Int.J.Curr.Microbiol.App.Sci (2020) *9*(6): 1437-1447

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 9 Number 6 (2020)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2020.906.179

Rainfall Trend Analysis by ANN Method

Bhupendra Dhankar[1*] and Gunja Dhruw[2]

[1]*SVCAET & RS IGKV Raipur, India*
[2]*Department of Farm Engineering, Inst. of Agri. Sci., BHU, Varanasi (UP), India*

**Corresponding author*

Keywords

Rainfall, ANN, Modelling, Variables, Accuracy

Article Info

Accepted:
18 May 2020
Available Online:
10 June 2020

A B S T R A C T

More than 60% of agricultural land in India is rainfed farming. But the amount of rainfall across India varies greatly with the location. (GKTODAY) Mawsynram receives 11690 mm of rainfall, While Eastern Rajasthan receives less than 150 mm of rainfall. In such a situation, proper management of rain water is very important so that proper amount of water is available at the right time for agriculture. Before doing the management of rain water, it is very important to know the rainfall trend of the area. It is very difficult to find the pattern of rainwater by human calculations. That is why many modeling methods have been developed. ANN is a modelling technique, which gives more accurate results than many physical and numerical modelling techniques. The study area is a macro-watershed of river Kelo, a tributary of Mahanadi River. In this paper, the results show that as the use of number of influencing variables increases, the accuracy increases. ANN technique gives accurate results in a very short period of time which makes the feasibility of this method. The results showed that the accuracy ranges from 93 to 98 percent with the different number of variables.

Introduction

India is an agricultural country with mainly rainfed farming. According to statistics, more than 60% of the agricultural land in the country is fully rainfed farming (GKTODAY)(1) The development issues of rainfed agriculture as some critical importance on account of the slow growth and its implications on livelihood security of significant number of rural poor in India (Rao 2004).(2) So it is very important to manage the rainfall water during the rainy season for taking a proper production in rainy season.(3) The amount, intensity and areal distribution of precipitation are essential in many hydrological studies. (WMO, 1983) (4) Detecting changes in climate is a prerequisite for a better understanding of the climate and developing adaptation and mitigation measures at a regional and local scale.(5)

Spatial-temporal variability of meteorological variables in the framework of changing climate is predominant. At the same time, if agriculture in those areas is depending on rainfall, then variables especially rainfall plays a vital role to assess climate-induced changes. Such types of studies will suggest feasible adaptation strategies of those particular areas (Rushi *et al.*, 2019)[6]

A major concern about the rainfed agriculture in India is the low level of productivity, in fact one among the lowest in dry and rainfed regions in the world (GoI, 2011).[7] There are various types of modelling techniques recommended by the scientists for solving the difficulty to forecast the rainfall trends in a very accurate manner.[8]

Venkateswarlu said that a number of economically viable rainfed technologies have been developed over the years in the country to address the problems of food production in rainfed agriculture through CRIDA and its network centres for the last three decades.[9]

Materials and Methods

Study area

The Kelo river (major tributary of river Mahanadi), is a perennial river, flows western direction from its origin about 4 km and deflected in south direction up to 35 km and leave the hilly track and enters plains near Milupara village. It flows 78 km, in plain area.[10]

The location of Kelo macro-watershed is in between 21.43^{0}N and 21.9^{0}N and between 83.4°E and 83.49°E longitude at an average elevation of 215 m above the mean sea level. The Kelo river flows through the Raigarh city and is prime source of water. Kelo River is 112.60 km., joins Mahanadi near village Mahadeopali, district Sambalpur (Odisha).[11]

Climatic characteristics

The study area has a sub-tropical climate. There are considerable variations in the rainfall, temperature and humidity. The climate is characterized by oppressive hot summer, a mild winter, and well distributed rainfall during south- west monsoon. The year can be divided into four seasons. Summer season lasts from March to middle of June, Monsoon season from middle of June to September, post monsoon season from October to November and cold season from December to February.

Topographical characteristics

The topographical characteristics of the study area were analysed by using the combination of survey of India toposheets No. 64-N and 64-O on 1:250,000 scale. The toposheets were procured from the office of the Director, Chhattisgarh Geo-Spatial Data Centre, Survey of India, Chhattisgarh. The analysis revealed that river Kelo drains from south to north and north-east. The Kelo river rises at an elevation of about 767 m above MSL about 40 km North of Ghargoda (v) in Raigarh district.

