Journal of Pharmacognosy and Phytochemistry
Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2018; 7(5): 658-662
Received: 19-07-2018
Accepted: 21-08-2018

Tushar Rathod
Ph. D Student, Department Soil and Water Conservation Engineering, VIAET, SHUATS, Allahabad, Uttar Pradesh, India

Dr. Vikram Singh
Assistant Professor, Department Soil and Water Conservation Engineering, VIAET, SHUATS, Allahabad, Uttar Pradesh, India

Dr. SK Srivastava
Associate Professor, Department Irrigation and Drainage Engineering VIAET, SHUATS, Allahabad, Uttar Pradesh,, India

Er. CJ Wesley
Assistant Professor, Department of Remote Sensing & GIS, SIET, SHUATS, Allahabad, Uttar Pradesh, India

Dr. MA Alam
Assistant Professor, Department Irrigation and Drainage Engineering VIAET, SHUATS, Allahabad, Uttar Pradesh, India

Correspondence
Tushar Rathod
Ph. D Student, Department Soil and Water Conservation Engineering, VIAET, SHUATS, Allahabad, Uttar Pradesh, India

Rainfall prediction using co-active neuro fuzzy inference system for umargaon watershed Nagpur India

Tushar Rathod, Dr. Vikram Singh, Dr. SK Srivastava, Er. CJ Wesley and Dr. MA Alam

Abstract
Rainfall forecasting represents a tremendously significant matter in field of hydrology. On other hand, rainfall is one of the most complicated effective hydrological processes in rainfall prediction. This study was undertaken to develop and evaluate the applicability of Co-Active Neuro Fuzzy Interface System (CANFIS) models to simulate rainfall from a watershed. The performance of the developed models, on the basis of training and testing, was judged on the basis of three statistical measures such as root mean squared error (MSE), coefficient of efficiency (CE) and correlation coefficient (r) during monsoon period (June to September) for Umargaon area in Nagpur, Maharashtra, India. The daily data of rainfall, sunshine hours, minimum temperature, maximum temperature and evaporation data were used for rainfall simulation. The appropriate parameter combination of input variables for CANFIS was used to simulate rainfall. The Neuro Solution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation of developed models, respectively. The architecture of CANFIS was designed with Gaussian membership function, Takagi-Sugeno-Kang fuzzy model, hyperbolic tangent activation function and Delta-Bar-Delta learning algorithm. The result indicated that the predicted rainfall using CANFIS model was found to be in close agreement with the observed one for the Umargaon. Therefore, according to CANFIS model, the rainfall can be simulated using the data of maximum temperature, evaporation and sunshine hours. The result indicates that the CANFIS model is suitable for rainfall prediction in Nagpur.

Keywords: Rainfall, development of Model, CANFIS

Introduction

The rainfall is very complex in nature involving a number of variables pertaining to rainfall, physiography, soil type, cropping system and management practices, and therefore needs in-depth conceptualization and formulation. In recent years, artificial neural network (ANN), fuzzy theory, genetic algorithm and chaos theory have been widely applied in the sphere of hydrology and water resource. The studies have demonstrated that, these forecasting approaches are not very satisfactory in precision because of their considering only some aspects of its property. ANN have been recently accepted as an efficient alternative tool for modeling the complex hydrologic systems and widely used for prediction. Many researchers have shown the applicability of ANN to hydrology which includes probability of precipitation (PoP) is important for many decision makers who are sensitive to the occurrence of precipitation. An accurate quantitative precipitation forecast (QPF) can identify the potential for heavy precipitation and possible associated flash flooding, as well as providing information for hydrologic interests.

The neuro-fuzzy techniques have been successfully applied to solve a number of problems in water resource planning and management, including: stream flow reconstruction (Chang *et al.* 2001) [1, 6]; modeling of hydrological time series prediction (Nayak *et al.* 2004, Keskin *et al.* 2004) [2, 3]. Co-Active Neuro-Fuzzy Inference System (CANFIS), which is integration of neural networks and fuzzy logic, has the potential to capture the benefits of both these fields in a single framework. CANFIS utilizes linguistic information from the fuzzy logic as well learning capability of an ANN. Fuzzy mapping algorithm that is based on Tagaki-Sugeno-Kang (TSK) fuzzy inference system (Jang *et al.*, 1997; Loukas, 2001) [4, 10, 5]; CANFIS has many applications such as, database management, system design and planning/forecasting of the water resources (Chang *et al.*, 2001; Nayak *et al.*, 2004; Chen *et al.*, 2006; Chang *et al.*, 2006; Firat *et al.*, 2009) [1, 6, 2, 7, 8, 9]; Analysis of rainfall and determination of annual maximum daily rainfall would enhance the management of water resources applications as well as the effective utilization of water resources (Subudhi, 2007) [12].

Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances. In this study, the rainfall models have been developed using CANFIS technique for the Umargaon watershed in the Nagpur District of Maharashtra, India. Appropriate input variables for these models were selected on the basis of statistical indices such as Mean Square Error (MSE), coefficient of efficiency (CE), coefficient of determination and coefficient of correlation (r).

Materials and Methods

Study area

The study area, the Umargaon watershed, lies between the longitudes 21.058º N Latitude and 78.058º E Longitude with an area of about 345 ha. It is a part of Nagpur (rural) of Nagpur district of Maharashtra, India. The elevation of the watershed varies from 310 to 315 m above the mean sea level. The catchment lies in sub-humid tropical climatic zone. The slope ranges from 1 to 5% with an average slope of 2%. The rainfall occurs mainly occurs during monsoon months from June to September with average annual rainfall of about 1240 mm. Watershed experiences hot summer and mild winter with mean summer temperature of 44.2 ^{0}C and mean winter temperature of 7.6 ^{0}C. Daily rainfall, sunshine hour, minimum and maximum temperature and evaporation were collected from IMD, Pune.

Methodology

Co-Active Neuro Fuzzy Inference System:

CANFIS (Co-active neuro-fuzzy inference system) belongs to a general class of adaptive neuro-fuzzy inference systems (ANFIS) (Jang *et al.* 1997) [4, 10]. This may be regarded as a universal approximator of any nonlinear function. The characteristics of CANFIS give the advantage of integrating neural networks (NNs) with fuzzy inference systems (FIS) in the same topology. The fundamental component for CANFIS is a fuzzy neuron that applies membership functions (MF) to the inputs. The number of different membership functions (MF) such as piece wise linear (triangular and trapezoidal), Gaussian, generalized bell shaped, and sigmoidal have been proposed by fuzzy control system.

For simplicity, the fuzzy inference system is assume with two inputs *x* and *y* and one outputs *f.* The first-order Sugeno fuzzy model, a typical rule set with two fuzzy IF-THEN rules for CANFIS architecture, can be expressed as follows (Saemi and Ahmadi 2008) [11]:

Rule 1: IF x is A_1 and y is B_1 THEN $f_1=p_1x+q_1y+r_1$ (1)

Rule 2: IF x is A_2 and y is B_2 THEN $f_2=p_2x+q_2y+r_2$ (2)

Where, A_1, A_2 and B_1, B_2 are the MFs for inputs x and y respectively; p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters in the THEN-part of the first-order Sugeno fuzzy model (Fig. 1).

The architecture of five layers CANFIS model is shown in Fig. 1. Each node in layer 1 is the membership grades of a fuzzy set (A1, A_2, B_1, or B_2) and specifies the degree to which the given input belongs to one of the fuzzy sets. The fuzzy sets are defined by three membership functions. Layer 2 receives input in the form of the product of all output pairs from the first layer. The third layer has two components. The upper component applies the membership functions to each of the inputs, while the lower component is a representation of the modular network that computes, for each output, the sum of all the firing strengths. The fourth layer calculates the weight normalization of the output of the two components from the third layer and produces the final output of the network (Ishak and Trifiro, 2007) [12]. The Hyperbolic tangent (Tanh) activation function was used for input data normalization (−1 to 1) and Delta-Bar-Delta algorithm was used as learning algorithm. Gaussian membership function was used as a membership function. The membership function is used to handle the complexity of the model and therefore in the present study membership function ranging from 2 to 6 have been tried to find the appropriate number of membership function for rainfall modeling for the study area. The Neuro Solution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation of developed models, respectively.

