Economic Affairs 2014, 59(3) : 479-423

New Delhi Publishers

WORKING PAPER

Economic Affairs

59(3): 2014: DOI 10.5958/0976-4666.2014.00015.1

Stochastic Model for Sticklac Forecasting in India

Lawrence Lepcha[1], Bishal Gurung[2]✉, Ranjit Kumar Paul[3], and Kanchan Sinha[4]

[1]Uttar Bangla Krishi Vishwavidyalaya, Coochbehar, India.

[2, 3, 4]Indian Agricultural & Statistical Research Instititute, New Delhi, India.

✉Corresponding author: vsalrayan@gmail.com

Paper No. 154 Received: 12th June, 2014 Accepted: 27 September, 2014

Abstract

In the present study, we aim to devise most appropriate prediction model for India's annual sticklac production data based on Exponential Autoregressive (EXPAR) model. Statistical modelling and forecasting of agricultural time-series data plays a vital role in comprehending the underlying relationships among statistically significant variables and helping the planners in policy making. Accordingly, in this paper, a promising methodology of EXPAR family of models has been employed to describe India's annual sticklac production data that depict such cyclical fluctuations. The fitted EXPAR model captured the data in a satisfactory manner. Further, the performance of the model is compared by computing various measures of goodness-of-fit and forecast performance. We conclude that EXPAR model performs quite well for modelling as well as forecasting of the cyclical data under consideration.

Keywords: Linear, stochastic model, ARIMA, Statistical modelling, EXPAR, Forecasting, Sticklac production

Introduction

Lac is obtained from numerous insect-species as a resinous secretion. The tiny insects which number in thousands colonize the branches of appropriate host trees and give out the resinous pigment.The branches which are coated with resins secreted by insects are cut and finally harvested as sticklac. The impurities of the harvested sticklac are removed by crushing and filtering. The filtered material is then repetitively washed to get rid of insect and other soluble parts to obtain seedlac. Jharkhand is the leading producer of Lac in India. Chattisgarh, West

Bengal and Maharashtra are the other states which produces Lac in abundance. Country wise India, Thailand, and China are the leading producers of sticklac. India approximately produced 50,000 tonnes in 1950s and 12,000 tonnes during 1980s.

In agriculture, data are, usually, collected sequentially over time. For analysing such data collected sequentially in time; statisticians and econometricianshave a well-established procedure based on linear time-series models, called the Autoregressive Integrated Moving Average (ARIMA) methodology. It is the most commonly used time-series model. But in many practical cases, the ARIMA models appear insufficient as it is not able to take into account important features of many observed time-series. Nonlinear models are required in many real-world applications,as it accurately describes the dynamics of the series and also makes better multistep-ahead forecasts.

One such family ofnonlinear time-series model of the parametric form is that of Exponential Autoregressive (EXPAR). It is capable of producingtime-series data with variousforms of marginal distributions by altering the parametric space into specific sections.A very important distinguishing feature of this model is that it is capable of capturing the non-Gaussian characteristics of the time-series and further, itdescribes those data sets that depict cyclical variations. As an illustration, the fitting ofEXPAR model to India's total Sticklac production is carried out. The importance of the EXPAR for modelling and forecasting purpose is studiedby carrying out comparative study with ARIMA.

Materials and Methods

Description of EXPAR model

An EXPAR (*p*) model is given as

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

with $\gamma > 0$, some scaling constant and $\{\epsilon_t\}$denotes a white noise process and has mean zero and variance σ_η^2. The range ofγchosen are such that$\exp(-\gamma X_t^2)$ varies reasonably widely over the range (*0*,*1*).

Estimation of parameters

A brief description of the procedure for estimating the parameters of (1) is as follows (Baragona*etal*., 2002). The algorithm requires that the interval (a,b), $a \geq 0$, be pre-specified for the γ values in (1). It is split in M sub-intervals to obtain a grid of candidate values forγ. Let $\delta = (b-a)/M$and $\gamma = a$. Then, for M times, the following steps are performed:

Set $\gamma = \gamma + \delta$

Estimate φ_j and π_j by ordinary least squares regression of X_t on $(X_{t-j}, X_{t-j}\, exp(-\gamma X_{t-1}^2)), j = 1, \ldots, p$.

Compute the NAIC and repeat step (II) for p=1, . . . , P, where P is a pre-specified integer greater than 1.

