International Journal of Information and Computation Technology.
ISSN 0974-2239 Volume 3, Number 4 (2013), pp. 261-266

http://www. irphouse.com /ijict.htm

An Application of Exponential Autoregressive (EXPAR) Nonlinear Time-series Model

Bishal Gurung

Indian Agricultural Statistics Research Institute, New Delhi 110012.

Key words: Cyclical fluctuations, EXPAR, Genetic Algorithm, Forecasts.

*** Short note**

[1]Scientist (Division of Biometrics and Statistical Modelling)

Most processes encountered in the real world are nonlinear to a great extent. So in many practical applications nonlinear models are required in order to achieve an acceptable predictive accuracy and to provide reasonably close forecasts. Various such nonlinear time-series models, viz. Bilinear models, Threshold autoregressive models, Exponential autoregressive models, State-dependent models and Doubly stochastic models have been developed. Nonlinear models not only provide a better fit to the data but can also reveal rich dynamic behaviour like volatility, chaos, limit cycle behaviour, and cyclical fluctuations, which are not captured by linear models. (Fan and Yao, 2003).

One such parametric family of nonlinear time-series model is that of Exponential Autoregressive (EXPAR). The EXPAR model, introduced for modelling and forecasting of cyclical data, is a kind of useful model that has properties similar to those of nonlinear random vibrations. It is capable of generating time-series data with different types of marginal distributions by restricting the parametric space in various specific regions. It also accounts for amplitude-dependent frequency, jump phenomena and limit cycle. A heartening feature of this model is that it captures the non-Gaussian characteristics of the time-series and is also seen to have a marginal distribution belonging to the exponential family (Ozaki, 1993). However, there are only few approaches available in the literature to estimate the parameters of EXPAR family of models. One such approach is the local polynomial regression technique to estimate the parameters of EXPAR (2) given by Cai *et al*. (2000). It was not only an approximate method but was applicable only for EXPAR of order 2. As the literature on parameter estimation of EXPAR model is vague, so a promising and powerful optimization technique of Genetic Algorithm (GA) is considered for fitting EXPAR model.

An EXPAR (p) model may explicitly be written as

$$X_{t+1} = \{\varphi_1 + \pi_1 \exp(-\gamma X_t^2)\}X_t + \ldots + \{\varphi_p + \pi_p \exp(-\gamma X_t^2)\}X_{t-p+1} + \eta_{t+1} \quad (1)$$

with $\gamma > 0$, some scaling constant and $\{\eta_{t+1}\}$ is white noise process with mean zero and variance σ_η^2 . The values of γ are selected such that $\exp(-\gamma X_t^2)$ varies reasonably widely over the range (*0,1*). Also, (1) may be thought of as a threshold autoregressive (TAR) model, in the sense that, if $|X_t|$ is large, then (1) is similar to an autoregressive model with parameters approximately equal to$(\varphi_1, \ldots, \varphi_p)$, while if $|X_t|$ is small, then the autoregressive parameters switch to$(\varphi_1 + \pi_1, \ldots, \varphi_p + \pi_p)$.

The Genetic Algorithm (GA), modelled on the theory of biological evolutionary process, is a method for solving optimization problems. It combines the "natural selection" and "survival of the fittest" principle of Darwin with computer-constructed evolution mechanism to select better offspring from the original population. Over successive generations, the population "evolves" toward an optimal solution. In the Binary coded GA, the coded parameters form a string (chromosome), where each string represents a solution to the problem. Better solutions are subsequently evolved within a population of chromosomes over a number of generations. GA uses three main types of rules, viz. Selection, Crossover and Mutation, at each step to create the next generation from the current population. Selection selects good individuals from a population and forms a mating pool. The essential idea is that above-average individuals are picked from the current population and duplicates of them are inserted in the mating pool. In crossover operator, two individuals are picked from the mating pool at random and some portion of the string is exchanged between the individuals. It is mainly responsible for the search aspect of GA. Mutation rules apply random changes to individual parents to form children and keep the diversity in the population. Furthermore, mutation is useful for local improvement of a solution. GA works iteratively by successively applying these three operators in each generation. The algorithm stops if the relative gain in the fitness value is very small for two successive generations or when the specified maximum number of generations is reached. A good description of Genetic Algorithm is given in Goldberg (2009).

