ISSN:0976-4933
Journal of Progressive Science
Vol.09, No.01 & 02, pp 24-34 (2018)

Autoregressive time series modeling for prediction of rainfall, runoff and silt load in Jharkhand state of central India

Shivi Garg, Munendra Pal[1] and Arpan Sherring [2]
Department of Soil and Water Conservation Engineering, SHIATS, Prayag Raj, India
[2]Department of Irrigation and Drainage Engineering, SHIATS, Prayag Raj, India
[1]Department of Soil Conservation, Shri Murli Manohar Town PG College, Ballia, India
Email- munendapal@gmail.com

Abstract

India is characterized by marked difference in climate and catchment properties across the country. The study was autoregressive time series model, for prediction of annual rainfall, runoff and silt load in karkara nala watershed Barakar catchment Jharkhand. The area of the watershed is 1751 ha and the data of 9 years from 1993 to 2001 were used to develop the model. The goodness of fit and adequacy of models were tested by Box-Pierce Portmonteau test, Akaike Information Criterion and by comparison of measured and predicted data correlogram. The graphical representation between measured and predicted correlogram. Where in the cases of rainfall, runoff and silt load there is a very close agreement between them. The lower values of error indicate the adoptability of the model for forecasting of rainfall, runoff and silt load. The comparison of different developed models, statistical characteristics and graphical representations, Autoregressive model AR (1) is proposed for generation of silt load and AR (2) is proposed for generation of rainfall and runoff in karkara nala watershed Jharkhand, Clearly shows that the developed model can be use efficiently for the prediction of rainfall, runoff and silt load for the Damodar catchment which is benefit for the farmers and research workers for water harvesting, ground water recharge, flood control and development of their water management strategies.

Keywords: Autoregressive, Time series, Barakar Catchment, Rainfall, Runoff, silt load

Introduction

The Indian climate comprises of a wide range of weather condition across varied topology and large topographical area. The agriculture and the other allied activities and in turn the prosperity and economic growth of a country depend on the soil and water resources. About 80% of the world's population (5.6 billion in 2011) lives in areas with threats to water security. The water security is a shared threat to human and nature and it is pandemic. Human water-management strategies can affect detrimentally to wildlife, such as migrating fish. Fresh water today is a scarce resource, the reality of water crisis cannot be ignored (Nature, 2010). Climatic factors such as rainfall and surface temperature determine the availability of moisture for physical, biological and chemical activities in plants that ultimately lead to a healthy plant (Houghton *et al*., 2001) gowth. A watershed is an area from which runoff, resulting from precipitation flows past a single point into a water body. A watershed may be a few hectares as in the case of a small stream or hundreds or square kilometers as in the case of a large river in the watersheds of large reservoirs, it becomes impractical to plan the soil conservation work treating the entire area as a single hydrological unit due to the limited availability of resources. In hydrological modeling, the main theme is to determine the deposition of rainfall, how much of it becomes runoff, infiltration, ground water recharge, evaporation and water storage. A problem is encountered, in that the data have three dimensions, spatial, time, and magnitudional. Modeling efforts are generally hampered by limitations in the representation of these three dimensions. It is most common to treat watersheds as lumped systems by spatially averaging the properties in order to accommodate them in programming languages. Most hydrological models are based on a set of underlying assumptions about time, space and randomness (Maidment, 1993). The stochastic time series models, such as Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) are widely used to predict the annual runoff. These models are suitable for prediction of annual, monthly or daily rainfall as well as short term prediction. Autoregressive (AR) model of order 0, 1 and 2 proposed by Kottegoda and Horder (1980). Iyenger (1982), Rai and Sherring (2007), Sherring *et al.* (2009) and Stadnytska *et al.* (2008) reported that a developed autoregressive model is applicable to various Indian climatic regions such as south Kashmir. First part of the study presents development of AR time series model of order of 0, 1 and 2 to predict the Autoregressive time series model for prediction in barakar catchment for rainfall, runoff and siltload for Karaka nala watershed of Jharkhand. Second part involved in comparing the predicted and measured rainfall, runoff and siltload to evaluate the performance of developed model.