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

$$NAIC = \{Nlog(\hat{\sigma}^2) + 2(2p + 1)\}/(\text{Effectivesamplesize})$$

Results and Discussion

As an illustration India's total Sticklac production for the period 1930-2007, obtained from www.indiastat.com, is considered. Out of total 78 data points, first 70 data points are used for model building, and the remaining 8 data points are used for validating the fitted model.

To justify the use of EXPAR model, preliminary exploratory data analysis is carried out. In Fig. one, the directed scatter diagrams show asymmetry in the joint distribution of (X_t, X_{t-j}). This indicates the non-Gaussianity of the joint distributions of (X_t, X_{t-i}), as two–dimensional normal distribution can never be asymmetric.

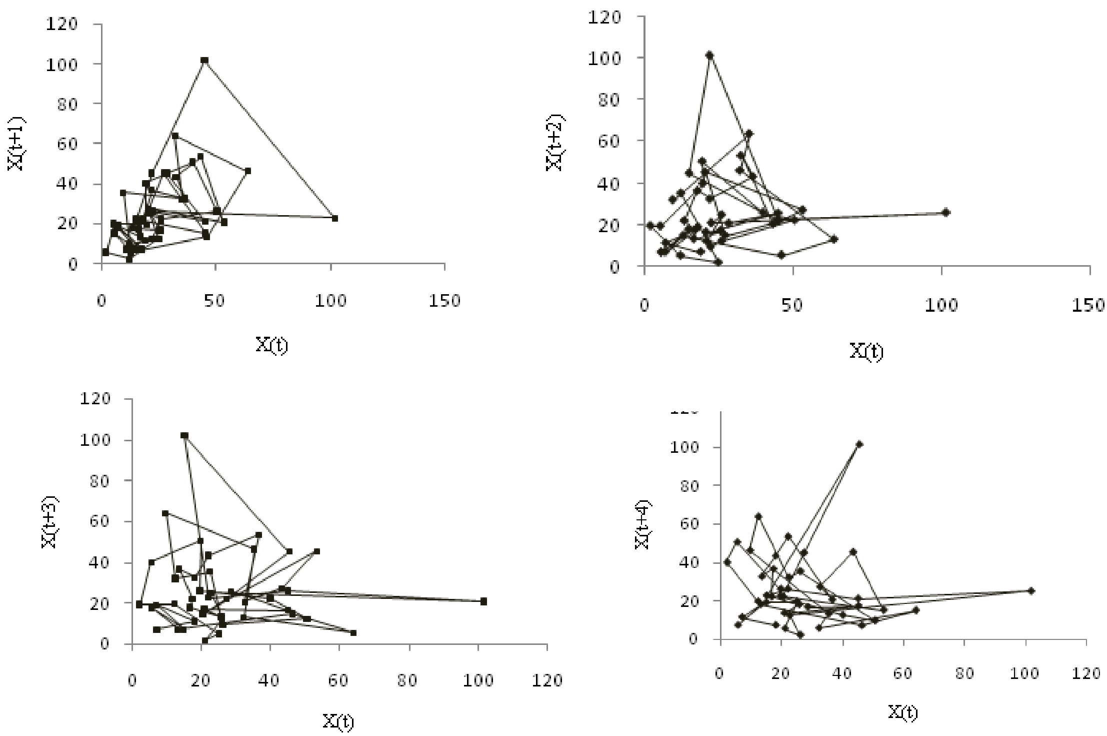

Fig. 1. Directed scatter diagram of India's total Sticklac production

We first fit the ARIMA models to the data considered. The best model which is identified on the basis of minimum NAIC value is ARIMA (1,1,0) modelThe fitted model is

$$X_{t+1} = 3.55 + 0.91X_t + \varepsilon_{t+1}$$

$$\text{Var}\{\varepsilon_t\} = 36.45$$

Subsequently, the algorithm proposed by Baragona*et al*. (2002) is applied for estimation of the parameters. On the basis of minimum NAIC criterion, the EXPAR(1) model is selected. The fitted model is given as

$$X_{t+1} = \{0.98 + 0.9\exp(-1.01X_t^2)\}X_t + \epsilon_{t+1}$$

with $\text{Var}\{\epsilon_t\} = 30.64$

A mechanistic interpretation that can be drawn from the fitted EXPAR model is that $\theta = \{0.98 + 0.9\exp(-1.01X_t^2)\}$is large (small) when $|X_{t+1}|$ is small (large). The interpretation means India's total Sticklac production tends to be stationary after some epoch of nonstationary and vice-versa. This spectacleis the reason behind cyclicity in the observed data. Figure 2 shows the graph of fitted EXPAR (1) model along with the data points. Aperusal indicates that EXPAR has properly captured the dynamics of the fluctuations present in the data sets.

