

E-ISSN: 2278-4136

P-ISSN: 2349-8234

www.phytojournal.com

JPP 2021; Sp 10(1): 211-216

Received: 09-11-2020

Accepted: 15-12-2020

Sonali Kumari

PG Scholar, SHUATS

Prayagraj, Uttar Pradesh, India

Dr. Vikram Singh

Assistant Professor, SHUATS,

Prayagraj, Uttar Pradesh, India

Tushar Rathod

Ph.D Scholar, SHUATS,

Prayagraj, Uttar Pradesh, India

Comparison of linear and non-linear methods for runoff prediction using, co-active neuro-fuzzy inference system (CANFIS) and multi linear regression (MLR) technique for Narmada river basin, Gujarat

Sonali Kumari, Dr. Vikram Singh and Tushar Rathod

Abstract

Runoff prediction is one of the most important and challenging task in the modern world. In this study, we attempted to forecast the daily runoff on the basis of Coactive Neuro-Fuzzy Inference System (CANFIS) and Multi Linear Regression (MLR) techniques. The performance of the developed models, on the basis of training and testing, was judged on the basis of four statistical measures such as Root Mean Squared Error (RMSE), Coefficient of Efficiency (CE), Correlation Coefficient (r) and Coefficient of Determination (R^2) during monsoon period (June to September) for Chhota Udaipur area in Gujarat, India. The daily data of rainfall, minimum temperature, maximum temperature and wind speed were used for runoff prediction. The appropriate parameter combination of input variables for CANFIS was used to predict runoff. 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. Ten CANFIS models and MLR were selected based on the performance evaluation indices during testing period. CANFIS models were found to be much closer to the observed values of runoff as compared to MLR.

Keywords: Soft computing, CANFIS, MLR, runoff prediction

Introduction

Runoff can be described as the part of the water cycle that flows over land as surface water instead of being absorbed into groundwater or evaporating. According to the U.S. Geological Survey (USGS), runoff is that part of the precipitation, snow melt, or irrigation water that appears in uncontrolled surface streams, rivers, drains, or sewers. An accurate long-term runoff prediction is necessary for water resources management, food production and maintaining flood risks. Runoff models play a significant role in water resource management, planning and hydraulic design. The main purpose of this paper is to compare and analyze the performance of the CANFIS and MLR to see their applicability in runoff forecasting.

ANN was first developed in the 1940s and the development has experienced a renaissance in iterative auto-associable neural networks (Hung *et al.*, 2008). An ANN provides the user a model free tool, which can generate input output mapping for any set of data as complex pattern recognition can be attempted without making any initial assumptions. In addition, ANN could learn and generalize from examples to produce meaningful solution even when the input data contain errors or is incomplete. In recent years, ANNs have been used intensively for prediction and forecasting in a number of water-related areas, including water resource study (Najah *et al.*, 2009; Ahmed *et al.*, 2009; El-Shafie *et al.*, 2007(a), 2009(b); 2010) [11, 2, 6], hydrograph simulator (Deka and Chandramouli, 2005 [5]; rainfall estimating (Lin and Chen 2005; Luk *et al.*, 2001) [7, 8].

The conventional form of CANFIS model is an extension of the original ANFIS model. The underlying concept of the CANFIS model can be extended to consider any number of input and output pairs. The essential component of CANFIS is generally similar to the components of ANFIS model where a fuzzy neuron that represents a parameter called the membership function (MF) is used to construct the modelling framework. There are several types of MFs that can be utilized, including the triangular, trapezoidal, sigmoidal, Gaussian, z-shape functions, pi, the general bell and the Gaussian equations. CANFIS structure includes the normalize the output variables in the range of 0-1.

Corresponding Author:**Sonali Kumari**

PG Scholar, SHUATS

Prayagraj, Uttar Pradesh, India

The architectural network of CANFIS model has the combined axon applied the target (or output) of the MFs to the target for the neuronal network which is still similar to the common ANFIS procedure.

