Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 65(2) 2011 249-256

Methodology for Combining Linear and Nonlinear Time-Series Models for Cyclical Data

Himadri Ghosh, Bishal Gurung and Prajneshu*
Indian Agricultural Statistics Research Institute, New Delhi

Received 03 January 2011; Revised 16 May 2011; Accepted 30 May 2011

SUMMARY

For modelling and forecasting of cyclical time-series data, linear time-series models, like *Autoregressive integrated moving average* (ARIMA) model of order more than one, and *Nonlinear time-series* models, like *Exponential autoregressive* (EXPAR) and *Self-exciting threshold autoregressive* (SETAR) models are generally employed. In practical situations, exact data generating process of time-series observations is not known. Therefore, fitted values from linear and nonlinear models may be used as explanatory variables to empirically describe the same. In this paper, the ARIMA, EXPAR and SETAR models, which are capable of capturing the cyclical behaviour are studied. In order to improve modelling and forecasting capabilities of the models, these are combined by using the *Constant coefficient regression method* (A.E.S.(C)) as well as the *Time-varying coefficient regression method* (A.E.S.(T.V.)) through *Kalman filter* (KF) technique. As an illustration, the models are then applied to describe annual Mackerel catch time-series data of Karnataka. Performance of fitted models is examined by computing various measures of goodness of fit, viz. *Normalized Akaike information criterion* (NAIC), *Bayesian information criterion* (BIC) and *Mean square error* (MSE). Finally, forecasting performance of fitted models is evaluated by *Mean square prediction error* (MSPE) criterion. It is found that the combined model fitted by using the A.E.S.(T.V.) has performed best for the data under consideration.

Keywords : Exponential autoregressive model, Self-exciting threshold autoregressive model, Kalman filter, Normalized Akaike information criterion, Bayesian information criterion, Mean square error, Combined models.

1. INTRODUCTION

Evidently, it is not possible to identify exactly the underlying true process. However, a combination of 'linear' and 'nonlinear' models may approximate the data generating process quite well. As pointed out by Terui and van Dijk (2002), a linear model generally performs well for short-term forecasting whereas nonlinear model can do so for long-term forecasting. Another reason for combining the models is that it is possible for a process to switch its structure over the observation period between a *linear* and a *nonlinear* structure. Several schemes for combining forecasts of different models have been studied by various authors (see e.g. Granger and Ramanathan 1984). The models can be combined either by employing *Constant coefficient regression method* (A.E.S.(C)) or the *Time-varying coefficient regression method* (A.E.S.(T.V.)). In the latter, state space equations, viz. *Transition equation* and *Measurement equation* are obtained. Subsequently, an efficient computational procedure, called the *Kalman filter* (KF), may be employed for prediction and smoothing purposes. The KF is used for computing optimal prediction of state vector. By using A.E.S.(C), contribution of each component can be evaluated for the entire time horizon whereas A.E.S.(T.V.) can be used to evaluate contribution of each component at every time-point.

*Corresponding author : Prajneshu
E-mail address : prajnesh@iasri.res.in

In this paper, *Linear* time-series model, viz. the *Autoregressive integrated moving average* (ARIMA) model of order more than one and *Nonlinear* time-series models, like *Exponential autoregressive* (EXPAR) and *Self-exciting threshold autoregressive* (SETAR) models (Fan and Yao 2003) are considered for combining them as all of these are capable of describing cyclical pattern of time-series data. As an illustration, annual Mackerel catch time-series data of Karnataka during 1961 to 2008 is considered for modelling and forecasting purposes. Individual fitted values and their forecasts from marginal models are also obtained. These are used as explanatory variables to obtain estimates of weights in A.E.S.(C) and to obtain predicted values of changing coefficients in A.E.S.(T.V.). For more than one-step-ahead forecasts, naive forecasts from individual fitted linear and nonlinear models are used to obtain forecasts for fitted combined models. Finally, forecasts of observations for some hold-out data is carried out. It is shown that the combined model fitted through A.E.S.(T.V.) has performed the best.

2. DESCRIPTION OF MODELS

A brief description of ARIMA, EXPAR, and SETAR models along with their estimation procedures is given below.