Materials and methods

This watershed lies in the eastern part of the Jharkhand in between 24°13′50''(N) longitude and 35°12'20'' (E) latitude with geographical area about 1751 ha. The maximum temperature during peak summer goes up to 46°C and minimum temperature drops down to 40°C. The wind velocity was normally 4.5 km per hour except in summer which was recorded 25 to 30 km per hour. The general slope of land is 1-5%. The average elevation varies from 150-200 m. The average annual rainfall is 967.2 mm. Soil conservation works in the upper Damodar-Barakar Catchment areas through multidisciplinary and integrated watershed management programmes. Rainfall, Runoff and Siltload data was collected from the Soil Conservation Department, Barakar catchment Jharkhand of Hazaribag District from the duration 1993 to 2001.

Stochastic Time Series Model

A mathematical model representing stochastic process is called stochastic time series model. A sample time series model could be represented by simple probability distribution f (X:Θ) with parameters $\Theta = (\Theta_1, \Theta_2, \ldots)$ valid for all positions t = 1,2,….N and without any dependence between $X_1, X_2, \ldots X_n$.

A time series model with dependence structure can be formed as

$$\varepsilon_t = \phi\, \varepsilon_{t-1} + \eta_t \tag{1}$$

where η_t is an independent series with mean zero and variance $(1-\phi^2)$, ε_t is dependent series and ϕ is the parameter of the model.

Time series modeling can be organized in five stages i.e. identification of model composition, Selection of model type, identification of model form, estimation of model parameters and testing of goodness of fit for validation of the model. (Salas and Smith, 1986)

Autoregressive (AR) Model

In the Autoregressive model, the current value of a variable is equated to the weighted sum of a pre assigned no. of part values and a variate that is completely random of previous value of process and error. The pth order Autoregressive model, AR (p) is estimated by following equation.

$$Y_t = \overline{Y} + \sum_{j=1}^{p} \Phi_j\,(Y_{t-j} - \overline{Y}) + \varepsilon_t \tag{2}$$

Where, Y_t is the time dependent series (variable), ε_t is the time dependent series which is independent of Y_t and is normally distributed with mean zero and variance σ_ε^2, $\overline{Y}$ is the mean of annual flow and rainfall data and $\Phi_1, \Phi_2, \ldots \Phi_p$ are the Autoregressive parameter

Estimation of Autoregressive parameter (Φ)

$$AR\ (1) : \Phi_1 = \frac{D_{1,2}}{D_{2,2}} \tag{3}$$

$$AR(2):\Phi_1 = \frac{D_{1,2}D_{3,3} - D_{1,3}D_{2,3}}{D_{2,2}D_{3,3} - D^2{}_{2,3}} \tag{4}$$

$$\Phi_2 = \frac{D_{1,3}D_{2,2} - D_{1,2}D_{2,3}}{D_{2,2}D_{3,3} - D^2{}_{2,3}} \tag{5}$$

For estimation of model parameters, method of maximum likelihood was used (Box and Jenkins, 1970)

$$Di, j = \frac{N-}{N+j-i-j} \sum_{i=0, j=0}^{N+i-(i+j)} Z_{i+l}\, Z_{j+l}, \tag{6}$$

where D is the difference operator, n is the sample size, i & j maximum possible order and l is the autocorrelation function.

The autocorrelation function r_k of of lag k was estimated as proposed by Kottegoda and Horder (1980)

$$r_k = \frac{\sum_{t=1}^{N-K}(Y_t - \overline{Y})(Y_{t+k} - \overline{Y})}{\sum_{t=1}^{N}(Y_t - \overline{Y})}, \tag{7}$$

where r_k is the Autocorrelation function of time series ($Y_{t)}$ at lag k, Y_t is stream flow series (observed data), $\overline{Y}$ is mean of time series Y_t , N are the total number of discrete values of time series ($Y_{t)}$. The autocorrelogram was used for identifying the order of the model for given time series as well as for comparing the sample correlogram with model correlogram. The 95% probability levels for the autocorrelation function was estimated (Anderson, 1942).

Partial Autocorrelation function

The partial autocorrelation PC_{kk} was calculated to identify both the type and order of the model. (Durbin, 1960**).**

$$PC_{k,k} = \frac{r_k - \sum_{j=1}^{k-1} PC_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} PC_{k-1,j}\, r_j}, \qquad \text{...... (8)}$$

where PC_{kk} is the Partial autocorrelation function at lag K and r_k is the autocorrelation function at lag K. The 95 percent probability limit for partial autocorrelation function was also calculated (Anderson 1942).

