Geomorphology-wavelet based approach to rainfall runoff modeling for data scarce semiarid regions, Kolar River catchment, India

Deepak Kumar Tiwari*, H. L. Tiwari and Raman Nateriya

Dept. of Civil Engineering, Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India
Corresponding Author: deepaktiwaricivil7sept@gmail.com

ABSTRACT

The conceptual and physical mathematical model of rainfall-runoff modeling uses various parameters such as land use land cover, soil type classification, rainfall, atmospheric data such as temperature, evapotranspiration, solar radiation, and wind speed. But these data may not be available for developing countries and data scares semiarid watershed. Also, the problem is even more critical for ungauged catchments and where manual record is maintained of water level and rainfall data. To address this issue, trend analysis is performed using Mann-Kendall test and Sen's slope test, which shows significant trend change stressing the need for new method for runoff prediction for better water resource management. In this study, a total of four models were used, namely, nonlinear autoregressive model with exogenous inputs lumped (LNARX), nonlinear autoregressive model with exogenous geomorphometrically processed inputs (GNARX), wavelet nonlinear autoregressive model with exogenous inputs (WLNARX), and nonlinear autoregressive model with exogenous geomorphometrically processed inputs (WGNARX). Ten models with different input combinations were selected based on their performance and analyzed for all the four networks. The best performing model for these networks is model no. 6 with WGNARX network with NSE 0.97 and RMSE 0.97 and with least value of RMSE. This method can be applied to data scarce region, where data available are available for shorter duration and helpful for ungauged catchments also.

Keywords: Rainfall-runoff modeling; Wavelet-transform; Geomorphology; Trend-analysis; Climate change.

1. INTRODUCTION

Precipitation is a complex phenomenon, but with advanced forecasting tool, it can be predictable (Amiri *et al.* 2018). An efficient forecasting model for rainfall-runoff can address the issue and provide a guidance and inefficient use of water leads to considering use of wastewater (Aleisa & Alshayji, 2019) and design of sedimentation basin (Khudair & Abdulhasen, 2020). The main classification of these models in hydrology is in three subdomains of empirical or mathematical, conceptual, and geomorphological. Artificial Neural Network based models (Badrzadeh *et al.* 2015) are the numerical or empirical models. Geomorphology based models associate the hydrological responses of a watershed to the morphometric parameters of catchment (Saharia and Bhattacharjya, 2012; Zhang and Govindaraju, 2003). Kolar River is one of the main tributaries of Narmada River basin and not much work on it in the field of runoff modeling has been done for the catchment. During last two decades, there is increase in the

number of researches towards the use of ANN as a tool for rainfall-runoff modeling (Modarres and Ouarda, 2013; Nourani and Komasi, 2013; Nourani, Alami and Vousoughi, 2015). ANN identifies the complex nonlinear relationship between rainfall and runoff. But these models cannot be relied on in a blindfolded way and require better data processing techniques like wavelet decomposition. Most of the models are of static nature and do not fit into the analysis involving temporal variability. This situation can be addressed by using dynamic neural networks, which can change its structure dynamically as per the outcome. Almost all of the models except a few had not considered morphometric parameters as input parameters for the analysis. In this paper, firstly, trend analysis was performed using Mann-Kendall test for ascertaining the climate change impact on the input parameters to enforce the necessity of rainfall-runoff modeling. The main objective of the present work is to develop a ANN model for rainfall-runoff prediction combining morphometric characteristics, wavelet decomposition data processing technique, and a dynamic NARX neural network. These four networks, namely, nonlinear autoregressive model with exogenous inputs lumped (LNARX), nonlinear autoregressive model with exogenous geomorphometrically processed inputs (GNARX), wavelet nonlinear autoregressive model with exogenous inputs (WLNARX), and nonlinear autoregressive model with exogenous geomorphometrically processed inputs (WGNARX), have been compared based on performance evaluation criteria (NSE, RMSE, and R^2) generally used for semidistributed and lumped models.