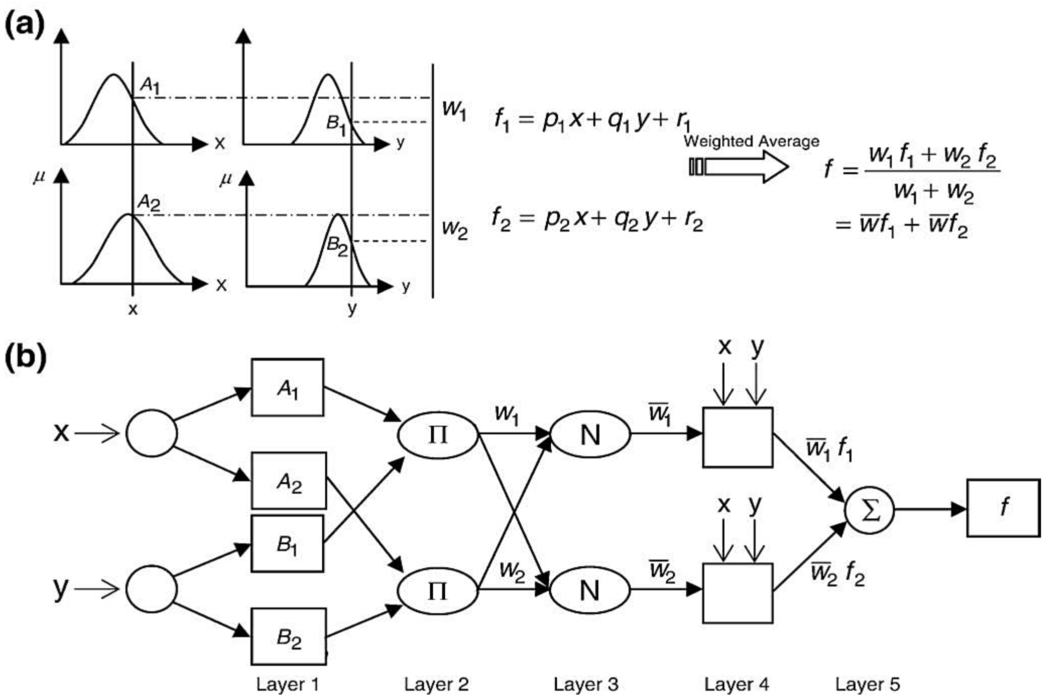

Fig 1: (a) Sugeno’s fuzzy IF-THEN rule; and (b) the equivalent CANFIS architecture

Development of CANFIS models

Identification of input and output variables is the first step for developing the CANFIS model. The output from the models is the Rainfall R_t based on respective input variables. The various input-output combinations of CANFIS models for the study area are listed in Table 1.

Table 1: Input-output combination for CANFIS models for rainfall simulation at Umargaon watershed.

Model No.	Input-Output Variables	Model No.	Input-Output Variables
CANFIS-1	$R_t = f(T_{min})$	CANFIS-9	$R_t = f(T_{max}, S_n)$
CANFIS-2	$R_t = f(T_{max})$	CANFIS-10	$R_t = f(E_p, S_n)$
CANFIS-3	$R_t = f(E_p)$	CANFIS-11	$R_t = f(T_{min}, T_{max}, E_p)$
CANFIS-4	$R_t = f(S_n)$	CANFIS-12	$R_t = f(T_{min}, T_{max}, S_n)$
CANFIS-5	$R_t = f(T_{min}, T_{max})$	CANFIS-13	$R_t = f(T_{min}, E_p, S_n)$
CANFIS-6	$R_t = f(T_{min}, E_p)$	CANFIS-14	$R_t = f(T_{max}, E_p, S_n)$
CANFIS-7	$R_t = f(T_{min}, S_n)$	CANFIS-15	$R_t = f(T_{min}, T_{max}, E_p, S_n)$
CANFIS-8	$R_t = f(T_{max}, E_p)$		

Training and Testing of CANFIS models

The daily data of rainfall, minimum temperature (T_{min}), maximum temperature (T_{max}), Evaporation (E_p) and sun shine hours (S_n) were split into two sets: a training data set from 2006 to 2013 for Umargaon watersheds and a testing data set from 2014 to 2015 for Umargaon watersheds. The input pairs in the training data set were applied to the network of a selected architecture and training was performed using Gaussian and generalized bell membership functions for CANFIS models.

