International Journal of Agriculture Sciences
ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 10, Issue 17, 2018, pp.-7070-7073
Available online at https://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000217

Research Article

DEVELOPMENT OF AUTOREGRESSIVE TIME SERIES MODEL FOR RAINFALL OF AURAIYA DISTRICT

SINGH B.[1], YADAV M.*[2], SRIVASTAVA S.K.[1] AND DENIS D.M.[1]
[1]Water Land Engineering and Management, Vaugh School of Agriculture Engineering & Technology, Sam Higginbottom Institute of Agriculture, Technology & Sciences, Allahabad, 211007, Uttar Pradesh, India
[2]Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, 221 005, Uttar Pradesh, India
*Corresponding Author: Email - mangalchamps@gmail.com

Received: September 01, 2018; Revised: September 11, 2018; Accepted: September 12, 2018; Published: September 15, 2018

Abstract: The present study was conducted with the prime objective to develop a stochastic time series models, capable of predicted rainfall data of a department of agriculture meteorological station Kakor in Auraiya district which has located on the southern portion of Uttar Pradesh 26° 21" and27° 01" north latitudeand 78° 45" and 79° 45" east longitude and forms a part of the Kanpur Division. The rainfall data of Auraiya District from the year 1999-2013 was collected and used for mean monthly flow of rainfall data using auto regressive model. Autoregressive (AR) models of orders 0, 1 and 2 were tried. The goodness of fit and adequacy of models were tested by Box-Pierce Portmonteau test, Akaike Information Criterion and by comparison of historical and predicted correlogram. The AIC values for AR (1) is lying between AR (0) and AR (2) which is satisfying the selection criteria. The mean forecast error is also very less in case of rainfall by AR (1) model. On the basis of the statistical test, Akaike Information Criterion, the AR (1) model with estimate model parameters was estimates for the best future predictions in Auraiya District. The comparison of the observed rain-fall flow and the synthetically generated data by AR (1) model shows that the developed model can be used efficiently for the further prediction of rainfall which can benefited the farmer and research workers for water harvesting.

Keywords: Stochastic time series model, Autoregressive (AR) models

Citation: Singh B., *et al*., (2018) Development of Autoregressive Time Series Model for Rainfall of Auraiya District. International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 10, Issue 17, pp.- 7070-7073.

Introduction

India is an agrarian country, due to its favourable monsoon climate and vast area of fertile culturable land. The Indian has been traditionally dependent on agriculture as 70% of its population is engaged in farming. Fresh water is an essential resource and rainfall is the primary source of fresh water supply. hydrological data such as rainfalls are the basic information used for the design of water resources systems. Rainfall is the principal phenomenon driving many hydrological extremes such as floods, droughts, landslides, debris and mud-flows; its analysis and modeling are typical problems in applied hydro-meteorology and it also directly or indirectly affects all the sectors like agriculture, insurance, industry and other allied field. Precipitation, especially rainfall has a dramatic effect on agriculture. Drought can kill crops and increase erosion, while overly wet weather can cause harmful fungus growth. These problems are closely linked with the behaviors of the monsoon rain in India. It is estimated that 69% of worldwide water use is for irrigation, with 15-35% of irrigation withdrawals being unsustainable. At the same time climatic changes are beginning to affect the global pattern climatic events like of rainfall and will probably lead to even higher temperatures and lower rainfall in tropical areas. Rainfall continues to serve as a reliable source of water for a Varity of purposes including industrial and domestic uses and irrigation. Rain plays a major role in hydrology that finds its greatest applications in the design and the operations of water resources, engineering works as well as agricultural systems [1]. There are a vast range of rainfall models proposed for different applications. Following [2] and [3] four broad types in the rainfall models may be classified, namely, models of dynamic metrology, multi-scaling models, empirical statistical models, and point process models time series models, such as Autoregressive (AR), Moving average (MA), Autoregressive and Moving average (ARMA) and (ARIMA) autoregressive and Integrated Moving average are widely used to developed and generate the annual stream flow. Autoregressive (AR) models have been extensively used in hydrology and water resources since the early 1999's, for modeling annual and periodic hydrologic time series. The application of these models has been attractive in hydrology mainly because (i) the autoregressive form has an intuitive type of time dependence (the value of a variable at the present time depends on the values at previous times), and (ii) they are the simplest models to use. Advanced approaches for rainfall modeling with sub-daily time steps are the point process models like Newman-Scott or Bartlett-Lewis rectangular pulse models [3], the most commonly used class of stochastic models is the autoregressive (AR) and autoregressive moving average (ARMA) models. Also known as Box -Jenkins type models, this class of stochastic model has been widely used in research and industry for hydrological time series simulation. The second stage is to use data generation methods to verify that the synthetic time series produced by the calibrated model reproduces appropriate. The basic statistical characteristics of the historical data (refer as mean, variance, etc.), dependent parameter (correlogram).

