"RAINFALL SIMULATION USING ANN BASED GENEREALIZED FEED FORWARD AND MLR TECHNIQUE"

Karuna Singh[1], Dr. Vikram Singh[2], Dr. S.K. Srivastava[3], Er. C.J Wesley[4]

[1]PG Scholar SHUATS Allahabad U.P
[2,4]Assistant Professor SHUATS Allahabad U.P
[3]Associate Professor SHUATS Allahabad U.P

Abstract - *Rainfall modeling is one of the most important topics in water resources planning, development and management on sustainable basis. In this study an effort has been carried out for the development of Generalized Feed Forward (GFF) and Multi Linear Regression (MLR) technique for daily monsoon rainfall prediction of Satna (M.P.). The daily data of monsoon period from (1st June to 30th September) of year 2004-2011 were used for training of models and data of remaining years 2012-2013were used for testing of the models. The NeuroSolution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation indices for developed models, respectively. The best input combination was identified using the input-output combination for the simulation. On the basis input combination, 10 best combinations. The input pairs in the training data set were applied to the network of a selected architecture and training was performed.*

Key Words: (Rainfall Simulation, Ann Based GFF Techniques, MLR Techniques, Validation)

1. INTRODUCTION

Rainfall prediction is one of the most important and challenging tasks in the modern world. In general, climate and rainfall are highly non-linear and complicated phenomena, which require advanced computer modeling and simulation for their accurate prediction. An Artificial Neural Network (ANN) can be used to predict the behavior of such nonlinear systems. ANN has been successfully used by most of the researchers in this field for the last twenty-five years survey of the available literature of some methodologies employed by different researchers to utilize ANN for rainfall prediction. The survey also reports that rainfall prediction using ANN technique is more suitable than traditional statistical and numerical methods, there are two main approaches in rainfall forecasting, numerical and statistical methods. The performance of the numerical method depends on the initial condition, which is inherently incomplete. The method is poor for long-range prediction. On the other hand, the statistical method is widely used for long-term rainfall prediction. In their studies stated those statistical method performances were successful in normal monsoon rainfall but fail in extreme monsoon years. In addition, the statistical method is useless for highly nonlinear relationship between rainfall and its predictors and there is no ultimate end in finding the best predictors.

2. REVIEW AND LITERATURE

ElShafie et al. (2011) have developed two rainfall prediction models i.e. Artificial Neural Network model (ANN) and Multi Regression model (MLR) and implemented in Alexandria, Egypt. They have used statistical parameters such as the Root Mean Square Error, Mean Absolute Error, Coefficient Of Correlation and BIAS to make the comparison between the two models and found that the ANN model shows better performance than the MLR model.1.2 Sub Heading 2

Saha et al. (2012) develop suitable Regression and Artificial Neural Network (ANN) models using identified 144 randomly selected indicators data sets over nine years historical time periods, collected from a successful case study namely "Semi micro watershed, Sehore District, Madhya Pradesh, India". Regression and ANN decision support system prediction models have been developed with eight most dominating parameters which have found most significant effect on livelihood security. The comparison study of these two models have indicated that, the statistical yield predicted through ANN models performed better than that predicted through regression models. The study has recommended the use of such models for improvement of similar degraded watershed for future reference.

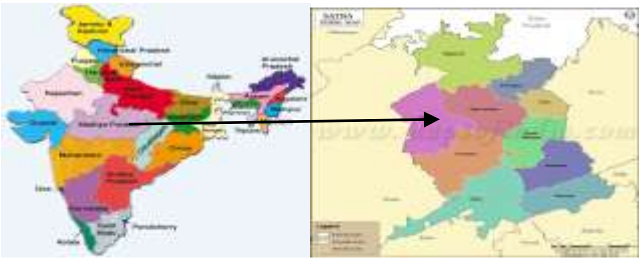
Chua et al. (2013) have employed several soft computing approaches for rainfall prediction. They have considered two aspects to improve the accuracy of rainfall prediction: (1) carrying out a data-preprocessing procedure and (2) adopting a modular modeling method. The proposed preprocessing techniques included moving average (MA) and singular spectrum analysis (SSA). The modular models were composed of local support vector regression (SVR) models or/and local artificial neural network (ANN) models. The ANN was used to choose data- preprocessing method from MA and SSA. Finally, they have showed that the MA was superior to the SSA when they were coupled with the ANN.