A Multi Linear Regression (MLR) is the simplest, statistical and well-developed representation of the time-invariant relationship between an input function and corresponding output function. MLR models are considered as benchmark in reservoir inflow forecasting (Chavan and Ukrande, 2012)^[4].

Literature Review

Sharad Patel *et al.*, (2016)^[12] Multiple Linear Regression based technique is used in determining the rainfall-runoff relations. Without considering the temperature, topography or other parameters of the study area, simply using the data of rainfall and runoff to predict the future runoff was the key characteristic of this technique. Multiple Linear Regression was used to analyze 14 years data on rainfall and runoff at the Balaghat district of Madhya Pradesh and to develop regression models for stream flow estimation with rainfall as input and different regression model tested with varying input length of data record

Abbas *et al.*, (2017)^[11] Applied and compared of artificial neural network (ANN) with back propagation algorithm and adaptive Neuro-Fuzzy Inference System (ANFIS) for predicting the daily discharge in the Tigris River in Qurnah, Basrah, south of Iraq. He constructed and developed three models for both ANNs and ANFIS techniques to forecast the daily discharge of the river depending on observed data was taken in earlier years. In all models, statistical parameters and graphical results showed that the convergence between observed and predicted data is very good by using ANFIS models as compared to ANNs models.

Tushar Rathod *et al.*, (2018)^[13] Evaluation of Co-Active Neuro Fuzzy Interface System (CANFIS) model to simulate rainfall from a watershed. The performance of 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 Neuro solution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation of developed models respectively. 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.

Anurag Malik *et al.* (2019)^[9] The heuristic approaches including Co-Active Neuro fuzzy Inference system (CANFIS), multilayer Perceptron Neural Network (MLPNN) and Multiple Linear Regression (MLR) were utilized to predict the hydrologic draught based on Multi-Scalar Stream Flow Draught Index (SDI) at Naula and Kedar stations located in upper Rāmgangā River basin, Uttarakhand state, India. The predicted values of multi-scalar SDI using CANFIS, MLPNN and MLR models were compared with the calculated values, based on root mean squared error (RMSE), Nash-Sutcliffe efficiency (NSE), Coefficient of Correlation (COC) and Wilmot Index (WI). The visual interpretation was also made using line diagram, scatter diagram and Taylor Diagram (TD). The MLR model was found to be the best at 24-month time scale for Naula station only. The result is helpful in prediction of hydrological drought on multiple time scales and decision making for remedial schemes to cope with hydrological drought at Naula and Kedar stations.

Md. Ajaz Alam *et al.* (2019)^[13] Modelled to forecast the daily rainfall on the basis of Co-Active Neuron Fuzzy Inference System (CANFIS) and Multi linear Regression (MLR) techniques. Fifteen CANFIS models and MLR were selected based on the performance evaluation indices during testing period. CANFIS models were found to be much closer to the observed values of rainfall as compared to MLR.

Materials and Method

The area selected for the study is Chhota Udaipur a city and municipality in Chhota Udaipur district in the state of Gujarat, India, located at 22°18'20" N, 74°0'50"E. The daily data of runoff for monsoon season June 2004 to September 2010 were used. Total data of monsoon season for 7 years were divided into two sets: (i) training data set consisting of first 5 years data from 1 June 2004 to 31 September 2008; and (ii) testing data set consisting of remaining 2 years data from 1 June 2009 to 31 September 2010.