2.1 Autoregressive Integrated Moving Average (ARIMA) Model

A univariate time-series (X_t) satisfies an ARIMA model, denoted by ARIMA (p, d, q), if

$$\varphi(L)\Delta^d X_t = \theta(L)\xi_t, \tag{2.1}$$

where $\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \ldots - \varphi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q$ are respectively the autoregressive and moving average terms, ξ_t is a white-noise process with mean zero and variance σ^2, and Δ^d indicates that the series is first-differenced d times. The ARIMA methodology is carried out in three stages, viz. Identification, estimation and diagnostic checking. Parameters of the tentatively selected model at the identification stage are estimated at the estimation stage and adequacy of selected model is tested at the diagnostic checking stage. If the model is found to be inadequate, the three stages are repeated until satisfactory ARIMA model is selected for describing the time-series under consideration. An excellent description of this methodology is given in Box *et al.* (2007). The ARIMA models, however, appear insufficient and beyond the *linear* domain, there are many *nonlinear* forms to be explored because the former are not able to take into account important features of many observed data sets.

2.2 Exponential Autoregressive (EXPAR) Model

The EXPAR parametric model, introduced for modelling and forecasting of cyclical data, is a useful *Nonlinear time-series* 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 (Ghazal and Elhassanein 2009). An EXPAR (p) model may explicitly be written as

$$X_t = \{\varphi_1 + \pi_1 \exp(-\gamma X_{t-1}^2)\} X_{t-1} + \ldots$$
$$+ \{\varphi_p + \pi_p \exp(-\gamma X_{t-1}^2)\} X_{t-p} + \varepsilon_t \tag{2.2}$$

where φ_i and π_i represent the autoregressive and exponential autoregressive parameters at lag i, $\gamma > 0$ is some scaling constant and $\{\varepsilon_t\}$ is white noise process with mean zero and variance σ_ε^2. The value of γ is selected such that $\exp(-\gamma X_t^2)$ varies reasonably widely over the range (0,1). Also, note that (2.2) has a regime-switching behaviour with respect to delayed observation, in the sense that, if $|X_{t-1}|$ is large, eq. (2.2) is similar to an autoregressive model with parameters approximately equal to (φ_1, ..., φ_p), while if $|X_{t-1}|$ is small, the autoregressive parameters switch to ($\varphi_1 + \pi_1$, ..., $\varphi_p + \pi_p$).

A brief description of the procedure for estimating parameters of (2.2) is as follows (Baragona *et al.* 2002). The algorithm requires that an interval (a, b), $a, b \geq 0$, be pre-specified for the γ values in (2.2). This interval is split in M sub-intervals, so that a grid of candidate γ values is built. Let $\delta = \dfrac{b-a}{M}$ and $\gamma = a$. Then, for M times, following steps are performed:

(i) Replace γ by ($\gamma + \delta$)

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

(iii) Compute the NAIC, defined as

$$\text{NAIC} = \frac{N \log(\hat{\sigma}^2) + 2(2p+1)}{\text{Effective sample size}} \quad (2.3)$$

and repeat step (ii) for $p = 1, \ldots, P$, where P is a pre-specified integer greater than 1. Final estimates of parameters are obtained by minimizing the NAIC.

2.3 Self-Exciting Threshold Autoregressive (SETAR) Model

This model was proposed to describe a particular data generating process by a piecewise linear autoregressive model. The SETAR model is governed by a known variable which determines as to whether each of the sub-models is active or not and belongs to *Threshold autoregressive* (TAR) family, which is defined as

$$X_t = \sum_{j=1}^{k} \left[\varphi_0^{(j)} + \sum_{i=1}^{p_i} \varphi_i^{(j)} X_{t-i} + \varepsilon_t^{(j)} \right] I^{(j)}(q_t) \quad (2.4)$$

where $\varepsilon_t^{(j)}$ is white noise process with mean zero and variance $\sigma^{2(j)}$ and $I^{(j)}(.)$ is an indicator function, i.e.

$$I^{(j)}(q_t) = \begin{cases} 1, & \text{if } q_t \in R_j \\ 0, & \text{otherwise} \end{cases}$$

$R_j = (r_{j-1}, r_j]$ is a partition of the real line R defined by a linearly ordered subset of the real numbers $\{r_0, r_1, \ldots, r_k\}$, such that $r_0 < r_1 < \ldots < r_k$, $r_0 = -\infty$, and $r_k = \infty$. The k autoregressive sub-models will be active or not depending on the values of q_t.