2. STUDY AREA AND DATA COLLECTİON

The study area is Kolar River catchment of Sehore district in of Madhya Pradesh. The salient features for the catchment are given in Table 1, and Figure 1 presents the Kolar catchment geographical map. Data of average monthly discharge in cubic meter per second was available from October 1999 to October 2018, for nineteen years and three months, and the flow of data has been provided by Kolar Dam Authority, a state government body associated with the management of dam.
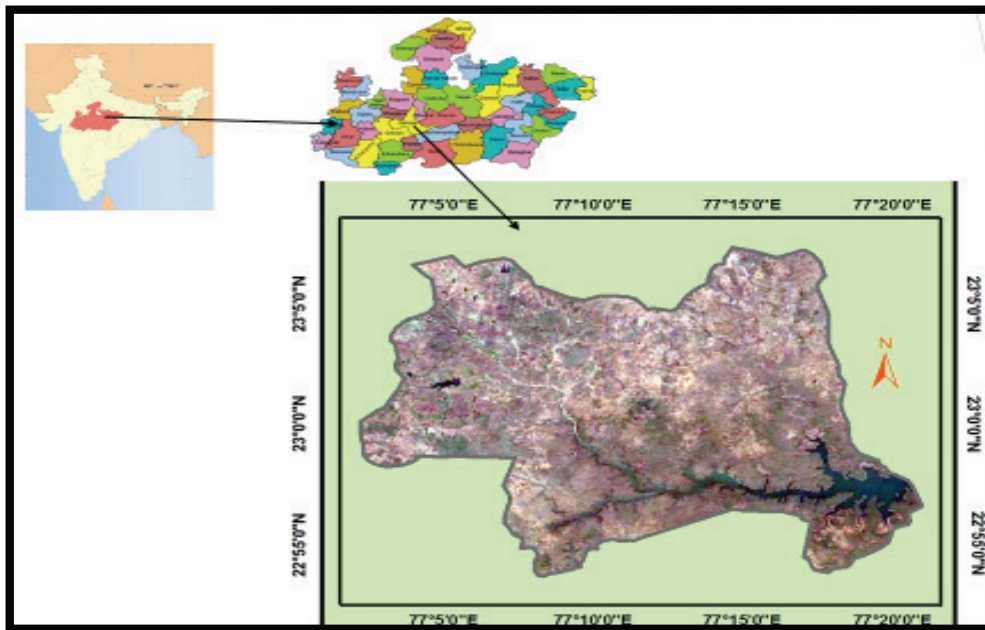

Figure 1. Kolar River watershed Map.

Table 1. Characteristics of the three weather stations.

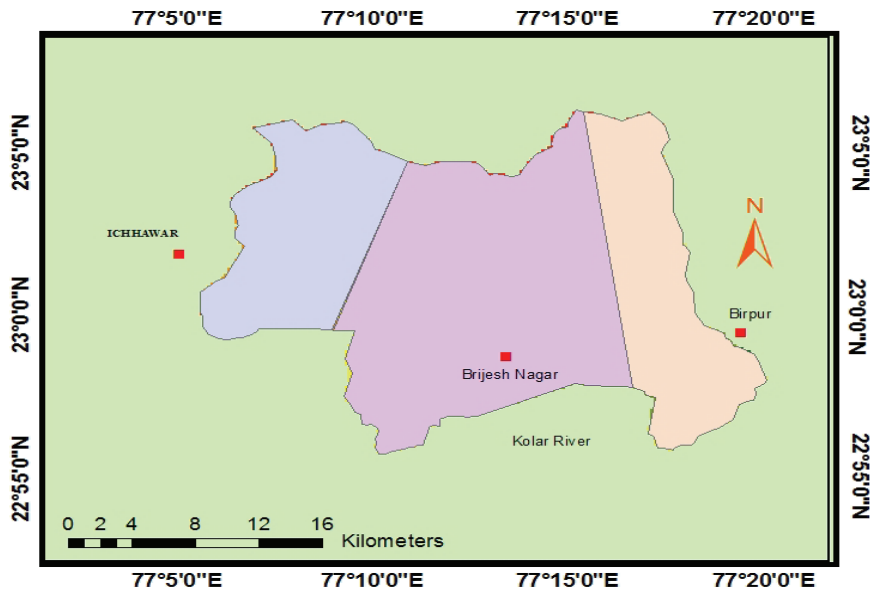| Station	Latitude	Longitude	Elevation	Avg Tmax	Avg Tmin	Rainfall	Climate
Birpur	22^0 58'	$77^0$20'	441	31.88	18.76	940	Semi-arid
Brijesh Nagar	22^0 57'	$77^0$08'	505	30.66	16.45	1370	Humid
Ichhawar	23^0 01'	77^0 01'	515	32.24	17.8	1032	Dry - Sub humid

Figure 2. Rain Gauge Stations at Kolar River watershed.