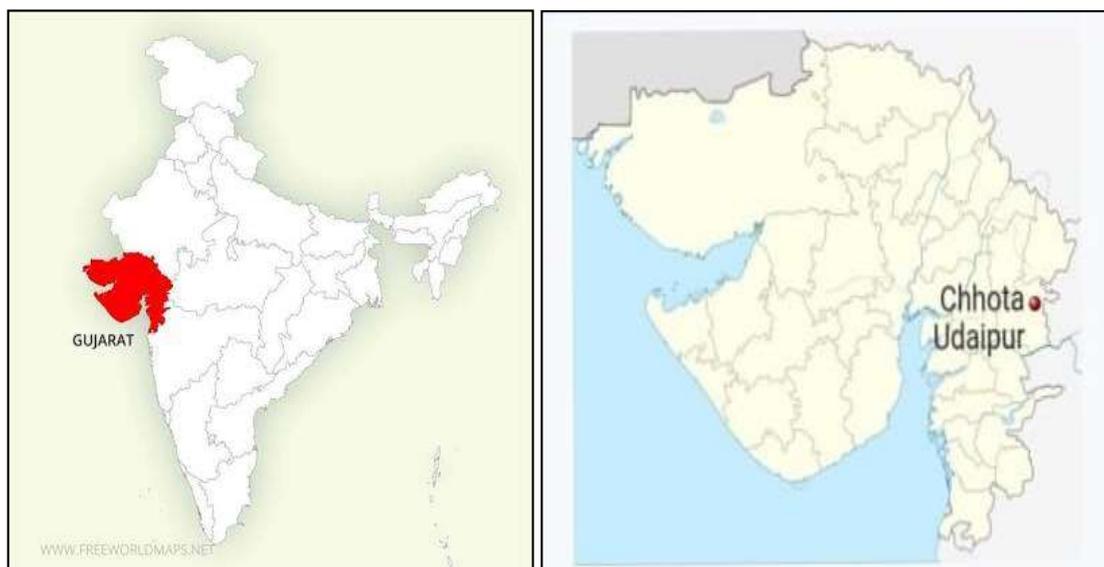


Fig 1: Map of Gujarat in India and map of Chhota Udaipur in Gujarat.

In this study, the soft computing technique based on ANN, Co-Active Neuro Fuzzy Interface System (CANFIS) and Multi Linear Regression (MLR) has been developed for predicting the total runoff in Chhota Udaipur, Gujarat. The

methodology of developing the CANFIS models along with training and testing of developed models, the Neuro Solution 5.0 software and Microsoft Excel were used in analysis and the performance evaluation indices for developed models.

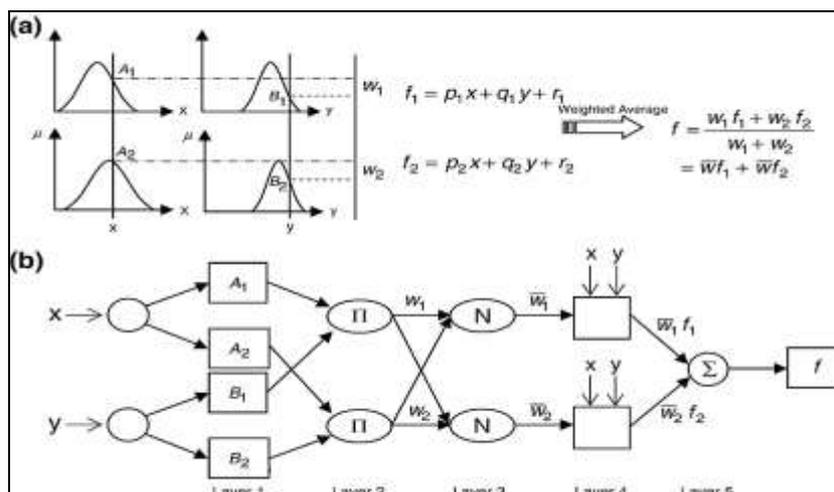


Fig 2: A basic overview of CANFIS structure

In this study, membership functions (Gaussian) is used. Choosing the number of membership functions for each input reflects the complexity of CANFIS model for selecting parameters. In each application, a different number of membership functions were tried, and the best one that gives the minimum errors was selected.

Table 1: Input-output combination for CANFIS models for runoff prediction at Chhota Udaipur city of Gujarat

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

The daily data of runoff and meteorological data (maximum temperature, minimum temperature, wind speed, rainfall) on daily basis were split into two sets: a training data set from 2004 to 2008 and a testing data set from 2009 to 2010 for Chhota Udaipur city.

The input pairs in the training data set were applied to the network of a selected architecture and training was performed using CANFIS.

