction of the Annual Rainfall in Tamilnadu are non linear model based on Hougen – Watson, Autoregressive Integrated Moving Average Method and algorithms of artificial neural networks.

Meaza Demissie [8] studied trend and applied Autoregressive Integrated Moving Average Model for rainfall prediction for Melksa observatory Nazreth rainfall series. Uruya Weesakul and Sudajai Lowanichchai [15] used both ARMA and ARIMA models to fit the time series of annual rainfall in Thailand and established that ARIMA model is more suitable to describe inter – annual variation of annual rainfall. Separate ARIMA models were constructed by Seyed Amir Shamsnia et. al. [13] for modeling precipitation, monthly average temperature and relative humidity for Adadeh region, Iran. Bambang Widjanarko Otok and Suhartono [1] found a best method to model the rainfall index data in Indonesia by comparing the forecast accuracy among ARIMA, ASTAR, Single input Transfer function and Multi input Transfer Function models. Artificial Neural Networks have gained significant attention in past two decades. Hung, et. al. [3] and Khaing Win Mar, et. al. [6] applied Neural Network algorithms for rainfall forecasting. Somvanshi et. al. [14] introduces two fundamentally different approaches for designing a model, the statistical method based on autoregressive integrated moving average and the emerging computationally powerful techniques based on ANN. Najmeh Khalili et. al., [11] used a three-layer feed-forward perceptron network with back propagation algorithm to realize the hidden dynamics of rainfall through the past information of the system.

2 Study Area and Materials

The state Tamilnadu is located in the southeastern portion of the Deccan in India, which extends from the Vindhya Mountains in the north to Kanyakumari in the south. Tamilnadu stretches between 8^o 5' N and 13^o 35' N latitudes and between 78^o 18' E and 80^o 20' E longitudes. Tamilnadu normally receives 979 mm of rainfall every year.

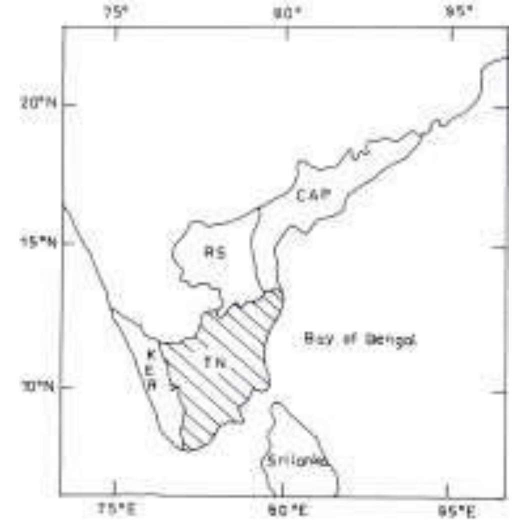


Fig.1: Geographical location of Tamilnadu

Approximately 33% is from the southwest monsoon (June to September) and 48 % is from the Northeast monsoon (October to December). The agricultural sector consumes about 85 percent total water available in Tamilnadu. High water requirement crops like rice (1200 mm), Sugarcane (2000 mm) and Banana (2000 mm) are largely grown in Tamilnadu from canal-irrigated area, tank irrigated area and well-irrigated area.

3 3

Rainfall information is an important input in the hydrological modeling, predicting extreme precipitation events such as droughts and floods, for planning and management of irrigation projects and agricultural production such as reservoir operation, irrigation area, irrigation water requirement and estimating quantity and quality of surface water and ground water. An understanding of rainfall variability in the state helps the agricultural management in planning and decision- making process.

The main objective of the current study is to find a suitable time series-forecasting model for the prediction of the amount of rainfall. For this purpose, a dataset containing a total of 144 years (1871 - 2014) monthly rainfall totals of Tamilnadu was obtained from Indian Institute of Tropical Meteorology, India.

3 Non-linear Test and Model

A powerful tool for detecting serial dependence in time series is BDS test. It is based on the estimation of the correlation integral. The procedure involved is discussed in the following steps. The null hypothesis is that the series is independently and identically distributed and the alternate hypothesis is the series is not independently and identically distributed. Hence, the time series is non – linearly dependent. The BDS test results for the rainfall series for embeddings dimensions from 2 to 5 are summarized in Table 2. At 5% level of significance, the BDS test statistic values are greater than the critical values significantly for different m values. The test analysis suggests that the rainfall series is non –linearly dependent.