3. METHODOLOGY

Firstly, trend analysis is done using nonparametric test of Mann-Kendall (M-K) test, and the slope of trend is obtained by Sen's method for all three weather stations, monthly rainfall and temperature of the region. In the second step, rainfall and temperature are used as an input for the neural network for the prediction, and normalization of the input is calculated. In the third step, geomorphological attributes of subwatershed are calculated. Moreover, wavelet decomposition of the inputs after morphometric treatment is carried as data preprocessing techniques. In the fourth step, various models of proposed NARX model are evaluated based on the data reprocessing techniques coupled with and morphology-based parameters. The evaluation criteria are Root Mean Square Error (RMSE), Coefficient of Correlation (R), Nash-Sutcliffe Efficiency (E), and Average Absolute Relative Error (AARE).

3.2 Normalization

Normalization of input and output data safeguards enhanced computation efficiency, data integrity, and redundancy elimination (Nawi *et al.* 2013). Equation 1 give formula for calculating normalized values:

$X_n = 0.1 + 0.8 \frac{X_i - X_{min}}{X_{max} - X_{min}}$ (1), where xn is the normalized value of data set, Xi is the actual value of the data set, and Xmin and Xmax are the minimum and maximum values of the data set, respectively.

3.3 Morphometric Parameters Analysis

Integration of morphometric parameters to neural network architecture is attained through applying weighted factor depending on various subwatershed specific parameters. Watershed is divided into three subwatersheds, Birpur, Brijesh Nagar, and Ichhawar, and morphometric parameters as given in Table 2 are determined. Here, seven attributes such as area of the watershed along with other parameters are calculated.

Weighted Factor

The response of drainage basin to runoff is directly proportional to relief ratio (Rr), drainage length (Dl), and inversely proportional to elongation ratio (Rl). A multiplying factor considering these parameters is obtained and given by Equations 2 and 3.

$A_i = \frac{B_i}{\sum_i^l B_i}$ (2), where l is the number of subcatchments;

$B_i = \frac{L_i R_{ri}}{R_{li}}$ (3), where Li, R_{ri} and R_{li} are length of watershed, relief ratio, and elongation ratio.

Table 2. Morphometric attributes of subwatersheds.

S No.	Attributes	Birpur	Birjish nagar	Ichhawar
1	Area	121,05	278,48	122
2	Drainage Length	406,47	922,63	413,87
3	Length of watershed	25,2	27,65	12,6
4	High Elevation	490	482,92	537,03
5	Low elevation	390	390	430,023
6	Relef ratio	3,97	3,36	8,49
7	Elongation ratio	0,493	0,681	0,989
8	Multiplying ratio	0,29	0,4	0,31

3.4. NARX ANN Model

The nonlinear autoregressive model with exogenous inputs (NARX) belongs to the class of RNNs, and they are well suited for predicting nonlinear relationship of hydrological parameters. It has several feedback connections enfolding numerous layers of networks. The defining equation (4) for the NARX model is

$$Y = f(y(t-1), y(t-2), y(t-3), \dots, y(t-n_y),\ u(t-1), u(t-2), \dots, u(t-n_u)) \tag{4}$$

where the next value of the dependent output signal y(t) is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal. A diagram of the resulting network is shown below, where a two-layer feedforward network is used for the approximation. As shown in Figure 3, the resultant of output layer 2 is used as feedback to input in hidden layer 1.
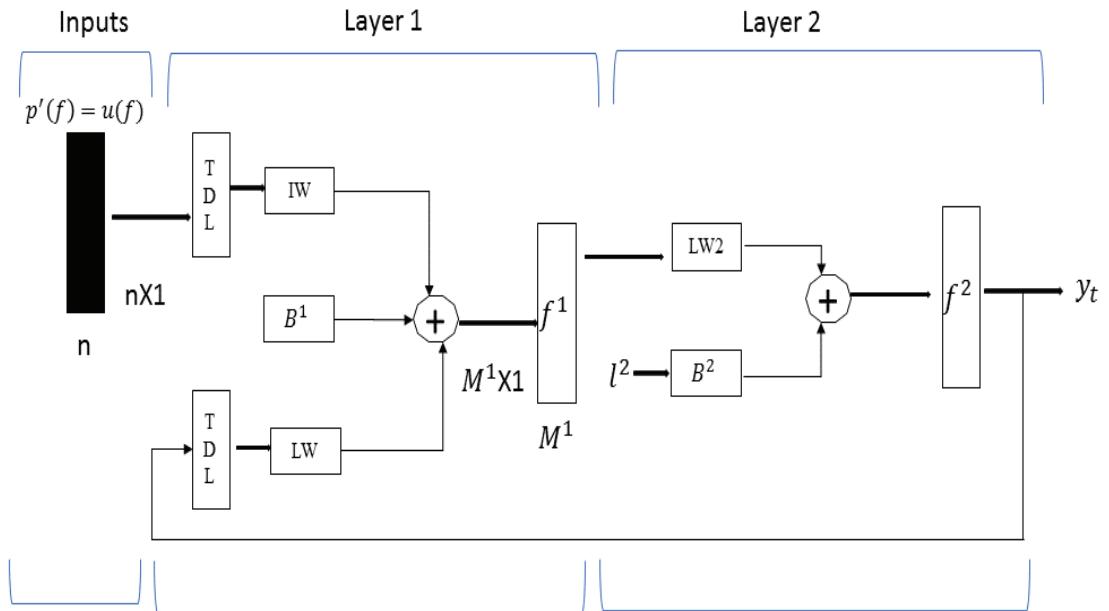