The model attempts to reproduce the outcome based on the learning or training on data input information. In ANN model of the biological neurons, there are three basic components such as:

- I. **Membership function:** The function that specifies the degree to which a given input belongs to a set.
- II. **Learning algorithm:** It trains the Multi-Layer Neuron Network.

III. **Activation function:** The activation function controls the amplitude of the output of the neuron.

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) [10].

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3.1)$$

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.

Table 2: Input-output combination for MLR models for runoff prediction at Chhota Udaipur city of Gujarat.

Model No.	Input-Output Variables
MLR-1	$Q_t = e_1 + g_1 T_{min}$
MLR-2	$Q_t = e_2 + g_2 T_{max}$
MLR-3	$Q_t = e_3 + h_1 Ws$
MLR-4	$Q_t = e_4 + h_2 Rt$
MLR-5	$Q_t = e_5 + g_3 T_{min} + g_3 T_{max}$
MLR-6	$Q_t = e_6 + g_4 T_{min} + h_3 Ws$
MLR-7	$Q_t = e_7 + g_5 T_{min} + h_4 Rt$
MLR-8	$Q_t = e_8 + g_6 T_{max} + h_5 Ws$
MLR-9	$Q_t = e_9 + g_7 T_{max} + h_6 Rt$
MLR-10	$Q_t = e_{10} + h_7 Ws + h_8 Rt$
MLR-11	$Q_t = e_{11} + g_8 T_{min} + g_9 T_{max} + h_9 Ws$
MLR-12	$Q_t = e_{12} + g_{10} T_{min} + h_{10} Ws + h_{11} Rt$
MLR-13	$Q_t = e_{13} + h_6 T_{max} + k_7 Ws + I_6 Rt$
MLR-14	$Q_t = e_{14} + g_7 T_{min} + h_7 T_{max} + I_7 Rt$
MLR-15	$Q_t = e_8 + g_8 T_{min} + h_8 T_{max} + k_8 Ws + I_8 Rt$

Performance evaluation

Performance of the models developed in this study will be evaluated by using qualitative performance and quantitative performance. Qualitative performance of the models will be checked by the visual observation, whereas, quantitative performance will be verified by estimating the values of statistical and hydrological indices such as Correlation Coefficient (CC), Root Mean Square Error (RMSE), Coefficient of Efficiency (CE) and Coefficient of Determination (R^2).

Correlation Coefficient (r)

This is a number between -1.0 and +1.0, which means the degree to which two variables are linearly related. If there is a perfect linear relationship with a positive slope between the two variables, the correlation coefficient is equal to 1. The measure, however, is very intensive to derive from larger observations. Equation 3.2 denotes Karl Pearson's Correlation Coefficient between observed and predicted discharge.

$$r = \frac{\sum_{i=1}^N (X_{oi} - \bar{X}_o)(Y_{pi} - \bar{Y}_p)}{\sqrt{\sum_{i=1}^N (X_{oi} - \bar{X}_o)^2 \sum_{i=1}^N (Y_{pi} - \bar{Y}_p)^2}} \quad (3.2)$$

Where, \bar{X}_o and \bar{Y}_p are the mean of observed and predicted values, respectively.

A positive r indicates that the observed and predicted values tend to go up and down together.

Root Mean Square Error (RMSE):

This measure gives an overall agreement between the observed and modelled datasets. It has no upper bound with zero as the value for a perfect model. RMSE is a good measure of accuracy, but only to compare prediction errors of different models for a particular variable. Equation 3.3 shows RMSE between observed and predicted values.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_{oi} - Y_{pi})^2} \quad (3.3)$$

Where,

X_{oi} and Y_{pi} are the observed and predicted values for i^{th} datasets and N is the total number of observations.