Data organization

Instrumentation and data collection

In 1975, gauging of the Kelo river at Raigarh site was started by the Central Water Commission, Ministry of Water Resources, Govt. of India. Non-recording rain-gauges were installed for measuring daily rainfall data at Raigarh, Gharghoda and Lailunga.

The daily Rainfall data for the years from 2002 to 2013 of the rain gauge stations in and around the study area were collected from State Data Centre, Department of Water Resources, Govt. of Chhattisgarh.

Mean areal precipitation

Theissen Polygon method is used to determine the mean areal precipitation. In this method, the mean areal precipitation is computed by weighing the rainfall depth with the area of polygon of the respective raingauge station. Due to this reason, sometimes this method is also known as weighed mean method. Thiessen method estimates more accurate value of mean areal precipitation, when size of the watershed is between 500 to 5000 Km^2 area and Topography is flat (Suresh R.,2005).

The ANN model

Artificial neural networks (ANNs) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains. Such systems "learn" (i.e. progressively improve performance on) tasks by considering examples, generally without task-specific programming. An ANN is based on a collection of connected units or nodes called artificial neurons (a simplified version of biological neurons in an animal brain). Each connection (a simplified version of a synapse) between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it and then signal artificial neurons connected to it.

Formulation of ANN model

It is very important to select proper input variables for formulation of ANN model. The input variables will be selected based on correlation matrix. This correlation matrix is prepared between the input variables viz. previous week rainfall (P_{t-1}), previous second week rainfall (P_{t-2}) and previous third week rainfall (P_{t-3}), previous fourth week rainfall (P_{t-4}). After finding the correlation matrix, the value of P_{t-1}, P_{t-2}, and P_{t-3} are found more significant. The output variable is P_t. The final input variables and output variables considered in ANN model are shown in Table 1.

Components of an artificial neural network

Neurons

A neuron with label j receiving an input Pj (t) from predecessor neurons consists of the following components: an activation aj (t), depending on a discrete time parameter, possibly a threshold j, which stays fixed unless changed by a learning function, an activation function f that computes the new activation at a given time t+1 from aj (t), j, and the net input Pj (t) giving rise to the relation

$$aj\,(t+1) = f\,(aj\,(t), Pj\,(t), j)$$

and an output function fout computing the output from the activation

$$oj\,(t) = fout\,(aj\,(t))$$

Often the output function is simply the Identity function.

An input neuron has no predecessor but serves as input interface for the whole network. Similarly, an output neuron has no successor and thus serves as output interface of the whole network.

Connections and weights

The network consists of connections, each connection transferring the output of a neuron i to the input of a neuron j. In this sense i is the predecessor of j and j is the successor of i. Each connection is assigned a weight wij.

Propagation function

The propagation function computes the input Pj (t) to the neuron j from the outputs oi (t) of predecessor neurons and typically has the form.

Pj (t) = Σ i oi (t) wij

Learning rule

The learning rule is a rule or an algorithm which modifies the parameters of the neural network, in order for a given input to the network to produce a favored output. This learning process typically amounts to modifying the weights and thresholds of the variables within the network.

Performance function

The performance function or the goal ascertains the level of error. Functions generally chosen are sum squared error (sse), mean squared error (mse) and mean squared error with regularization (msereg). In this study "sse" is adopted. It is a network performance function and measures performance according to the sum of squared errors.

Stopping criteria

Training stops when any of these conditions occurs:

The maximum number of epochs (repetitions) is reached.

The maximum amount of time is exceeded.
Performance is minimized to the goal.
The performance gradient falls below min-grad.
Mu exceeds mu-max.

Performance evaluation criteria

The network is trained on the training or calibration data set and its performance is evaluated both in calibration and in verification data sets. The training stops when there is no more improvement both in training and in verification. The statistical performance evaluation criteria considered in this study are mean absolute deviation (MAD), root mean square error (RMSE), correlation coefficient (CC), and coefficient of efficiency (CE).

