International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Modelling of Daily Streamflows using Linear Regression and Artificial Neural Network Models in Krishna Basin

Jyothi Pandla[1] , Shanmukha Srinivas Gorantla[2*], S. Vijaya Bhaskar[3], M. Guruprasad[4], N. Rushwanth Chowdary[5]

[1]Assistant Executive Engineer, Water Resources Department, Govt. of A.P., Tirupati, Andhra Pradesh, India.

[2,3,4,5] Department of Civil Engineering, Sri Venkateswara University College of Engineering, Tirupati, Andhra Pradesh, India.

ABSTRACT. Estimation of stream flow is a primary requirement in water resources development and planning. The use of conceptual methods in preference to the simpler regression and time series methods has not been found to be justifiable in terms of improved forecasting performance. Moreover, simpler regression and time series methods are easier to use and are the most commonly recommended for practical applications. Correlation and regression analysis is one of the commonly used statistical techniques in hydrology. Most of the hydrologic processes exhibit a high degree of temporal and spatial variability and are further plagued by issues of non-linearity of physical processes and uncertainty in parameter estimates. As Artificial Neural Networks (ANNs) are able to extract the relation between the inputs and outputs of a process, even if the data are noisy and contaminated with errors, they are well-suited for the estimation and prediction of hydrologic parameters. In the present investigation, an attempt has been made to correlate the stream flows at a gauging site with that at upstream gauging sites of Krishna basin. Performance of regression and ANN models is compared based on performance evaluation criteria. The gauging sites of K.Agraharam, Huvinhedgi, Kurundwad, Sadalga, Terwad, Arjunwad, Samdoli, Karad, Warunji and Koyna on Krishna basin have been selected for the present study. It has been observed from the correlation analysis that the daily stream flows at Karad gauging site depend on those of Warunji; Arjunwad streamflows on those of Samdoli. Streamflows at Kurundwad depend on streamflows at Arjunwad, Samdoli and Terwad gauging sites. Streamflows at K.Agraharam gauging site depend on the stream flows at the gauging site of Huvinhedgi in the month of June, July, and August, respectively. Based on the dependence of stream flows, linear regression models relating daily stream flow at a gauging site with that of gauging sites upstream of it have been developed. The relations so developed may be adopted for prediction of daily stream flows with a reasonable degree of accuracy at the gauging sites of Krishna basin. ANN models with optimal architectures employing Log-Sigmoid function have also been developed to model daily stream flows at the gauging sites taking the stream flows of upstream gauging sites as input.

1. INTRODUCTION

Water is the principal constituent of all living things on the earth. Hydrology, which treats all phases of the earth's water, is a subject of great importance for people and their environment. Analysis of water resources system often requires long and continuous streamflow records to assess the possible extreme hydrology conditions. Many of the systems are large in spatial extent and have a hydrometric data collection network that is very sparse. It may not be possible to assess the reliability of water resource system if the available hydrologic data do not cover high, normal and low flow conditions. Estimation of streamflows is a primary requirement in water resources development and planning. Importance of accurate estimation of flows has been well recognized in planning of irrigation, hydropower, water supply, pollution control etc. Methods available for estimation of streamflows can be broadly classified as regression methods, time-series methods and conceptual methods. However, use of conceptual methods in preference to the simpler regression and time series methods has not been found to be justifiable in terms of improved forecasting performance. Moreover, simpler regression and ANN methods are easier to use and are the most commonly recommended methods for practical applications. Theoretical conceptual or black-box models based on the importance of the water resources systems are generally adopted for the simulation and forecasting of streamflows. The black box and time series models are more practical despite their inherent inability to understand the internal structure of the process. Streamflow models are generally developed using linear regression techniques based on the cross-correlation analysis. Artificial Neural Networks (ANNs) which are capable of representing the complex and non-linear process effectively are used in recent times as a successful soft computing tool in the Streamflow modeling. Although ANNs belong to the class of data driven approach, it is important to determine the dominant network model inputs, as this not only reduces the training time but also increases the generalization ability of the network for a given data set.

2. DOMAIN AREA AND REVIEW OF LITERATURE

2.1 STUDY AREA

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Krishna River rises in the Western Ghats at an altitude of 1337 m North of Mahabaleeshwar, about 64 km from the Arabian Sea and flows from West to East through the States of Maharashtra, Karnataka and Andhra Pradesh and finally joins the Bay of Bengal. Total length of the river from the source to its outfall in the sea is about 1,400 km of which 612 km is in Andhra Pradesh, 305 km in Maharashtra and 483 km in Karnataka its basin extends over an area of 2,58,948 sq.km out of which 27.01% falls in Maharashtra, 44.06% in Karnataka and 28.93% in Andhra Pradesh. The basin lies between East Longitudes 73° 21' to 81° 09' and North latitudes 13° 07' to 19° 25' in the Deccan plateau. In the present investigation, an attempt is made to correlate the daily streamflow at a gauging site with that at upstream gauging sites of Krishna basin. Present study focuses on the development of simulation and forecasting models of daily streamflows at K.Agraharam, Huvinhedgi, Kurundwad, Sadalga, Terwad, Arjunwad, Samdoli, Karad, Warunji and Koyna sites of Krishna basin. These are shown in Fig 2.1. A brief description of the gauging sites is presented in Table 2.1. The streamflows at the gauging sites are modeled using Linear regression and Artificial Neural Networks. Performance of models is compared based on performance evaluation criteria.

Table 2.1: Details of Gauging Sites

Gauging Site	Latitude/ Longitude	Catchment Area (km²)	Period of Data		Mean Seasonal Streamflow (m³/s)	
			Training period	Testing period	Training period	Testing period
K.Agraharam	16°15'N/77°52'E	132920	1981-1999	2000-2006	2096.59	1546.79
Huvinhedgi	16°27'N/76°51'E	55150	1981-1999	2000-2006	1424.90	699.82
Kurundwad	16°36'N/75°12'E	15190	1972-1998	1999-2010	937.36	1021.21
Sadalga	16°34'N/74°31'E	2322	1969-1998	1999-2010	241.07	194.30
Terwad	16°44'N/74°35'E	2425	1980-2001	2002-2010	372.72	410.50
Arjunwad	16°33'N/74°38'E	12660	1969-2010	1999-2010	673.87	627.09
Samdoli	16°51'N/74°30'E	1948	1967-1997	1998-2010	279.36	240.76
Karad	17°18'N/74°18'E	5462	1965-1996	1997-2010	365.46	372.99
Warunji	17°16'N/74°10'E	1890	1967-2010	1998-2010	222.68	237.15
Koyna	17°23'N/73°44'E	920	1973-1996	1997-2006	73.38	91.28

Fig. 2.1: Location Map of Study Area

2.2 REVIEW OF LITERATURE

Vedula and Reddy (1981) compared the applicability of different streamflow generating models using the historic data of monthly streamflows into Hemavathy and Krishnasagar reservoirs in upper Cauvery River Basin and concluded that the Thomas-Fiering modelwas found to best preserve the mean, standard deviation and lag-one correlation of historic streamflows. Nagesh Kumar *et al.* (2004) demonstrated the use of ANNs to forecast monthly river flows using feed-forward and recurring neural networks the study indicated that the recurrent neural networks perform better than the feed forward networks. Jain and Sudheer (2008) demonstrated that Nash-Sutcliffe efficiency index alone is not adequate in describing the performance of a hydrologic model. It was shown that relatively poor models can give a high value of index and vice-versa suggested that other statistical tools such as a scatter plot which may reveal important information about the ability of the model to reproduce the dependent variable in different ranges need to be employed to arrive at a definite conclusion about the model performance. Nanda *et al.* (2013) illustrated the use of ANN for modelling the extreme rainfall events of Banswara and Visakhapatnam. The series of annual 1-day maximum rainfall (referred as extreme rainfall) is used in the study. Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) networks are applied to train the network data. Model performance indicators such as correlation coefficient, model efficiency and root mean square error

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

are used to evaluate the performance of the MLP and RBF networks. The paper presents the MLP network is better suited for modelling the extreme rainfall events recorded at Banswara and Visakhapatnam stations.

3. METHODOLOGY

Hydrologic phenomena are highly erratic and are amenable to statistical interpretation. Streamflow records of very long durations are required in order to assess the possible extreme hydrologic conditions. In case, where long sequences of historic data are not available, simulated hydrological sequences are generated and found to be very useful. These generated sequences are very close to the historic sequences in terms of their statistical properties. The present study deals with a brief review of multiple correlation, regression analysis and artificial neural networks.

3.1 REGRESSION MODELS

Objective of simple regression methods is the transfer of information between several variables observed simultaneously and the prediction of a dependent variable from several other observed independent variables. Linear and non-linear regression model structures are generally developed to predict the dependent variable.

Multiple Correlation and Regression Analysis: The association of three or more variables is investigated by multiple regression and correlation analysis. General multiple regression relation is expressed in the form

$$X_1 = f(X_2, X_3, X_4\text{------------},X_m) \quad (1)$$

in which X_1,X_2,X_3------------,X_m are m variables. Since linear equations are easier to treat than nonlinear multiple relations, the nonlinear relations in hydrology are often transformed to linear for the multiple regression analysis.

Linear Regression of Several Variables: If there are m variables to correlate, including one dependent and (m-1) independent, the general equation for the multiple linear regression is

$$X_1=b_1+b_2X_2+ \text{------------------}+ b_iX_i + \text{-------}+ b_mX_{m-1} \quad (2)$$

in which b_1 is the intercept, and b_i (i = 1 to m) is the multiple regression coefficient of the dependent variable X_1 on the independent variable X_i, with all other variables kept constant.

Applying least squares method for the sum of the residuals, $e_1 = X_1-b_1-b_2X_2$ -----b_mX_{m-1}, the m partial differential equations in b_1, b_2, ------- and b_m give m linear equations. The solution of these equations facilities the determination of m parameters.

3.2 STREAMFLOW MODELLING

The daily simulation and forecasting models during monsoon period at a gauging site are developed using the streamflow data of the upstream gauging sites. The steps in the modeling include i) identification of influencing gauging sites, ii) development of regression model, iii) performance evaluation of the models and iv) comparison of statistical parameters. The streamflows at the gauging sites of K.Agraharam, Huvinhedgi, Kurundwad, Sadalga, Terwad, Arjunwad, Samdoli, Karad and Warunji, Koyna of Krishna basin are modeled using the correlation and regression analysis.

Simulation Models: The influencing gauging sites are identified through cross correlation analysis. The correlation coefficients between the daily streamflows of monsoon months of the upstream gauging sites and those at the gauging sites on the downstream of Krishna basin. The daily streamflows at Karad and Arjunwad gauging sites respectively depend on those of Warunji and Samdoli gauging sites during monsoon period. The streamflows at Kurundwad gauging site are highly correlated with those of Terwad, Arjunwad and Samdoli in the month of June and also with Sadalga streamflows in the month of July and Sadalga, Arjunwad and Samdoli in the month of August. The daily streamflows at K.Agraharam depend on those of Huvinhedgi in the months of July and August while Sadalga streamflows depend on those of Terwad in the month of June and Warunji streamflows on those of Koyna in the month of August.

Table 3.1 Linear Regression Models (Simulation Mode)

Month: June

Gauging Site	Regression Model
Karad	$Q_t=3.77+1.65Q_{t(Warunji)}$
Arjunwad	$Q_t=7.63+2.002Q_{t(Samdoli)}$
Kurundwad	$Q_t = -0.42+1.02Q_{t(Terwad)}+0.67Q_{t(Arjunwad)}+0.38Q_{t(Samdoli)}$

Month: July

Gauging Site	Regression Model
Karad	$Q_t =5.49+1.64Q_{t(Warunji)}$
Arjunwad	$Q_t = -41.25+2.30Q_{t(Samdoli)}$
Kurundwad	$Q_t= 41.18+0.13Q_{t(Sadalga)}+1.35Q_{t(Terwad)}+0.56Q_{t(Arjunwad)}-0.02Q_{t(Samdoli)}$
K.Agraharam	$Q_t =239.24+1.12Q_{t(Huvinhedgi)}$
Sadalga	$Q_t =-17.13+0.65Q_{t(Terwad)}$

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Month: August

Gauging Site	Regression Model
Karad	$Q_t = 18.80+1.59Q_{t(Warunji)}$
Warunji	$Q_t =189.11+1.05Q_{t(Koyna)}$
Arjunwad	$Q_t =-38.99+1.90Q_{t(Samdoli)}$
Kurundwad	$Q_t = 125.61+0.93Q_{t(Sadalga)}+0.95Q_{t(Arjunwad)}+0.29Q_{t(Samdoli)}$
K.Agraharam	$Q_t = 618.63+1.16Q_{t(Huvinhedgi)}$

Month: September

Gauging Site	Regression Model
Karad	$Q_t = 33.32+1.43Q_{t(Warunji)}$
Arjunwad	$Q_t = 14.46+2.74Q_{t(Samdoli)}$

Based on these observations, regression equations for daily streamflow simulation at the gauging sites are developed and presented in Table 3.1. The streamflow at any time t, Q_t, at a gauging site is expressed as a simple linear model as

$$Q_t = C + a_1Q_{us1(t)} + a_2Q_{us2(t)} + a_3Q_{us3(t)} + \ldots.. \quad (3)$$

where a_1, a_2, a_3, …. and C are empirical constants and Q_{us1}, Q_{us2}, Q_{us3}, …. are streamflows at the upstream of gauging site. From equation (3) it is found that if we know the discharge of gauging site Warunji using the equation

$$Q_t = 3.77 + 1.65\ Q_{t(Warunji)} \quad (4)$$

Discharge of gauging site Karad can be calculated. Similarly for other gauging sites in the month of June, July, August, September using the equations discharges of gauging sites can be calculated.