Duong Tran Anh et al. (2019) Rainfall prediction is a fundamental process in providing inputs for climate impact studies and hydrological process assessments. In which we combined two pre-processing methods (Seasonal Decomposition and Discrete Wavelet Transform) and two feed-forward neural networks (Artificial Neural Network and Seasonal Artificial Neural Network). In detail, observed monthly rainfall time series at the Ca Mau hydrological station in Vietnam were decomposed by using the two pre-processing data methods applied to five sub-signals at four levels by wavelet analysis, and three sub-sets by seasonal decomposition. After that, the processed data were used to feed the feed-forward Neural Network (ANN) and Seasonal Artificial Neural Network (SANN) rainfall prediction models.

3. MATERIALS AND METHOD

3.1 Location of study area:

The study area located between 24°56'09.31"N, 25°10'53.19N latitude and 80°44'23.31"E, 80°52'44.53"E longitude (approx.) with an average elevation of 315 meters (1,352 feet).



3.2 Climatic Characteristics

The climate of Satna district is characterized by a hot summer with general dryness, except during the south-west monsoon season. The normal annual rainfall of Satna district is 1092.1 mm. The district receives maximum rainfall during south-west monsoon period (i.e. June to September) and about 87.7% of annual rainfall is received during this period. Only 12.3% of the annual rainfall takes place between periods October to May.

3.3 Data Collection

The weather data (rainfall, minimum and maximum temperature, relative humidity) of monsoon season (1st June to 30th September) during the years 2004-2013 were obtained from global weather data for SWAT website.

3.4 Methodology

In this study, the soft computing techniques such as Generalized Feed Forward (GFF) based ANN and statistical multiple linear regression (MLR) have been developed for simulating the rainfall in Satna district. The methodology of developing the GFF and MLR models along with training and testing of the developed models, the NeuroSolution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation indices for developed models.

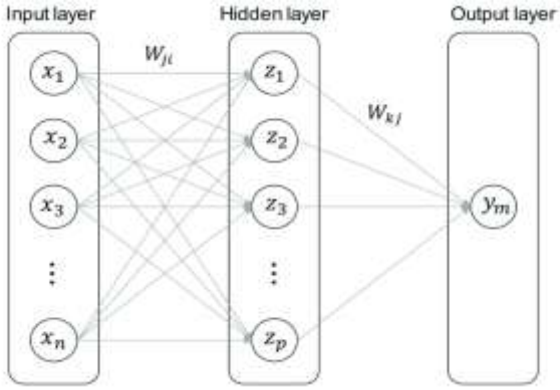


Figure 3.1 Basic structure of ANN

3.4.1 Generalized Feed Forward (GFF)

In this study, a different approach was used to obtain the network architecture. Instead of limiting the size of the networks, complex networks were developed with a high number of connections. The objective was to obtain a network with greater capacity for establishing generalized relationships between the parameters on which rainfall depends.

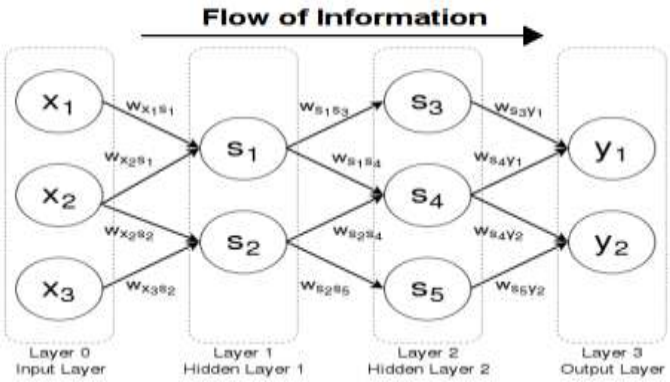


Fig 3.2 Four layer feed forward neural network

3.4.2 Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) is simply extended form of Simple regression in which two or more variables are independent variables are used and can be expressed as (Kumar and Malik, 2015):

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots\dots\dots\dots\dots \beta_p X_p$$

Where,

Y = Dependent variable;

α = Constant or intercept;

β1 = Slope (Beta coefficient) for X1;

X1 =First independent variable that is explaining the variance in Y;

β2 = Slope (Beta coefficient) for X2;

X2 = Second independent variable that is explaining the variance in Y;

p= Number of independent variables;

βp= Slope coefficient for Xp;

Xp= pth independent variable explaining the variance in Y.

