



## Note

# Analysis of cyclical fish landings through ESTAR nonlinear time-series approach

BISHAL GURUNG, K. N. SINGH, PRAJNESHU AND AVNISH GROVER

ICAR-Indian Agricultural Statistics Research Institute, New Delhi - 110 012, India

e-mail: vsalrayan@iasri.res.in

## ABSTRACT

Exponential smooth transition autoregressive (ESTAR) is a family of parametric nonlinear time-series models capable of capturing the non-Gaussian characteristics of the time-series along with cyclical fluctuations. The present study is based on the time-series data on oilsardine landings from Kerala during the period 1961 to 2008. The parameters of the model were estimated by genetic algorithm (GA). From the analysis of data it was concluded that ESTAR model fitted through GA has performed better than ARIMA model.

Keywords: Exponential smooth transition autoregressive model, Fish landings data, Forecast performance, Genetic algorithm, Out-of-sample forecast

Time-series is a sequence of observations taken serially in time. Understanding, description of the generating mechanism and forecasting of future values based on past values are primary objectives of time-series. To this end, statisticians possess a well-established methodology based on linear time-series models, called the ARIMA methodology. It is a well-known fact that numerous economically important marine fish species, like the mackerel, oilsardine and Bombayduck exhibit fluctuations of cyclical nature. Noble (1980) had examined ten year cycle in mackerel fishery. ARIMA methodology was employed by Noble and Sathianandan (1991) for investigating the trend analysis in all-India mackerel catches. Sathianandan and Srinath (1995) carried out the analysis of marine fish landings in India using time-series technique. The three statistical methodologies, viz., regression, univariate and multivariate time-series methods for modelling and forecasting fish catches was compared by Venugopalan and Srinath (1998). The ARIMA models are quite popular certainly due to their comparative ease and also due to fact that there exists number of computer softwares incorporating the same. The ARIMA model, however, is insufficient as it is not able to capture many important features. Nowadays, time-series analysis has moved towards the nonlinear domain, which is generally more appropriate for accurately describing dynamics of a time-series, for making better multi-step-ahead forecasts and also in terms of fitting as well as forecasting when the data is nonlinearly related with its past values (Fan and Yao, 2003). An elaborate study of the all-India landings of oil sardine, mackerel and Bombayduck using the more advanced technique of spectral decomposition was done by Sathianandan and

Alagaraja (1998). Nampoothiri and Balakrishna (2000) applied the threshold autoregressive model for a time-series data.

One important family of parametric nonlinear time-series model is the exponential smooth transition autoregressive (ESTAR) family of parametric nonlinear time-series models, which was introduced for modelling and forecasting of "cyclical" data. A new method proposed for parameter estimation of time-series data is through the minimisation of the objective function through the genetic algorithm (GA) technique. This approach gives a simple and efficient algorithm to handle a nonlinear model. Sathianandan and Jayasankar (2009) recently used the technique of binary coded genetic algorithm (BCGA) for efficient management of marine fishery in Kerala.

The smooth transition autoregressive (STAR) model proposed by Terasvirta (1994) is a generalisation of the threshold models. It avoids discontinuities in the autoregressive parameters as the transition from one regime to the next depends on a continuous function.

The STAR model may be written as:

$$y_t = G(Z_t, y_t; \Psi) = \phi' w_t + (\theta' w_t) G_k^L(y, c; y_{t-d}) + \varepsilon_t \quad (2.1)$$

where,  $\{\varepsilon_t\}$  is a sequence of normal  $(0, \sigma^2)$  independent errors,  $\phi = \phi_0, \phi_1, \dots, \phi_p$  and  $\theta = \theta_0, \theta_1, \dots, \theta_p$  are  $(P+1)$  de parameter vector,  $w_t = (1, Y_{t-p}, \dots, Y_{t-p})^1$  is the vector consisting of an intercept and the first  $p$  lags of  $y_t$  and  $G_k^L(y, c; y_{t-d})$  is known as transition function. Depending upon the forms of transition function, different forms of STAR models are defined.

If the function  $G_k^L$  in equation 2.1 is written as:

$$G_k^L(y, c; y_{t-d}) = 1 - \exp\{-y(y_{t-d} - c)^2\}, y > 0 \quad (2.2)$$

Terasvirta (1994) called this model as Exponential smooth transition autoregressive (ESTAR) model. The transition function for ESTAR allows  $y_t$  to move smoothly between very small and very large values for which local dynamics are stable. Van Dijk *et al.* (2002) surveyed recent developments related to the STAR time series model and several of its variants. Lundbergh *et al.* (2003) have given a variant of STAR model to study the nonlinearity and structural change.