3.3 PERFORMANCE EVALUATION CRITERIA

The criteria used to evaluate model performance should be able to reflect the relative merits of various modeling approaches. A variety of verification criteria for the evaluation and comparison of different models are proposed by World Meteorological Organization (WMO, 1975) and other investigators (Nash and Sutcliffe, 1970). These are grouped as numerical and graphical performance indices.

Of the several numerical indices, the suitable ones chosen for the present study are the Coefficient of Determination (R^2), the Root Mean Square Error (RMSE), the Efficiency Coefficient (EC) and the Volumetric Error (EV).

The performance indices shown in Table 3.2 indicate that R^2 values and efficiency coefficients at most of the gauging sites are high depicting satisfactory performance of models. The RMSE values are reasonably low as there exists fairly good correlation. The statistical parameters such as mean, standard deviation and skewness coefficient of observed and simulated flows for the training and testing periods are tabulated in Table 3.3. The parameters are also satisfactory indicating the fair resemblance of statistical structure.

Table 3.2. Performance Indices of Linear Regression Models (Simulation Mode)

Month: June

Gauging Site	Training Period				Testing Period			
	R^2	RMSE (m^3/s)	EC (%)	EV (%)	R^2	RMSE (m^3/s)	EC (%)	EV (%)
Karad	0.97	47.59	97.15	1.44	0.93	66.06	93.17	11.70
Arjunwad	0.89	137.83	89.77	-0.01	0.92	114.35	90.27	-0.34
Kurundwad	0.96	84.17	96.8	-0.92	0.95	110.05	95.94	-0.54

Month: July

Gauging Site	Training Period				Testing Period			
	R^2	RMSE (m^3/s)	EC (%)	EV (%)	R^2	RMSE (m^3/s)	EC (%)	EV (%)
Karad	0.94	152.82	94.01	-0.06	0.97	120.67	97.56	5.55
Arjunwad	0.88	362.80	88.10	-0.17	0.90	340.25	90.81	-13.76
Kurundwad	0.96	253.75	96.58	-0.53	0.96	306.84	96.14	-6.38
K.Agraharam	0.81	1102.76	81.85	-0.49	0.76	790.95	76.59	20.24
Sadalga	0.86	140.61	86.13	-0.94	0.84	123.32	84.47	30.33

Month: August

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Gauging Site	Training Period				Testing Period			
	R^2	RMSE (m^3/s)	EC (%)	EV (%)	R^2	RMSE (m^3/s)	EC (%)	EV (%)
Karad	0.96	126.04	96.63	-0.57	0.98	126.81	98.57	1.80
Warunji	0.76	214.45	76.15	-0.31	0.86	259.3	86.02	-1.55
Arjunwad	0.89	344.57	89.26	-0.68	0.9	534.62	90.90	-13.13
Kurundwad	0.95	296.24	95.11	-0.19	0.96	426.63	96.02	-7.34
K.Agraharam	0.80	1123.58	80.19	-0.36	0.82	2245.78	82.23	-9.80

Month: September

Gauging Site	Training Period				Testing Period			
	R^2	RMSE (m^3/s)	EC (%)	EV (%)	R^2	RMSE (m^3/s)	EC (%)	EV (%)
Karad	0.96	68.94	96.63	-0.53	0.96	52.66	96.21	4.8
Arjunwad	0.79	248.38	79.92	-0.23	0.92	142.87	92.19	5.46

Table 3.3 Comparison of Statistical Parameters of Linear Regression Models (Simulation Mode)

Month: June

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	108.34	282.29	2.60	7.93	116.92	253.48	2.16	6.16
	Simulated	109.91	280.26	2.54	8.29	130.60	262.42	2.0	5.82
Arjunwad	Observed	171.78	431.35	2.51	5.14	152.70	367.29	2.40	4.16
	Simulated	171.77	408.70	2.37	4.27	141.85	313.64	2.21	4.22
Kurundwad	Observed	201.27	471.57	2.34	4.16	276.95	547.69	1.97	3.30
	Simulated	119.42	459.59	2.30	4.47	275.43	533.43	1.93	3.29

Month: July

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	511.08	625.23	1.22	2.56	511.63	773.57	1.51	3.57
	Simulated	510.72	605.83	1.18	2.95	540.02	808.07	1.49	4.34
Arjunwad	Observed	984.84	1052.53	1.06	1.82	855.91	1123.92	1.31	2.23
	Simulated	983.09	986.26	1.60	1.00	738.11	925.59	1.25	1.86
Kurundwad	Observed	1432.74	1374.17	0.95	1.2	1394.24	1566.32	1.12	1.71
	Simulated	1425.12	1342.51	0.94	1.51	1305.16	1361.15	1.04	1.92
K.Agraharam	Observed	2406.89	2591.13	1.07	1.94	857.58	1638.79	1.91	2.51
	Simulated	2395.01	2331.53	0.97	1.50	1031.16	1597.53	1.54	2.56
Sadalga	Observed	380.30	377.95	0.99	1.53	261.82	313.60	1.19	2.30
	Simulated	376.11	347.63	0.92	1.28	340.67	329.49	0.96	1.54

Month: August

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	589.79	687.64	1.17	2.96	688.15	1063.71	1.54	2.84
	Simulated	600.60	698.73	1.16	2.82	720.44	1104.96	1.53	3.24
Warunji	Observed	365.91	439.45	1.2	2.82	441.28	694.94	1.57	3.24
	Simulated	364.74	380.97	1.04	3.36	434.4	531.07	1.22	3.19
Arjunwad	Observed	1124.78	1021.99	0.91	1.77	1145.13	1657.28	1.45	2.56
	Simulated	1099.62	986.80	0.89	1.48	1038.72	1407.35	1.35	2.35
Kurundwad	Observed	1647.29	1341.11	0.81	1.30	1835.7	2143.11	1.17	1.86
	Simulated	1644.11	1305.34	0.79	1.55	1700.96	2093.50	1.23	2.37
K.Agraharam	Observed	3541.92	2526.63	0.71	1.61	3704.96	5340.32	1.44	1.55
	Simulated	3529.06	2252.69	0.63	1.15	3341.54	3502.79	1.04	1.36

Month: September

Gauging Site	Data	Training Period Mean (μ) (m^3/s)	Training Period Standard Deviation (σ) (m^3/s)	Training Period Coefficient of Variation (C_v)	Training Period Coefficient of Skewness (C_s)	Testing Period Mean (μ) (m^3/s)	Testing Period Standard Deviation (σ) (m^3/s)	Testing Period Coefficient of Variation (C_v)	Testing Period Coefficient of Skewness (C_s)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Simulated	251.27	367.61	1.46	4.32	183.67	248.45	1.35	4.87
Arjunwad	Observed	414.07	554.74	1.33	3.49	354.61	512.33	1.44	2.97
	Simulated	413.10	494.77	1.19	3.04	373.99	486.02	1.29	3.21

It may be observed from Table 3.2 and Table 3.3 that the magnitude of deviation from the agreement line has significantly decreased towards the downstream gauging sites. The possible reason for this is that the flows from the intercepted catchments which may be mostly linear are predominant compared to flows from free catchments at the downstream gauging sites. The closeness of the values at these gauging sites therefore reflects the appropriateness of the analysis.

Forecasting Models: The influencing guaging sites are identified through cross corellation analysis, using correlation coefficients between the daily stream flows of monsoon months of the upstream guaging sites and those at the guaging sites on the downstream. The analysis similar to that of simulation models but using the data pertained to previous periods is performed for the forecasting mode. The streamflows of the upstream gauging sites of the previous periods during monsoon period that influence the streamflows at the gauging sites are determined. Table 3.4 presents the ralationships developed through regression analysis applicable to the range of streamflow data collected at the gauging sites for the streamflows in the forecasting mode.

Table 3.4 Linear Regression Models (Forecasting Mode)

Month: July

Arjunwad	$Q_t=99.62+0.89Q_{t-1(Arjunwad)}$
Kurundwad	$Q_t= 109.23+0.92Q_{t-1(Kurundwad)}$
Terwad	$Q_t=43.22+0.92Q_{t-1(Terwad)}$
Sadalga	$Q_t=40.86+0.90Q_{t-1(Sadalga)}$

Month: August

Arjunwad	$Q_t=95.14+0.91Q_{t-1(Arjunwad)}$
Samdoli	$Q_t= 51.09+0.88Q_{t-1(Samdoli)}$

Month: September

Karad	$Q_t= 68.81+1.20Q_{t-1(Warunji)}$
Warunji	$Q_t= 19.68-0.15Q_{t-2(Karad)}+0.68Q_{t-1(Karad)}$
Arjunwad	$Q_t=-64.91+1.233Q_{t-1(Samdoli)}+0.37Q_{t-2(Karad)+}0.33Q_{t-2(Arjunwad)}$
Kurundwad	$Q_t=27.51+0.94Q_{t-1(Kurundwad)}$

The streamflow at any time t, Q_t, at a gauging site is expressed as a simple linear model as

$$Q_t = C + a_1Q_{us1(t)} + a_2Q_{us2(t)} + a_3Q_{us3(t)} + \ldots \quad (5)$$

where a_1, a_2, a_3, …. and C are empirical constants and Q_{us1}, Q_{us2}, Q_{us3}, …. are streamflows at the upstream of gauging site. From equation (5) it is found that if we know the discharge of gauging site Arjunwad using the equation

$$Q_t=99.62+0.89Q_{t-1(Arjunwad)} \quad (6)$$

Discharge of gauging site can be calculated. Similarly for other gauging sites in the month of June, July, August, September using the equations discharges of gauging sites can be calculated.

The performance indices shown in Table 3.5 and 3.6 indicate that the streamflows are fairly simulated at the gauging sites of Kurundwad, Arjunwad, Terwad, Sadalga, Karad, the R^2 values and Efficiency coefficients depict that the performance of the models developed is satisfactory. The RMSE values are low as three exists fairly good correlation.

Table 3.5 Performance Indices of Linear Regression Models (Forecasting Mode)

Month: July

Gauging Site	Training Period				Testing Period			
	RMSE	R^2	EC (%)	EV (%)	RMSE	R^2	EC (%)	EV (%)
Arjunwad	459.97	0.80	80.88	-0.92	466.98	0.82	82.72	0.08
Kurundwad	531.97	0.84	84.99	-0.15	606.41	0.85	85.00	-0.88
Terwad	201.02	0.86	86.04	-0.76	192.73	0.85	85.06	-0.73
Sadalga	157.28	0.81	81.06	-0.001	115.63	0.86	86.14	4.78

Month: August

Gauging Site	Training Period				Testing Period			
	RMSE	R^2	EC (%)	EV (%)	RMSE	R^2	EC (%)	EV (%)
Arjunwad	426.37	0.83	83.50	-0.17	602.59	0.88	88.47	-0.75
Samdoli	174.05	0.79	79.87	-0.98	189.15	0.87	87.33	0.88

Month: September

Gauging Site	Training Period				Testing Period			
	RMSE	R^2	EC (%)	EV(%)	RMSE	R^2	EC (%)	EV (%)
Karad	92.24	0.93	93.38	-0.36	78.38	0.91	91.62	11.26
Warunji	148.02	0.66	66.85	0.80	96.18	0.69	69.42	6.87
Arjunwad	378.64	0.87	87.45	-1.2	474.59	0.92	92.83	-9.51
Kurundwad	40.08	0.99	99.63	-0.12	44.1	0.99	99.63	-1.23

Table 3.6 Comparison of Statistical Parameters of Linear Regression Models (Forecasting Mode)

Month: July

Gauging Site	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	985.86	1052.64	1.06	1.82	857.54	1125.02	1.31	2.22
	Simulated	976.74	937.07	0.95	1.82	858.29	999.86	1.16	2.24
Kurundwad	Observed	1430.74	1374.19	0.96	1.20	1396.68	1568.61	1.12	1.71
	Simulated	1428.54	1264.79	0.88	1.19	1384.35	1438.03	1.03	1.73
Terwad	Observed	607.90	538.52	0.88	1.26	548.10	499.59	0.91	1.58
	Simulated	603.23	495.45	0.82	1.26	544.05	457.33	0.84	1.61
Sadalga	Observed	404.34	361.68	0.89	1.45	263.83	311.07	1.17	2.08
	Simulated	404.33	325.64	0.80	1.45	276.45	278.80	1.00	2.10

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Month: August

Gauging Site	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	1104.86	1050.45	0.95	1.76	1194.78	1777.80	1.48	2.54
	Simulated	1102.96	958.12	0.86	1.75	1185.82	1617.19	1.36	2.55
Samdoli	Observed	469.48	388.24	0.82	1.33	417.92	532.12	1.27	2.29
	Simulated	464.87	343.23	0.73	1.34	421.63	469.78	1.11	2.26

Month: September

Gauging Site	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Simulated	251.27	370.18	1.47	4.32	183.20	250.19	1.36	4.87
Warunji	Observed	150.68	257.27	1.70	4.32	105.48	174.21	1.65	4.85
	Simulated	153.25	211.55	1.38	4.58	112.73	151.34	1.34	5.10
Arjunwad	Observed	1107.22	1052.35	0.95	1.75	1195.82	1775.52	1.48	2.54
	Simulated	1093.92	1014.39	0.92	1.61	1082.04	1557.01	1.43	2.45
Kurundwad	Observed	468.13	668.55	1.42	3.63	577.94	727.04	1.25	2.50
	Simulated	467.55	628.44	1.34	3.63	570.77	683.42	1.19	2.50

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

3.2 ARTIFICIAL NEURAL NETWORKS (ANN)

An ANN is a massively parallel – disturbed information processing system that has certain performance characteristics resembling biological neural networks of the human brain. ANNs are developed as a generalization of mathematical models of human cognition. Their development is based on the following rules:

1. Information processing occurs at many single elements called nodes, also referred as units, cells or neurons.
2. Signals are passed between nodes through connection links.
3. Each connection link has an associated weight that represents its connection strength.
4. Each node typically applies a transformation called an activation function to its net input to determine its output signal. A neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function. A typical ANN consists of a number of nodes that are organized according to a particular arrangement. One way of classifying neural networks by its number of layers: single, bilayer and multilayer. ANNs are also categorized based on the direction of the information flow and processing. In a feed-forward network, the nodes are generally arranged in layers starting from the first input layer and ending at the final output layer. There can be several hidden layers, having one or more nodes in each layer. Information passes from the input to the output side. The nodes in one layer are connected to those in the next, but not to those in the same layer. Thus, an output of node in a layer is only a dependent on the inputs it receives from previous layers and the corresponding weights.