A non-linear model for forecasting the annual rainfall in Tamilnadu with three predictors x_1, x_2, and x_3 is proposed. Let x_1 denotes the year, x_2 denotes the Northeast monsoon rainfall (October - December), x_3 denotes the Southwest monsoon rainfall (June - September) and y denotes the annual rainfall in Tamilnadu. The Hougen – Watson non - linear model is given by

$$y = \frac{\beta_1 x_2 - \dfrac{x_3}{\beta_5}}{1 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_3} \tag{1}$$

The parameter estimates for non-linear models [P.Prabhakaran et.al [16] are found using least squares method. Hougen – Watson model uses the Gauss-Newton algorithm with Levenberg- Marquardt modifications for global convergence. The best iterations are obtained for the parameter's estimates with MAPE value 5.0576. The fitted model is given as

$$y = \frac{1.129 x_2 + \dfrac{x_3}{0.000006}}{1 - 0.000098 x_1 + 0.000015 x_2 + 0.000062 x_3} \tag{2}$$

4 4

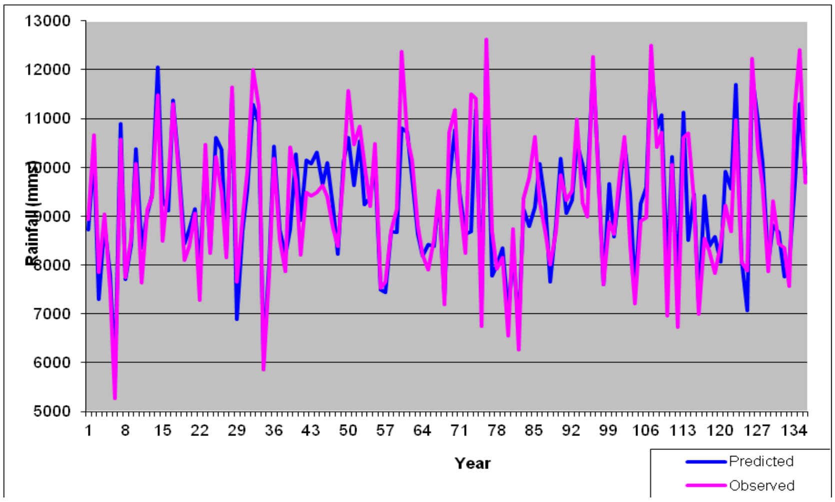


Fig. 2: Graph of Observed vs Fitted values of Annual rainfall in Tamilnadu

Figure2 shows the plot of predicted values vs. observed values of the annual rainfall data from year 1871 to 2014 by the non – linear model. Both the observed and predicted annual rainfall values are quite closer. Also the predicted values through the non-linear model follow the trend of the observed annual rainfall in most of the cases and the model fits extremely well with the actual data values.

4 Autoregressive Integrated Moving Average Method

In a time series data the past values in the series may influence the future values, depending on the presence of underlying deterministic forces. These forces may be characterized by trends, seasonality, cycles and nonstationarity behaviour in the time series. The ARIMA model offers an effective alternative technique for predicting time series data [8] and [10]. A powerful model for describing stationary and non - stationary time series is Autoregressive Integrated Moving Average process of order (p, d, q) where p is the order of the Autoregressive part, d is the order of differencing and q is the order of Moving Average process [2]. The ARIMA (p,d,q) model is expressed as :

$$\phi(B)(1-B)^d(Z_t-\mu)=\theta(B)a_t \tag{3}$$

where Z_t denotes the observed value at time t, B represents the backward shift operator. $\phi(B) and \theta(B)$ have the following representations:

$$\phi(B)=1-\phi_1 B-\phi_2 B^2-\dots\dots\dots\dots-\phi_p B^p \tag{4}$$

$$\theta(B)=1-\theta_1 B-\theta_2 B^2-\dots\dots\dots\dots-\theta_q B^q \tag{5}$$

5 Artificial Neural Network

An Artificial Neural Network is a computational model inspired by biological neural networks both structurally and functionally. It consists of a group of interconnected computation units called neurons. Different types of Artificial Neural Networks take variant forms of activation functions and inter-neuron connections to model the underlying complex mathematical relationships between the inputs and the outputs of different types of systems. Numerous Artificial Neural Network models have been developed for time series prediction. Among them, the two major types are Multi-Layer Perceptron (MLP) networks and Radial Basis Function (RBF) networks. Artificial Neural Networks do not have any stationarity constraint on the time series to be learned and predicted. The neural network training produces appropriate output patterns for corresponding input patterns. Different learning rules form the basis of different training algorithms and their applicability is dependent on the neural network architecture and the learning category that is being used. The original procedure used in the gradient descent algorithm is to adjust the weights towards convergence using the gradient. It uses a gradient search technique to minimize a cost function equal to the sum square difference between desired and estimated net outputs. Derivatives of error are calculated for the network's output layer, and then back propagated through the network until delta vectors are available for each hidden layer of the network. In Scaled Conjugate Gradient Algorithm, the time-consuming line search at each iteration is avoided, by combining the model trust region approach in conjugate Gradient Algorithm. A search is made along conjugate directions rather in steepest descent directions for faster convergence in Conjugate Gradient Algorithm. Then a line search is performed to determine the optimal distance to move along the current search direction.