As the literature on parameter estimation of ESTAR model is vague, a promising and powerful optimisation technique of genetic algorithm (GA) is considered for fitting ESTAR model. The genetic algorithm (GA) is modelled on the theory of biological evolutionary process and is employed for solving optimisation problems. It combines both the principles of “natural selection” and “survival of the fittest” 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 genetic algorithm (BCGA), 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 main 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 a little portion of the string is exchanged between the individuals. It is mainly responsible for the search aspect of GA.

Mutation allows random changes to individual parents to form children and keep the diversity in the population. In addition, mutation is also 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). Wu and Chang (2002) used GA to estimate the parameters of threshold autoregressive models. Further, Baragona *et al.* (2004) fitted threshold subset autoregressive moving-average models by GA. Taking this in view, in this paper GA is applied to estimate the parameters of ESTAR model by minimising the Akaike information

criterion (AIC), which is the required objective function in this problem.

Oilsardine landings time-series data in Kerala, India for the period form 1961 to 2008, sourced from ICAR-Central Marine Fisheries Research Institute, Kochi, India, was used for the study. From the total 48 data points denoted as  $\{X_t, t = 0.1, \dots, 47\}$ , first 45 data points corresponding to fish landings for the period 1961 to 2005 were used for building the model and remaining 3 data points for validation purpose. The entire data analysis was carried out using MATLAB, Ver. 7.4 software package.

The performance of fitted models is 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:

$$MSPE = 1/3 \sum_{i=0}^2 \{Y_{T+i} - \hat{Y}_{T+i}\}^2$$

$$MAPE = 1/3 \sum_{i=0}^2 \{|Y_{T+i} - \hat{Y}_{T+i}|\}$$

$$RMAPE = 1/3 \sum_{i=0}^2 \{|Y_{T+i} - \hat{Y}_{T+i}| / Y_{T+i}\} \times 100$$

Fig. 1 shows the directed scatter diagram, which indicates asymmetry in the joint distribution of the observations. This asymmetry indicates that the joint distributions of  $(X_t, X_{t-j})$  is non-Gaussian, as two-dimensional normal distribution cannot be asymmetric. Obviously, conventional ARIMA model for the given time-series data may not be able to describe the dataset satisfactorily.

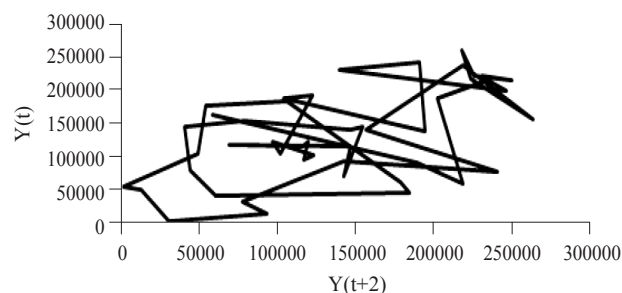


Fig. 1. Directed scatter diagram of oilsardine landings data at lag 2

To this end, several ESTAR models were accordingly fitted to the data and the best model was identified on the basis of minimum AIC criterion (Table 1). The best ESTAR model is given by:

$$Y_t = 42.93 + 0.56Y_{t-1} + 0.11Y_{t-2} + [29.04 + 0.39Y_{t-1} + 0.03Y_{t-2}] [1 - \exp\{-0.4(y_{t-2} - 0.82)^2\}]$$

Table 1. ESTAR models fitted and corresponding AIC values

Model	AIC value
ESTAR (1)	228.67
ESTAR (2)	220.62
ESTAR (3)	249.13
ESTAR (4)	272.60
ESTAR (5)	282.81
ESTAR (6)	288.61

Fig. 2 shows that the fitted model is able to properly capture the cyclical fluctuations present in the data set. Subsequently, portmanteau test was also carried out and it was inferred that the fitted ESTAR model was properly specified. The fitted model has further been validated by carrying out one-step ahead forecasts, which also shows that it is able to capture the underlying nonlinear and cyclicity phenomena quite satisfactorily.

Further, a comparative study intended to evaluate the ARIMA and ESTAR models on their ability to produce

Table 2. Performance of one-step ahead forecasts

Model / Criterion	ARIMA	ESTAR
MSPE	126.15	41.81
MAPE	32.05	17.24
RMAPE	13.50	7.31

Table 3. One-step ahead forecasts of oilsardine landings ('000 t)

Year	Actual	Forecast
2009	167.43	161.53
2010	218.94	218.85
2011	99.99	99.17
2012	211.40	214.82

time-series model capable of capturing the non-Gaussian characteristics, through GA technique is described. Illustration is also provided based on the time-series data on oilsardine landings from Kerala. From the analysis of data it was concluded that ESTAR model fitted through GA has performed better than ARIMA model. It is suggested