The SETAR model is a special case of TAR model when $q_t = X_{t-d}$. It is defined as

$$X_t = \sum_{j=1}^{k} \left[\varphi_0^{(j)} + \sum_{i=1}^{p_i} \varphi_i^{(j)} X_{t-i} + \varepsilon_t^{(j)} \right] I^{(j)}(X_{t-d}) \quad (2.5)$$

where d is length of the threshold and is called "delay parameter". Ghosh *et al.* (2006) applied SETAR model for forecasting of annual lac export data of India through Tong's "Search algorithm" procedure. Subsequently, optimal out-of-sample forecast formulae were also developed for SETAR two-regime model by recursive use of conditional expectation. However, one limitation of this algorithm is that the number of models to be searched becomes very large. So, Wu and Chang (2002) proposed an efficient stochastic search optimization procedure based on *Binary-coded genetic algorithm* (GA) for estimation of parameters of SETAR models. The GA combines Darwin's principle of "natural selection" and "survival of the fittest" and uses selection, crossover and mutation operators for finding optimal solution. Primary objective of selection operator is to make duplicates of good solutions which are used for applying crossover operator to produce better offspring. The pre-specified value for probability of crossover (p_c) decides as to how many individuals are used in the crossover operation. Mutation operator is applied for random changes to individual parent to form children and keep diversity in the population. The mutation probability (p_m) controls number of individuals to be mutated. Success of GA is also based on crossover probability, mutation probability and ability of crossover operator to maintain spread between offspring proportional to that of parent along with more chance of creating near (far) offspring when the parents are near (far) (see Deb 2005).

Iquebal *et al.* (2010) applied the SETAR model to describe annual lac export time-series data of India by employing the Real-coded genetic algorithm (RCGA). Superiority of this methodology over "Search algorithm" procedure was demonstrated for the time-series data under consideration. The best model was identified on the basis of minimum NAIC criterion, defined as

$$\text{NAIC}(\theta) = \frac{\sum_{i=1}^{2} n_i \log\left(\frac{S_i}{n_i}\right) + 2\sum_{i=1}^{2}(p_i + 1)}{\text{Effective sample size}} \quad (2.6)$$

where

$$\theta = (d; r; \{p_i; \varphi_1^{(i)}, \ldots, \varphi_{p_i}^{(i)}, i = 1, 2\})' \quad (2.7)$$

n_i is the number of observations that belong to regime i and S_i is residual sum of squares for i^{th} subset SETAR model.

3. KALMAN FILTER AND STATE SPACE MODEL

State space model includes two classes of variables, the state variable and the observation variable, which are modelled as:

$$\alpha_{t+1} = \mathbf{F}_t\alpha_t + \mathbf{G}_t\varepsilon_t \quad (3.1)$$

$$X_t = \mathbf{H}_t'\alpha_t + v_t \quad (3.2)$$

where eq. (3.1) is called the *State transition equation*. It allows the state variable to change through time. Eq. (3.2) is called the *Measurement equation* and it relates state variable to an observation. It is assumed that $\{\varepsilon_t\}$ of eq. (3.1) and $\{v_t\}$ of eq. (3.2) are independent, zero mean, Gaussian white noise processes with

$$\mathrm{E}[v_t v_t'] = R_t \text{ and } \mathrm{E}[\varepsilon_t\varepsilon_t'] = \mathbf{Q}_t \quad (3.3)$$

After the model has been put in state space form, an efficient computational procedure, obtained through conditional mean and variance of state and measurement, can optimally predict the state. Also, measurement update equation filters the state at time-epoch t using *Minimum mean square error* criterion. This procedure, called Kalman filtering (KF), is applied recursively through time where the next state can be predicted followed by filtering on advancing the information χ_t to next time-epoch, i.e. $\chi_{t+1} = \{\mathbf{X}_{t'} : t' = t_0, t_0 + 1, ..., t_0 + t + 1\}$. Thus, knowledge of state variable at time $t = t_0$ together with the information at $t \geq t_0$ completely determines the behaviour of the process at time t. The recursive equations for implementing KF are discussed below:

Denote

$$\hat{\alpha}_{t|t-1} = \mathrm{E}\{\alpha_t | \chi_{t-1}\} \text{ and } \hat{\alpha}_{t|t} = \mathrm{E}\{\alpha_t | \chi_t\},\ t = t_0, t_0 + 1, ... \quad (3.4)$$

Assume $\hat{\alpha}_{t_0|t_0-1} = \mathrm{E}\{\alpha_0\}$ and $\mathbf{\Sigma}_{t_0|t_0-1} = \mathbf{P}_0$. Using eqs. (3.1) and (3.2), the recursive filter equations are as follows:

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + \mathbf{\Sigma}_{t|t-1}\mathbf{H}_t(\mathbf{H}_t'\mathbf{\Sigma}_{t|t-1}\mathbf{H}_t + R_t)^{-1}(X_t - \mathbf{H}_t'\hat{\alpha}_{t|t-1}) \quad (3.5)$$