Mean Absolute Deviation (MAD)

It is measure of the mean absolute deviation of the observed values from the estimated values. It has a unit and is not a normalised criterion. It is expressed as,

$$\text{MAD} = \sum_{j=1}^{n} \frac{|Oj - Sj|}{n}$$

Root Mean Squared Error (RMSE)

The Root mean squared error is the difference between observed (Yu, 1994) and the estimated values of runoff. The RMSE is compared as follows:

$$\text{RMSE} = \left(\frac{variance}{n}\right)^{1/2} = \left(\frac{\sum_{j=1}^{n}(Oj - Sj)2}{n}\right)^{1/2}$$

Correlation Coefficient (CC)

The correlation between the observed and simulated values is described by the correlation statistic, called the correlation coefficient. It is estimated by the equation:

$$\text{CC} = \frac{\sum_{j=1}^{n}(Oj - \bar{O})\,(Sj - \bar{S})}{\sum_{j=1}^{n}(Oj - \bar{O})2\,(Sj - \bar{S})2}$$

Coefficient of efficiency (CE)

Nash and Sutcliffe (1970) proposed the criterion on the basis of standardization of the residual variance with initial variance and named it as the coefficient of efficiency. The dimensionless criterion of coefficient of efficiency is estimated as follows:

$$CE = \{1 - \frac{\sum_{j=1}^{n}(Oj-Sj)2}{\sum_{j=1}^{n}(Oj-\bar{O})2}\}$$

Where,
O= observed runoff in mm,
S= Simulated runoff in mm,
$\bar{S}$= Mean simulated runoff in mm,
$\bar{O}$= Mean observed runoff in mm.

Thus, a perfect agreement between the observed and estimated values yields the CE value as 100 percent while for a zero agreement, all the estimated values must be equal to the observed mean. A negative efficiency represents that the estimated values are less than the observed mean. As the efficiency depends strongly upon the initial variance of the observed records.

Results and Discussion

Estimation of weekly rainfall (Pt) using Artificial Neural Network (ANN)

For ANN modelling, the previous weeks rainfall data have been considered as inputs while the observed current week rainfall data (Pt) has been taken as target. Two models have been developed according to the input data. Performance evaluation of the model have been carried out by calculating mean absolute deviation (MAD), root mean square error (RMSE), coefficient of correlation (CC) and Nash - Sutcliffe coefficient efficiency (CE). The network are selected based on maximized CC and CE value and minimized MAD and RMSE values both in training and testing.

Performance of ANN models

ANN models have been named as A1 andA2. These models have been developed using 183 patterns. Out of total number of 183 patterns, 137 (75%) has been used for training the network while remaining 46 (25%) has been used for testing the model. In this way, observed and simulated Pt was obtained for calibration and verification period separately. Now as per the performance evaluation criteria discussed, values of MAD, RMSE, CC and CE were obtained.

Performance of A1model

This model was designed with 2 input (P_{t-1}& P_{t-2}) and one output (P_t)The value of CC and CE were found to be 84.35 and 82.17 during training and highest CC and CE value as 79.82 and 76.92 and minimum MAD and RMSE value as 6.53 and 7.39 during training and minimum MAD and RMSE value as 8.52 and 9.21 during testing of the model. Performance of M3 model during Training and Testing period is given in Table-2. Scatter plots between observed Pt and simulated values using ANN showed that most of the values lie near 45° lines for model as explained in fig. 4.

Table.1 Input variables and output variables considered in developing ANN model

Model	Input variables	Output variables
1.	Previous week rainfall (P_{t-1}), Previous second weeks rainfall (P_{t-2}) and Previous third weeks rainfall (P_{t-3})	**Current week rainfall (P_t)**
2.	Previous week rainfall (P_{t-1}). Previous second weeks rainfall (P_{t-2}), Previous third weeks rainfall (P_{t-3}) and Previous fourth week rainfall (P_{t-4})	**Current week rainfall (P_t)**

Table.2 Performance of GA model during training and testing

Model Name	Training				Testing			
	MAD	RMSE	CC	CE	MAD	RMSE	CC	CE
A1	6.53	7.39	84.35	82.17	8.52	9.21	79.82	**76.92**
A2	**5.19**	**5.57**	**90.52**	**87.79**	**6.23**	**7.02**	**84.52**	**80.69**