Performance evaluation criteria

In this study, a number of networks were constructed and each of them was trained separately, and the best network was selected based on the accuracy of predictions in the testing phase. Different performance evaluation indices were applied to test the performance of the developed CANFIS models. The visual observation based on the graphical comparison between observed and predicted values was made to evaluate and compare the performance of the models. Since, the comparison by visual observations may have personal bias, the following statistical indices such as mean square error (MSE), coefficient of efficiency (CE), coefficient of determination (R^2) and coefficient of correlation (r) were used to evaluate the performance of the CANFIS models for comparison between observed and predicted values.

The mean squared error (MSE) was determined to measure the prediction accuracy. The MSE is zero for perfect fit and increased values indicate higher discrepancies between predicted and observed values. The MSE between observed and predicted values is computed by the following equation:

$$MSE = \frac{\sum_{i=1}^{n}(X_{oi}-X_{pi})^2}{n} \qquad \text{Eq. (3)}$$

Where, X_{oi} and X_{pi} are the observed and predicted values for i^{th} dataset and n is the total number of observations.

The correlation coefficient measures the degree to which two variables are linearly related. A positive correlation coefficient indicates that the observed and predicted values tend to go up and down together. If the variables go in opposite directions, it results in a negative correlation coefficient. The correlation coefficient (r) is computed by the following equation:

$$r = \frac{\sum_{i=1}^{N}(X_{oi}-\overline{X_o})(X_{pi}-\overline{X_p})}{\sqrt{\sum_{i=1}^{N}(X_{oi}-\overline{X_o})^2\sum_{i=1}^{N}(X_{pi}-\overline{Y_p})^2}} \qquad \text{Eq. (4)}$$

where, $\overline{X_o}$ and $\overline{Y_p}$ are the mean of observed and predicted values, respectively.

To assess the goodness of fit between observed and predicted values of the output, the CE was suggested by The coefficient of efficiency is computed by the following equation;

$$CE = \left[1 - \frac{\sum_{i=1}^{N}(X_{oi}-X_{pi})^2}{\sum_{i=1}^{N}(X_{oi}-\overline{X_o})^2}\right] X\ 100 \qquad \text{Eq. (5)}$$

Where, X_{oi} and X_{pi} are the observed and predicted values for i^{th} dataset and n is the total number of observations.

Results and Discussion

Canfis

CANFIS models (Table 1) were used to simulate rainfall as output based on various input combinations of minimum and maximum temperature, evaporation and sunshine hours. CANFIS model Nos. 14, 11, 10, 15, 8 were selected for further analysis and comparison based on the statistical indices, such as mean squared error (MSE), coefficient of efficiency (CE) coefficient of determination (R^2) and correlation coefficient (r). The values of statistical indices for the selected CANFIS models during testing are presented in Table 2 respectively.

As observed from Table 4.8 indicates that the MSE for the selected CANFIS models varied from 0.0057 to 0.0094 during testing. The CE values ranged from 0.4217 to 0.6469 for testing. The correlation coefficient (r) values ranged from 0.8580 to 0.9085 for testing. The increased values of CE and r by CANFIS models during testing period indicate good generalization capability of the selected CANFIS models. On the basis of lower values of MSE (0.0063) and higher CE (0.6125) and r (0.9069) in the testing phase, the CANFIS – 14 models were found to be the best performing model. Therefore, according to this model, the rainfall can be simulated using the data of maximum temperature, evaporation and sunshine hours. It was closely followed by CANFIS-11 models according to which the Rainfall depends on minimum and maximum temperature and evaporation.

Table 2: Statistical indices for selected CANFIS models during testing phase for Umargaon watershed

Model No.	MF per input	Testing			
		MSE	CE	r	R^2
CANFIS – 14	Gauss – 3	0.0063	0.6125	0.9069	0.8224
CANFIS – 11	Gauss – 4	0.0080	0.5111	0.9085	0.8254
CANFIS – 10	Gauss – 3	0.0057	0.6469	0.9012	0.8122
CANFIS – 15	Gauss – 3	0.0094	0.4217	0.8939	0.7991
CANFIS – 8	Gauss – 6	0.0068	0.5849	0.8873	0.7872