Coefficient of Efficiency (CE)

To assess the goodness of fit between observed and predicted values of runoff simulation, the CE was suggested by Nash and Sutcliffe (1970). The coefficient of efficiency is computed by the following equation;

$$CE = \left[1 - \frac{\sum_{i=1}^N (X_{oi} - Y_{pi})^2}{\sum_{i=1}^N (X_{oi} - \bar{X}_o)^2} \right] \quad (3.4)$$

Coefficient of Determination (R^2)

It is a measurement used to explain how much variability of one factor can be caused by its relationship to another related factor. This correction, is known as "goodness of fit" is represented as a value between 0.0 and 1.0.

$$R^2 = \frac{SSR}{SST} \quad (3.5)$$

Where,

SSR = Sum of squared regression

SST = Total variation in data

Results and Discussion

The performances of models were evaluated qualitatively and quantitatively by visual observation and various statistical and hydrological indices viz. Root Mean Square Error (RMSE), Correlation Coefficient (r), Coefficient of Efficiency (CE) and Coefficient of Determination (R^2). The model having higher values of Correlation Coefficient and Coefficient of Efficiency and low value of Root Mean Square Error is consider as the best fit model.

Runoff modelling using CANFIS

CANFIS models (Table 3) were used to predict daily runoff as output based on various input combinations of minimum and maximum temperature, wind speed and rainfall. CANFIS-3 Gauss-2, CANFIS-3 Gauss-3, CANFIS-4 Gauss-3, CANFIS-10 Gauss-4, CANFIS-3 Gauss-6, CANFIS-6 Gauss-3, CANFIS-3 Gauss-5, CANFIS-12 Gauss-6, CANFIS-12 Gauss-5 and CANFIS-7 Gauss-6 were selected for further analysis and comparison based on the statistical indices, such as root mean squared error (RMSE), coefficient of efficiency (CE), correlation coefficient (r) and coefficient of determination (R^2). The values of statistical indices for the selected CANFIS models during testing are presented in Tables 3 respectively.

Table 3: Statistical Indices for selected CANFIS models during testing phase for CHHOTA UDAIPUR

S. No.	Testing					
	Model	Combination	RMSE	r	CE	R^2
1	N3	(1-2-1)	0.0148	0.9565	0.8656	0.9150
2	N3	(1-3-1)	0.0158	0.9464	0.8458	0.8957
3	N4	(1-3-1)	0.0177	0.9013	0.8068	0.8124
4	N10	(2-4-1)	0.0201	0.9348	0.7510	0.8739
5	N3	(1-6-1)	0.0216	0.9488	0.7131	0.9002
6	N6	(2-3-1)	0.0200	0.8682	0.7534	0.7538
7	N3	(1-5-1)	0.0213	0.8905	0.7208	0.7930
8	N12	(3-6-1)	0.0219	0.8897	0.7047	0.7915
9	N12	(3-5-1)	0.0222	0.8682	0.6971	0.7538
10	N7	(2-6-1)	0.0220	0.8529	0.7018	0.7274