Materials and Methods

Location of description area

Auraiya district is the state of Uttar Pradesh. Auraiya is located on the southern portion of Uttar Pradesh 26° 21" and 27° 01" north latitude and 78° 45" and 79° 45" east longitude and forms a part of the Kanpur Division. It is bounded on the north by the districts of Kannauj, western border adjoins tehsil Bharthana of the Etawah district and the district of Gwalior. The total area is calculated to be 2054 square kilometre.

Stochastic Time series Model

A mathematical model representing stochastic process is called stochastic time series model. It has a certain mathematical form or structure and set of parameters. A sample time series model could be represented by simple probability distribution f(X: θ) with the parameters θ = (θ1, θ2…) valid for all

positions t = 1, 2... N and without any dependence between X1, X2 ....Xn.
If X is normal with mean μ and variance σ2 the time series can be written as

$$X_t = \mu + \sigma \varepsilon_t, \quad t = 1, 2 \ldots N \qquad \ldots\ldots (1)$$

Where εt is also normal with mean zero and variance one and ε1, ε2 ..... are independent. In equation (3.1) the model parameters μ and σ are constants. The structure of the model is simple since the variable Xt is a function of the independent variable εt and so Xt is also independent.

A time series model with dependence structure can be formed as:

$$\varepsilon_t = \phi\, \varepsilon_{t-1} + \eta_t \ldots\ldots (2)$$

Where,

ηt= An independent series with mean zero and variance (1- ϕ2)

εt= Dependent series

ϕ = Parameter of the model

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 shock. The pth order autoregressive model AR (p), representing the variable Yt is generally written as.

$$Y_t = Y + \sum \phi_j (Y_{t-j} - Y) + \varepsilon_t \ldots\ldots (3)$$

Where,

Yt = The time dependent series (variable)

εt= The time dependent series which is independent of Yt and is normally distributed with mean zero and variance σε2

Y, σ2, σε2 = Estimated from data

Estimation of Autoregressive parameter (ϕ) maximum likelihood estimate

For estimation of model parameter method of maximum likelihood will be used (Box and Jenkins, 1970)

$$z_i z_j + z_{i+1} z_{j+1} + \ldots\ldots\ldots + z_{N+1-J}\, z_{N+1-I'}$$

and define

$$D_{ij} = D_{ji} = N / (N+2-i-j) \ldots\ldots (4)$$

Where

D = difference operator

N = sample size

i,j = maximum possible order

$$AR(1): \phi = \frac{D1,2}{D2,2} \ldots\ldots (5)$$

$$AR(2): \phi 1 = \frac{D1,2\ D3,3 - D1,3\ D2,3}{D2,2\ D3,3 - D^2\ 2,3} \ldots\ldots (6)$$

$$AR(2): \phi 2 = \frac{D1,3\ D2,2 - D1,2\ D2,3}{D2,2\ D3,3 - D^2\ 2,3} \ldots\ldots (7)$$

Autocorrelation Function

The autocorrelation function rk of the variable Yt of equation (3.2) is obtained by multiplying both sides of the equation (3.2) by Yt+kand taking expectation term by term.

$$rk = \frac{\sum_{t=1}^{N-K} (Yt - \bar{Y})(Yt + k - \bar{Y})}{\sum_{t=1}^{N} (Yt - \bar{Y})} \qquad \ldots\ldots (8)$$

Where, rk= Autocorrelation function of time series Yt at lag k

Yt= Stream flow series (historical data)

$\bar{Y}$ = Mean of time series Yt

k = Lag of K time unit

N = Total number of discrete values of time series Yt

The following equation was used to determine the 95 percent probability levels, (Anderson, 1942).

$$rk\ (95\%) = \frac{-1 \pm 1.96\sqrt{N-K-1}}{N-K} \qquad \ldots\ldots (9)$$

Where, N = Sample size

Partial Autocorrelation function

The following equation was used to calculate the partial autocorrelation function of lag K (Durbin, 1960)

$$PCk,k = \frac{rk - \sum PCk - 1, jrk - j}{1 - \sum PCk - 1, jrj} \qquad \ldots\ldots (10)$$