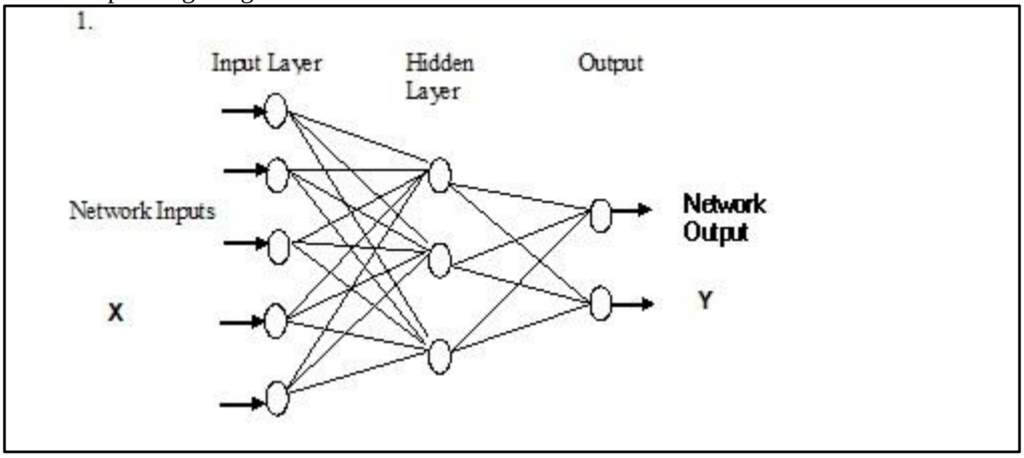


Fig. 3.1: Configuration of Feed Forward Three-Layer ANN

In most networks, the input (first) layer receives the input variables for the problem and the output (last) layer consists of values predicted by the network. The number of hidden layers and the number of nodes in each hidden layer are usually determined by trial and error procedure. The nodes within the neighboring layers of the network are full connected by links. A synaptic weight is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input-output relationship. Figure 3.1 shows the configuration of a feed-forward three-layer ANN. In this figure, X is a system input vector composed of a number of causal variables that influence system behavior, and Y is the system output vector composed of a number of resulting variables that represent the system behavior.

Mathematical Aspects: A schematic diagram of a typical j^{th} node is displayed in Fig 3.2. The inputs to such a node may come from outputs of other nodes, depending on the location of layer. These inputs form an input vector $X = (x_{1,....}\ x_i\ ,........x_n)$. The sequence of weights leading to the node form a weight vector $w_j = (w_{ij}\ ,\ldots,\ w_{ij}\ ,..\ w_{nj})$, where w_{ij} represents the connection weight from the i^{th} node in the preceding layer to this node.

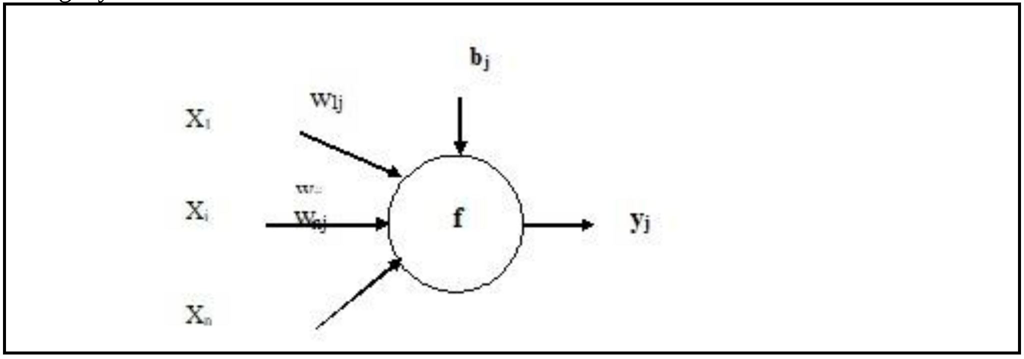


Fig. 3.2: Schematic diagram of node J

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

The output of node j, Y_j, is obtained by computing the value of function f with respect to the inner product of vector X and W_j minus b_j, where b_j is the threshold value, also called the bias, associated with this node. The following equation defines the operation

$$y_j = f(X.W_j - b_j) \quad (7)$$

The function f is called an activation function. Its functional form determines the response of a node to the total input signal it receives. The most commonly used form is the sigmoid function, given as

$$f(t) = 1/\ 1+e^{-t} \quad (8)$$

Sigmoid function is a bounded, monotonic, non-decreasing function that provides a graded non-linear response. This function enables the network to map any nonlinear processes. The popularity of the sigmoid function is partially attributed to the simplicity of its derivative that is used during the training process. A number of such nodes are organized to form an artificial neural network.

Network Training

The primary goal of network training is to minimize the error function by searching for a set of connection strengths and threshold values that cause the ANN to produce outputs that are equal or close to targets. In order for an ANN to generate an output vector $Y=(y_1,y_2,\text{---},y_p)$ that is as close as possible to the target vector $T=(t_1,t_2,\text{---},t_p)$, a training process, also called learning is employed to find optimal weight matrices W and bias vectors V, that minimize a predetermined error function that usually has the form

$$E = \sum_P \sum_P (y_i - t_i)^2 \quad (9)$$

where, t_i is a component of the desired output T, y_i = corresponding ANN output, p = number of output nodes and P = number of training patterns. The network training is a process by which the connection weights of an ANN are adopted through a continuous process of stimulation by the environment in which the network is embedded. There are primarily two types of training, supervised and unsupervised. A supervised training algorithm requires an external teacher to guide the training process. This typically implies that a large number of examples (or patterns) of inputs and outputs are required for training. The inputs are the cause variables of a system and the outputs are the effect variables. This training procedure involves the iterative adjustment and optimization of connection weights and threshold values for each of the nodes. After training has been accomplished, it is hoped that the ANN is capable of generating reasonable results given new inputs. In contrast, an unsupervised training algorithm does not involve a teacher. During training, only an input data set is provided to the ANN that automatically adapts its connection weights to cluster those input patterns into classes with similar properties. A combination of these two training strategies generally leads to reinforced learning.

Back Propagation: It is essentially a gradient descent technique that minimizes the network error function. Each input pattern of the training data set is passed through the network from the input layer to the output layer. The network output is compared with the desired target output, and an error is computed. This error is propagated backward through the network to each node, and correspondingly the connection weights are adjusted based on the equation.

$$\Delta w_{ij}(n) = -\epsilon\ \partial E/\partial w_{ij} + \alpha\ \Delta w_{ij}(n-1) \quad (10)$$

where $\Delta W_{ij}(n)$ and $\Delta W_{ij}(n-1)$ = weight increments between node i and j during the n^{th} and $(n-1)^{th}$ pass, or epoch and, ϵ and α are called learning rate and momentum factor respectively. The momentum factor speeds up training in very flat regions of the error surface and helps to prevent oscillations in the weights. A learning rate is used to increase the chance of avoiding the training process being trapped in local minima instead of global minima. The back propagation algorithm involves two steps. The first step is a forward pass in which the effect of the input is passed forward through the network to reach the output layer. After the error is computed, a second step starts backward through the network. The errors at the output layer are propagated back toward the input layer with the weights being modified.

Selection of Input and Output Variables: The goal of an ANN is to generalize a relationship of the form

$$Y^m = f\ (X^n) \quad (11)$$

where X^n is an n-dimensional input vector consisting of variables $x_1,\ldots,x_i,\ldots,x_n$; and Y^m is an m-dimensional output vector of resulting variables of interest $y_1,\ldots y_i,\ldots y_m$.

The selection of an appropriate input vector that allows an ANN to successfully map to the desired output vector is not a trivial task. Unlike physically-based models, the set of variables that influences the system is not known a priori. In this sense of nonlinear process identification, an ANN cannot be considered as a mere black box. The physical insight into the problem being studied leads to better choice of input variables for proper mapping. This helps in avoiding loss of information that may result if key input variables are omitted, and also prevents inclusion of spurious inputs that tend to confuse the training process.

Collection and Processing of Data: Most hydrologic data are obtained either from gauges that are placed on site or through remote sensing instruments. Also, either the existing models or laboratory experiments are used to generate the data patterns for specific applications. The number of data pairs used for training is equal to or greater than the number of parameters (weights) in the network. An optimal data set is representative of the probable occurrence of an input vector and facilitates the mapping of the underlying

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

nonlinear process and, an insufficient data set leads to poor learning. Routine procedures such as plotting and examining the statistics are effective in judging the reliability of the data.

Designing of ANN: An optimal architecture is considered the one yielding the best performance in terms of error minimization, while retaining a simple and compact structure. No unified theory exists for determination of an optimal ANN architecture. Often, more than one ANN can generate similar results. The numbers of input and output nodes are problem dependent. The flexibility lies in selecting the number of hidden layers and in assigning the number of nodes to each of these layers.

Training: The purpose of training is to determine the set of connection weights and nodal thresholds that cause the ANN to estimate outputs that are sufficiently close to the target values. The fraction of the complete data to be employed for training should contain sufficient patterns so that the network can mimic the underlying relationship between input and output variables adequately. The weights and threshold values are assigned small random values initially. During training, these are adjusted based on the error, or the difference between ANN output and the target responses. This adjustment is continued recursively until a weight space is found, which results in the smallest overall prediction error. However, there is the danger of overtraining a network in this fashion, also referred as over fitting. This happens when the network parameters are too fine-tuned to the training data set. It is as if the network, in the process of trying to 'learn' the underlying rule, has started trying to fit the noise component of the data as well. In other words overtraining results in a network that memorizes the individual examples, rather than trends in the data set as a whole. When this happens, the network performs very well over the data set used for training, but shows poor predictive capabilities when supplied with data other than the training patterns. The simplest way to prevent overtraining is to stop training when the mean square error ceases to decrease significantly.

Model Validation: The performance of a trained ANN is fairly evaluated by subjecting it to new pattern other than the patterns used during training. It is determined by computing the percentage error between predicted and desired values. In addition, plotting the model output versus desired response is also used to assess ANN performance. Since finding optimal network parameters is essentially a minimization process, it is advisable to repeat the training and testing processes several times to ensure satisfactory results.

3.2.1 ARTIFICIAL NEURAL NETWORK MODELLING

Traditional hydrologic models require significant effort of assessing model parameters and performing model calibration and verification. ANN modeling methodology offers a promising alternative to the traditional regression and time series approaches. Artificial Neural Networks (ANNs) have been successfully used in the area of streamflow simulation and forecasting. Its ability to identify a relationship from given patterns make it possible for ANNs to solve complex problems. This study presents the development of daily streamflow models using ANNs and comparison of regression and ANN models.

A standard multilayer feed forward ANN is adopted for the present study. The modeling of ANN is initiated with the normalization of all inputs and output as given below reducing the data in the range of 0.1 to 0.9 to avoid any saturation effect that may be caused by the use of sigmoid function.

$$x_{norm} = 0.1+0.8\,(x_i/x_{max}) \qquad (12)$$

A constant value of 0.1 for learning rate and 0.9 for momentum factor are considered. The criteria selected to avoid over learning was through generalization of ANN for which the developed model is simultaneously checked for its improvement on verification of data on each iteration. The training is continued until there is an improvement in the performance of the model in calibration and verification periods. The performance of the models developed is tested through selected evaluation criteria. The entire process is carried out using MATLAB routines.