Parameter estimation of AR (p) model

The mean ($\overline{Y}$) and variance (σ^2_ε) of the observed data Y_t was estimated. After computation of $\overline{Y}$ and σ^2_ε, the remaining parameters Φ_1, Φ_2,…. Φ_p of the AR models were estimated by

$$Z_t = Y_t - \overline{Y}, \text{ for } t = 1, 2, \ldots N \quad (9)$$

The parameters Φ_1, Φ_2 ,…. Φ_p were estimated by solving the p system of linear equations (Yule and Walker, 1927)

Statistical characteristics

The MFE, MAE, MRE, RMSE and ISE. Were estimated to evaluate the performance of the model for prediction of annual rainfall and runoff and siltload.

Goodness of fit of autoregressive (AR) model

The goodness of fit tests of AR model fitted to annual hydrologic series was accomplished by testing the residuals of a dependence model for co-relation and the order of the fitted model is compared with the order of the dependence model. The Box-Piece Portmonteau lack of fit test and Akaike Information Criterion (Akaike, 1974) was done for checking the order of the fitted model and its adequacy.

Results and discussion

Identification of model

The selected models were using Autocorrelation and partial autocorrelation for identification of the proper type and order of the models. The annual rainfall and runoff and siltload series was modelled through the autoregressive model. The various steps involved in are preliminary analysis and identification, estimation of parameters and diagnostic checking for the adequacy of the selected models (Sharma *et al.,* 2003). The autocorrelation functions and partial autocorrelation functions were determined for the 95 percent probability limits (Anderson, 1942). The autocorrelation functions and partial autocorrelation functions with 95 percent probability limits upto12 lags of the series were computed and the autoregressive model of first order, AR(1) model was selected for further analysis.

Models of Autoregressive (AR) Family

The parameters of AR models were computed for annual rainfall and runoff and siltload. The predicted values of annual rainfall and runoff and siltload were computed with the observed values. It was observed that AR (p) model upto order 2 has shown the good fit and correlation between the observed and predicted values and given in Figs. 1, 2 and 3.

AR (p) models for the prediction of Rainfall

AR (1): $Yt = 967.2-0.04624\ (Y_{t-1}-967.2) + \varepsilon_t$

AR (2): $Yt = 967.2-\ 0.04624\ (Y_{t-1}-967.2) +.00037\ (Y_{t-2}-967.2) + \varepsilon_t$

AR (p) models for the prediction of Runoff

AR (1): $Yt = 510.40-0.32581\ (Y_{t-1}-510.40) + \varepsilon_t$

AR (2): $Yt = 510.40-\ 0.32581\ (Y_{t-1}-510.40) +0.00365\ (Y_{t-2}-510.40) + \varepsilon_t$

AR (p) models for the prediction of Siltload

AR (1): $Yt = 4.63 +0.9962(Y_{t-1}-4.63) + \varepsilon_t$

AR (2): $Yt =4.63 +0.9962\ (Y_{t-1}-4.63) -0.6832(Y_{t-2} -4.63) + \varepsilon_t$

Box Pierce Portmonteau test for AR model

The Box-Pierce Portmonteau lack of fit test to check the adequacy of autoregressive models for both annual rainfall, runoff and siltload for AR(0), AR(1) and AR(2) models were estimated . The results revealed that all three models viz. AR (0), AR(1) and AR(2) were giving good fit and were acceptable.

Akaike Information criterion test

Akaike Information Criterion (AIC) for all three models were estimated and values of AIC for annual rainfall and runoff and siltload are given in table 1, 2 and 3. Results shows that AIC value of AR(1) in all three cases are lying in between AR(2) model and AR(0) model which was suitable model for further prediction of rainfall, runoff and siltload.

Comparison of the observed and predicted annual rainfall, runoff and siltload

The correlogram of observed and predicted series for annual rainfall, runoff and siltload were developed by plotting autocorrelation functions against lag K. A graphical comparison of observed and predicted annual rainfall, runoff and siltload with the selected model are shown in Figurs. The graphical representation of the data shows a closer agreement between observed and predicted annual rainfall, runoff and siltload by selected model. It reveals that developed model for rainfall, runoff and siltload can be utilized for the prediction of future trends with minimum chance of error.