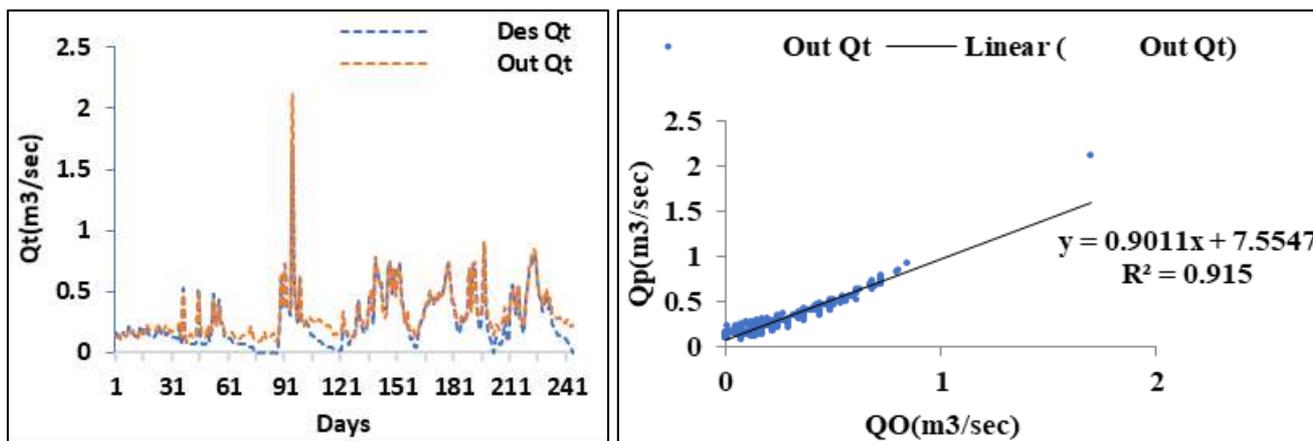


Fig 3: Comparison of observed (Qo) and predicted (Qp) runoff and corresponding scatter plot by CANFIS -3, Gauss-2 during testing period

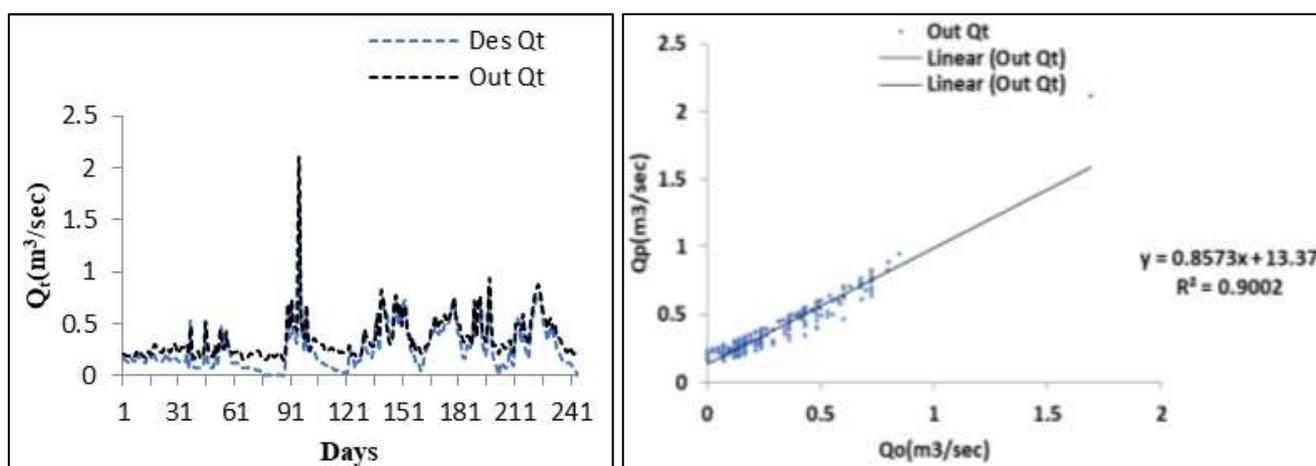


Fig 4: Comparison of observed (Qo) and predicted (Qp) runoff and corresponding scatter plot by CANFIS -3, Gauss-3 during testing period

Runoff modelling using MLR

Universal processes of forecasting runoff amount involve Data collection, data pre-processing and data selection,

Reduction of explanatory predictor, building model using regression and at the last validity check.

Table 4: Statistical indices for selected MLR models during testing phase for CHHOTA UDAIPUR