Figure 3. A two-layer feedforward NARX model network.

3.5 Wavelet Decomposition

Wavelet decomposition is an advanced tool in data processing technique. Its use has increased nowadays in data handling, signal processing, image processing, and optical engineering *(*Seo *et al.*, 2015). The wavelet transform of a time series f(t) is defined as

$$f(x, y) = \frac{1}{\sqrt{x}} \int_{-\infty}^{\infty} f(t) t \Psi\left(\frac{t-y}{x}\right) dt \tag{5}$$

where $\Psi(t)$ is the basic wavelet having time series length t. The variables are x and y, where x is the scale or dilation factor that determines the characteristic frequency, so that its variation gives rise to a 'spectrum'; and y is the translation in time, so that its variation represents the 'sliding' of the wavelet over f(t) (Krishna et al. 2011).

3.5 A New Modeling Approach

Various networks with different permutation and combination of input parameters and differing in number of hidden layers are selected, and the optimum design is chosen based on optimum evaluation criteria. The detailed methodology is stepwise given in Figure 4.
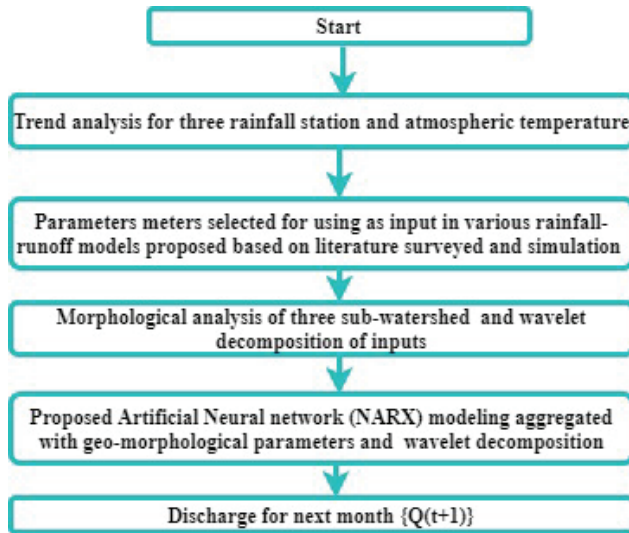

Figure 4. Methodology adopted for rainfall runoff modeling.

In this study, a total of four models were used, namely, nonlinear autoregressive model with exogenous inputs lumped (LNARX), nonlinear autoregressive model with exogenous geomorphometrically processed inputs (GNARX), wavelet nonlinear autoregressive model with exogenous inputs (WLNARX), and nonlinear autoregressive model with exogenous geomorphometrically processed inputs (WGNARX). All the four models are trained using Levenberg–Marquardt (LM) algorithm used for training of neural networks. All the models used the input combinations are given in Table 3.

Table 3. Input structure of neural network for rainfall runoff modeling.