The reliability of model results does not only depend on the ANN structure which has to be carefully chosen through the training-validation process, but also on the input data. To get reliable results, the input data always need to be trustworthy. The selection of an appropriate and parsimonious input vector that allows the ANN to map the desired output vector successfully is not a trivial task. In most of the applications reported, it has been done either by trial and error or through statistical procedure. In the present study, the statistical procedure is adopted to identify the appropriate input vector for a model.

The selection of an appropriate transfer function is based on the degree of noise and nonlinearity present in the data. The present study adopted log-sigmoid function.

The number of neurons in the input and output layers are fixed as the number of input and output parameters considered for the present study. However, the number of hidden layer neurons is more difficult to determine since no general methodology is available for its determination. Determining an appropriate architecture of a neural network for a particular problem is an important issue since the network topology directly affects its computational complexity and generalization capability. The number of neurons in the hidden layer is decided after a rigorous course of training and testing. In the present study, a method of progressively adding hidden neurons as needed to escape local minima in the error function is adopted. Care is taken to avoid too few and too many neurons which can respectively cause difficulties in mapping each input and output in the training set and increase training times unnecessarily, in the process of determination of optimal number of hidden layers and nodes in each hidden layer.

The following trial and error procedure is used for the determination of optimal network architecture. After deciding the input and output parameters, training size and learning algorithm, a network is chosen with a trial number of nodes in the hidden layer. Error back propagation which is an iterative nonlinear optimization process based on the gradient descent search method is used during calibration. The network is trained and the error gradient is observed. Then the network architecture is changed by increasing the number of hidden nodes. The training procedure is repeated for the new architecture. This procedure is continued for several architectures. The network architecture that resulted in the maximum error gradient is adopted as the optimal architecture.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

3.2.2 STREAMFLOW MODELING

It may be observed from the cross correlation analysis that the streamflows at Karad gauging site depend on those of warunji, during monsoon month Arjunwad streamflows on those of Samdoli, Streamflows at Kurundwad depend on streamflows at Arjunwad, Samdoli & Terwad gauging sites and streamflows at K.Agraharam gauging site depend on the Streamflows at the gauging site of Huvinhedgi in the months of June, July, and August.

The results of the simulation and forecasting ANN models evaluated through the performance indices viz., Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Efficiency Coefficient (EC) and Volumetric Error (EV) during training and validation periods are summarized in Table 3.7 and 3.8 for the most of gauging sites as it corresponds with low and high values of the RMSE and efficiency coefficient respectively. The low EV values also indicate that the models do not overestimate or underestimate the streamflows in both simulation and forecasting modes. The statistical parameters of observed and simulated/predicted streamflows as shown in Table 3.9 and 3.10 for the testing period are preserved satisfactorily at most of the gauging sites.

It may be concluded that a relatively Simple Neural Network, with an adequate choice of input variables and an appropriate learning procedure can achieve accuracy superior to regression models. The models developed show satisfactory performance based on evaluation criteria and statistical structure and therefore may be adopted with reasonable accuracy for streamflow simulation or forecasting at the gauging sites selected for the present study.

Table 3.7 Performance Indices of ANN Models (Simulation Mode)

Month : June

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	1-5-1-1	0.78	134.35	78.34	15.46	0.65	146.62	65.37	10.47
Arjunwad	1-5-1-1	0.87	157.82	87.69	2.81	0.84	132.01	84.8	-7.74
Kurundwad	3-6-1-1	0.95	102	95.36	0.66	0.93	130.72	93.83	-6.38

Month: July

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	1-6-1-1	0.94	131.25	94.87	0.084	0.97	113.16	97.58	0.11
Arjunwad	2-4-1-1	0.88	187.35	88.49	2.81	0.9	374.33	90.27	-7.17
Kurundwad	4-6-1-1	0.98	188.79	98.19	0.72	0.98	200.88	98.14	-2.7
K.Agraharam	1-4-1-1	0.82	1161.75	82.09	-4.74	0.83	720.54	83.37	5.56
Sadalga	1-6-1-1	0.85	146.08	85.55	-7.74	0.87	109.84	87.94	17.16

Month: August

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	2-5-1-1	0.94	145.24	94.34	-0.21	0.96	190.58	96.65	-1.52
Warunji	1-3-1-1	0.70	212.29	70.29	-3.42	0.9	203.78	90.95	3.78
Arjunwad	3-5-1-1	0.79	459.53	79.79	2.25	0.93	432.17	93.19	-2.29
Kurundwad	4-4-1-1	0.93	339.78	93.58	1.81	0.97	357.68	97.21	-4.36
K.Agraharam	2-3-1-1	0.88	784.71	88.81	1.95	0.92	1273.33	92.97	46.77

Month: September

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	1-4-1-1	0.92	80.84	92.98	-7.93	0.93	88.69	93.3	3.29
Arjunwad	3-5-1-1	0.81	160.57	81.81	-1.33	0.94	171.39	94.5	-2.92

Table 3.8 Performance Indices of ANN Models (Forecasting Mode)

Month: July

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Arjunwad	1-4-1-1	0.97	160.51	97.69	0.86	0.97	169.51	97.63	-0.81
Kurundwad	1-5-1-1	0.99	89.09	99.43	0.42	0.99	106.18	99.47	1.85
Terwad	1-5-1-1	0.98	66.54	98.46	1.24	0.98	66.01	98.24	1.92
Sadalga	1-5-1-1	0.98	48.09	98.22	-0.19	0.97	44.18	97.96	4.54

Month: August

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Arjunwad	1-4-1-1	0.97	167.14	97.48	-1.02	0.98	175.52	98.98	-2.19
Samdoli	1-5-1-1	0.97	66.32	97.09	-2.97	0.98	69.14	98.3	0.82

Month: September

Gauging Site	Architecture	Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC(%)	EV(%)
Karad	2-5-1-1	0.75	185.50	75.93	1.94	0.78	124.14	78.31	5.55
Warunji	2-4-1-1	0.72	157.69	72.8	-3.96	0.76	82.05	76.95	-1.35
Arjunwad	3-4-1-1	0.85	220.78	85.45	-0.6	0.84	174.7	84.41	0.64

Table 3.9 Comparison of Statistical Parameters of ANN Models during Testing Period (Simulation Mode)

Month: June

Gauging Site	Data	Training Period				Testing Period			
		Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	108.34	282.29	2.60	7.93	116.92	253.48	2.16	6.16
	Simulated	129.44	283.00	2.18	8.88	135.87	303.95	2.23	7.56
Arjunwad	Observed	171.78	431.35	2.51	5.14	152.70	367.29	2.40	4.16
	Simulated	368.82	612.04	1.65	3.02	267.32	458.65	1.71	3.19
Kurundwad	Observed	201.27	471.57	2.34	4.16	276.95	547.69	1.97	3.29
	Simulated	584.22	675.90	1.15	2.66	611.49	719.96	1.17	1.83

Month: July

Gauging site	Data	Training Period				Testing Period			
		Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variation (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	511.08	625.23	1.22	2.56	511.63	773.57	1.51	3.57
	Simulated	482.33	588.39	1.22	2.85	558.46	762.80	1.37	3.0
Arjunwad	Observed	984.84	1052.53	1.06	1.82	855.91	1123.92	1.31	2.23
	Simulated	984.23	1027.28	1.04	1.87	867.37	1130.41	1.30	2.04
Kurundwad	Observed	1432.74	1374.18	0.96	1.2	1394.25	1566.32	1.12	1.72
	Simulated	1478.66	1398.70	0.94	1.13	1311.62	1412.44	1.07	1.64
K.Agraharam	Observed	2406.89	2591.13	1.07	1.94	857.58	1638.79	1.91	2.51
	Simulated	2580.14	2435.80	0.94	2.06	1215.12	1704.13	1.40	1.91
Sadalga	Observed	380.30	377.95	0.99	1.53	261.82	313.60	1.19	2.30
	Simulated	362.81	351.04	0.97	1.28	317.39	318.13	1.00	1.61

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Month : August

Gauging site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	589.79	687.64	1.17	2.96	688.15	1063.71	1.54	2.84
	Simulated	573.46	651.21	1.14	2.63	671.92	979.32	1.46	2.63
Warunji	Observed	365.91	439.45	1.20	2.83	441.29	694.95	1.57	3.24
	Simulated	342.28	384.94	1.12	3.39	453.19	607.05	1.33	3.10
Arjunwad	Observed	1124.78	1021.99	0.91	1.77	1145.13	1657.28	1.45	2.56
	Simulated	1150.20	1094.63	0.95	1.83	1118.84	1503.00	1.34	2.57
Kurundwad	Observed	1647.29	1341.11	0.81	1.3	1835.70	2143.11	1.17	1.86
	Simulated	1677.13	1373.82	0.82	1.59	1755.61	2065.21	1.18	2.04
K.Agraharam	Observed	3541.92	2526.63	0.71	1.61	3704.96	5340.32	1.44	1.55
	Simulated	3612.67	2571.42	0.71	1.61	3703.14	4214.96	1.14	1.92

Month: September

Gauging site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Simulated	214.46	298.07	1.39	3.24	234.32	367.48	1.57	4.68
Arjunwad	Observed	414.7	554.74	1.33	3.49	354.61	512.33	1.44	2.97
	Simulated	296.79	350.42	1.18	2.49	419.1	687.6	1.60	3.7

Table 3.10 Comparison of Statistical Parameters of ANN Models during Testing Period (Forecasting Mode)

Month: July

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	985.86	1052.64	1.06	1.82	857.54	1125.02	1.31	2.22
	Simulated	1010.13	968.69	0.96	1.84	819.27	1003.17	1.22	2.02
Kurundwad	Observed	1430.74	1374.19	0.96	1.20	1396.68	1568.61	1.12	1.71
	Simulated	1533.63	1331.54	0.87	1.18	1315.99	1410.02	1.07	1.97
Terwad	Observed	607.90	538.52	0.88	1.26	548.10	499.59	0.91	1.58
	Simulated	615.11	496.57	0.81	1.25	558.67	461.83	0.83	1.39
Sadalga	Observed	404.35	361.68	0.89	1.46	262.06	309.78	1.18	2.11
	Simulated	403.54	329.05	0.82	1.32	273.97	279.48	1.02	1.93

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Month: August

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	1104.86	1050.45	0.95	1.76	1194.78	1777.80	1.48	2.54
	Simulated	1102.62	951.08	0.86	1.76	1139.99	1632.00	1.43	2.67
Samdoli	Observed	469.48	388.24	0.82	1.33	417.92	532.12	1.27	2.29
	Simulated	456.42	342.15	0.75	1.42	422.23	487.81	1.16	2.61

Month: September

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Simulated	259.58	380.26	1.46	4.18	181.49	257.12	1.41	5.04
Warunji	Observed	150.68	257.27	1.70	4.32	105.48	174.21	1.65	4.85
	Simulated	150.65	243.56	1.61	4.04	97.85	180.01	1.83	4.81
Arjunwad	Observed	1107.22	1052.35	0.95	1.75	1195.82	1775.52	1.48	2.54
	Simulated	1082.81	1013.58	0.82	1.55	1098.51	1688.42	1.32	2.31
Kurundwad	Observed	468.13	668.55	1.42	3.63	577.94	727.42	1.25	2.7
	Simulated	464.41	614.68	1.32	3.8	514.68	661.34	1.28	3.17

4. RESULTS AND DISCUSSIONS

4.1 Comparison of Regression and ANN Streamflow Models

The streamflow models at the gauging sites of Krishna basin are developed using regression and ANN models. The performance indices of the models in the simulation and forecasting modes are compared as shown in Tables 4.1 and 4.2. It may be observed from the results that the performance of regression models in terms of R^2, RMSE, EC and EV is satisfactory in simulation mode. The performance of ANN models for the upstream gauging sites, has improved significantly. The RMSE has also reduced considerably. This may be due to the fact that the streamflows at the upstream flow gauging sites depend significantly on flows from their free catchments which exhibit nonlinearity. The regression models in the forecasting mode have failed to perform satisfactorily at all the gauging sites irrespective of their location as the flows show weaker linear correlation, with those in the previous time periods. It may be noticed from Table 4.3 and 4.4 that the statistical parameters of streamflows are comparable with those of

ANN models. However, the ANN models have outperformed the regression models in the forecasting mode as the technique has inherent property of learning complex nonlinear relation between the flows. The comparison plots were drawn for the regression and ANN streamflow models as shown in Fig. 4.1 to 4.7.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Table 4.1 Comparison of Performance Indices of Regression and ANN Models (Simulation mode)

Month: June

Gauging Site		Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	Regression	0.97	47.59	97.15	1.44	0.93	66.06	93.17	11.70
	ANN	0.78	134.35	78.34	15.46	0.65	146.62	65.37	10.47
Arjunwad	Regression	0.89	137.83	89.77	-0.01	0.92	114.35	90.27	-0.34
	ANN	0.87	157.82	87.69	2.81	0.84	132.01	84.8	-7.74
Kurundwad	Regression	0.96	84.17	96.8	-0.92	0.95	110.05	95.94	-0.54
	ANN	0.95	102	95.36	0.66	0.93	130.72	93.83	-6.38