Recent studies focus on the problem of time series prediction using Radial Basis Function Network. These networks are used commonly in function approximation and time series prediction. Due to their non linear approximation properties, RBF networks are able to model complex mappings. The design of RBF network is an approximation problem in a high dimensional space and the learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data. The normalized Gaussian radial basis function network can be used to model the non linear noise. The first input layer feeds data to a hidden layer. The hidden layer processes the data and transports it to the output layer. The weights between the hidden layer and the output layer are modified during training. This kernel is centered at the point in the input space specified by the weight vector. If the input signal is closer to the current weight vector then the output of the neuron will be higher.

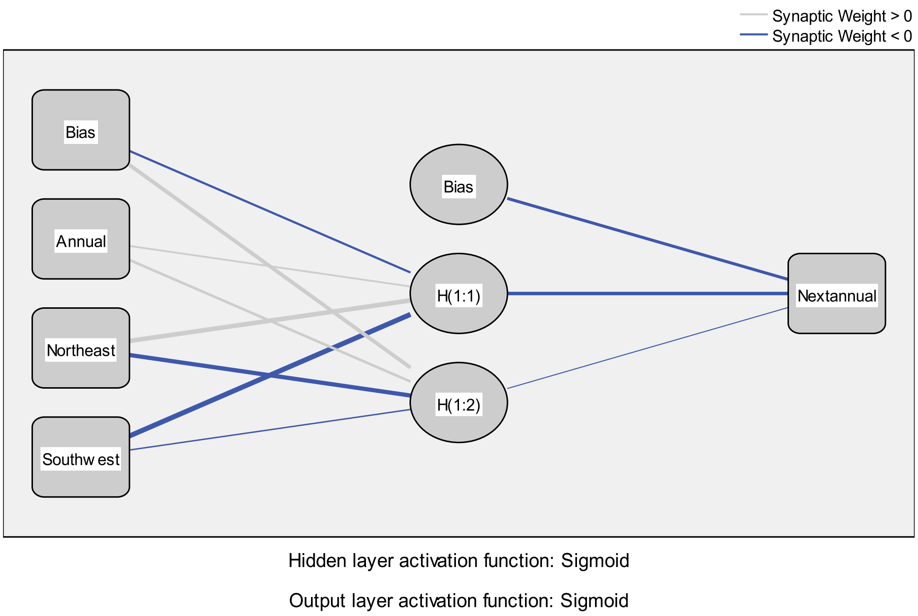
5.1. Training of Learning Algorithms of Neural Network

The main objective of the training is to find the set of weights between the neurons that determines the global minimum of the error function. Training involves the following steps: deciding the number of iterations, selection of learning rate and the momentum values. The predictors for the year Y that are used to forecast the annual rainfall in Tamilnadu in the year (Y + 1) are the Tamilnadu Rainfall amount in the Northeast Monsoon, Tamilnadu Rainfall amount in the Southwest Monsoon and the annual rainfall in Tamilnadu. For developing an Artificial Neural Network model, the whole dataset is smoothed by the powerful classical technique ARIMA. Then the smoothed data is used for developing artificial neural network model by dividing the data set into training set consisting of 110 years (1871 - 1980) and testing set consisting of 34 years (1981 - 2014).