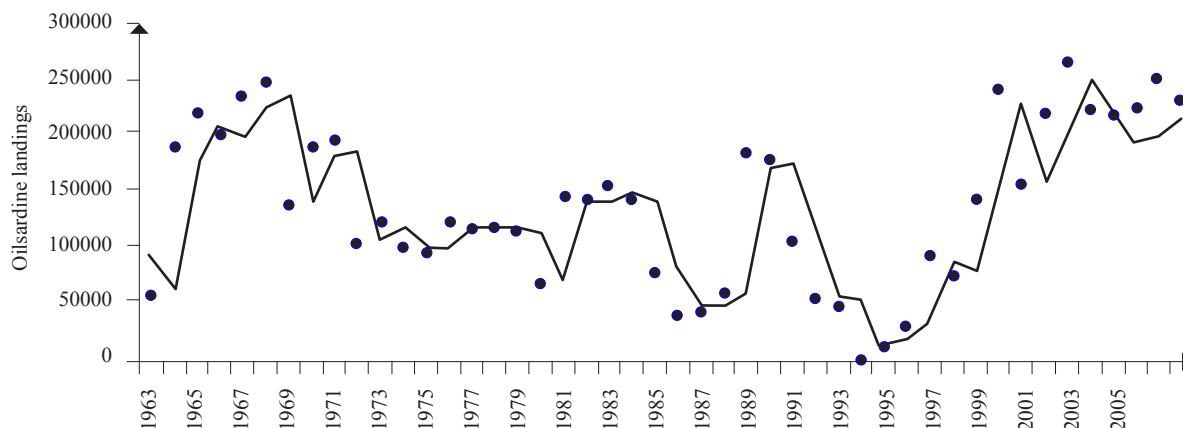


Fig. 2. Fitted ESTAR model along with data points

forecast was also 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:

$$X_{t+1} = 43.40 + 0.76X_t - 0.07X_{t-1} + \mathcal{E}_{t+1}$$

The MSPE, MAPE and RMAPE values for fitted ESTAR model are found to be lower than the corresponding ones for fitted ARIMA model. This indicates the superiority of ESTAR model over ARIMA model for forecasting purposes for the data under consideration. Further, the out-of-sample forecasts for the years 2009 up to 2012 were also computed and presented in Tables 2 and 3.

In this paper, methodology for fitting ESTAR nonlinear time-series model, a parametric nonlinear

that, for modelling and forecasting cyclical time-series data, this model rather than the ARIMA model, may be used.

**References**

Baragona, R., Battaglia, F. and Cucina, D. 2004. Estimating threshold subset autoregressive moving-average models by genetic algorithms. *Metron*, 6: 39-61.

Fan, J. and Yao, Q. 2003. *Nonlinear time series: Nonparametric and parametric methods*. Springer, New York.

Goldberg, D. E. 2009. *Genetic algorithms in search, optimisation and machine learning*. Dorling Kindersley, New Delhi.

Lundbergh, S., Terasvirta, T. and Van Dijk, D. 2003. Time varying smooth transition autoregressive models. *J. Busi. Econ. Stat.*, 21: 104-121.

- Nampoothiri, C. K. and Balakrishna, N. 2000. Threshold autoregressive model for a time-series data. *J. Indian Soc. Agr. Stat.*, 53: 151-160.
- Noble, A. 1980. Is there a ten year cycle in the mackerel fishery? *Seafood Exp. J.*, 12(4): 9-13.
- Noble, A. and Sathianandan, T. V. 1991. Trend analysis in all India mackerel catches using ARIMA models. *Indian J. Fish.*, 38(2): 119-122.
- Sathianandan, T. V. and Alagaraja, K. 1998. Spectral decomposition of the all India landings of oilsardine, mackerel and Bombayduck. *Indian J. Fish.*, 45(1): 13-20.
- Sathianandan, T. V. and Srinath, M. 1995. Time series analysis of marine fish landings in India. *J. Mar. Biol. Ass. India*, 37(2): 171-178.
- Sathianandan, T. V. and Jayasankar, J. 2009. Managing marine fishery in Kerala through simulation using surplus production model, genetic algorithm and spectral methods. *Indian J. Fish.*, 56(3): 163-168.
- Terasvirta, T. 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. *J. Amer. Statist. Ass.*, 89: 208-218.
- Van Dijk, D., Terasvirta, T. and Franses, P. H. 2002. Smooth transition autoregressive models - a survey of recent developments. *Econ. Rev.*, 21: 1-47.
- Venugopalan, R. and Srinath, M. 1998. Modelling and forecasting fish catches: Comparison of regression, univariate and multivariate time series methods. *Indian J. Fish.*, 45(3): 227-237.
- Wu, B. and Chang, C. L. 2002. Using genetic algorithms to parameters (d, r) estimation for threshold autoregressive models. *Comput. Stat. Data Anal.*, 38: 315-330.