Month: July

Gauging Site		Training Period				Training Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	Regression	0.94	152.82	94.01	-0.06	0.97	120.67	97.56	5.55
	ANN	0.94	131.25	94.87	0.084	0.97	113.16	97.58	0.11
Arjunwad	Regression	0.88	362.80	88.10	-0.17	0.90	340.25	90.81	-13.76
	ANN	0.88	187.35	88.49	2.81	0.9	374.33	90.27	-7.17
Kurundwad	Regression	0.96	253.75	96.58	-0.53	0.96	306.84	96.14	-6.38
	ANN	0.98	188.79	98.19	0.72	0.98	200.88	98.14	-2.7
K.Agraharam	Regression	0.81	1102.76	81.85	-0.49	0.76	790.95	76.59	20.24
	ANN	0.82	1161.75	82.09	-4.74	0.83	720.54	83.37	5.56
Sadalga	Regression	0.86	140.61	86.13	-0.94	0.84	123.32	84.47	30.33
	ANN	0.85	146.08	85.55	-7.74	0.87	109.84	87.94	17.16

Month: August

Gauging Site		Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	Regression	0.96	126.04	96.63	-0.57	0.98	126.81	98.57	1.8
	ANN	0.94	145.24	94.34	-0.21	0.96	190.58	96.65	-1.52
Warunji	Regression	0.76	214.45	76.15	-0.31	0.86	259.3	86.02	-1.55
	ANN	0.70	212.29	70.29	-3.42	0.9	203.78	90.95	3.78
Arjunwad	Regression	0.89	344.57	89.26	-0.68	0.9	534.62	90.9	-13.13
	ANN	0.79	459.53	79.79	2.25	0.93	432.17	93.19	-2.29
Kurundwad	Regression	0.95	296.64	95.11	-0.19	0.96	426.63	96.02	-7.34
	ANN	0.93	339.78	93.58	1.81	0.97	357.68	97.21	-4.36
K.Agraharam	Regression	0.80	1123.58	80.19	-0.36	0.82	2245.78	82.23	-9.8
	ANN	0.88	784.71	88.81	1.95	0.92	1273.33	92.97	46.77

Month: September

Gauging site		Training period				Testing period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	Regression	0.96	68.94	96.63	-0.53	0.96	52.66	96.21	4.8
	ANN	0.92	80.84	92.98	-7.93	0.93	88.69	93.3	3.29
Arjunwad	Regression	0.79	248.38	79.92	-0.23	0.92	142.87	92.19	5.46
	ANN	0.81	160.57	81.81	-1.33	0.94	171.39	94.5	-2.92

Table 4.2 Comparison of Performance Indices of Regression and ANN Models (Forecasting Mode)

Month: July

Gauging Site		Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

Arjunwad	Regression	0.8	459.97	80.88	-0.92	0.82	466.98	82.72	0.08
	ANN	0.97	160.51	97.69	0.86	0.97	169.51	97.63	-0.81
Kurundwad	Regression	0.84	531.97	84.99	-0.15	0.85	606.41	85.00	-0.88
	ANN	0.99	89.09	99.43	0.42	0.99	106.18	99.47	1.85
Terwad	Regression	0.86	201.02	86.04	-0.76	0.85	192.73	85.06	-0.73
	ANN	0.98	66.54	98.46	1.24	0.98	66.01	98.24	1.92
Sadalga	Regression	0.81	157.28	81.06	-0.01	0.86	115.63	86.14	4.78
	ANN	0.98	48.09	98.22	-0.19	0.97	44.18	97.96	4.54

Month: August

Gauging Site		Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Arjunwad	Regression	0.83	426.37	83.50	-0.17	0.88	602.59	88.47	-0.75
	ANN	0.97	167.14	97.48	-1.02	0.98	175.52	98.98	-2.19
Samdoli	Regression	0.79	174.05	79.87	-0.98	0.87	189.15	87.33	0.88
	ANN	0.97	66.32	97.09	-2.97	0.98	69.14	98.30	0.82

Month: September

Gauging Site		Training Period				Testing Period			
		R^2	RMSE	EC (%)	EV (%)	R^2	RMSE	EC (%)	EV (%)
Karad	Regression	92.24	0.93	93.38	-0.36	78.38	0.91	91.62	11.26
	ANN	0.75	185.50	75.93	1.94	0.78	124.14	78.31	5.55
Warunji	Regression	0.66	148.02	66.85	0.8	0.69	96.18	69.42	6.87
	ANN	0.72	157.69	72.8	-3.96	0.76	82.05	76.95	-1.35
Arjunwad	Regression	0.87	378.64	87.45	-1.2	0.92	474.59	92.83	-9.51
	ANN	0.85	220.78	85.45	-0.6	0.84	174.7	84.41	0.64
Kurundwad	Regression	0.99	40.08	99.63	-0.12	0.99	44.1	99.63	-1.23
	ANN	0.97	106.92	97.54	-0.56	0.96	123.78	96.81	-1.88

Table 4.3 Comparison of Statistical Parameters for Regression and ANN Models (Simulation Mode)

Month: June

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variantion (Cv)	Coefficient of Skewness (Cs)	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance Cv)	Coefficient of Skewness (Cs)
Karad	Observed	108.34	282.29	2.60	7.93	116.92	253.48	2.16	6.16
	Regression	109.91	280.26	2.54	8.29	130.60	262.42	2.01	5.82
	ANN	129.44	283.00	2.18	8.88	135.87	303.95	2.23	7.56
Arjunwad	Observed	171.78	431.35	2.51	5.14	152.70	367.29	2.40	4.16
	Regression	171.77	408.70	2.37	4.27	141.85	313.64	2.21	4.22
	ANN	368.82	612.04	1.65	3.02	267.32	458.65	1.71	3.19
Kurundwad	Observed	201.27	471.57	2.34	4.16	276.95	547.69	1.97	3.30
	Regression	119.42	459.59	2.30	4.47	275.43	533.43	1.93	3.29
	ANN	584.22	675.90	1.15	2.66	611.49	719.96	1.17	1.83

Month: July

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)
Karad	Observed	511.08	625.23	1.22	2.56	511.63	773.57	1.51	3.57
	Regression	510.72	605.83	1.18	2.95	540.02	808.07	1.49	4.34
	ANN	482.33	588.39	1.22	2.85	558.46	762.8	1.37	3
Arjunwad	Observed	984.84	1052.53	1.06	1.82	855.91	1123.92	1.31	2.23
	Regression	983.09	986.26	1.60	1.00	738.11	925.59	1.25	1.86
	ANN	984.23	1027.28	1.04	1.88	867.37	1130.41	1.3	2.04
Kurundwad	Observed	1432.74	1374.18	0.96	1.2	1394.25	1566.32	1.12	1.72
	Regression	1425.12	1342.51	0.94	1.52	1305.16	1361.16	1.04	1.93
	ANN	1478.66	1398.70	0.94	1.13	1311.62	1412.44	1.07	1.64
K.Agraharam	Observed	2406.89	2591.13	1.07	1.94	857.58	1638.79	1.91	2.51
	Regression	2395.01	2331.53	0.97	1.50	1031.16	1597.53	1.54	2.56
	ANN	2580.14	2435.80	0.94	2.06	1215.12	1704.13	1.40	1.91
Sadalga	Observed	380.30	377.95	0.99	1.53	261.82	313.60	1.19	2.30
	Regression	376.11	347.63	0.92	1.28	340.67	329.49	0.96	1.54
	ANN	362.81	351.04	0.97	1.28	317.39	318.13	1.00	1.61

Month: August

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)
Karad	Observed	589.79	687.64	1.17	2.96	688.15	1063.71	1.54	2.84
	Regression	600.60	698.73	1.16	2.82	720.44	1104.96	1.53	3.24
	ANN	573.46	651.21	1.14	2.63	671.92	979.32	1.46	2.63
Warunji	Observed	365.92	439.45	1.2	2.83	441.29	694.95	1.57	3.24
	Regression	364.75	380.98	1.04	3.37	434.41	531.08	1.22	3.2
	ANN	342.28	384.94	1.12	3.39	453.19	607.05	1.33	3.10
Arjunwad	Observed	1124.78	1021.99	0.91	1.77	1145.13	1657.28	1.45	2.56
	Regression	1099.63	986.81	0.9	1.48	1038.72	1407.35	1.35	2.35
	ANN	1150.2	1094.63	0.95	1.83	1118.84	1503	1.34	2.57
Kurundwad	Observed	1647.29	1341.11	0.81	1.3	1835.7	2143.11	1.17	1.86
	Regression	1644.11	1305.35	0.79	1.55	1700.96	2093.5	1.23	2.37
	ANN	1677.13	1373.82	0.82	1.59	1755.61	2065.21	1.18	2.04
K.Agraharam	Observed	3541.92	2526.63	0.71	1.61	3704.96	5340.32	1.44	1.55
	Regression	3529.06	2252.69	0.63	1.15	3341.54	3502.79	1.04	1.36
	ANN	3612.67	2571.42	0.71	1.61	3703.14	4214.96	1.14	1.92

Month: September

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (Cv)	Coefficient of Skewness (Cs)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Regression	251.27	367.61	1.46	4.33	183.67	248.45	1.35	4.87
	ANN	214.46	298.07	1.39	3.24	234.32	367.48	1.57	4.68
Arjunwad	Observed	414.07	554.74	1.33	3.49	354.61	512.33	1.44	2.97
	Regression	413.10	494.77	1.19	3.04	373.99	486.02	1.29	3.21
	ANN	296.79	350.42	1.18	2.49	419.10	687.6	1.60	3.70

Table 4.4 Comparison of Statistical Parameters for Regression and ANN Models (Forecasting Mode)

Month: July

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m³/s)	Standard Deviation (σ) (m³/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	985.86	1052.64	1.06	1.82	857.54	1125.02	1.31	2.22
	Regression	976.74	937.07	0.95	1.82	858.29	999.86	1.16	2.24
	ANN	1010.13	968.69	0.96	1.84	819.27	1003.17	1.22	2.02
Kurundwad	Observed	1430.74	1374.19	0.96	1.20	1396.68	1568.61	1.12	1.71
	Regression	1428.54	1264.79	0.88	1.19	1384.35	1438.03	1.03	1.73

	ANN	1533.63	1331.54	0.87	1.18	1315.99	1410.02	1.07	1.97
Terwad	Observed	607.90	538.52	0.88	1.26	548.10	499.59	0.91	1.58
	Regression	603.23	495.45	0.82	1.26	544.05	457.33	0.84	1.61
	ANN	615.11	496.57	0.81	1.25	558.67	461.83	0.83	1.39
Sadalga	Observed	404.34	361.68	0.89	1.45	263.83	311.07	1.17	2.08
	Regression	404.33	325.64	0.80	1.45	276.45	278.80	1.00	2.10
	ANN	403.54	329.05	0.82	1.32	273.97	279.48	1.02	1.93

Month: August

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Arjunwad	Observed	1104.86	1050.45	0.95	1.76	1194.78	1777.80	1.48	2.54
	Regression	1102.96	958.12	0.86	1.75	1185.82	1617.19	1.36	2.55
	ANN	1102.62	951.08	0.86	1.76	1139.99	1632.00	1.43	2.67
Samdoli	Observed	469.48	388.24	0.82	1.33	417.92	532.12	1.27	2.29
	Regression	464.87	343.23	0.73	1.34	421.63	469.78	1.11	2.26
	ANN	456.42	342.15	0.75	1.42	422.23	487.81	1.16	2.61

Month: September

Gauging Site	Training Period					Testing Period			
	Data	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)	Mean (μ) (m^3/s)	Standard Deviation (σ) (m^3/s)	Coefficient of Variance (C_v)	Coefficient of Skewness (C_s)
Karad	Observed	252.61	376.23	1.48	4.33	175.24	271.18	1.54	4.82
	Regression	251.27	370.18	1.47	4.32	183.20	250.19	1.36	4.87
	ANN	259.58	380.26	1.46	4.18	181.49	257.12	1.41	5.04
warunji	Observed	150.68	257.27	1.70	4.32	105.48	174.21	1.65	4.85
	Regression	153.25	211.55	1.38	4.58	112.73	151.34	1.34	5.1
	ANN	150.65	243.56	1.61	4.04	97.85	180.01	1.83	4.81
Arjunwad	Observed	1107.22	1052.35	0.95	1.75	1195.82	1775.52	1.48	2.54
	Regression	1093.92	1014.39	0.92	1.61	1082.04	1557.01	1.43	2.45
	ANN	1082.81	1013.58	0.82	1.55	1098.51	1688.42	1.32	2.31
Kurundwad	Observed	468.13	668.55	1.42	3.63	577.94	727.04	1.25	2.5
	Regression	467.55	628.44	1.34	3.63	570.77	683.42	1.19	2.5
	ANN	464.41	614.68	1.32	3.8	514.68	661.34	1.28	3.17