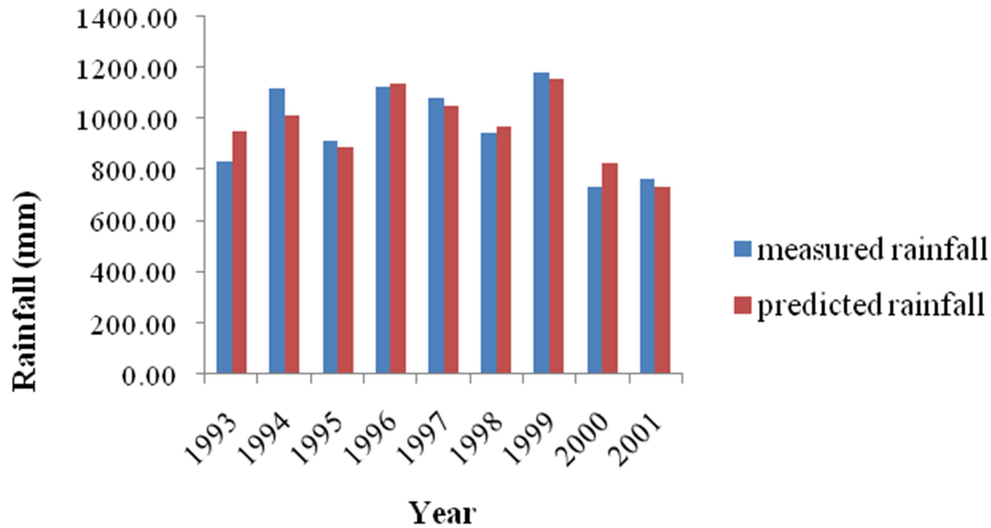


Fig-1: Comparison of correlogram of measured and predicted series for Rainfall

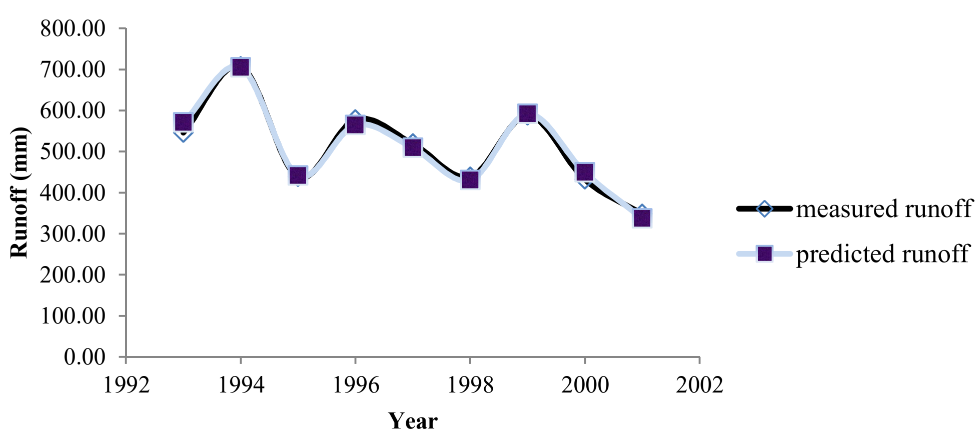
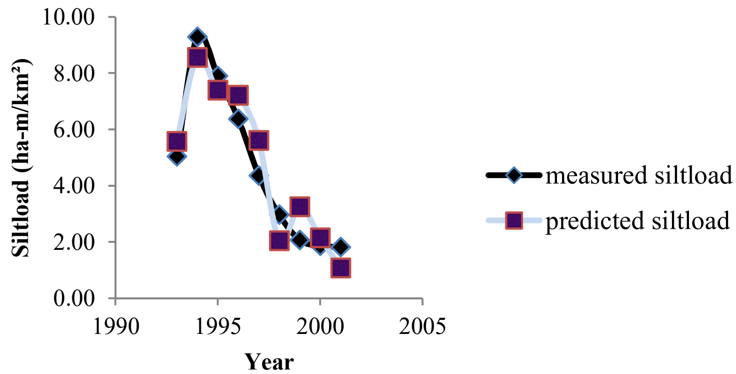


Fig-2: Comparison of correlogram of measured and predicted series for Runoff

The mean, standard deviation and skewness of observed and predicted data were calculated to evaluate the fitting of the model in moment preservation. The results are tabulated in Table 3.

Fig-3: Comparison of correlogram of measured and predicted series for siltload statistical Characteristics of Data



The results show that the skewness of predicted data by AR (1) model and observed data is lying between -1 to +1 and therefore AR (1) model preserved the mean and Skewness better.