Model No.	Input combination for prediction Q(t)	Model No.	Input combination for prediction Q(t)
1	Q(t-1), R(t-1)	6	Q(t-1), Q(t-2), Q(t-3), R(t-1)
2	Q(t-1), Q(t-2)	7	Q(t-1), R(t-1), T(t-1)
3	Q(t-1), Q(t-2), Q(t-3)	8	Q(t-1), Q(t-2), R(t-1), T(t-1)
4	Q(t-1), Q(t-2), Q(t-3), Q(t-4)	9	Q(t-1), Q(t-2), Q(t-3) R(t-1), T(t-1)
5	Q(t-1), Q(t-2), R(t-1)	10	Q(t-1), Q(t-2), Q(t-3) R(t-1), R(t-2) T(t-1)

Here, Q(t) is the predicted discharge for the current month, R is the rainfall, T is the temperature, and t is the time period in months. These ten models are used for analysis.

3.6 Evaluation Criteria

Nash Sutcliff Efficiency (NSE) and Root mean square error (RMSE) are efficiency parameters for validation and testing of model. Coefficient of determination (R^2) measures the dispersion between observed value and simulated value from Model.

4. RESULT AND DISCUSSION

4.1 Mann-Kendall (M-K) and Sen's Slope Analysis

The results for the M-K test for the three rainfall stations and average temperature are shown in Figure 5. In Brijesh Nagar, the significant negative trend is observed, and it is one of the reasons for the rainfall-runoff modeling. Ichhawar rainfall and temperature do not have significant trends. Figure 5 shows the annual trends for three stations and temperature, which shows the trends and Sen's slope.
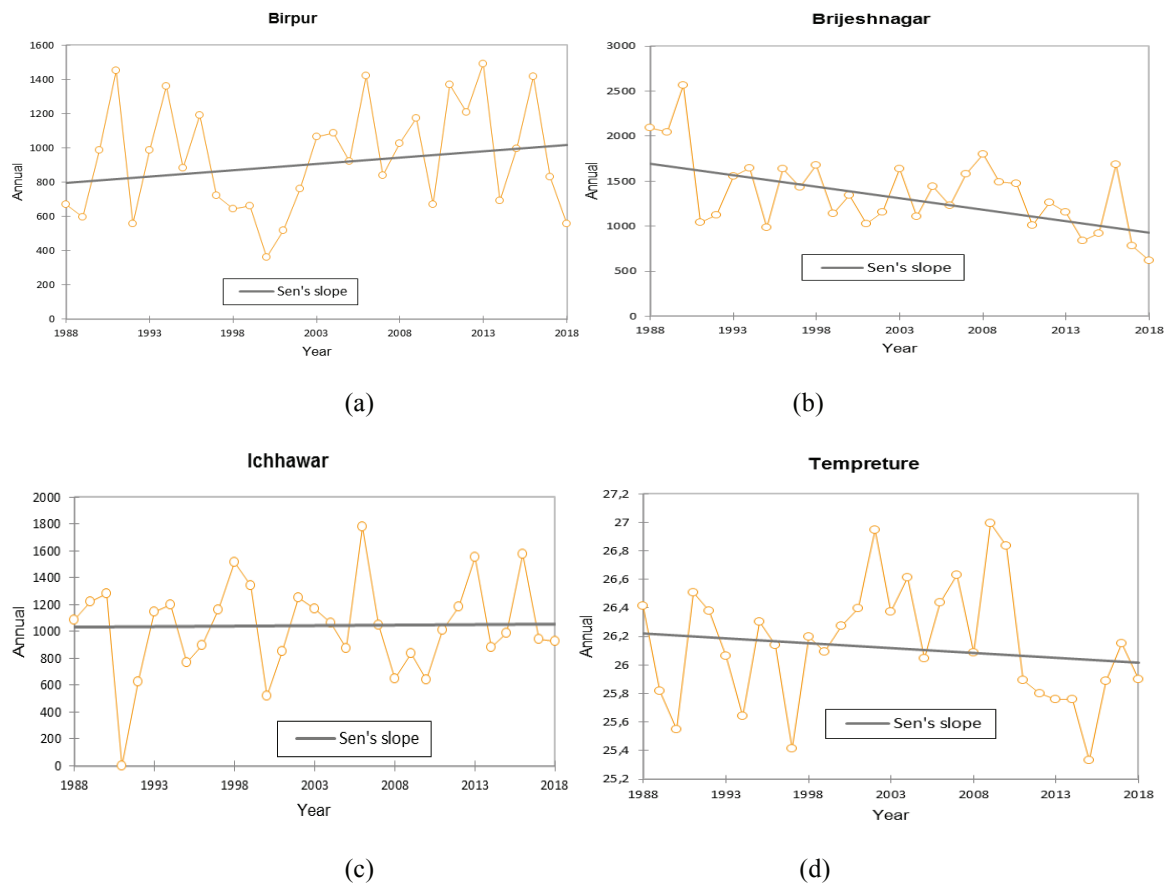

Figure 5. Trend analysis with Sen's slope graph of annual data for
(a) Birpur rainfall. (b) Brijesh Nagar. (c) Ichhwara. (d) Temperature.