Figure.1 Different river basins in Chhattisgarh

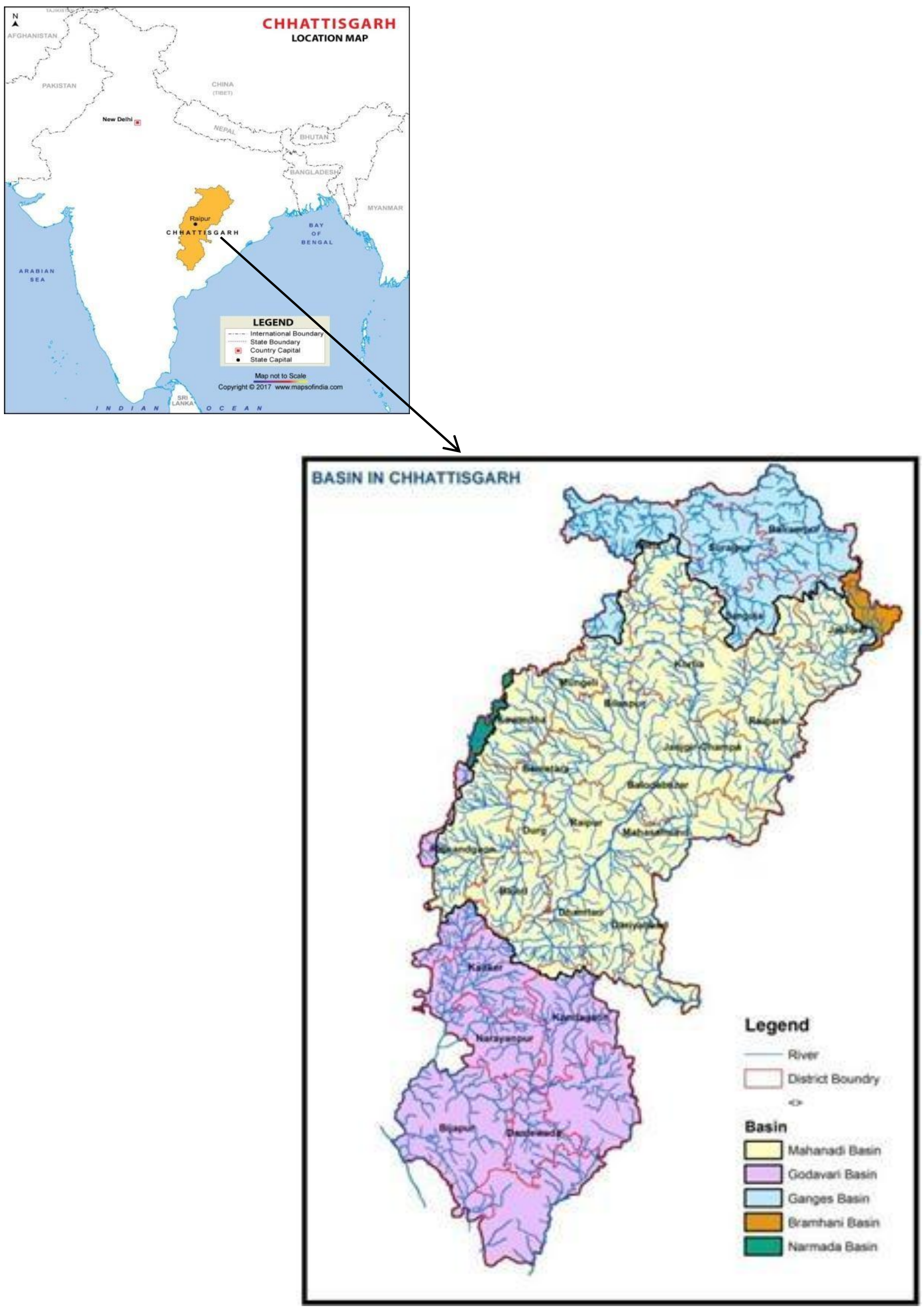


Figure.2 Location of the study area

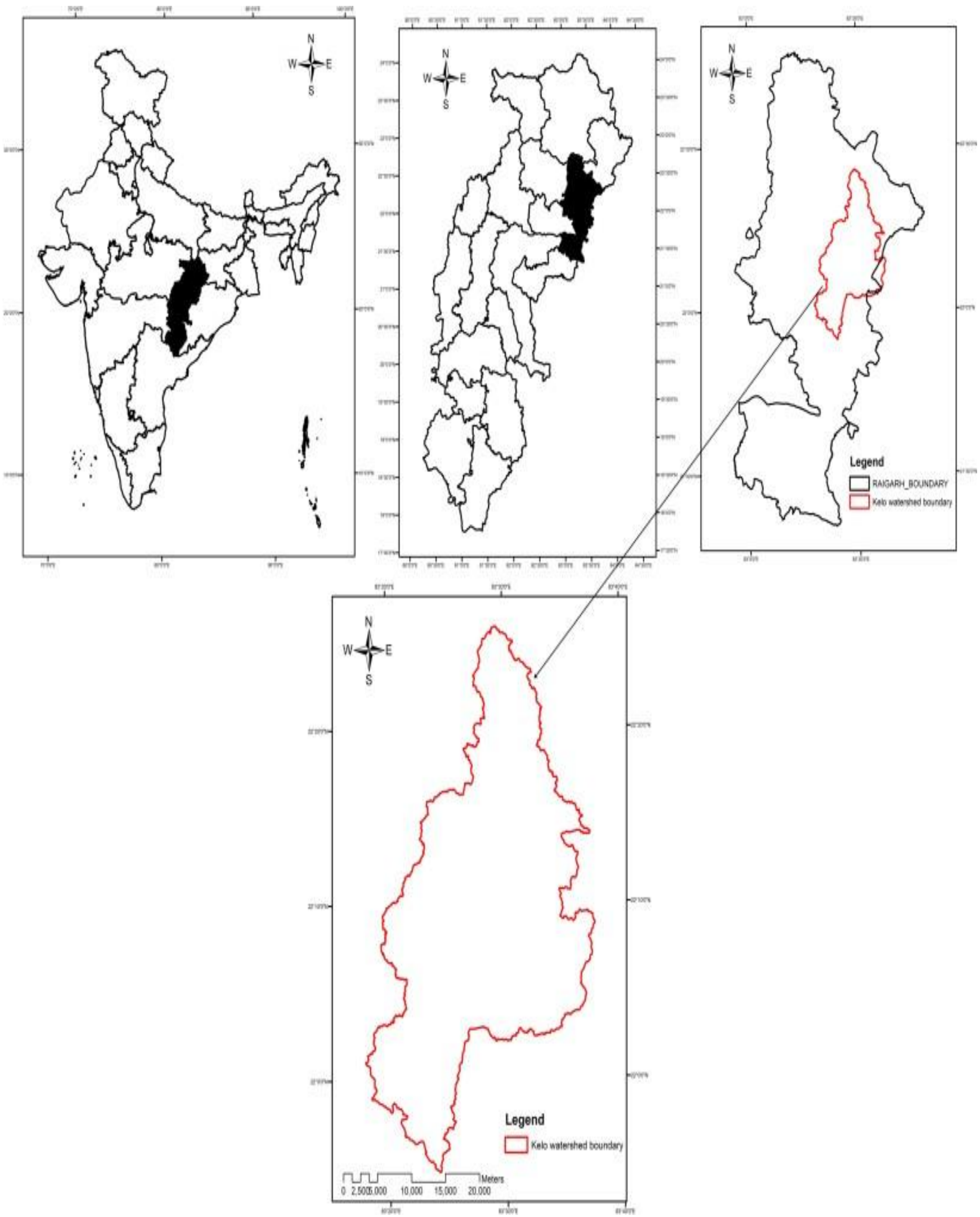