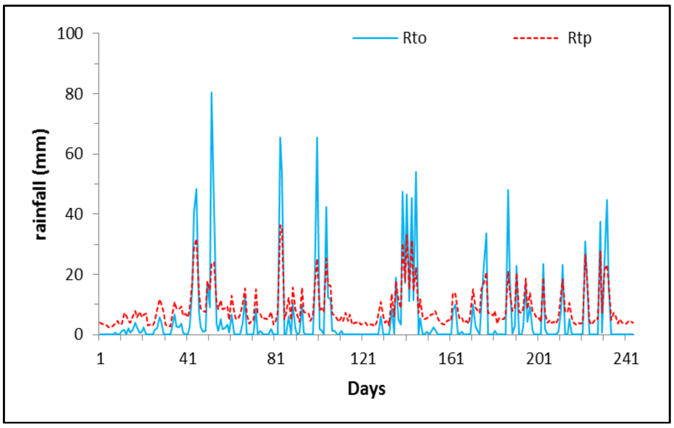

Fig 2: Comparison of observed and predicted rainfall by CANFIS-14, Gauss-3 during the validation period

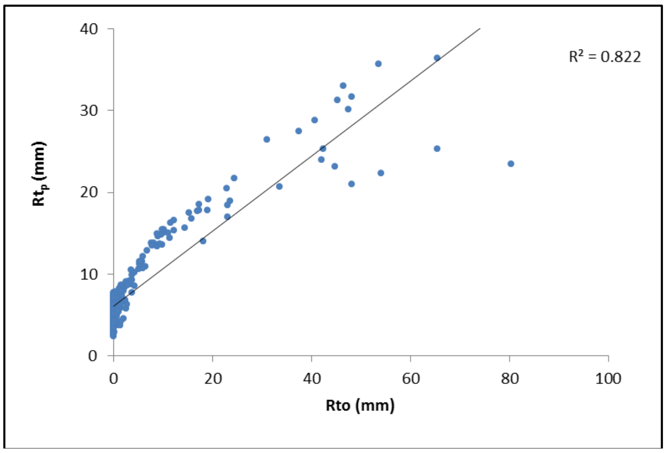

Fig 3: Correlation between observed and predicted rainfall by CANFIS- 14, Gauss 3 during the validation period

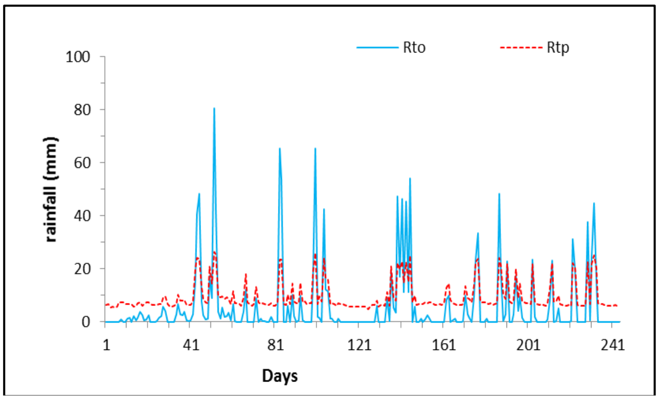

Fig 4: Comparison of observed and predicted rainfall by CANFIS- 11, Gauss-4 during the validation period.

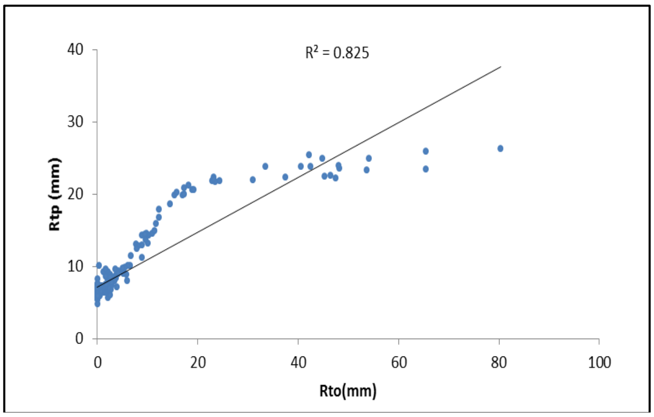

Fig 5: Correlation between observed and predicted rainfall by CANFIS- 11, Gauss-4 during the validation period.