4.2 ANN Models Performance Evaluation

ANN models LNARX, GNARX, WLNARX, and WGNARX were evaluated based on hydrology data for rainfall, runoff, and temperature at Kolar dam catchment and compared and evaluated based on the criteria given in Table 4 and Table 5. According to Table 4, Model no. 4 gives the best result for LNARX model with NSE 0.68 and R^2 value of 0.88. The best model performed for this network is Q(t-1), Q(t-2), Q(t-3), and Q(t-4). Also, the performance of models 1, 2, 5, 6, 8, and 9 is also satisfactory with higher NSE (>0.50) and low RMSE value. For GNARX, model no. 6 is the best performing model with NSE value 0.80 and RMSE value of 0.91. For this network, the best input combination is {Q(t-1), Q(t-2), Q(t-3), R(t-1)}. Other models also performed nicely except 7 and 8.

Table 4. Results for runoff modeling for LNARX and GNARX models.

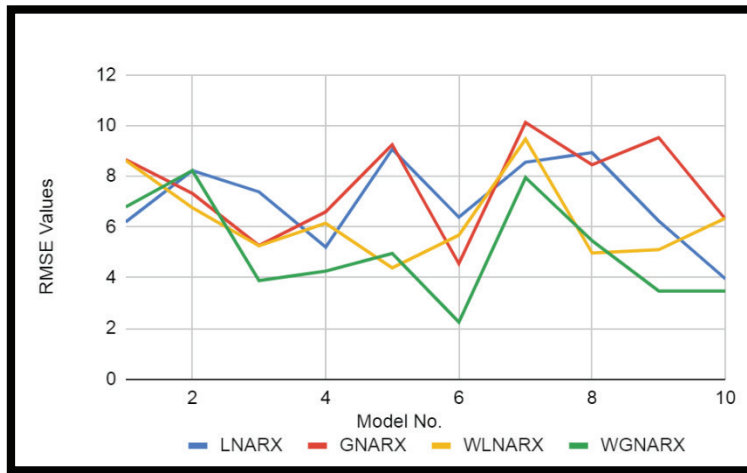| Model No	LNARX			GNARX		
	NSE	RMSE	R^2	NSE	RMSE	R^2
1	0.58	6.2	0.59	0.61	8.66	0.64
2	0.47	8.23	0.49	0.66	7.33	0.68
3	0.51	7.39	0.52	0.71	5.27	0.73
4	**0.68**	**5.21**	**0.88**	0.79	6.6	0.81
5	0.63	9.07	0.65	0.58	9.25	0.64
6	0.61	6.39	0.65	**0.8**	**4.56**	**0.91**
7	0.38	8.56	0.41	0.38	10.13	0.48
8	0.61	8.94	0.62	0.49	8.46	0.66
9	0.64	6.24	0.64	0.75	9.53	0.82
10	0.36	3.96	0.39	0.74	6.32	0.83

Table 5 gives the results for the models which uses wavelet transform of the input parameters to the LNARX and GNARX networks. In this Table, also, it is clearly observed that wavelet-GNARX model better performed than the WLNAX model. The best performing model for these networks is model no. 6 with NSE 0.97 and RMSE 0.97 and with least value of RMSE. Most of the models have performed better than the same models without wavelet transformed inputs.

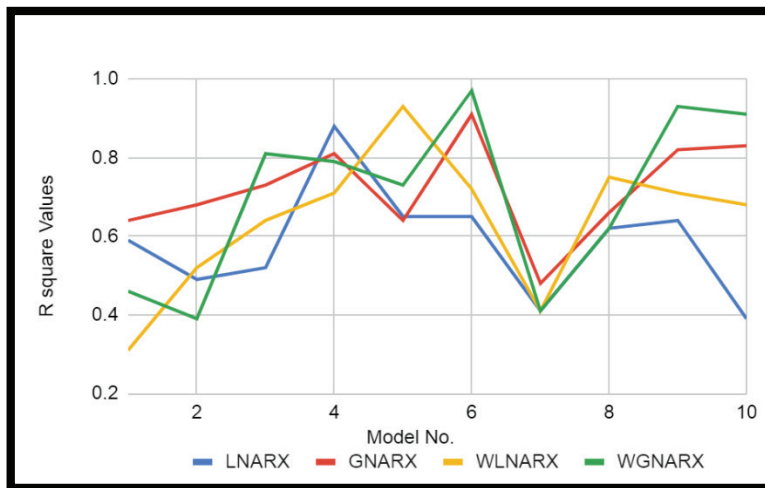**Table 5.** Results for runoff modeling for WLNARX and GNARX models.