Fig.3 Flow chart for development of artificial neural network models

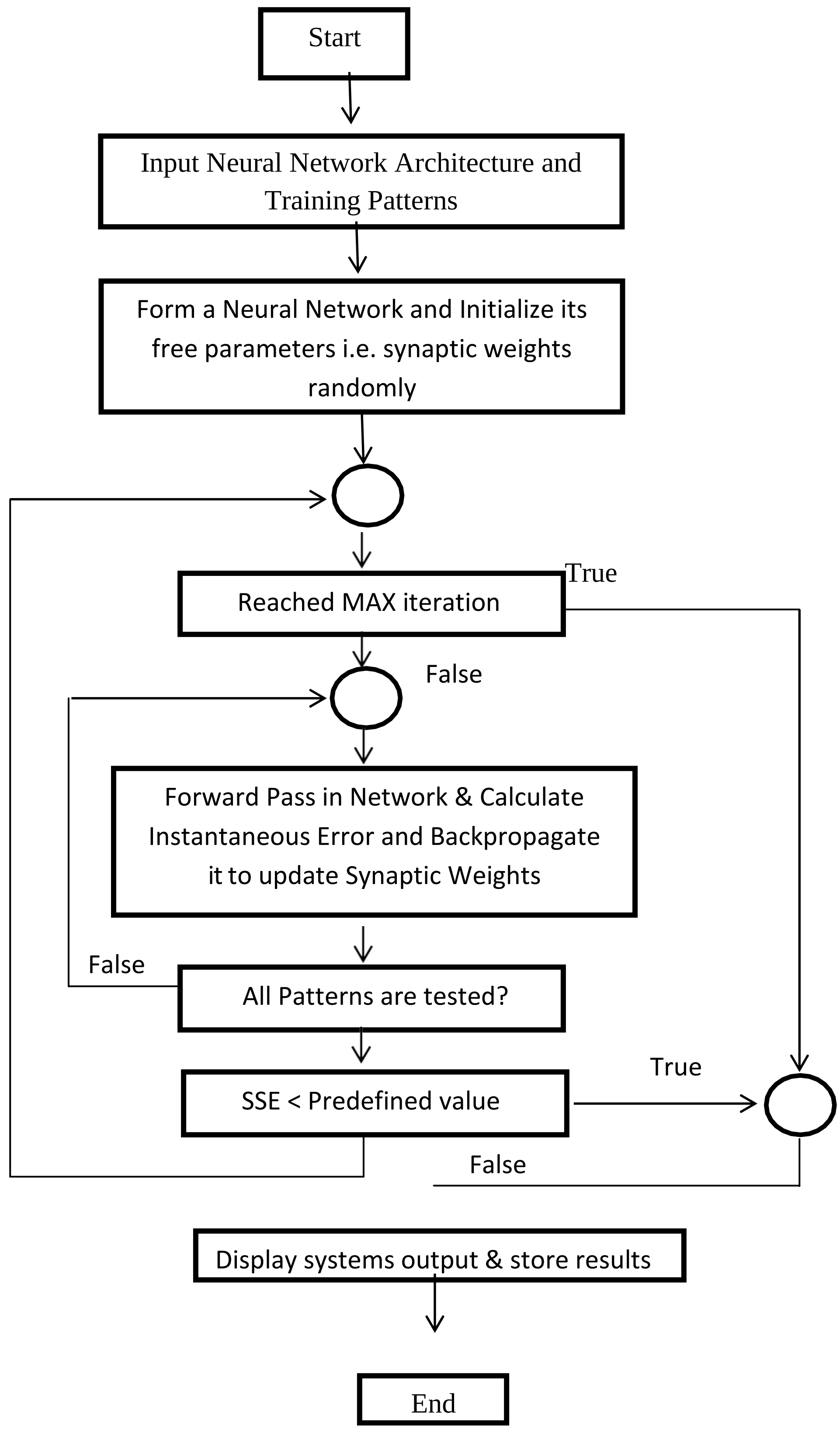


Fig.4 Relationship between observed P_t and the P_t estimated by ANN during training and testing period of model A1