Performance Evaluation of AR (1) model for prediction of annual rainfall, runoff and silt load-

The statistical characteristics such as MFE, MAE, MRE, MSE, RMSE and ISE were also used to test the adequacy of the model for future prediction with higher degree of correlation to previous measured observations. The statistical error for rainfall and siltload by using the AR (1) model. All the errors are less. It indicates that autoregressive model AR (2) is suitable for rainfall, AR (1) is suitable for runoff and AR (1) is suitable for siltload for the prediction in karkara nala watershed. The co-relation between measured and predicted values for rainfall ($R^2 = 0.839$), runoff ($R^2 = 0.985$) and siltload ($R^2 = 0.904$) lies between 0 to +1.

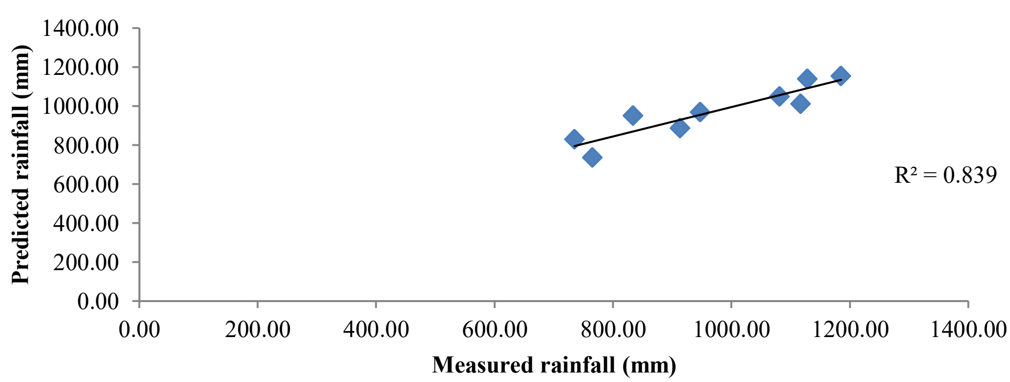


Fig-4: Comparison between measured and predicted annual rainfall of Karkara nala watershed of Damodar catchment in Jharkhand.

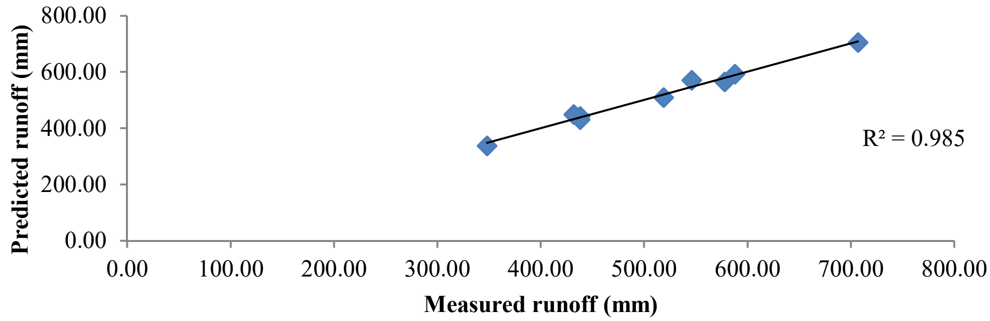


Fig-5: Comparison between measured and predicted annual runoff of Karkara nala watershed of Damodar catchment in Jharkhand.

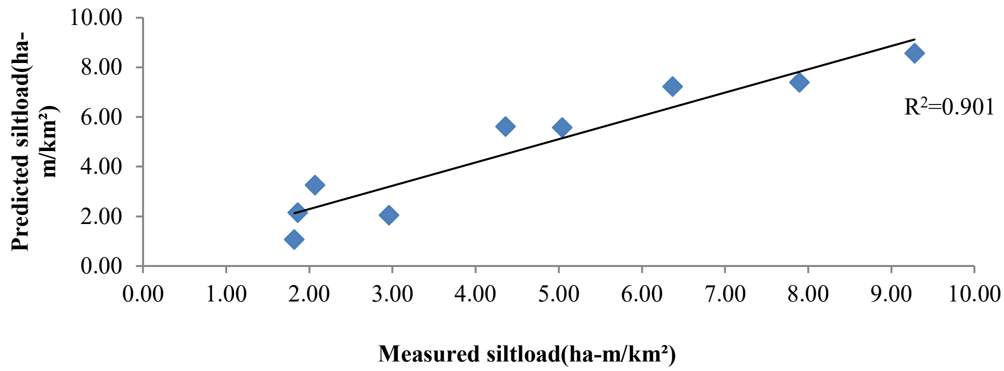


Fig-6: Comparison between measured and predicted annual siltload of Karkara nala watershed of Damodar catchment in Jharkhand