Where, PCk,k= Partial autocorrelation function at lag K

Rk = Autocorrelation function at lag K

$$PC_{k,j} = PC_{k-1,j} - PC_{k,k}.PC_{k-1,k-j} \ldots\ldots (11)$$

$$j = 1, 2 \ldots k\text{-}1$$

The 95 percent probability limit for partial autocorrelation function was calculated using the following equation (Anderson 1942)

$$PCk, k(95\%) = \frac{1.96}{\sqrt{N}} \ldots\ldots (12)$$

©Parameter estimation of AR (p) models

The average of sequence Yt was computed by following equation

$$Y = \frac{1}{N} \sum_{t=1}^{N} Yt \qquad \ldots\ldots (13)$$

The average σ2ε of Yt was computed by the following equation

$$\sigma^2_\varepsilon = \frac{1}{(N-1)} \sum_{t-1}^{N} (Yt - \bar{Y})^2 \qquad \ldots\ldots (14)$$

After computation of Y and σ2ε , the remaining parameters ϕ1, ϕ2,. ....ϕp of the AR models were estimated by using the sequence.

$$Z_t = Y_t - Y \quad \ldots\ldots (15)$$

$$t = 1, 2, \ldots\ldots N$$

The parameters ϕ1, ϕ ,.........ϕpwere estimated by solving the p system of following linear equation (Yule and Walker equation).

$$r_k = \phi_1 r_{k-1} + \phi_2 r_{k-2} + \ldots\ldots + \phi_p r_{k-p} \qquad K>0$$

$$r_k = \sum_{j=1}^{p} \phi j\ rk - 1 \quad \ldots\ldots (16)$$

Where, r_1, r_2 were computed from equation (8)

Statistical Characteristics

Mean Forecast Error

Mean forecast error was calculated to evaluate the performance off autoregressive models fitted to time series of annual stream flow of rainfall. The mean forecast error (MFE) was computed for the annual stream flow series by the following equation:

$$MFE = \frac{\sum_{i=1}^{n} \chi c(t) - \sum_{i=1}^{n} \chi o(t)}{\eta} \qquad \ldots\ldots (17)$$

Where, χc (t) = Computed rainfall value

χo (t) = Observed rainfall value

η = Number of observations.

Mean Absolute Error

$$MAE = \frac{\sum_{i=1}^{n} |\chi c\ (t) - \chi o\ (t)|}{\eta} \qquad \ldots\ldots (18)$$

Mean Relative Error

$$MRE = \frac{\dfrac{\sum_{i=1}^{n} |\chi c\ (t) - \chi o(\ t)|}{\chi o\ (t)}}{\eta} \qquad \ldots\ldots (19)$$

Mean Square Error

$$MSE = \frac{\sum_{i=1}^{n} [\chi c\ (t) - \chi o\ (t)]^2}{\eta} \qquad \ldots\ldots (20)$$

Root Mean Square Error

$$RMSE = \left[\frac{\sum_{i=1}^{n} [\chi c\ (t) - \chi o\ (t)]^2}{\eta}\right]^{½} \qquad \ldots\ldots (21)$$

Integral Square Error

$$ISE = \frac{\sqrt{\sum_{i=1}^{n} [\chi c\ (t) - \chi o\ (t)]^2}}{\sum_{i=1}^{n} \chi o\ (t)} \qquad \ldots\ldots (22)$$

Goodness of fit of Autoregressive (AR) models

The following tests were performed to test the goodness of fit of autoregressive (AR) models. Box-Piece Portmanteau lack of fit test

$$Q = N \sum_{k=1}^{L} rk^2 \qquad \ldots\ldots (23)$$

Where, N = Number of observation

rk = serial correlation or Autocorrelation of series Yt

The statistic Q follows χ^2 distribution with r = k-p degree of freedom. The estimated value of χ^2

Table-1 Statistical Parameters of Autoregressive (AR) models for rainfall of Auraiya District

Model	AR (0)	AR(1)	AR (2)
Autoregressive Parameters		$\Phi_1 = 0.195763$	$\Phi_1 = 0.176211$, $\Phi_2 = 0.126598$
White Noise Variances, σ_ε^2	71058.9	69375.049	74174.46
Akaike Information Criterion, AIC (P)	169.569	169.2092	170.2126
Value of Port monteau Statistics, Q	3.36396	2.92535	1.70697
Degree of Freedom up to 5 lags	5	4	3
Table value of χ^2 at 5% level of significance	11.07	9.48	7.81