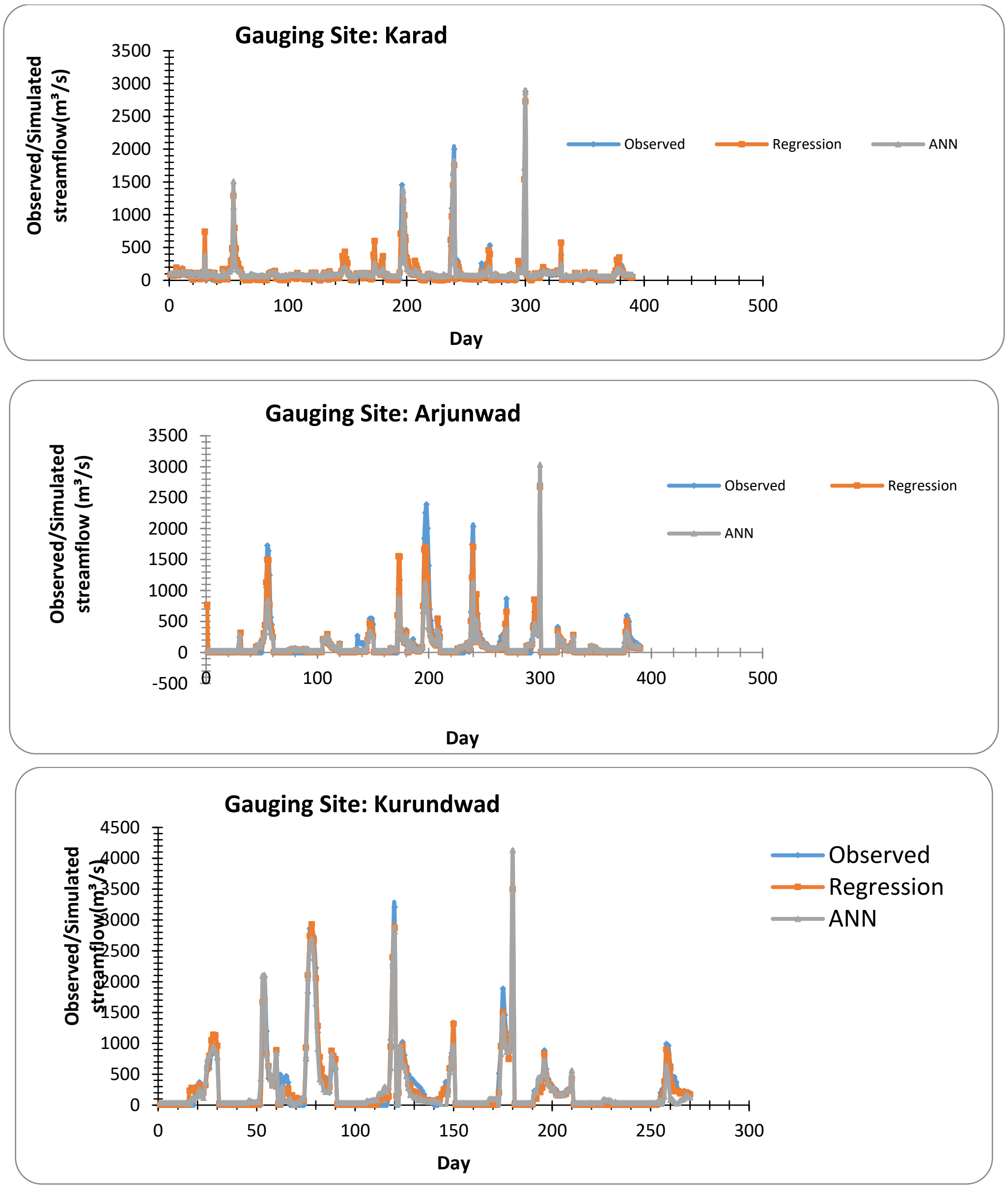
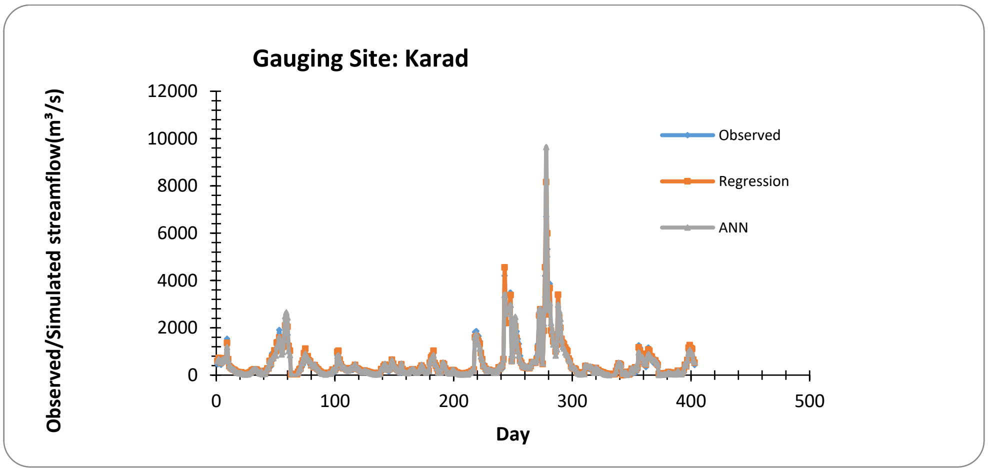


Fig.4.1 Comparison plots of Observed and Simulated Streamflows for Regression and ANN Models (Month: June)



International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

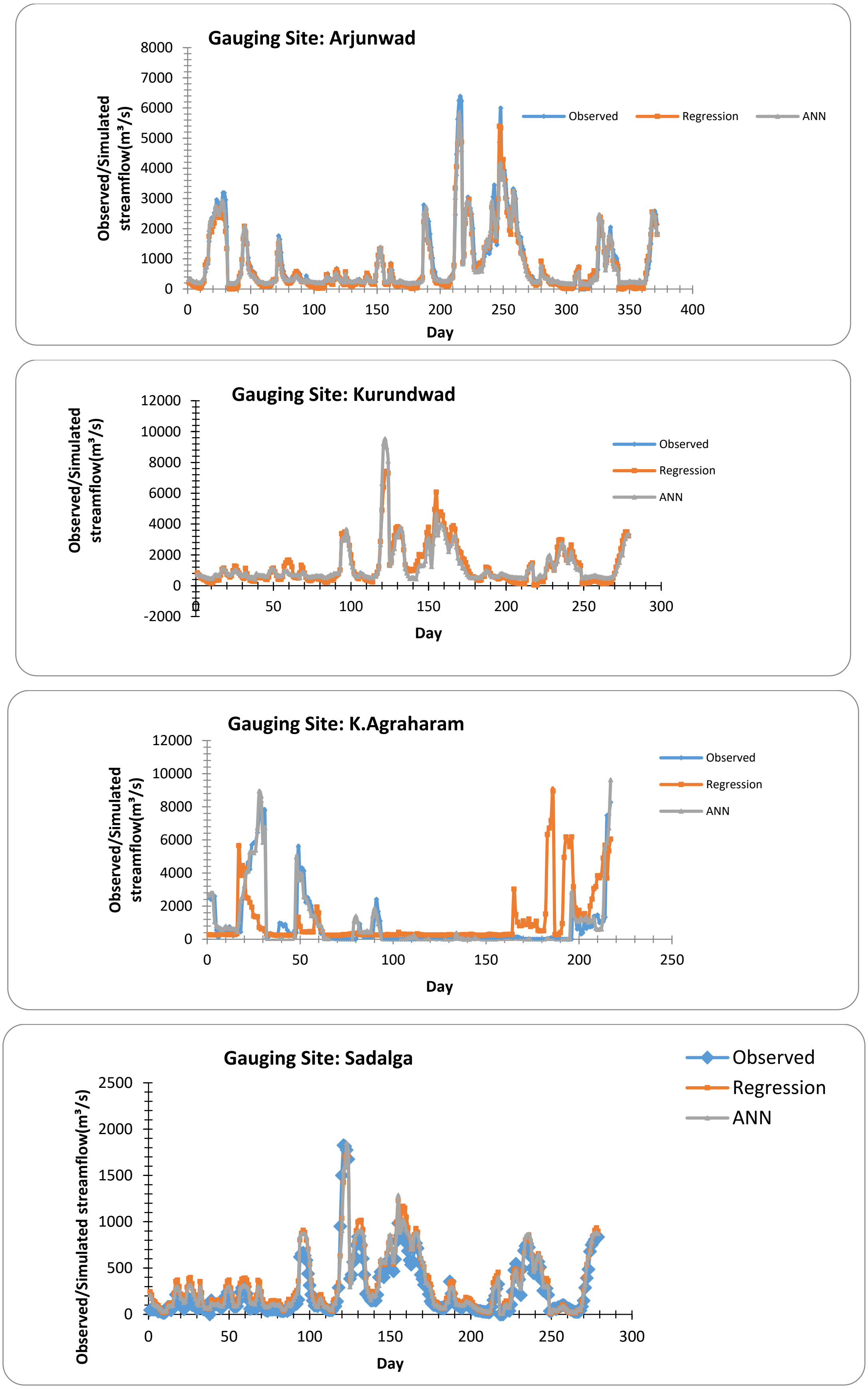
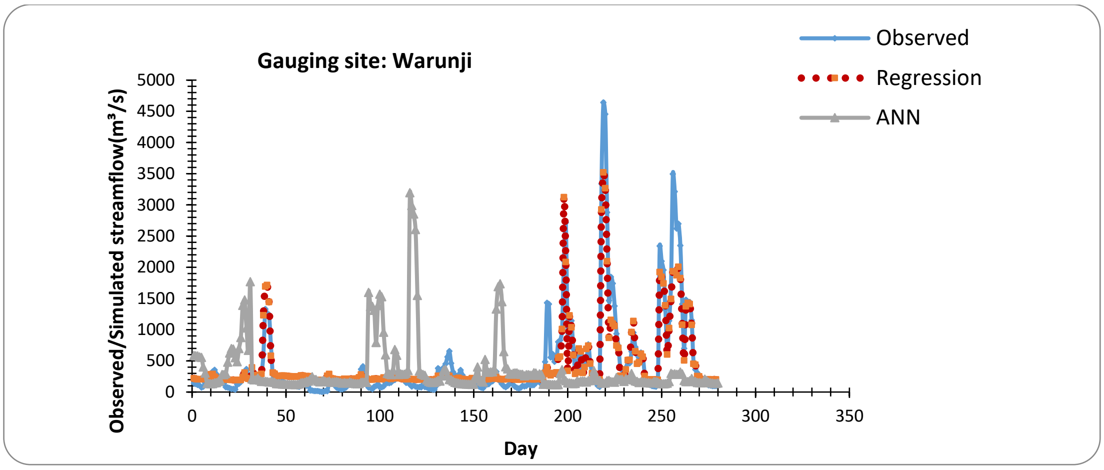
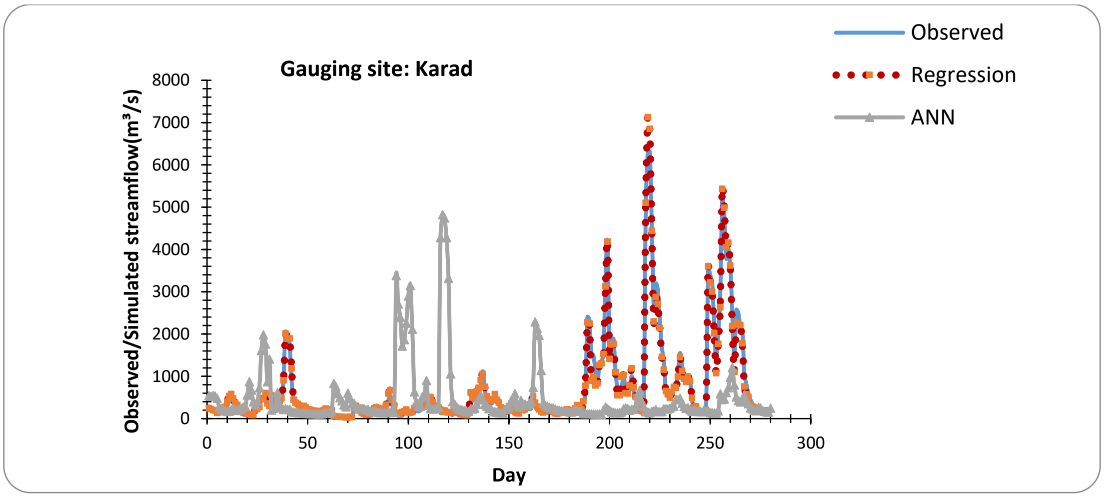
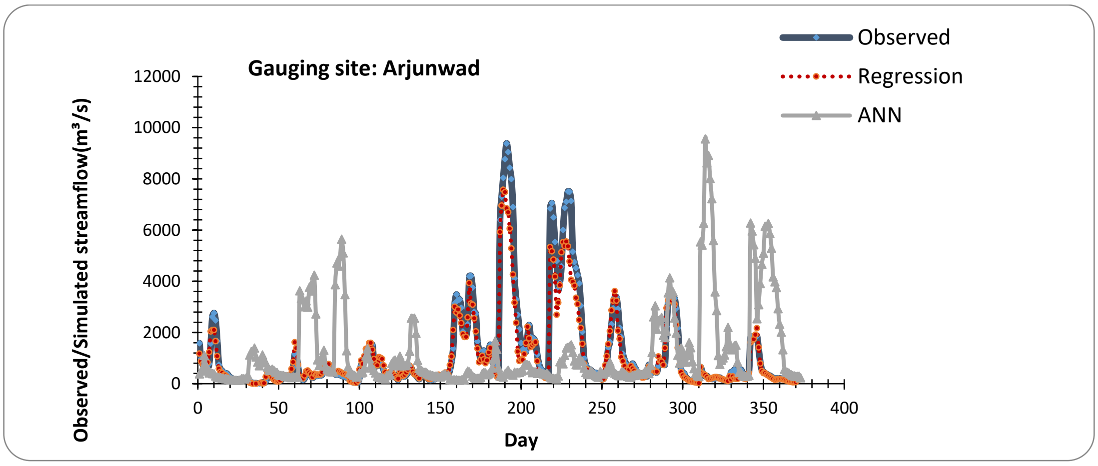