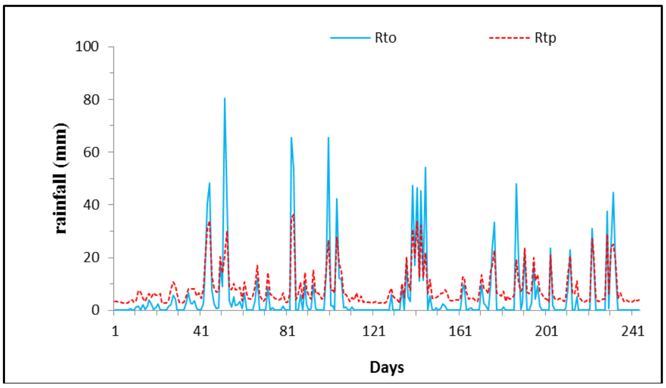

Fig 6: Comparison of observed and predicted rainfall by CANFIS- 10, Gauss-3 during the validation period

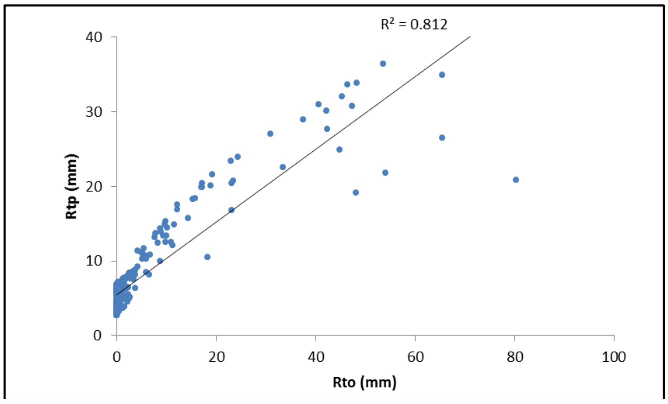

Fig 7: Correlation between observed and predicted rainfall by CANFIS- 10, Gauss-3 during the validation period

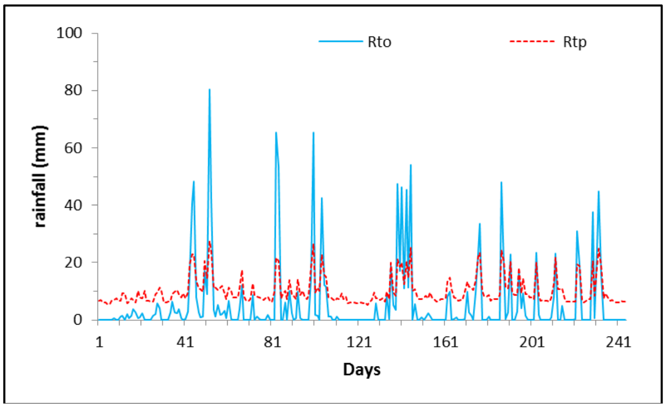

Fig 8: Comparison of observed and predicted rainfall by CANFIS- 15, Gauss-3 during the validation period

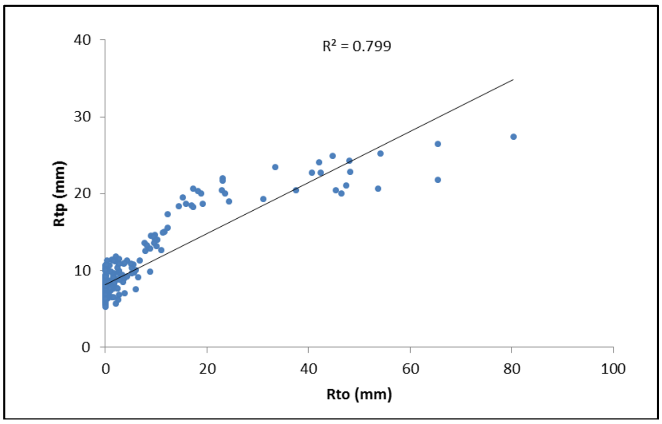

Fig 9: Correlation between observed and predicted rainfall by CANFIS- 15, Gauss-3 during the validation period.