Model No.	Input-Output Variables
GFF-1	$R_t = f(T_{max})$
GFF-2	$R_t = f(T_{min})$
GFF -3	$R_t = f(W_S)$
GFF -4	$R_t = f(RH)$
GFF -5	$R_t = f(RS)$
GFF -6	$R_t = f(T_{max}\ T_{min,})$
GFF -7	$R_t = f(T_{max,}\ W_S)$
GFF -8	$R_t = f(T_{max,}\ RS)$
GFF -9	$R_t = f(T_{max,}\ RH)$
GFF-10	$R_t = f(T_{min,}\ W_S)$
GFF -11	$R_t = f(T_{min,}\ RH)$
GFF -12	$R_t = f(T_{min,}\ RS)$
GFF -13	$R_t = f(W_{S,}\ RH)$
GFF -14	$R_t = f(W_{S,}\ RS)$
GFF -15	$R_t = f(RH, RS)$
GFF-16	$R_t = f(T_{max}\ T_{min,}\ W_{S,})$
GFF-17	$R_t = f(T_{max}\ T_{min,,}RH)$
GFF-18	$R_t = f(T_{max}\ T_{min,}RS)$
GFF-19	$R_t = f(T_{max,}\ W_S, RH)$
GFF-20	$R_t = f(T_{max,}\ W_S, RS)$
GFF-21	$R_t = f(T_{max,}\ RH, RS)$
GFF-22	$R_t = f(T_{min,}\ W_S, RH)$
GFF-23	$R_t = f(T_{min,}\ W_S, RS)$
GFF-24	$R_t = f(T_{min,},RH, RS)$
GFF-25	$R_t = f(W_{S,}\ RH,RS)$
GFF-26	$R_t = f(T_{max}\ T_{min,}\ W_{S,}\ RH)$
GFF-27	$R_t = f(T_{max}, T_{min,,}\ RH, RS)$
GFF-28	$R_t = f(T_{max}\ T_{min,}\ W_{S,}\ RS)$
GFF-29	$R_t = f(T_{max,}\ W_{S,}\ RH, RS)$
GFF-30	$R_t = f(T_{min,,}\ W_{S,}\ RH, RS)$
GFF-31	$R_t = f(T_{max}, T_{min,}\ W_{S,}\ RH, RS)$

Table -3.1: Input-output combinations for GFF models for rainfall simulation

Model No.	Input-Output Variables*
MLR-1	$S_t = a_1 + b_1T_{max}$
MLR-2	$S_t = a_2 + c_1W_s$
MLR-3	$S_t = a_3 + b_2T_{min}$
MLR-4	$S_t = a_4 + c_2RH$
MLR-5	$S_t = a_5 + d_1R_s$
MLR-6	$S_t = a_6 + b_3T_{max} + c_3W_s$
MLR-7	$S_t = a_7 + b_4T_{max} + b'_4T_{min}$
MLR-8	$S_t = a_8 + b_5T_{max} + c_4RH$
MLR-9	$S_t = a_9 + b_6T_{max} + d_2R_s$
MLR-10	$S_t = a_{10} + b_7T_{min} + c_5W_s$
MLR-11	$S_t = a_{11} + c_6W_s + c'_6RH$
MLR-12	$S_t = a_{12} + c_7W_s + d_3R_s$
MLR-13	$S_t = a_{13} + b_8T_{min} + c_8RH$
MLR-14	$S_t = a_{14} + b_9T_{min} + d_4R_s$
MLR-15	$S_t = a_{15} + c_9RH + d_5R_s$
MLR-16	$S_t = a_{16} + b_{10}T_{max} + c_{10}W_s + b'_{10}T_{min}$
MLR-17	$S_t = a_{17} + b_{11}T_{max} + c_{11}W_s + + c'_{11}RH$
MLR-18	$S_t = a_{18} + b_{12}T_{max} + c_{12}W_s + d_6R_s$
MLR-19	$S_t = a_{19} + b_{13}T_{max} + b'_{13}T_{min} + c_{13}RH$
MLR-20	$S_t = a_{20} + b_{14}T_{max} + b'_{14}T_{min} + d_7R_s$
MLR-21	$S_t = a_{21} + b_{15}T_{max} + c_{14}RH + d_8R_s$
MLR-22	$S_t = a_{22} + b_{16}T_{min} + c_{15}W_s + c'_{15}RH$
MLR-23	$S_t = a_{23} + b_{17}T_{min} + c_{16}W_s + d_9Rs$
MLR-24	$S_t = a_{24} + c_{17}W_s + c'_{17}RH + d_{10}R_s$
MLR-25	$S_t = a_{25} + b_{18}T_{min} + c_{18}RH + d_{11}R_s$
MLR-26	$S_t = a_{26} + b_{19}T_{max} + b'_{19}T_{min} + c_{19}W_s + c'_{19}RH$
MLR-27	$S_t = a_{27} + b_{20}T_{max} + b'_{20}T_{min} + c_{20}W_s + d_{12}R_s$
MLR-28	$S_t = a_{28} + b_{21}T_{max} + c_{21}W_s + c'_{21}RH + d_{13}R_s$
MLR-29	$S_t = a_{29} + b_{22}T_{max} + b'_{22}T_{min} + c_{22}RH + d_{14}R_s$
MLR-30	$S_t = a_{30} + b_{23}T_{min} + c_{23}W_s + c'_{23}RH + d_{15}Rs$
MLR-31	$S_t = a_{31} + b_{24}T_{max} + b'_{24}T_{min} + c_{24}W_s + c'_{24}RH + d_{16}R_s$