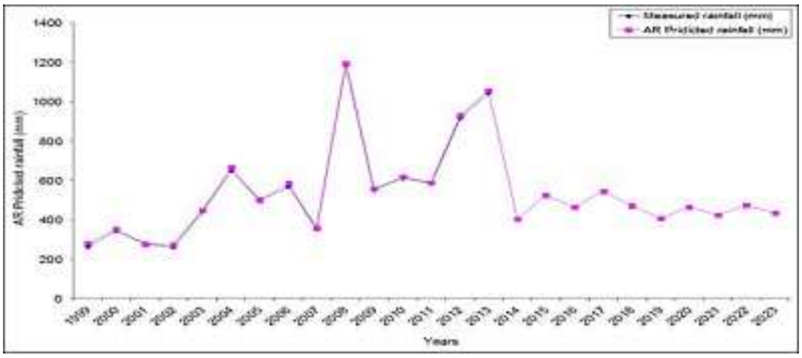

Fig-1 Comparisons of Correlogram of measured and Predicted Rainfall of Auraiya District

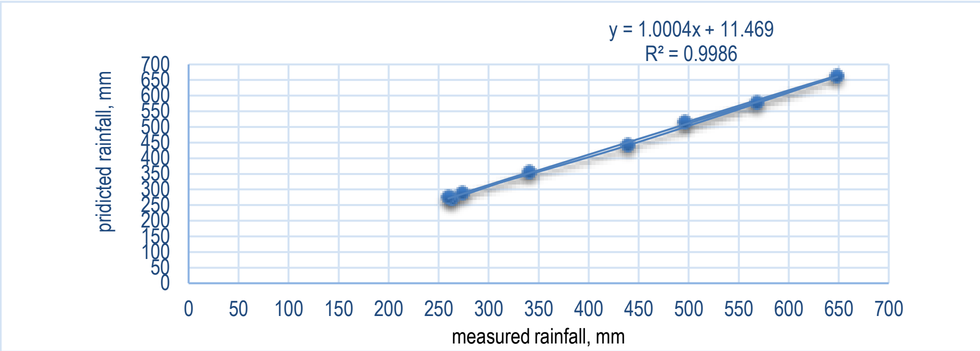

Fig-2 Comparison between Measured and Predicted Rainfall of Auraiya District

Alkaike Information Criterion

Akaike Information Criterion for AR (p) models was computed using the following equation.

$$AIC(P) = N \ln(\sigma E^2) + 2(p) \quad ......(24)$$

Where, N = Number of observations

σE^2 = Residual Variance

A comparison was made between the AIC (p) and the AIC (p-1) and AIC (p+1). If the AIC (p) is less than both AIC (p-1) and AIC (p+1), then the AR (p) model is best otherwise, the model which gives minimum AIC value was the one to be selected model.

Table-2 *Statistical Characteristics of Measured and Predicted Rainfall for Auraiya District*

SN	Statistical Characteristics	MeasuredRainfall (mm)	PredictedRainfall (mm)
1	Mean	390.465	402.3131
2	Standard Deviation	150.0065	148.5455
3	Skewness	0.460143	0.454793

Table-3 *Evaluation of Regeneration performance with statistical errors*

SN	Statistical Errors	Autoregressive AR(1) model Rainfall (mm)
1	Mean forest error	7.66916
2	Mean absolute error	-0.2092
3	Mean relative error	-7.66916
4	Mean square error	882.2403
5	Root mean square error	29.7053
6	Integral square error	1.244836

Result and Discussion

The study annual rainfall series were modeled through the autoregressive model. The modeling procedure of the data series involved various steps like preliminary analysis and identification , estimation of model parameters and diagnostic checking for the adequacy of the model.(Salas and Smith 1980 b) Autocorrelation and partial autocorrelation are used for identification of the proper type and order of the model .The identification generally depends on the characteristics of overall water resources system, the characteristics of time series and the models input.(Salas and Smith 1981) demonstrated these of physical consideration of the type of model. The autocorrelation functions and partial autocorrelation functions were determined for the 95% probability limits. The autocorrelation function and partial autocorrelation functions with 95% probability limits upto 5 lag of the series (lag k) were computed and the autoregressive model of first order AR(1) was selected for further analysis.