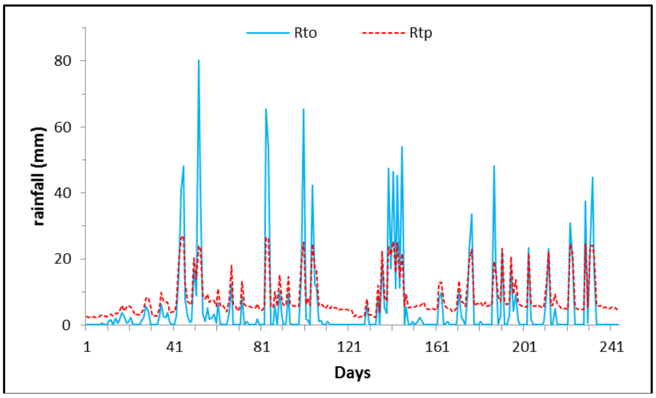

Fig 10: Comparison of observed and predicted rainfall by CANFIS- 8, Gauss-6 during the validation period

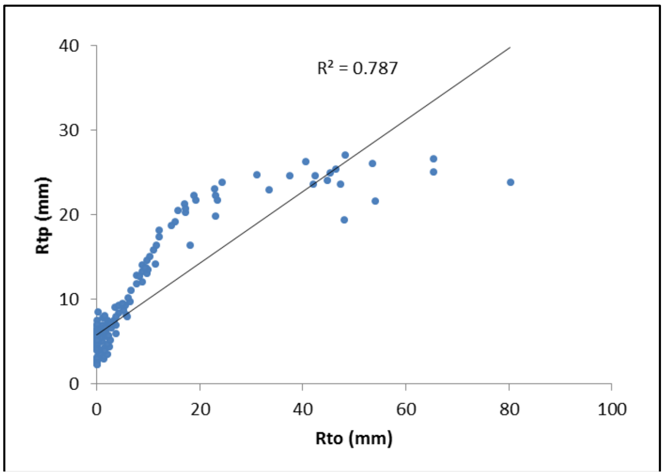

Fig 11: Correlation between observed and predicted rainfall by CANFIS- 8, Gauss-6 during the validation period

Conclusions

The main objective of this study was to investigate the applicability of CANFIS models to simulate rainfall for the Umargaon watershed in Nagpur district of Maharashtra, India. The Neuro Solution 5.0 software and Microsoft Excel were used in the analysis and the performance evaluation of the developed models.

References

1. Chang FJ, Hu HF, Chen YC. Counter propagation fuzzy-neural network for stream flow reconstruction. Hydrology Process. 2001; 15(2):219-232.
2. Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS. A Neuro-fuzzy computing technique for modelling hydrological time series. J Hydro. 2004; 291:52-66.
3. Keskin ME, Terzi O, Taylan D. Fuzzy logic model approaches to daily pan evaporation estimation in Western Turkey. Hydrology. Sci. J. 2004; 49(6):1001-1010.
4. Jang JSR, Sun CT, Mizutani E. Neuro fuzzy and soft computing: a computational approach to learning andmachine intelligence. Prentice-Hall, NJ, 1997, 607.
5. Loukas YL. Adaptive neuro-fuzzy inference system: an instant and architecture-free predictor for improved QSAR studies. J Med Chem. 2001; 44(17):2772-2783.
6. Chang FJ, Hu HF, Chen YC. Counter propagation fuzzy-neural network for streamflow reconstruction. Hydrolog. Process. 2001; 15(2):219-232.
7. Chen, Jieyun, Adams, Barry J. Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. Journal of Hydrology. 2006; 318:232-249.
8. Chang FJ, Chang YT. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Adv. Water Resoure. 2006; 29:1-10.
9. Firat M, Turan M, Yurdusev MA. Comparative analysis of fuzzy inference systems for water consumption time series prediction. J Hydrol. 2009; 374:235-241.
10. Jang JSR, Sun CT, Mizutani E. Neuro fuzzy and soft computing: A computational approach to learning andmachine intelligence. Prentice-Hall, NJ, 1997, 607.
11. Saemi M, Ahmadi M. Integration of genetic algorithm and a coactive neuro-fuzzy inference system for permeability prediction from well logs data. Transp Porous Media. 2008; 71:273-288.
12. Ishak S, Trifiro F. Neural networks, transportation research circular EC113. Artif Intell Transp, 2007, 17-32.