The EXPAR model is applied to describe India's annual rainfall data for the period 1901-2003, obtained from www.indiastat.com. From the total 103 data points denoted as $\{X_t, t = 0,1, \ldots, 102\}$, first 95 data points corresponding to the period 1901 to 1995 are used for building the model and remaining are used for validation purpose.

Preliminary Exploratory data analysis is carried out to justify the choice of EXPAR model. The directed scatter diagrams exhibited in Fig. 1 show asymmetry in the joint distribution of the observations, indicating thereby that the joint distributions of (X_t, X_{t-j}) is non-Gaussian, as two–dimensional normal distribution cannot be asymmetric.

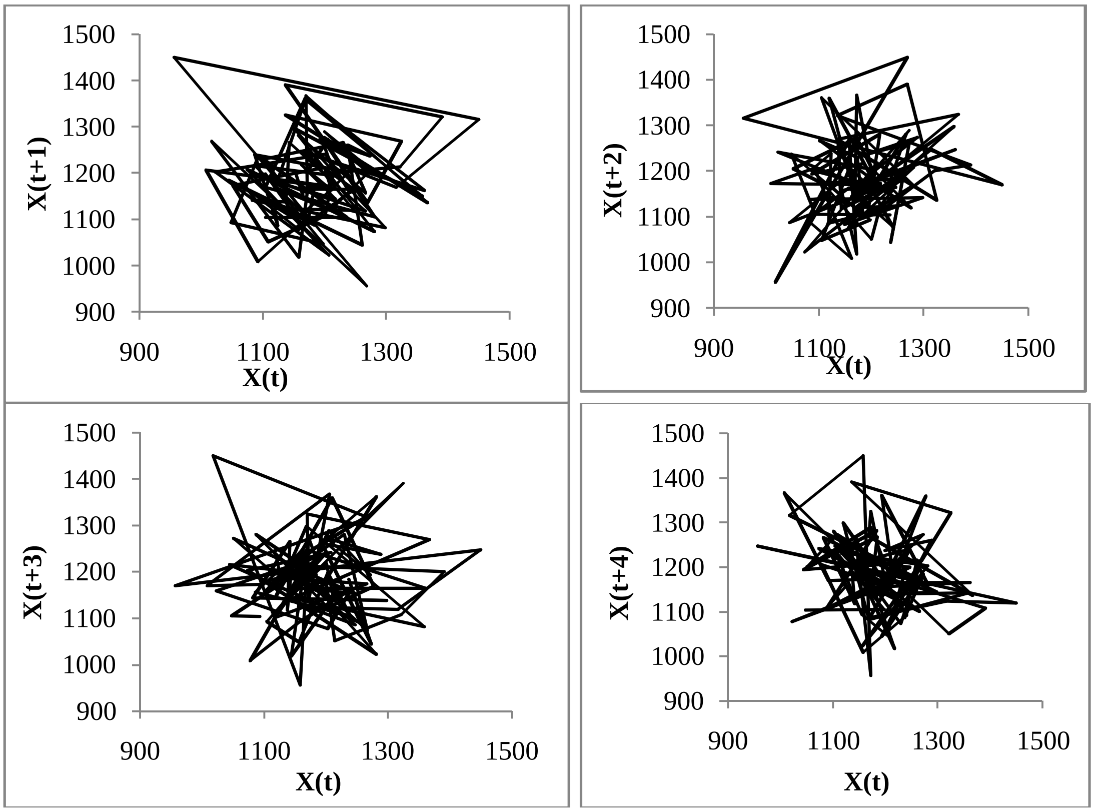

Fig. 1: Directed scatter diagram of India's annual rainfall data.