The measure of performance to predict α_t using χ_t is

$$\mathbf{\Sigma}_{t|t} = \mathbf{\Sigma}_{t|t-1} - \mathbf{\Sigma}_{t|t-1}\mathbf{H}_t(\mathbf{H}_t'\mathbf{\Sigma}_{t|t-1}\mathbf{H}_t + R_t)^{-1}\mathbf{H}_t'\mathbf{\Sigma}_{t|t-1} \quad (3.6)$$

Using recursive filter equations (3.1) and (3.2), we can obtain $\hat{\alpha}_{t+1|t}$ as

$$\hat{\alpha}_{t+1|t} = \mathbf{F}_t\hat{\alpha}_{t|t} \quad (3.7a)$$

and

$$\mathbf{\Sigma}_{t+1|t} = \mathbf{F}_t\mathbf{\Sigma}_{t|t}\mathbf{F}_t' + \mathbf{G}_t\mathbf{Q}_t\mathbf{G}_t' \quad (3.7b)$$

Eq. (3.7a) can also be written as

$$\hat{\alpha}_{t+1|t} = \mathbf{F}_t\hat{\alpha}_{t|t-1} + \mathbf{F}_t\mathbf{\Sigma}_{t|t-1}\mathbf{H}_t(\mathbf{H}_t'\mathbf{\Sigma}_{t|t-1}\mathbf{H}_t + R_t)^{-1}(X_t - \mathbf{H}_t'\hat{\alpha}_{t|t-1}) \quad (3.8)$$

which implies that time update rules for each forecast of state are weighted average of previous forecast $\hat{\alpha}_{t|t-1}$ and forecast error $(X_t - \mathbf{H}_t'\hat{\alpha}_{t|t-1})$. After obtaining $\hat{\alpha}_{t|t-1}$, one may predict X_t by the optimal predictor $\hat{X}_{t|t-1}$, where

$$\hat{X}_{t|t-1} = \mathbf{H}_t'\hat{\alpha}_{t|t-1} \quad (3.9)$$

and the conditional error variance due to predictor $\hat{X}_{t|t-1}$ is

$$\mathbf{H}_t'\mathbf{\Sigma}_{t|t-1}\mathbf{H}_t + R_t \quad (3.10)$$

A good description of Kalman filtering is given in Durbin and Koopman (2001).

4. THE COMBINED MODELS

Best ARIMA model is identified on the basis of minimum NAIC, which is defined as

$$\text{NAIC} = \frac{N\log(\hat{\sigma}^2) + 2(p+q+1)}{\text{Effective sample size}} \quad (4.1)$$

Further, for fitting of EXPAR and SETAR models, expressions for NAIC are given respectively in (2.3) and (2.6). Two methods are considered to obtain the combined models. The first one, viz. *Constant coefficient regression method* (A.E.S.(C)) is given by the model:

$$Y_t = \beta^0 + \beta^a X_t^a + \beta^e X_t^e + \beta^s X_t^s + u_t \quad (4.2)$$

where X_t^a, X_t^e and X_t^s are the mean marginal fitted values generated respectively by ARIMA, EXPAR and SETAR models. The β's are regression coefficients and $\{u_t\}$ is a white noise process. Here, parameters are estimated by ordinary least squares method to get fitted values of Y_t in eq. (4.2). In the second method, viz. *Time-varying coefficient regression method* (A.E.S.(T.V.)), the β's are taken to be time-varying stochastic process denoted as

$$Y_t = \beta_t^0 + \beta_t^a X_t^a + \beta_t^e X_t^e + \beta_t^s X_t^s + v_t \quad (4.3)$$

Noticing that eq. (4.3) is univariate, it can also be written as

$$Y_t = \mathbf{X}'_t \beta_t + v_t;\ v_t \sim \mathrm{N}(0, R) \qquad (4.4)$$

$$\beta_t = \beta_{t-1} + \mathbf{e}_t;\ \mathbf{e}_t \sim \mathrm{N}(0, \mathbf{Q}) \qquad (4.5)$$

where $\mathbf{X}_t = [1, X_t^a, X_t^e, X_t^s]'$ and $\beta_t = (\beta_t^0, \beta_t^a, \beta_t^e, \beta_t^s)'$. The $\{v_t\}$ and $\{\mathbf{e}_t\}$ are independent white noise processes. Eqs. (4.4) and (4.5) can be interpreted as a state space model given in eqs. (3.1) and (3.2). Here eq. (4.4) is the *measurement equation* which defines the distribution of Y_t, $t \geq 1$, and eq. (4.5) is the *state equation* which defines the distribution of β_t for every $t \geq 1$ (Terui and van Dijk 2002). