

L E T T E R S

551.509.314

AN EXPONENTIAL AUTOREGRESSIVE (EXPAR) MODEL FOR THE FORECASTING OF ALL INDIA ANNUAL RAINFALL

1. In agricultural research, data are generally collected sequentially. For modelling and forecasting of such data, Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) methodology is the most popular, powerful and widely used time-series models. However, the ARIMA models are insufficient in many practical situations as they are not able to take into account important features of observed time-series data in practical situations (Fan and Yao, 2003).

India is an agrarian country as agriculture provides the principal means of livelihood for over 58.4% of India's population. It contributed 13.7% to the total gross domestic product (GDP) for the year 2013-14. Further, agriculture also accounts for about 10 per cent of the total export earnings and provides raw material to a large number of industries. The agricultural growth of a country depends to a large extent on the quantity and spread of rainfall. So if forecasting of the rainfall of our country is done properly then it may be helpful for proper planning and policy making. Accordingly, various authors have attempted to forecast rainfall data [Chandran and Prajneshu (2007); Ghosh *et al.* (2010); Ghosh and Prajneshu (2011); Chattopadhyay and Chattopadhyay (2013)]. A review of multiple and power regression models employed in India since 1988 along with various modifications made in these models from time to time has been provided in Rajeevan *et al.* (2004).

In this paper, a versatile parametric nonlinear time-series model, *viz.*, Exponential autoregressive (EXPAR) model (Ozaki, 1993), is studied for forecasting the Indian annual rainfall time-series data which showed cyclical fluctuations (Fig. 1). In addition, a comparison of EXPAR model with an ARIMA model is also made to see its utility over the usual linear time-series models.

2. The EXPAR nonlinear time-series model can generate time-series data with different types of marginal distributions by restricting the space of the parameters in various specific regions (Ozaki, 1993). It is also capable of capturing the non-Gaussian characteristics of the time-series and also can have a marginal distribution belonging to the exponential family. It can also account for jump phenomena, amplitude-dependent frequency and limit

cycle. Thus, EXPAR model is capable of dealing quite satisfactorily with cyclical data of various natures. EXPAR (p) nonlinear time-series model is given as :

$$X_{t+1} = \left\{ \varphi_1 + \pi_1 \exp(-\gamma X_t^2) \right\} X_t + \dots + \left\{ \varphi_p + \pi_p \exp(-\gamma X_t^2) \right\} X_{t-p+1} + n_{t+1}$$

where $\gamma > 0$ is a scaling constant and $\{n_t\}$ is white noise process with mean zero and variance σ_n^2 . The connection between EXPAR model and well-known Autoregressive conditional heteroscedastic (ARCH) model, proposed by Noble Laureate R. F. Engle in 1982, was thoroughly discussed by Ozaki (1993). A brief description of the procedure for estimating the parameters of EXPAR model is as follows (Baragona *et al.*, 2002).

The algorithm requires that an interval (a, b) , $\alpha \geq 0$, be pre-specified for the γ values. This interval is split in M sub-intervals, so that a grid of candidate γ values is built. Let $\delta = (b - a)/M$ and $\gamma = a$. Then, for M times, the following steps are performed:

- (i) Set $\gamma = \gamma + \delta$
- (ii) Estimate φ_j and π_j by ordinary least squares regression of X_t on $X_{t-j}, X_{t-j} \exp(-\gamma X_{t-1}^2), j = 1, \dots, p$.
- (iii) Compute the NAIC and repeat step (ii) for $p = 1, \dots, P$, where P is a pre-specified integer greater than 1.

Final estimates of parameters are obtained by minimizing the AIC, defined as:

$$AIC = N \log(\hat{\sigma}^2) + 2(2p + 1)$$

where, p is the number of parameters to be estimated.

3.1. The above discussed model is applied to Indian annual rainfall time-series data during the period 1901 to 2012, obtained from the website (www.tropmet.res.in) of the Indian Institute of Tropical Meteorology, Pune. Out of total 112 data points, first 102 data points are used for model building and the remaining 10 data points are used for validation purpose. A perusal of dataset in Fig. 1 indicates sudden burst and fluctuations at several time-epochs.