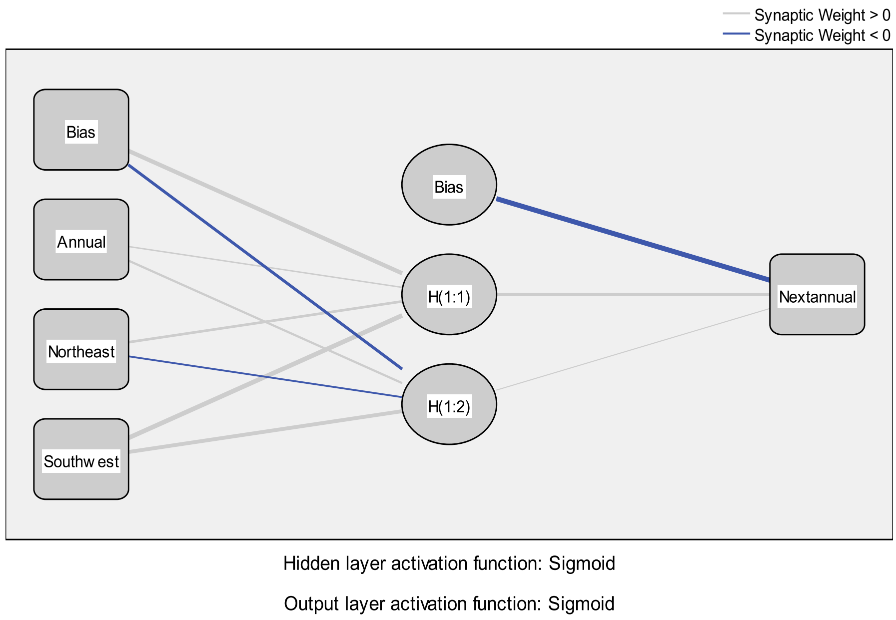
In this research article, three predictors for the year Y are used to predict the Annual Rainfall in Tamilnadu in the year (Y + 1). The three predictors are the Tamilnadu Rainfall amount in the Northeast Monsoon, Tamilnadu Rainfall amount in the Southwest Monsoon and the Annual Rainfall in Tamilnadu

After the training, it has been tested with the test data set. Error correction learning rules in supervised learning uses the difference between the correct and actual output patterns to adjust connection weights with the aim of reducing this error.

The observed annual rainfall data was first fitted by ARIMA model of order (0, 0, 4). Also the predictors Northeast and Southwest monsoon rainfall series were fitted through the same model of order (4, 0, 4) and (1, 0, 1) respectively. Training was carried over using three predictors for the year Y to predict the annual rainfall in Tamilnadu in the year (Y + 1). The three predictors are the fitted Northeast Monsoon Rainfall series, fitted Southwest Monsoon Rainfall series and the fitted annual Rainfall series in Tamilnadu. The various learning algorithms GDA, SCG and RBF were used for training the annual rainfall in Tamilnadu. The architecture of these algorithms was given in figure 3. The various parameters used for training the rainfall series is given in the table 1. The error measure, MAPE values were tabulated in the table 2.



6



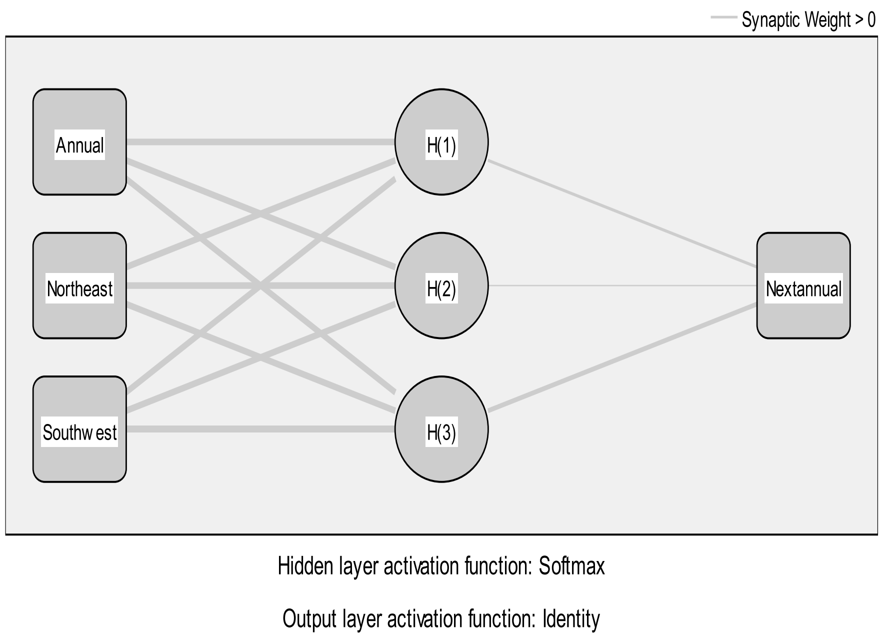


Fig.3: Architecture of ARIMA-ANN model using GDA, SCG and RBF

Table 1: Various Parameters used for ANN training

Learning Algorithms	GDA	SCG	RBF
No. of Nodes in the Input Layer	3	3	3
No. of Hidden Layers	1	1	1
No. of Nodes in the Hidden Layer	2	2	3
Activation Function in the Hidden Layer	Sigmoid	Sigmoid	Softmax
No. of Nodes in the Output Layer	1	1	1
Activation Function in the Output Layer	Sigmoid	Sigmoid	Identity