| Model No.	WLNARX			WGNARX		
	NSE	**RMSE**	**R²**	**NSE**	**RMSE**	**R²**
1	0.38	8.63	0.31	0.44	6.8	0.46
2	0.51	6.76	0.52	0.37	8.23	0.39
3	0.62	5.26	0.64	0.81	3.89	0.81
4	0.68	6.15	0.71	0.78	4.26	0.79
5	**0.8**	**4.39**	**0.93**	0.71	4.96	0.73
6	0.72	5.68	0.72	**0.97**	**2.25**	**0.97**
7	0.41	9.47	0.41	0.39	7.95	0.41
8	0.76	4.98	0.75	0.61	5.46	0.62
9	0.72	5.11	0.71	0.91	3.48	0.93
10	0.7	6.35	0.68	0.9	3.48	0.91
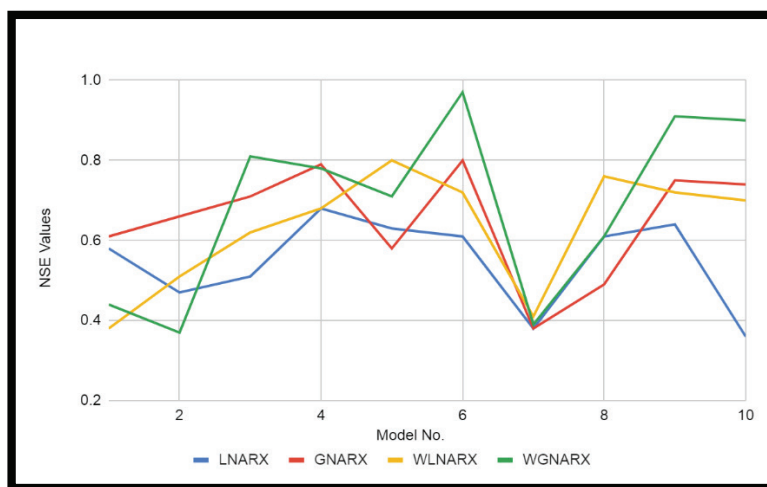
(a)

(b)

(c)

Figure 6. (a) NSE values plot. (b) RMSE values plot.
(c) R^2 values plot of four models LNARX, GNARX, WLNARX, and WGNARX

Figure 6 has three graphs comparing the three evaluation criteria, namely, NSE, RMSE, and R^2. Among all the ten model combinations of four models LNARX, GNARX, WLNARX, and WGNARX. Model no. 6 of WGNARX model outperformed the other models and input combinations. Figure 7 gives the runoff prediction for one month ahead for Kolar River catchment. Scatter plot shows the accuracy of regression analysis of observed and predicted data. Time series plot gives the details about temporal variation during the months.
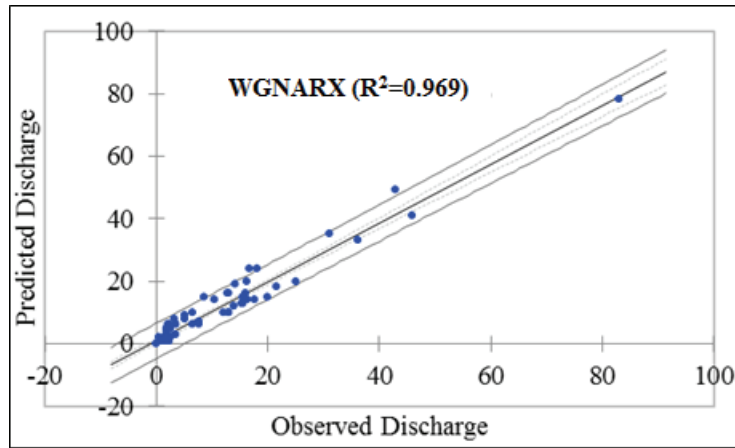
(a)
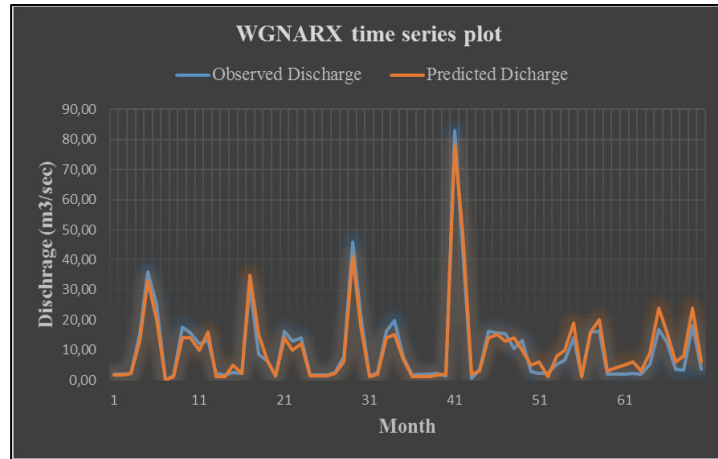
(b)