Fig. 4.2 Comparison plots of Observed and Simulated Streamflows for Regression and ANN Models (Month: July)

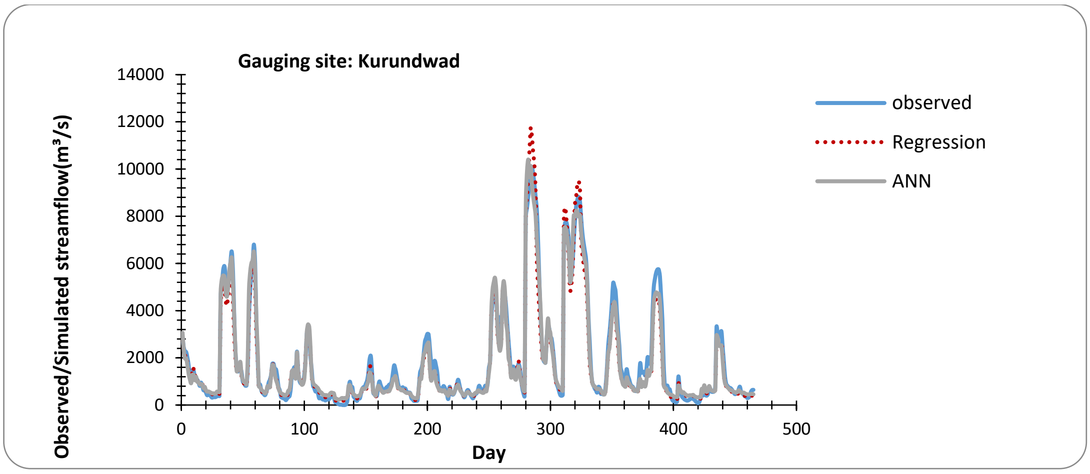
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026







International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026



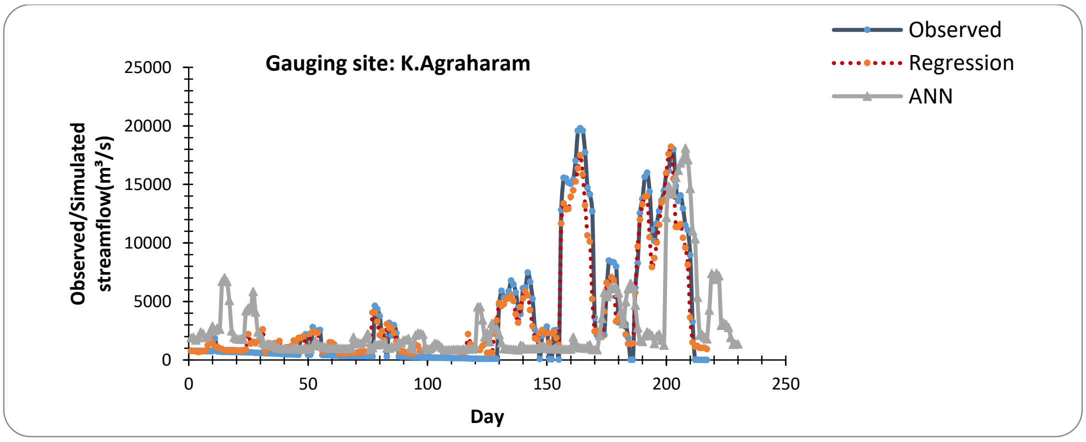
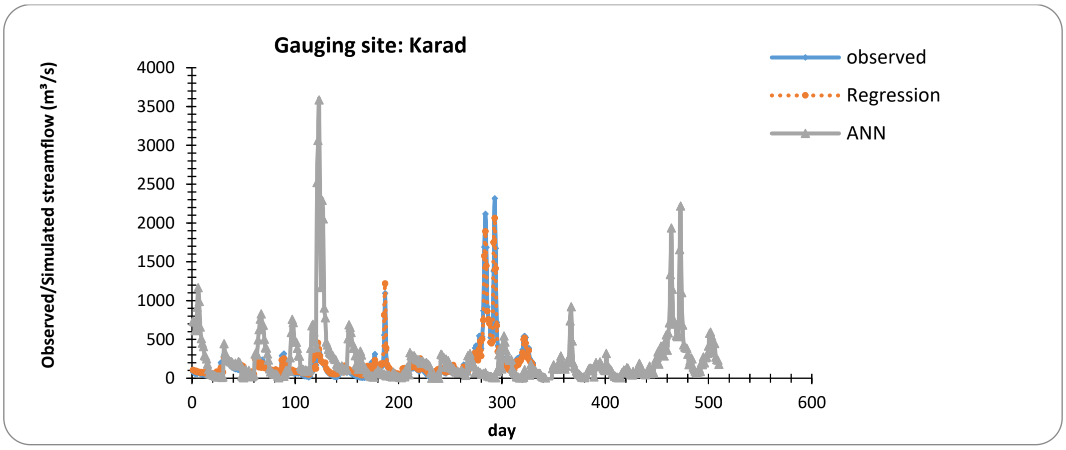


Fig. 4.3 Comparison plots of Observed and Simulated Streamflows for Regression and ANN Models (Month: August)



International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

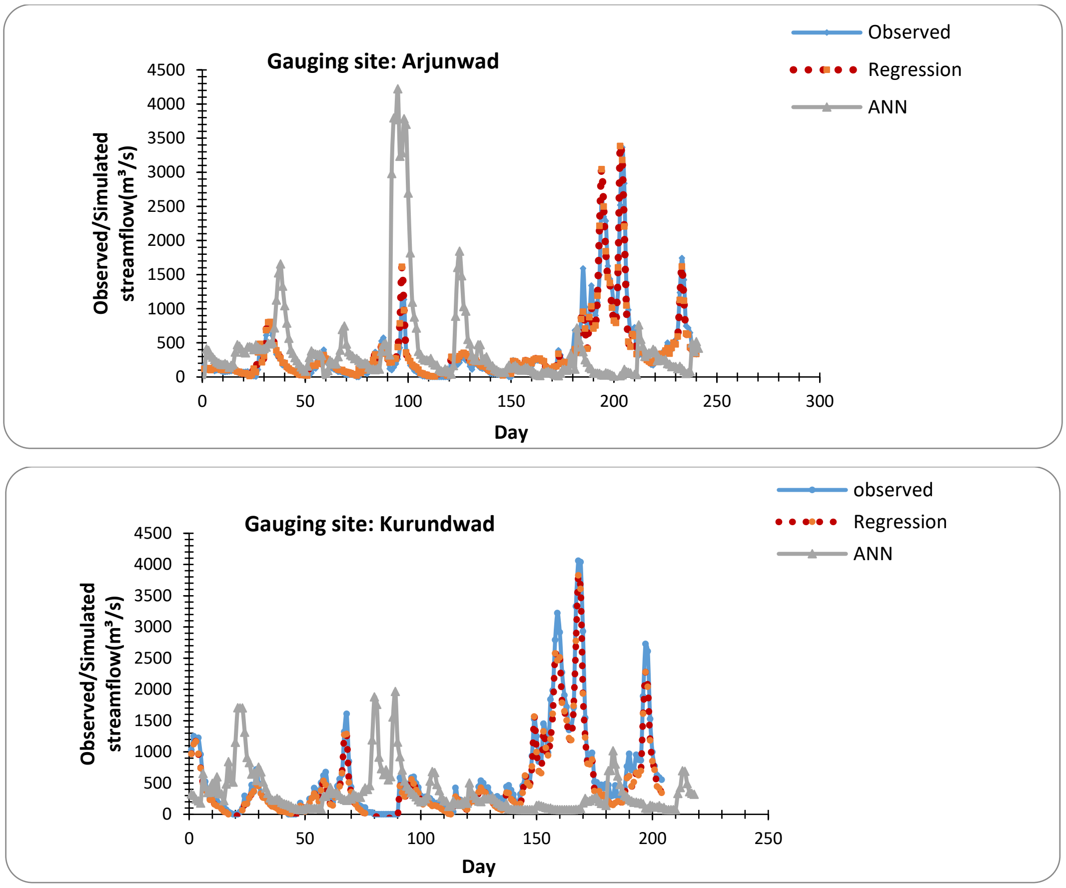
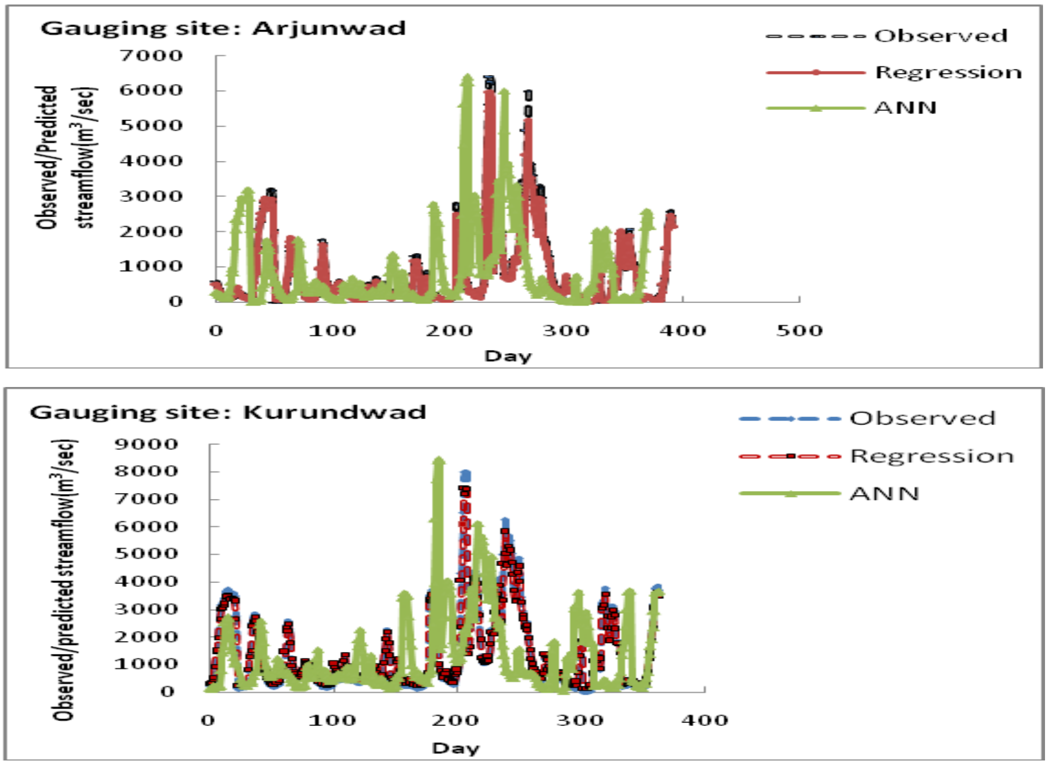


Fig. 4.4 Comparison plots of Observed and Simulated Streamflows for Regression and ANN Models (Month: September)



International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

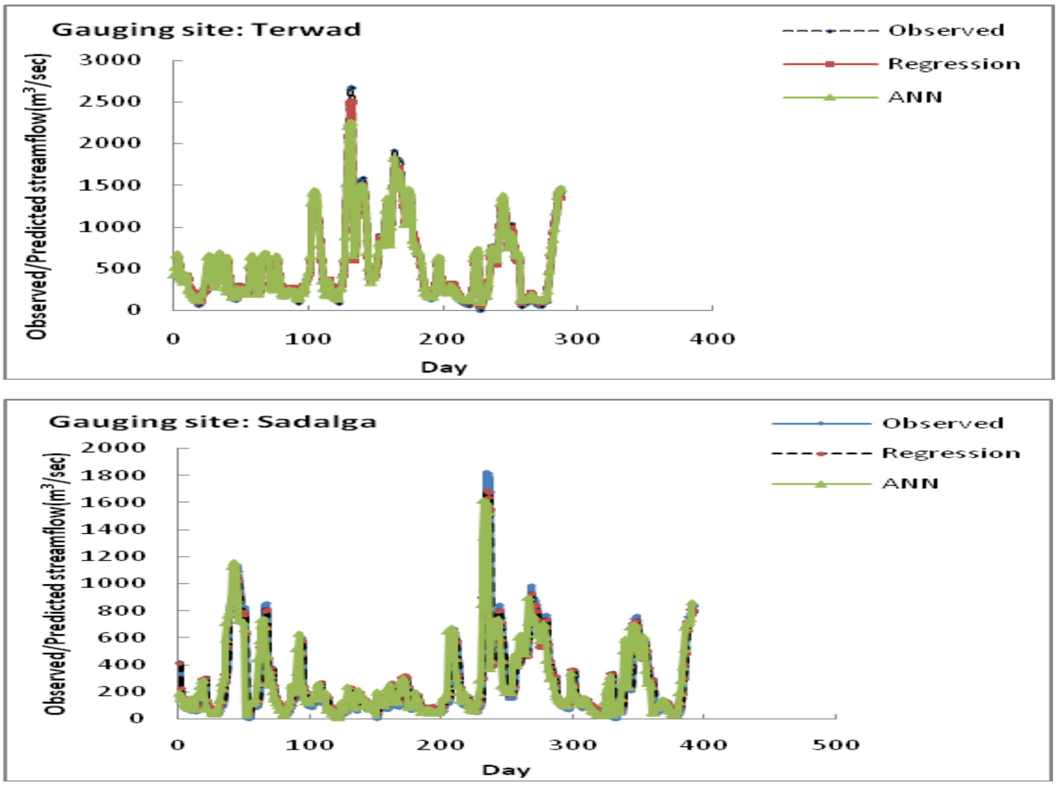


Fig. 4.5 Comparison plots of Observed and Predicted Streamflows for Regression and ANN Models (Month: July)