In the first instance, ARIMA models are fitted using EViews software package, Ver. 3. On the basis of minimum AIC criterion, the ARIMA model of order (2, 0, 0) is selected. The fitted model is given by

$$X_{t+1} = \underset{(9.82)}{1186.47} - \underset{(0.10)}{0.08}X_t + \underset{(0.10)}{0.16}X_{t-1} + \varepsilon_{t+1}$$

with $\mathrm{Var}\{\varepsilon_{t+1}\} = 15626.61$.

The roots of the characteristic polynomial, viz. (1-0.08L+0.16 L^2) are complex as $0.16 > (-0.08)^2/4$, which lends support to cyclical pattern of the data (Box *et al.*, 2008).

Subsequently, the GA procedure is applied for estimation of the parameter by using the AIC as the objective function to be minimized. The best identified model on the basis of minimum AIC value is EXPAR (1) model. The fitted model is given as

$$X_{t+1} = \{0.99 + 0.82\exp(-0.57X_t^2)\}X_t + \eta_{t+1}$$

with $\mathrm{Var}\ \{\eta_{t+1}\} = 12163.64$.

The standard errors for $\{\hat{\gamma}_1, \hat{\varphi}_1, \hat{\pi}_1\}$ by generating 1000 bootstrap samples are (0.07, 0.03, 0.05) respectively. A mechanistic interpretation of fitted EXPAR model is that $\Theta = \{0.99 + 0.82\exp(-0.57X_t^2)\}$ is large (small) when $|X_{t+1}|$ is small (large). This means that the annual rainfall data X_{t+1} tends to be nonstationary after a period of stationarity and vice-versa. This phenomenon leads to cyclicity, which is in agreement with observed data.

To get a visual idea, fitted EXPAR model along with data points is exhibited in Fig. 2. A visual inspection shows that the fitted model is able to capture properly the cyclical behaviour present at various time epochs in the data set.

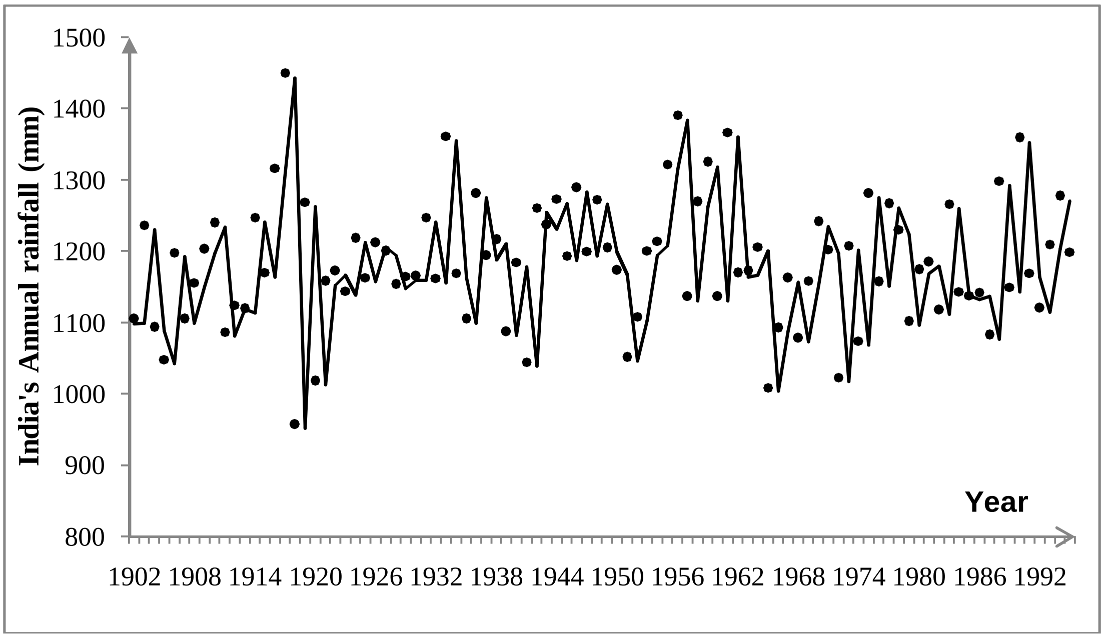

Fig. 2: Fitted EXPAR nonlinear time-series model along with data points.