Models of Autoregressive (AR) Family

The parameters of AR model were computed for annual rainfall and the predicted values of annual rainfall and runoff were compared with the observed values . It was observed that AR(p) model up to order 2 has shown the good fit and correlation between the observed and predicted values and given in [Fig-1 and Fig-2] AR (p) models for prediction of annual rainfall

$$AR(1): Y_t = 569.33 + 0.195763 (Y_{t-1})$$

$$AR(2): Y_t = 569.33 + 0.176211 (Y_{t-1})$$

Box Pierce Port monteau test for AR model

The Box–Pierce Port monteau lack of fit test was used to check the adequacy of autoregressive models for both river flow and rainfall. The values of statistical tests for AR (0), AR (1) and AR (2) models were computed by equation (3.15).

The test statistics was compared with table value of $\chi 2$ and tabulated in [Table-1] and [Table-2]. The table data reveals that the values of test statistics for all three models are found to be less than tabulated value which is the condition for acceptance. Therefore all 3 models *viz*. AR (0), AR (1) and AR (2) were giving fit and were acceptable.

Comparison of the observed and predicted annual rainfall and runoff

The correlogram of observed and predicted series for annual rainfall were developed by plotting autocorrelation function against lag k. A graphical comparison of observed and predicted annual rainfall with the selected model are shown in [Fig-1 and Fig-2]. The graphical representation of the data shows a closer agreement between observed and predicted annual rainfall selected model. It reveals that developed model for rainfall and runoff can be utilized for the prediction of future trends with the minimum chance of error.

Statistical Characteristics of Data

The mean, standard deviation and skewness of historical and generated data was calculated to evaluate the fitting of the model in moment preservation. The results are tabulated in [Table-2]. The results clearly show that the skewness of generated data by AR (0) model and historical 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 rainfall

To evaluate the performance of the model beside the comparison of historical and generated values some other statistical characteristics such as MFE, MAE, MRE, MSE, RMSE and ISE were also estimated to prove the adequacy of the model for future prediction with higher degree of correlation to previous measured observations. The table also indicates that in case of rainfall, the MFE for AR (1) model is 1.67 mm which is lower the error, higher the quality of predicted rainfall. As the error is minimum, so the AR (1) model can be best suited for rainfall prediction in Auraiya district

Conclusion

The present study was conducted with the prime objective to develop a stochastic time series model, capable of predicting rainfall in Auraiya District. The annual rainfall data of the area from the year 1999-2013.were collected and used for the development of model. Autoregressive (AR) models of orders 0, 1, and 2, were tried for time series and different parameters were estimated by the general recursive formula proposed by [4]. The goodness of fit and adequacy of models were tested by Box-Pierce Portmonteau test, Akaike Information Criterion (AIC) and by comparison of historical and predicted correlogram. The AIC value for AR (1) model (169.21) is lying between AR (0) (169.57) and AR (2) (170.213) which is satisfying the selection criteria

The proposed Autoregressive AR(1) model for prediction of rainfall

$$AR\ (1):\ Yt = 569.33 + 0.195763\ (Yt - 1)$$

$$AR\ (2):\ Yt = 569.33 + 0.176211\ (Yt - 1)$$

In case of rainfall generation there is an effective agreement between observed and measured data with mean forecast error, mean absolute error, mean relative error, mean square error and integral square error.

Application of research: On the basis of estimated errors, statistical characteristics and correlation between observed and predicted values it is concluded that the proposed autoregressive AR(1) model can be used to predict the annual rainfall in Auraiya district of Uttar Pradesh state.

Research Category: Rainfall study

Acknowledgement / Funding: Author thankful to Sam Higginbottom Institute of Agriculture, Technology & Sciences, Allahabad, 211007, Uttar Pradesh, India

***Research Guide or Chairperson of research: Dr D.M. Denis**

University: Sam Higginbottom Institute of Agriculture, Technology & Sciences, Allahabad, 211007, Uttar Pradesh, India

Research project name or number: M Tech Thesis

Author Contributions: All author equally contributed

Author statement: All authors read, reviewed, agree and approved the final manuscript

Conflict of Interest: None declared

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Srikanthan R. and Mcmohan T. A. (1982) Journal of Hydrology,66, 231-226.

[2] Cox D.R. and Isham V. (1994) *Stochastic models of Precipitation. In Statistics for the enviroment, Water Issues, V. Barnett and K.F.Turkmanan (Eds.), Wiley New York,* Pg: 3-18.

[3] Onof C., Chandler R. E., Kakou A., Northrop P., Wheater H.S., and Isham V. (2000) *Stochastic Enviromental Research and Risk Assessment,* 14, 384-411.

[4] Lungu E.M. and Sefe F.T.K. (1991) *Journal of Hydrology*, 126,171-182.