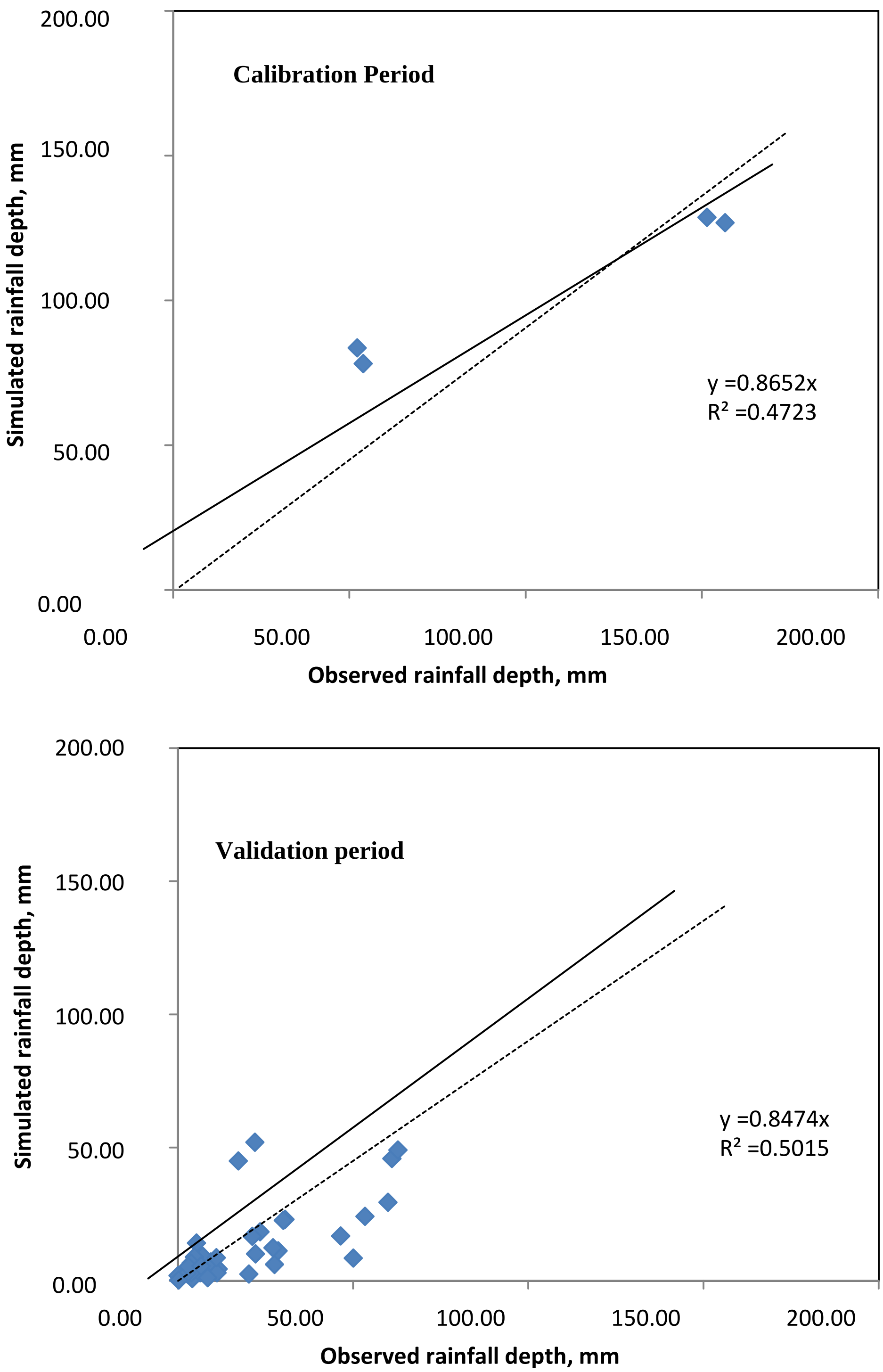
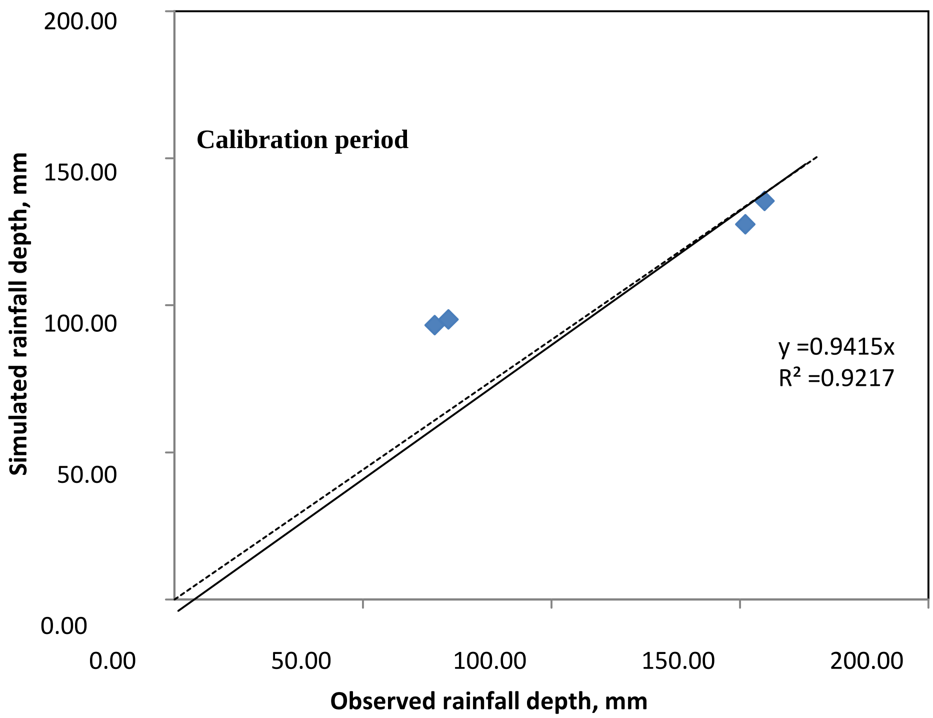