*ai, bi, b'_i, ci, c'_i and di are regression coefficients (i = 1, 2,…., 31)

Table 3.2 Input-output combinations MLR models for rainfall simulation at Satna (M.P.)

4. RESULTS AND DISCUSSION

The performance of the development model were evaluated qualitatively and quantitatively by the visual observation, and based on various statistical and hydrological indices such as correlation coefficient (r), coefficient of efficiency (CE) and mean squared error (MSE). The thirty one model having high values of rand CE and lower values of MSE is considered as the better fit model.

4.1 Rainfall Modeling using GFF

The increased values of CE and r by GFF models during testing period indicate good generalization capability of the selected GFF models. It is clear from table 4.1 GFF-2 model with 1-2-1 architecture (one inputs; four hidden neurons; one output) had lower MSE (0.00140) and higher CE (0.8336) and r (0.9926) values in the testing phase.

Table 4.1 Statistical indices for GFF models for rainfall simulation during testing

S.NO	MODEL	STRUCTURE	TESTING			
			MSE	r	R2	CE
1	M1	(1-2-1)	0.00140	0.9926	0.9853	0.8336
2	M2	(1-2-1)	0.00143	0.9944	0.9889	0.8297
3	M1	(1-4-1)	0.00149	0.9892	0.9786	0.8226
4	M9	(2-8-1)	0.00153	0.9736	0.9479	0.8179
5	M8	(2-6-1)	0.001681	0.9761	0.9529	0.8006
6	M15	(2-8-1)	0.001701	0.9612	0.9239	0.7981
7	M2	(2-8-1)	0.00173	0.9950	0.9902	0.7939
8	M13	(2-10-1)	0.00178	0.9632	0.9279	0.7882
9	M13	(2-8-1)	0.00181	0.9589	0.9197	0.7842
10	M4	(1-2-1)	0.001843	0.96984	0.9406	0.7813

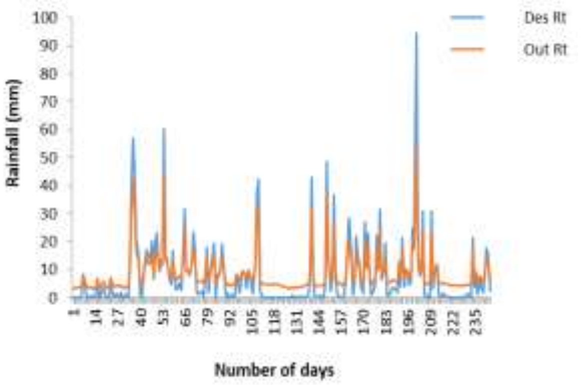


Fig 4.1 Comparison of observed and predicted rainfall by model 1, Neuron 2, during the validation period

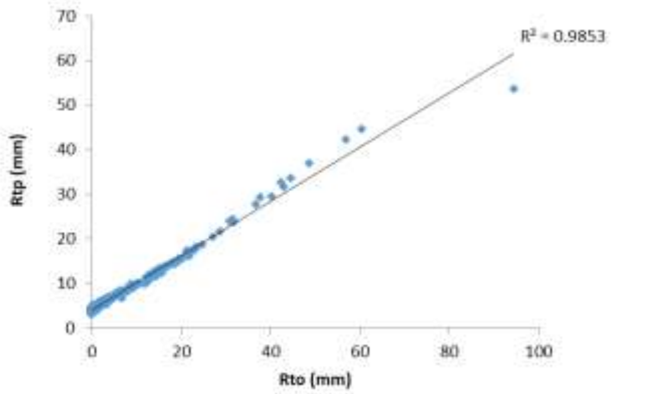


Fig 4.2 Correlation between observed and predicted rainfall by model 1, Neuron 2 during the validation period.

4.2 Rainfall Modeling using MLR

On the basis of the lowest value of MSE and the highest values of r and CE, the MLR-31 model was found to be the best performing model. Therefore, according to MLR-31 model, the current day's rainfall depends on minimum temperature, relative humidity and rainfall of current day.