Model No.	Regression equation	Statistical index			
		RMSE	CE	r	R ²
MLR-15	$Q_t = 213.6775 + (T_{max} * 2.92753) + (T_{min} * -3.73447) + (W_s * 0.991803) + (R_t * 1.204708)$	75.3207	0.1666	0.4082	0.1666
MLR-14	$Q_t = 209.9411 + (T_{min} * -3.39579) + (T_{max} * -2.97665) + (R_t * 1.207769)$	75.3257	0.1665	0.4081	0.1665
MLR-13	$Q_t = 160.0035 + (T_{max} * -3.86126) + (W_s * -1.35477) + (R_t * 1.16748)$	75.4524	0.1637	0.4046	0.1637
MLR-9	$Q_t = 158.0361 + (T_{max} * 3.9231) + (R_t * 1.156284)$	75.4644	0.1634	0.4043	0.1634
MLR-12	$Q_t = 245.1614 + (T_{min} * 9.14803) + (W_s * 2.965122) + (R_t * 1.379027)$	75.9174	0.1534	0.3916	0.1534
MLR-7	$Q_t = 235.2392 + (T_{min} * -8.3815) + (R_t * 1.397541)$	75.9637	0.1523	0.3903	0.1523
MLR-10	$Q_t = 46.36337 + (W_s * -4.85877) + (R_t * 1.4627729)$	77.3632	0.1208	0.3476	0.1208
MLR-4	$Q_t = 32.19349 + (R_t * 1.438534)$	77.5199	0.1172	0.3424	0.1172
MLR-11	$Q_t = 221.8175 + (T_{min} * 0.90881) + (T_{max} * 5.20962) + (W_s * 2.601608)$	78.4818	0.0952	0.3086	0.0952
MLR-1	$Q_t = 264.7419 + (T_{min} * -9.15482)$	30.1469	0.0420	0.2049	0.0420

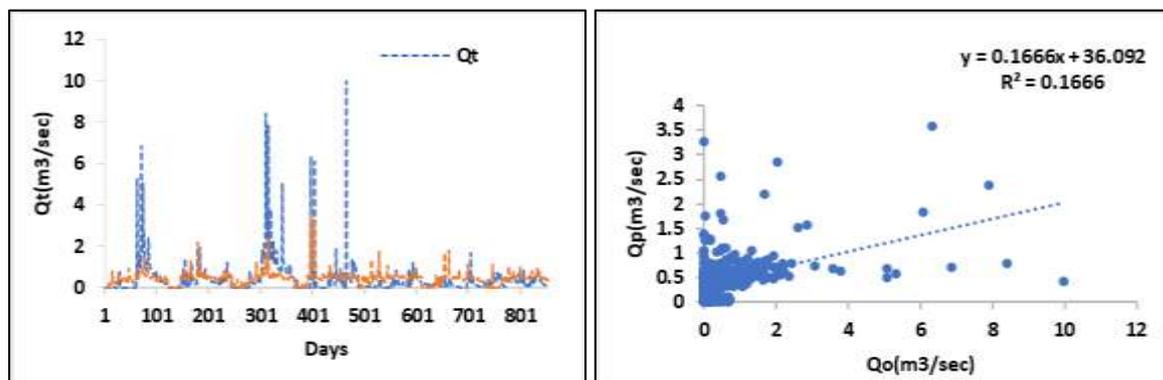


Fig 5: Comparison of observed (Q_o) and predicted (Q_p) runoff and corresponding scatter plot by MLR-15 during the testing period

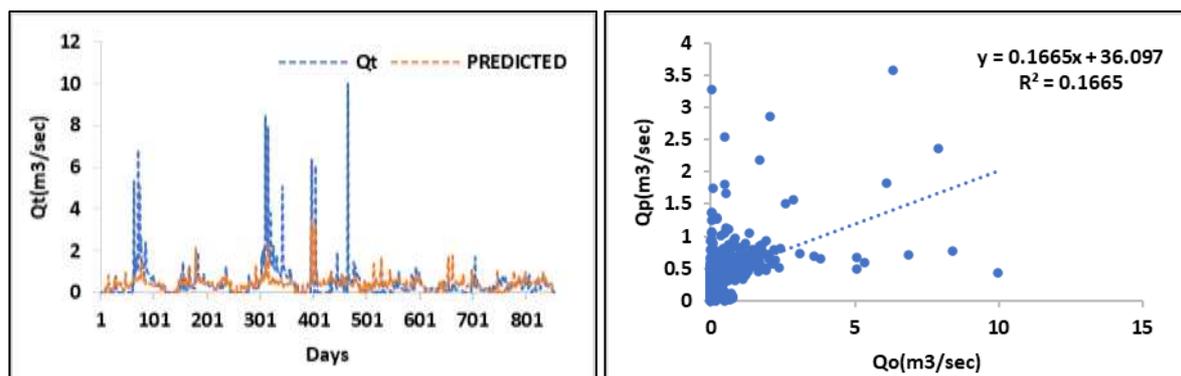


Fig 6: Comparison of observed (Q_o) and predicted (Q_p) runoff and corresponding scatter plot by MLR-14 during the testing period

Conclusion

In this study, we attempted to forecast the daily runoff on the basis of Co-Active Neuro-Fuzzy Inference System (CANFIS) and Multi Linear Regression (MLR) techniques for Narmada basin at Chhota Udaipur District Gujarat. Daily weather data were collected from the site swat.tanu.edu.