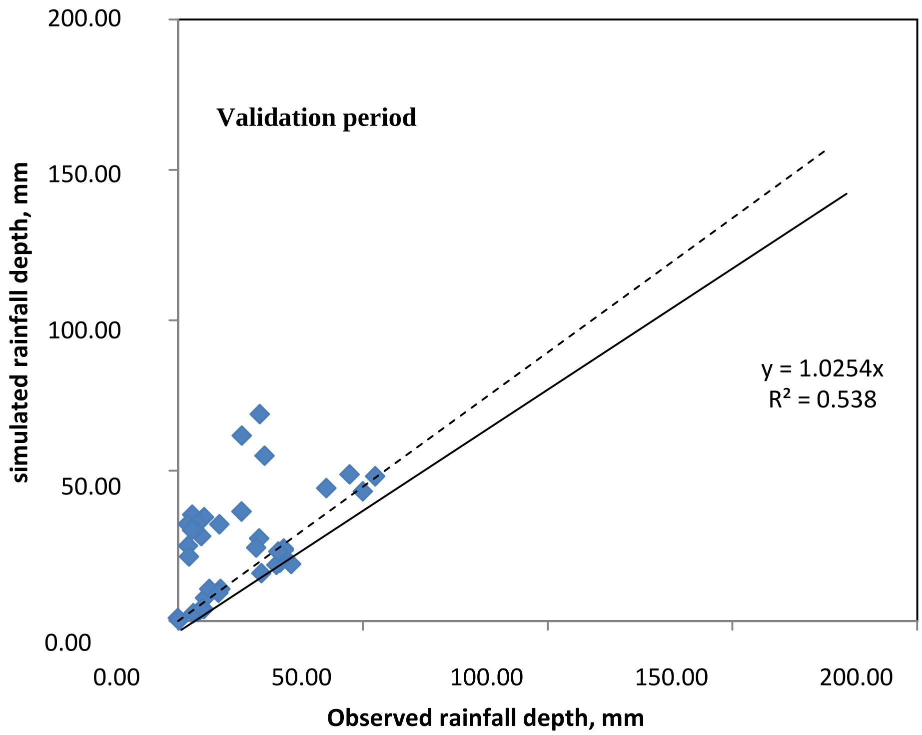