Table 2: MAPE values of various models

Model	MAPE	
ARIMA – ANN	GDA	2.2359
	SCG	2.1687
	RBF	2.0918
Non Linear Model		5.0576

6 Conclusion

Artificial Neural Networks do not have any stationarity constraint on the time series to be learned and predicted. It has highly flexible nonlinear regressive structure to fit the target pattern space. Traditional statistical time series forecasting methods, including moving average, exponential smoothing, and Auto-Regressive Moving Average, all assume stationarity and linearity of the time series. In this research article, a comparison between the traditional statistical nonlinear model and the algorithms of Artificial Neural Network were made. Through the error measure, it is concluded that the Hybrid based ANN performs better than the Statistical non-linear model. For some real world problems, Artificial Neural Network will never replace the existing conventional techniques but because of the fast growing applications it can be an alternative to those existing techniques. From the error measure, MAPE it can be concluded that ARIMA-ANN model was optimal based on the model performance and the complexity of modeling.

6

References

[1]. Bambang Widjanarko Otok and Suhartono,:Development of rainfall forecasting model in Indonesia by using ASTAR, Transfer Function and ARIMA methods, European Journal of Scientific Research 38(3), 386 – 395(2009).

[2]. C. Chatfield: The analysis of time series: Theory and Practice, Chapman and Hall Co(1983).

[3]. Hung N Q, Babel M S, Thailand, Weesakul S, and Tripathi N K: An artificial neural network model for rainfall forecasting in Bangkok, Hydrology and Earth System Sciences discussions 5, 183 –218(2008).

[4]. J.A. Anderson: An Introduction to Neural Networks, Prentice Hall(2003).

[5]. Karim Solaimani: Rainfall-Runoff Prediction based on Artificial Neural Network (A case study: Jarahi Watershed), American – Eurasian Journal of Agriculture and Environmental Sciences 5, 856 – 865(2009).

[6]. Khaing Win Mar and Thinn Thu Naing: Optimum Neural Network architecture for Precipitation Prediction of Myanmar, Proceedings of World Academy of Science, Engineering and Technology, 36, 130-134(2008).

[7]. M.D. Gorka Landeras, Amaia Ortiz-Barredo, and José Javier López: Forecasting Weekly Evapotranspiration with ARIMA and Artificial Neural Network Models, Journal of Irrigation and Drainage Engineering 135, 323 – 334(2009).

[8]. Meaza Demissie: Characterizing trends and identification of rainfall predicting model: An application to Melksa observatory Nazreth rainfall series, Proceedings of Biometric society,132 – 134(2003).

[9]. M. Nirmala: Computational Models for Forecasting Annual Rainfall in Tamilnadu, Applied Mathematical Sciences, 9(13), 617-621(2015).

[10]. M. Nirmala : A Comparative study of ARIMA model with Radial Basis Function for Forecasting Monsoon Rainfall in Tamilnadu, International Journal of Applied Engineering Research 10(1),1045-1054(2015).

[11]. Najmeh Khalili, Saeed Reza Khodashenas, Kamran Davary and Fatemeh Karimaldini :Daily Rainfall forecasting for Mashhad synoptic station using Artificial Neural Networks, Proceedings of International Conference on Environmental and Computer Science, 19(2011).

[12]. S.Makridakis, and S.C.Wheelwright,:Forecasting: Methods and Applications, John Wiley and Sons, New York(1978).

[13]. Seyed Amir Shamsnia, Naeem Shahidi, Ali Liaghat, Amirouya Sarraf, and Seyed Farnood Vahdat: Modeling of Weather Parameters using Stochastic Methods, International Conference on Environment and Industrial Innovation,12, 282 – 285(2011).

[14]. Somvanshi V.K., Pandey O.P., Agrawal P.K., Kalanker N.V., Ravi Prakash M. and Ramesh Chand, Modelling and prediction of rainfall using artificial neural network and ARIMA techniques, Journal of Indian Geophysical Union,10(2),141-151(2006).

[15].Uruya Weesakul and Sudajai Lowanichchai: Rainfall Forecast for Agricultural water allocation planning in Thailand, Thammasat International Journal of Science and Technology, 10(3), 18 – 27(2005).

[16] P.Prabhakaran and S.Saravanan, “Parameter Estimation And Classification Using Gaussian mixture Model”, International Journal of Innovations in Scientific and Engineering Research (IJISER), ISSN: 2347-971X (online), ISSN: 2347-9728(print) Vol.2, No.4, pp.73-76, 2015, http://www.ijiser.com/.