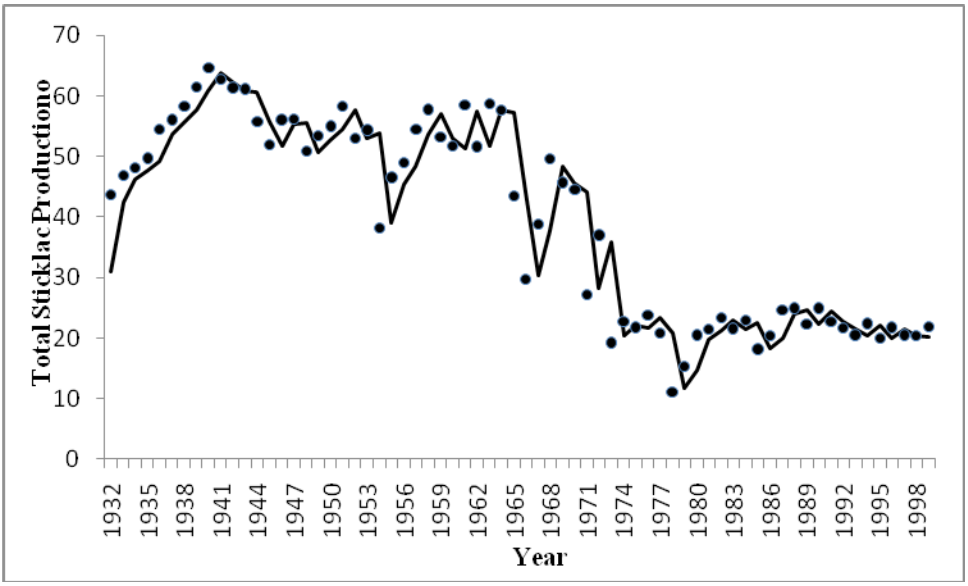

Figure 2. Graph ofEXPAR model along with data points

Table 3. Goodness of fit of models

Model / Criterion	Arima	Expar
AIC	843.34	783.39
BIC	872.23	759.92
MSE	14873.42	10582.91

The AIC, BIC and MSE of the EXPAR model as well as the ARIMA model is computed and reported in Table 1. A perusal indicates that EXPAR model has performed better than ARIMA from modelling point of view.

Evaluation of forecasting Performance

In this section,we compare the models in order to evaluate the forecasting performance of ARIMA and EXPAR. On the basis of one-step ahead Mean square prediction error (MSPE), Mean absolute prediction error (MAPE) and Relative mean absolute prediction error (RMAPE), we evaluate the forecasting performance of the fitted models. The MSPE, MAPE and RMAPE values for fitted EXPAR model are computed as 7.83, 2.36 and 11.71 respectively. It is found lower than the corresponding values, i.e. 14.98, 3.59, and 18.06 respectively, for the fitted ARIMA model. This shows that EXPAR performs better than ARIMA for forecasting purposes for the data sets under consideration.

In this paper, significance of using EXPAR nonlinear time-series model for fitting cyclical dataset is highlighted. Further, superiority of EXPAR model over ARIMA model from modelling and forecasting point of view is also validated.We hope that researchers would start applying EXPAR for cyclical time-series data.

References

Baragona, R., Battaglia, F., and Cucina, D. (2002). A note on estimating autoregressive exponential models. Quad. di Stat., 1-18.

Box, G., Jenkins, G.,and Reinsel, G. (2008). *Time Series Analysis: Forecasting and Control*, Prentice Hall, New Jersey.

Cai Z, Fan J and Li R. (2000). Efficient estimation and inferences for varying-coefficient models. *J. Amer. Statist. Assoc.*, : 888-902.

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

Ghosh, H, Gurung, Bishal, and Prajneshu. (2011). Methodology for combining linear and nonlinear time-series models for cyclical data. *J. Ind. Soc. Agril. Stat.*, 249-256.

Ghosh, Himadri, Gurung, Bishal, and Prajneshu (2014). Fitting EXPAR models through the extended Kalman filter.*Sankhya B.* DOI: 10.1007/s13571-014-0085-8

Ozaki T. (1993). Non-Gaussian characteristics of exponential autoregressive processes. "*Developments in time-series analysis*". Ed. *T. SubbaRao*. Chapman and Hall, 257-73.