Table 1. Statistical parameters of autoregressive (AR) model for Rainfall

Model	AR(0)	AR (1)	AR (2)
Autoregressive parameter	-	Φ1= -0.04625	Φ1= -0.04624 Φ2= 0.00037
White noise variance	151.59	160.62	151.43
Akaike Information Criteria	90.38	93.42	94.36
Value of port monteaue statistics	1.05435	0.98685	0.872775
Degree of freedom upto 5 lags	5	4	3

Table 2. Statistical parameters of autoregressive (AR) model for Runoff

Model	AR(0)	AR (1)	AR (2)
Autoregressive parameter	-	Φ_1= -0.32546	Φ_1= -0.32518 Φ_2 = 0.003652
White noise variance	101.28	97.85	103.79
Akaike Information Criteria	83.122	85.56	86.50
Value of port monteaue statistics	2.57892	2.49778	1.56352
Degree of freedom upto 5 lags	5	4	3

Table 3. Statistical parameters of autoregressive (AR) model for Siltload

Model	AR(0)	AR (1)	AR (2)
Autoregressive parameter	-	Φ_1=0.9961	Φ_1= 1.4526 Φ_2 = -0.6832
White noise variance	2.588	1.535	1.863
Akaike Information Criteria	17.117	9.717	15.205
Value of port monteaue statistics	8.38728	8.311437	7.875954
Degree of freedom upto 5 lags	5	4	3

Table 4. Evaluation of Regeneration performance with Statistical errors

Sr. No.	Autoregressive AR (1) model		
Statistical error			
Rainfall (mm)		Runoff (mm)	Siltload (ha-m/km^2)
1. Mean Forecast Error	2.833058	0.4850944	0.131018
2. Mean Absolute Error	-2.83306	-0.485094	-0.131018
3. Mean Relative Error	-0.02673	-0.008726	-0.301728
4. Mean Square Error	72.23595	2.1178496	0.154493
5. Root Mean Square Error	8.499173	1.4552833	0.393056
6. Integral Square Error	0.002929	0.0009503	0.02830

Conclusion

According to the estimated errors, statistical characteristics and correlation between the measured and predicted values, it was concluded that Autoregressive model AR (1) is proposed for generation of siltload and AR (2) is proposed for generation of rainfall and runoff in Karkara nala watershed Jharkhand.

References

1. Akaike, H. (1974). A new look at the statistical model identification on IEEE Transactions on automatic control, AS-19, pp. 716-723.
2. Anderson, R. L. (1942). Distribution of the serial correlation coefficients. Annals of Math. Statistics, 13(1):1-13.
3. Box, G.E.P. and Jenkins, G. (1970). Time series Analysis, Forecasting and control. First Ed. Holden – Day Inc., San Francisco.
4. Durbin, J. (1960). The fitting of time series model. Rev. Int. Inst. stat., 28: 23-33.
5. Iyengar, R. N. (1982). Stochastic modeling of monthly rainfall. J.Hydrolog,57:375-387
6. Kottegoda, N. T. and Horder, M. A. (1980). Daily flow model rainfall occuerces using pulse and a transfer function. Jounal of Hydrology, 47: 215-234.
7. Rai, Eno and Sherring, Arpan (2007). Development of autoregressive time series model for prediction of rainfall and runoff for Mansahara watershed of lower Gomti catchment. Journal of Agricultural Engineering. 44(4):38-42
8. Salas, J.D. and Smith, R.A. (1981). Physical basis of stochastic models of annual flows. Water Resources Research, 17 (2): 428-430.
9. Sharma, A., Singh, R. V., Narulkar, S. M. and Dashora, P. K. (2003). Identification of appropriate stochastic model for forecasting monsoonal river inflows. Journal of Agricultural Engineering, 12(3-4):191-208.
10. Sherring, Arpan, Manish Dwivedi and Mishra, A.K. (2009). Development of Autoregressive Time Series Model for Prediction of Rainfall and Runoff in Toriya Watershed. Environment and Ecology. 27(1A), pp.354-359
11. Yule, G. V. (1927). On a method of investing periodicities in disturbed series with special reference to Wolfer's sunspot number. Phill. Tran., A,226 pp.267

Received 13.08.2018 on and accepted on 28.11.2018