Fig.5 Relationship between observed P_t and the P_t estimated by ANN during training and testing period of model A2





Performance of A2model

This model was designed with 3 inputs (P_{t-1}, P_{t-2}& P_{t-3}) and one output (P_t). The value of CC and CE were found to be 90.52 and 87.79 during training and highest CC and CE value as 84.52 and 80.69 and minimum MAD and RMSE value as 5.19 and 5.57 during training and minimum MAD and RMSE value as 6.23 and 7.02 during testing of the model. Performance of A2 model during Training and Testing period is given in Table-2. Scatter plots between observed Pt and simulated values using A2 showed that most of the values lie near 45° lines for model as explained in fig. 5.

Comparison of ANN models

Comparison of the models A1 and A2 on the basis of highest CC and CE values and minimum values of MAD and RMSE shows that the model A2 is performing better than A1 model. It showed highest CE values as 87.79 and minimum RMSE value as 7.02 during testing period. Details of their comparison are presented in table 2.

Acknowledgement

The authors are immensely grateful to Parents, Advisor Dr J.K Sinha Associate Professor FAE IGKV Raipur and Ms Gayam Sujata Mam for providing rainfall data.

References

Abrahart, *et al.,* 2007. Neural network modelling of non-linear hydrological relationships. Hydrology and Earth System Science, 14: 2157–2172.

Babovic, V.: 1996. Emergence, Evolution, Intelligence: Hydroinformatics, Balkema, Rotterdam.

Dawson, *et al.,* 2001. Hydrological modelling using artificial neural networks, Progress in Physical Geography, 25, 1, 80–108.

Eeles, CWO, 1994. Parameter optimization of conceptual hydrological models, PhD Thesis, Open University, Milton Keynes, U.K.

Gowda, *et al.,* 2015. Determining the Sensitivity of Rainfall- Runoff Model using Artificial Neural Network Approach, International Journal of Advanced Structures and Geotechnical Engineering ISSN 2319-5347, Vol. 04, No. 01, 01-03.

Minns, A. W. and Hall, M. J., 1996. Artificial neural networks as rainfall-runoff models, Hydrological Sci. J. 41, 399–417.

Mishra, Satanand, Saravanan, C., and Dwivedi, V. K., 2015. Study of Time Series Data Mining for the Real Time Hydrological Forecasting: A Review, International Journal of Computer Applications (0975 – 8887), Volume 117– No. 23, 6-17.

Mislan, Haviluddin, Hardwinarto, Sumaryono, and Aipassa, 2015. Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan – Indonesia, International Conference on Computer Science and Computational Intelligence, Procedia Computer Science, 59, 142 – 151.

How to cite this article:

Bhupendra Dhankar and Gunja Dhruw. 2020. Rainfall Trend Analysis by ANN Method. *Int.J.Curr.Microbiol.App.Sci.* 9(06): 1437-1447. doi: https://doi.org/10.20546/ijcmas.2020.906.179