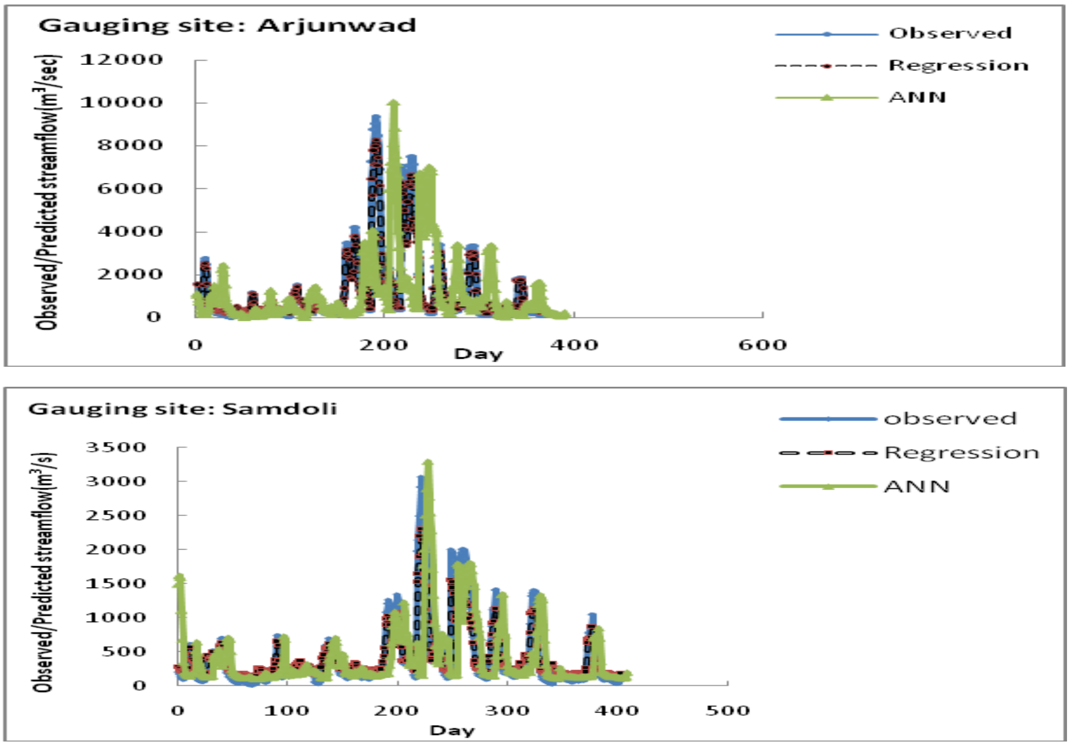


Fig. 4.6 Comparison plots of Observed and Predicted Streamflows for Regression and ANN Models (Month: August)

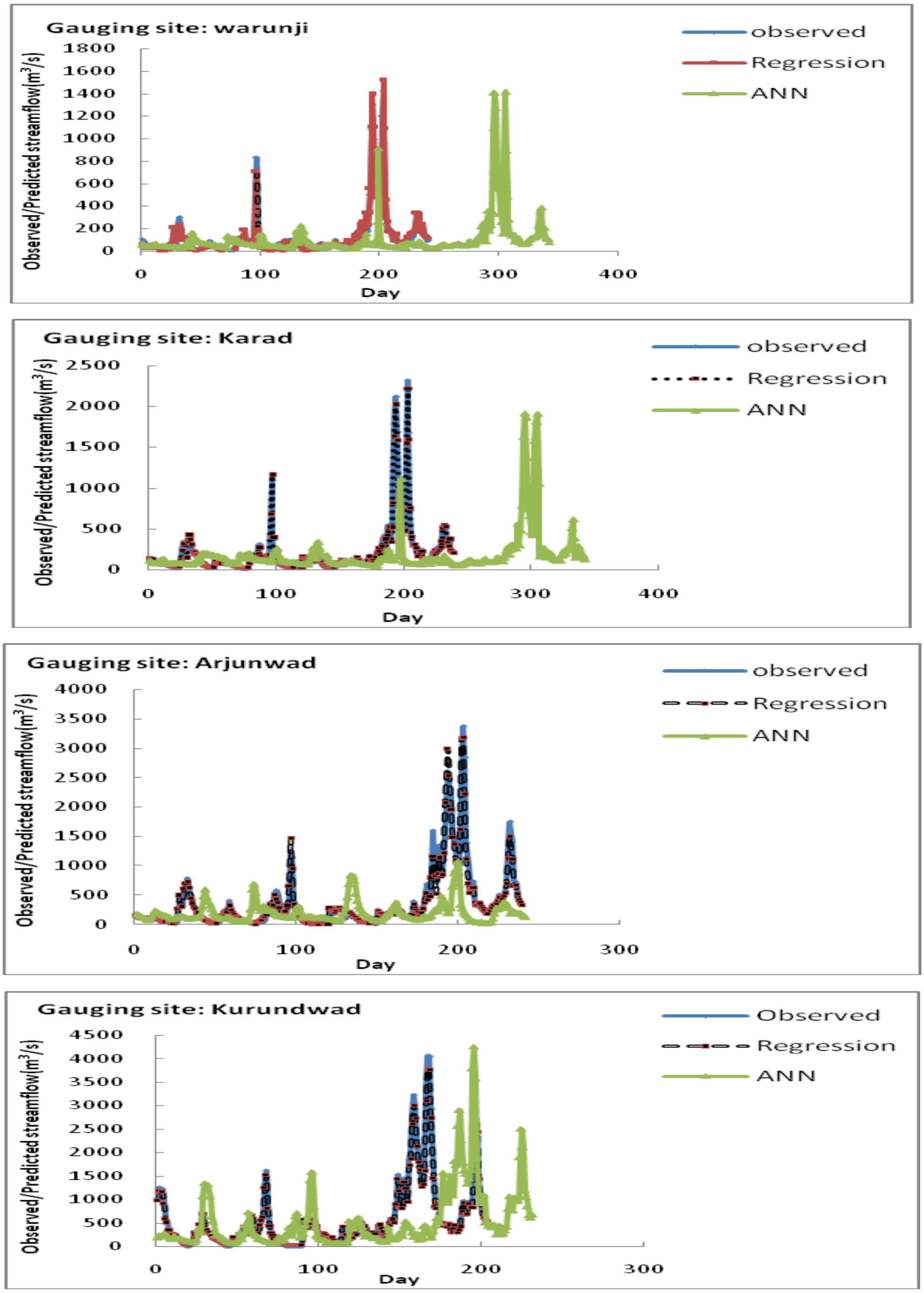


Fig. 4.7 Comparison plots of Observed and Predicted Streamflows for Regression and ANN Models (Month: September)

5. SUMMARY AND CONCLUSIONS

Modeling of streamflow is essential for the development, regulation and efficient management of water resources for different activities such as irrigation, water supply, navigation, hydropower development, land drainage and flood control. Simple linear regression techniques are adopted for streamflow modelling. Optimal Neural Network architectures with appropriate activation function are employed in the development of ANN models. Present study has focused on the development, comparison of simple regression and artificial neural network streamflow for the gauging sites of Krishna basin. Streamflow data collected from

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

the Irrigation Departments and Central Water Commission are used in the development of models. Data are divided into training and testing sets based on the statistical structure. Performance of the models is evaluated through numerical and graphical performance indicators. Numerical indicators suitable for the present study are selected as the coefficient of determination, root mean square error, efficiency coefficient and volumetric error. Scatter and comparison plots are adopted as graphical indicators. It is also verified whether the statistical structure of historical data in terms of mean, standard deviation and skewness coefficient is preserved in the simulated or predicted data. Based on the development, comparison of regression and artificial neural network models in the modelling of streamflow for the gauging sites of Krishna basin, the following conclusions are drawn.

Regression Modelling: The streamflow inputs influencing the flows at the gauging sites are identified through cross correlation analysis. Streamflow regression models based on the flows of the upstream gauging sites are developed for both simulation and forecasting modes. Linear regression models show fairly good performance in the simulation mode and their performance in the forecasting mode is observed to be unsatisfactory.

Artificial Neural Network Modelling: Optimal architecture for the streamflow modelling at some of the gauging sites both in simulation and forecasting modes exhibits satisfactory performance.

REFERENCES

[1]. Anctil, F., and Rat, A. (2005) “Evaluation of neural network stream flow forecasting on 47 watersheds.” Journal of Hydrologic Engineering, Vol. 10, pp. 85-88.
[2]. Awwad, M. Haitham, Juan B. Valdes, and Pedro J. Restrepo. (1994) “Streamflow forecasting for Han River Basin, Korea.” Journal of Water Resources Planning and Management, Vol.120, No.5, pp.651-673.
[3]. Birkundavyi, S., Labib, R., Trung, H.T., and Rousselle, J. (2002) “Performance of neural networks in daily streamflow forecasting.” Journal of Hydrologic Engineering, Vol.6, pp.553-555.
[4]. Jain Sharad, K., and Sudheer, K.P., (2008) “Fitting of hydrologic models : A close look at the Nash-Sutcliffe index”. Journal of Hydrologic Engineering, Vol.13, No.10, pp.981-986.
[5]. Jyothiprakash, V., Devamane, M.G., Mohan, S. (2006) “Comparative Study on Multisite Streamflow Generation Model hec-4 and ANN Model” Journal of Institution of Engineers (India), Vol. 87, pp. 9-14.
[6]. Kang, K. W., Kim, J. H., Park, C. Y., and Ham, K.J. (1993), “Evaluation of hydrological forecasting system based on neural network model”. Proc., 25th Congress of Int. Assoc. for Hydr. Res., International Assciation for Hydraulic Research, Delft, The Netherlands 257-264.
[7]. Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K. (1994) “Neural Networks for river flow prediction”, J. Computing in Civil Engineering, ASCE, Vol.8 (2), pp.201-220.
[8]. Markus, M., Salas, J.D., and Shin, H, K. (1995). “Predicting streamflows based on neural networks”, Proc. 1st Int. Conf. on Water Resour. Engg., ASCE, New York, 1641-1646.
[9]. Muttaih, R.S., Srinivasan, R., and Allen, P.M. (1997). Prediction of two-year peak stream discharges using neural networks”, J. Am. Water Resour. Assoc., 33(3), 625-630.
[10]. Nagesh kumar, D., Srinivasa Raju, K., and Sathish, T. (2004) “River flow forecasting using recurrent neural networks.” Journal of Water resources management, Vol.18, pp.143-161.
[11]. Nanda, S.K. Tripathy, D.P. Nayak, S.K. and Mohapatra, S. (2013). “Prediction of rainfall in India using artificial neural network models”, Journal of Intelligent Systems and Applications, Vol. 12, pp. 1-22, 2013.
[12]. Nilsson, P., Cintia, B.U., and Berndtsson, R. (2006) Monthly runoff simulation: “Comparing and combining conceptual neural network models.” Journal of Hydrology, Vol.321,pp. 344-363.
[13]. Pallavi, M. Swaptik, C. Sangeeta, R. Nikhil B. and Roshan, S. “Dual artificial neural network for rainfall-runoff forecasting, Journal of Water Resource and Protection, Vol. 4, pp. 1024-1028, 2012.
[14]. Poff, N.L., Tokar, S., and Johnson, P. (1996). “Steam hydrological and ecological responses to climate change assessed with an artificial neural network”, Limnol. And Oceanog., 41(5), 857-863.
[15]. Raman, H. and Sunilkumar, N. (1995) “Multivariate modeling of water resources time series using artificial neural networks”. Journal of Hydrological Sciences, Vol.40, pp.145-163.
[16]. Sudheer, K.P. Srinivasan, K. Neelakantan, T.R. and Srinivas, V. “A nonlinear data-driven model for synthetic generation of annual streamflows”, Journal of Hydrological Processes, Vol. 22, pp. 1831-1845, 2008.
[17]. Suribabu, C.R., Ramyya, V.V., and Kumar, R.P. (2005) “Application of Artificial neural network model for inflow prediction- a case study.” Journal of Water resources society Vol. 25,No.4,pp.-24.
[18]. Thirumalaiah, K., and Deo, M.C. (1998). “River stage forecasting using artificial neural networks”. J. Hydrologic Engrg., ASCE, 3(1),26-32.
[19]. Vedula, S and Reddy, D.V. (1981) “Stochastic models for monthly streamflows-A comparison of case study”. Journal of Institution of Engineers (India), Vol. 61, pp.271-276.