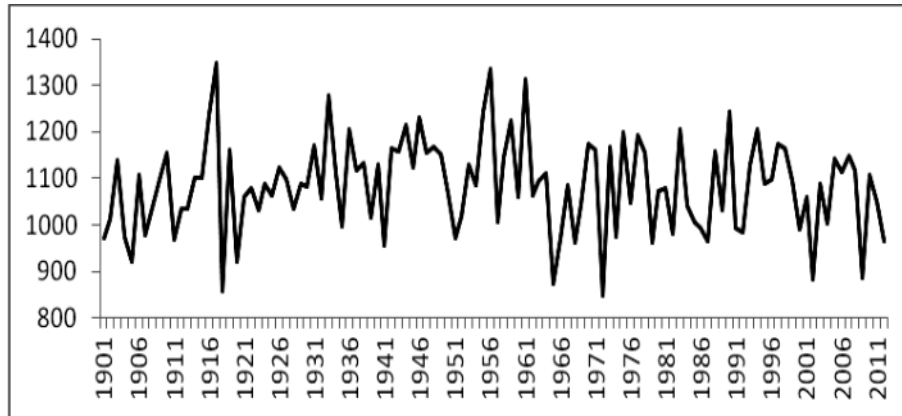


Fig. 1. Annual rainfall datalines

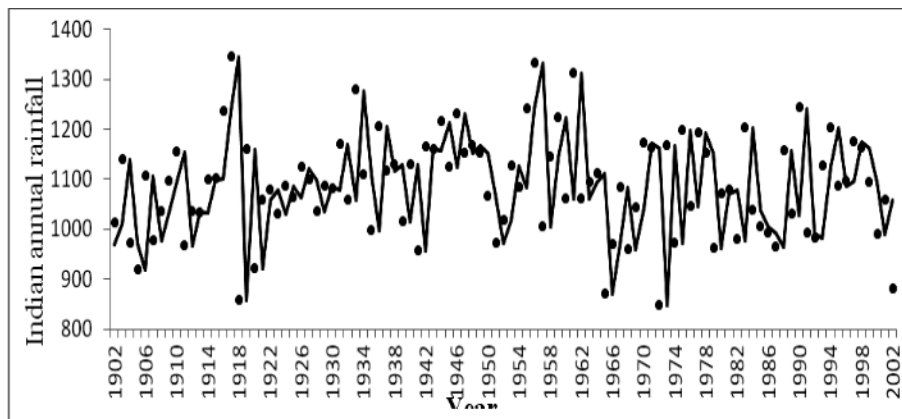


Fig. 2. Fitted EXPAR model along with data points

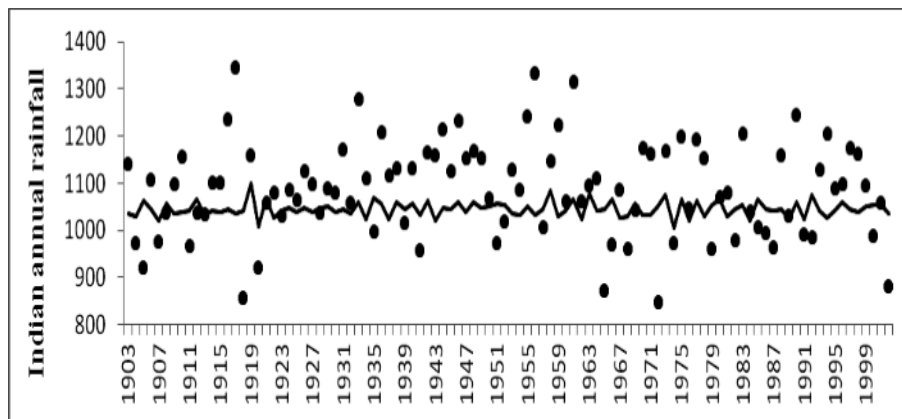


Fig. 3. Fitted ARIMA model along with data points

The entire data analysis is carried out using MATLAB, Ver. 7.4 software package. On the basis of minimum Akaike information criterion (AIC), EXPAR (1) is selected for modelling of the yearly data. To get a visual idea, the graph of fitted EXPAR (1) model along with data points are exhibited in Fig. 2. A visual inspection shows

that the fitted model is able to properly capture the cyclical fluctuations present in the data set compared to the ARIMA model shown in Fig. 3. To study the appropriateness of the fitted model, the autocorrelation function of the standardized residuals was computed. It was found that the autocorrelation function was

TABLE 1
One-step ahead forecasts of rainfall data

Year	Rainfall	ARIMA	EXPAR
2003	1088.5	706.67	881.01
2004	1003.4	582.51	1086.79
2005	1141.1	651.43	1001.82
2006	1112.7	572.93	1139.31
2007	1149.7	605.10	1110.95
2008	1118.3	583.27	1147.89
2009	886.4	603.51	1116.54
2010	1107.6	712.48	885.01
2011	1051.8	573.72	1105.86
2012	963.8	630.36	1050.15

insignificant at 5% level, thereby conforming that the model is correctly specified.

3.2. In this section a comparative study intended to evaluate the ARIMA and EXPAR models on their ability to produce forecast is carried out. To this end, ARIMA models are fitted to the data under consideration. On the basis of minimum AIC criterion, the ARIMA (2, 0, 0) model is selected, which is given by

$$X_{t+1} = 996.30 - 0.04X_t + 0.13X_{t-1} + \varepsilon_{t+1}$$

A visual inspection of Table 1 indicates that EXPAR model performs comparatively well. Further, the performance of fitted models is also compared on the basis of one-step-ahead Mean square prediction error (MSPE), Mean absolute prediction error (MAPE) and Relative mean absolute prediction error (RMAPE) given by

$$MSPE = 1/10 \sum_{i=0}^2 (Y_{T+i+1} - \hat{Y}_{T+i+1})^2$$

$$MAPE = 1/10 \sum_{i=0}^2 (|Y_{T+i+1} - \hat{Y}_{T+i+1}|)$$

$$RMAPE = 1/10 \sum_{i=0}^2 (|Y_{T+i+1} - \hat{Y}_{T+i+1}| / Y_{T+i+1}) \times 100$$

The MSPE, MAPE and RMAPE values for fitted EXPAR model are respectively computed as

18537.92 cm², 111.82 cm and 10.81 which are found to be lower than the corresponding ones for fitted ARIMA model, viz., 201275.60 cm², 440.12 cm and 41.12 respectively. This indicates the superiority of EXPAR model over ARIMA model from forecasting point of view.

4. In this paper, a parametric nonlinear time-series model (EXPAR) was described to forecast the Annual rainfall of India. It is seen that the cyclical fluctuations of the time-series data is properly captured by the EXPAR model. Superiority of the methodology over other competing methodology, ARIMA, was also demonstrated.

References

Baragona, R., Battaglia, F. and Cucina, D., 2002, "A note on estimating autoregressive exponential models", *Quad. di Statist.*, **4**, 1-18.

Chandran, K. P. and Prajneshu, 2007, "Nonparametric regression approach for modelling and forecasting country's meteorological sub-divisions rainfall data", *J. Agrometeorol.*, **8**, 254-259.

Chattopadhyay, S. and Chattopadhyay, G., 2013, "A supervised Neural Network model for predicting average summer monsoon rainfall in India", *J. Ind. Soc. Ag. Statist.*, **67**, 43-49.

Fan, J. and Yao, Q., 2003, "Nonlinear Time-series : Nonparametric and Parametric Methods", New York: Springer.

Ghosh, H. and Prajneshu, 2011, "Statistical learning theory for fitting multimodal distribution to rainfall data: an application", *J. Appl. Stat.*, **38**, 2533-45.

Ghosh, H., Paul, R. K. and Prajneshu, 2010, "Wavelet frequency domain approach for statistical modeling of rainfall time-series data", *J. Statist. Theo. Pract.*, **4**, 4, 813-825.

Ozaki, T., 1993, "Non-gaussian characteristics of exponential autoregressive processes, In Developments in Time-Series Analysis", T. Subba Rao, Ed., 257-273, London: Chapman and Hall.

Rajeevan, M., Pai, D. S., Dikshit, S. K. and Kelkar, R. R., 2004, "IMD's new operational models for long - range forecast of southwest monsoon rainfall over India and their verification for 2003", *Curr. Sci.*, **86** 422-431.

BISHAL GURUNG

*Indian Agricultural Statistics Research Institute,
New Delhi - 110 012, India
(Received 29 July 2013, Modified 29 August 2014)
email : vsalrayan@gmail.com*