The AIC, BIC and MSE of the ARIMA and EXPAR models are computed and reported in Table 1. It indicates that EXPAR model performs comparatively well for modelling the cyclical data under consideration.

Table 1: Goodness of fit of models.

Model Criterion	ARIMA	EXPAR
AIC	924.94	899.59
BIC	932.61	907.24
MSE	15373.23	1775.43

The fitted model has further been validated by carrying out one-step ahead forecasts which also shows that the model is able to capture the underlying nonlinear phenomena satisfactorily. To this end, forecasting performance for 8 data points corresponding to India's annual rainfall data for the period 1996-2003 as hold-out data is studied. For evaluation purpose, the one-step ahead forecasts are computed and reported in Table 2.

Table 2: One-step ahead forecasts of India's annual rainfall data.

Year	Actual	ARIMA	EXPAR
1996	1190	1287.37	1192.49
1997	1211.8	1275.65	1183.94
1998	1250.9	1272.44	1205.63
1999	1181	1272.58	1244.53
2000	1125.2	1284.72	1174.98
2001	1116.8	1278.36	1119.47
2002	1022	1270.22	1111.11
2003	1220.1	1276.94	1016.79

A perusal of the table indicates that EXPAR model is better than ARIMA for forecasting the data under consideration. 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

$$MSPE = 1/8 \sum_{i=0}^{7} \{Y_{T+i+1} - \hat{Y}_{T+i+1}\}^2$$

$$MAPE = {}^{1}/_{8} \sum_{i=0}^{7} \{|Y_{T+i+1} - \hat{Y}_{T+i+1}|\}$$

$$RMAPE = {}^{1}/_{8} \sum_{i=0}^{7} \left\{ {|Y_{T+i+1} - \hat{Y}_{T+i+1}|} \Big/ {Y_{T+i+1}} \right\} \times 100$$

The MSPE, MAPE and RMAPE values for fitted EXPAR model are respectively computed as 7328.65, 60.49 and 5.49, which are found to be lower than the corresponding ones for fitted ARIMA model, viz. 17350.77, 112.55 and 8.81 respectively.

To sum up, the EXPAR model has performed satisfactorily for modelling as well as forecasting of the cyclical data under consideration.

Summary

In this paper, importance of GA for fitting EXPAR nonlinear time-series model is highlighted. Superiority of EXPAR model over ARIMA model is also clearly demonstrated. This work could be of help to planners to take proper policy decision related to agriculture. It is also hoped that, applied statisticians would also start employing the GA for fitting other similar nonlinear time-series models.

Acknowledgements

The author is grateful to Dr. Prajneshu and Dr. Himadri Ghosh for providing assistance during the course of this research work.

References

[1] Box G E P, Jenkins G M and Reinsel G C. 2008. *Time Series Analysis: Forecasting and Control*, Prentice Hall.

[2] Cai Z, Fan J and Li R. 2000. Efficient estimation and inferences for varying-coefficient models. *Journal of the American Statistical Association* **95**: 888-902.

[3] Fan J and Yao Q. 2003. *Nonlinear Time-series*: *Nonparametric and Parametric Methods*. Springer, New York.

[4] Goldberg D E. 2009. *Genetic Algorithms in search*, *optimization and machine learning*. Dorling Kindersley, New Delhi.

[5] Ozaki T. 1993. Non-Gaussian characteristics of exponential autoregressive processes. In *Developments in time-series analysis*. Ed. *T. Subba Rao*. Chapman and Hall, 257-73.