Fifteen CANFIS models and MLR were selected based on the performance evaluation indices during testing period. The following conclusions were drawn from the results in this study:

1. On the basis of lower values of RMSE (0.0056) and higher CE (0.5793) and r (0.9224) in the testing phase, the CANFIS 4 model were found to be the best performing model. The results obtained on the basis of statistical indices (RMSE, CE, r and R^2) indicates that the CANFIS model in general gave consistently better performance than the MLR models.
2. The CANFIS model with input of maximum temperature, minimum temperature, wind speed and rainfall was found to be the best for prediction of runoff.
3. It was clearly evident the Multi Linear Regression Method is not suitable for the dataset under this study.

References

- 1 Abbas SA. A Comparative study of Artificial Neural Networks and Adaptive Neuro- Fuzzy Inference System for Forecasting Daily Discharge of a Tigris River, International Journal of Applied Engineering Research ISSN 0973-4562 2017;12:9.
- 2 Ahmed AN, Elashafie A, Karim O, Jaffar O. Evaluation the efficiency of Radial Basis Function Neural Network for Prediction of water quality parameters. Eng. Intelligent Syst 2009;17(4):221-231.
- 3 Alam Md. Ajaz, Singh Vikram. Rainfall simulation using co-active neuro fuzzy inference system (CANFIS) and multi linear regression method (MLR)International Research Journal of Engineering and Technology (IRJET) 2019;06(07):2237-2239.
- 4 Chavan PM, Ukarande SK. Comparison of Rainfall-Runoff Model using ANN and MLR. Water Res. Manage 2012;23(11):2289-2315.
- 5 Deka P, Chandramouli V. Fuzzy neural network model for hydrologic flow routing. J. Hydrol. Eng 2005;10(4):302-14.
- 6 El-Shafie A, Reda TM, Noureldin A. A Neuro-Fuzzy Model for inflow Forecasting of the Nile River at Aswan High Dam. Water Res. Manage 2007;21(3):533-556.
- 7 Lin GF, Chen LH. Application of an artificial neural network to typhoon rainfall forecasting. Hydrol. Process 2005;19:1825-37.
- 8 Luk KC, Ball JE, Sharma A. An application of artificial Neural networks for rainfall forecasting. Math. Compute. Modell 2001;33 683-93.
- 9 Malik Anurag, Kumar Anil, Singh P Rajesh. Application of heuristic approaches for hydrological drought using multi-scalar streamflow drought index 2019;22:4-9.
- 10 Malik A, Kumar A. Co-active Neuro-Fuzzy Inference System (CANFIS) and Multiple Linear Regression (MLR) based Suspended Sediment Modelling. Journal of Indian Water Resources Society 2015, 35.
- 11 Najah A, Elshafie A, Karim OA, Jaffar O. Prediction of johor river water quality parameters using artificial neural networks. Euro. J Sci. Res 2009;28(3):422-435.
- 12 Patel Sharad, Hardaha KM, Seetpal Mukesh K, Madankar KK. Multiple Linear Regression Model for Stream Flow Estimation of Wainganga River, American Journal of Water Science and Engineering 2016;2(1):1-5.
- 13 Rathod Tushar, Singh Vikram, Srivastava SK, Wesley CJ, Alam MA. Rainfall prediction using co-active neuro fuzzy inference system for Umargaon watershed Nagpur India, Journal of Pharmacognosy and Phytochemistry 2018;2018;7(5):658-662.