The MSE varied from 96.8136 to 105.2627 m3/s; the CE varied from 0.3231 to 0.2640; and r varied from 0.5684 to 0.5138.

Table 4.2 Statistical indices for selected MLR rainfall models during testing period (2012-2013)

Model No.	Statistical index			
	MSE	CE	r	R^2
M31	96.8136	0.3231	0.5684	0.3231
M30	96.9728	0.3219	0.5674	0.3219
M27	96.9379	0.3222	0.5258	0.322
M29	99.4299	0.304	0.552	0.304

M17	100.49	0.2977	0.5456	0.2977
M21	100.457	0.2976	0.5455	0.2976
M9	101.017	0.2937	0.5419	0.2937
M28	101.172	0.2765	0.525	0.2765
M15	102.705	0.2819	0.5309	0.2819
M6	105.26	0.2640	0.5138	0.2640

The observed (Rto) and predicted (Rtp) rainfall simulated by MLR models were compared in the form of graph and scatter-plot as shown in Figs. 4.11 to 4.20. The rainfall graphs indicate that the models under predict the peak rainfall as confirmed by the scatter plots also. This study gave clear indication of non-applicability of the MLR model to simulate rainfall for the study area due to low values of CE and r.

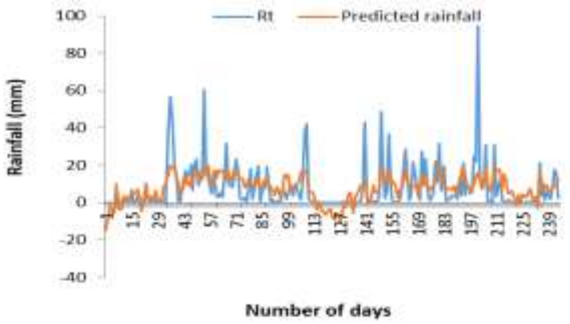


Fig 4.3 Comparison of observed and predicted rainfall by MLR-31, during the validation period.

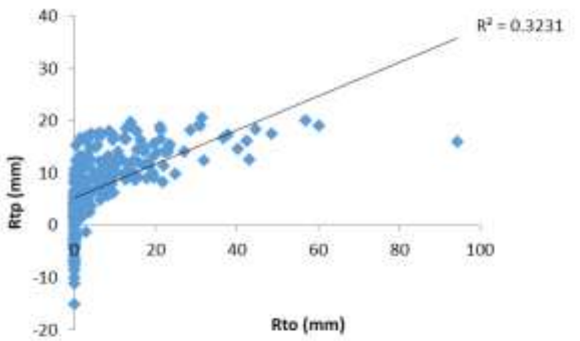


Fig 4.22 Correlation between observed and predicted rainfall by MLR-31, during the validation period.

3. CONCLUSIONS

1. Rainfall can be simulated by using GFF model with input parameters as maximum temperature, minimum temperature, relative humidity and rainfall.

2. The predicted values of rainfall using GFF (Generalized Feed Forward) were found to be much closer to the observed value of rainfall as compared to MLR.

3. On the basis of lower MSE value and higher CE and r values, GFF-1 model was found to be the best model.

4. It was clearly evident that MLR model fits very poorly for the dataset under study.

REFERENCES

[1] **C. L. Wu, K. W Chau, and C. Fan, 2013,** "Prediction of Rainfall Time Series Using Modular Artificial Neural Networks Coupled with Data Preprocessing Techniques", Journal of Hydrology, Vol. 389, No. 1-2, pp. 146-167.

[2] **Duong Tran Anh 1,3,* , Thanh Duc Dang 2,3** and Song Pham Van 3 (2019) Improved Rainfall Prediction Using Combined Pre-Processing Methods and Feed-Forward Neural Networks.

[3] **El-shafie, M. Mukhlisin, A. Najah Ali and M. R. Taha, 2011**, "Performance of artificial neural network and regression techniques for rainfall-runoff prediction", International Journal of the Physical Sciences, Vol. 6(8), pp. 1997-2003.

[4] **Saha, N., Jha, A. K., Pathak, K. K. 2012** "Comparison of Regression and Artificial Neural Network Impact Assessment Models: A Case Study of Micro-Watershed Management in India"

BIOGRAPHIES

[1]PG Scholar SHUATS Allahabad U.P

[2]Assistant Professor SHUATS Allahabad U.P

[3]Associate Professor SHUATS Allahabad U.P

[4]Assistant Professor SHUATS Allahabad U.P