Figure 7. The runoff prediction for one month ahead for the best performing network model:
(a) scatter plot WGNARX; (b) time series plot WGNARX.

5. CONCLUSIONS AND RECOMMENDATIONS

The results for the M-K test suggest that there is a need for management of water resources judiciously and new effective method for hydrological modeling. The simple yet innovative technique adopted in this study uses morphological parameters as input factors along with rainfall data and gives better predictability as compared to the same network where these parameters were not considered. Furthermore, the efficiency of the models further increased if the input is wavelet transform. The best performing model for these networks is model no. 6, having input combination of Q(t-1), Q(t-2), Q(t-3), R(t-1) with NSE 0.97 and RMSE 0.97 and with least value of RMSE. This method can be applied to data scarce region, where data are available for shorter duration and are manually recorded, which invites human errors. This model can also be recommended to be used in ungauged catchments to understand the model performance and evaluating its efficiency.

REFERENCES

Aleisa, E. & Alshayji, K. (2019) Analysis on Reclamation and Reuse of Wastewater in Kuwait. Journal of Engg. Research, 7(1), pp. 1–13.

Amiri, M. A., Conoscenti, C. & Mesgari, M. S., (2018). Improving the accuracy of rainfall prediction using a regionalization approach and neural networks. Vol-45(4), 66–75.

Badrzadeh, H., Sarukkalige, R. and Jayawardena, A. W., (2015). Hourly runoff forecasting for flood risk management: Application of various computational intelligence models', Journal of Hydrology. Vol-529, 1633–1643.

Khudair, K. M. & Abdulhasen, K. N. (2020) Design criteria for pre sedimentation basin treats: Shatt Al-Arab River water. Journal of Engg. Research, 8(3), pp. 31–49.

Krishna, B., Rao, Y. R. S. and Nayak, P. C., (2011). Time Series Modeling of River Flow Using Wavelet Neural Networks. Journal of Water Resource and Protection. 03(01), 50–59.

Mann, H. B., (1945). Nonparametric Tests Against Trend. Econometrica, 13(3), 245–259.

Modarres, R. & Ouarda, T. B. M. J. (2013) Modeling rainfall-runoff relationship using multivariate GARCH model', Journal of Hydrology. Elsevier B.V., 499, 1–18.

Kendall, M. G., (1975). Rank correlation methods. Oxford University Press.

Nawi, N. M., Atomi, W. H. & Rehman, M. Z. (2013). The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. Procedia Technology, 11, 32–39.

Nourani, V., Alami, M. T. & Vousoughi, F. D. (2015) Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling. Journal of Hydrology. Elsevier B.V., 524, 255–269.

Nourani, V. & Komasi, M. (2013) A geomorphology-based ANFIS model for multi-station modeling of rainfall-runoff process. Journal of Hydrology. Elsevier B.V., 490, 41–55.

Saharia, M. & Bhattacharjya, R. K. (2012) 'Geomorphology-based Time-Lagged Recurrent Neural Networks for runoff forecasting', KSCE Journal of Civil Engineering, 16(5), 862–869.

Seo, Y., Kim, S. & Singh, V. P., (2015). Multistep-ahead flood forecasting using wavelet and data-driven methods. KSCE Journal of Civil Engineering. 19(2), 401–417.

Zhang, B. & Govindaraju, R. S., (2003). Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. 273, 18–34.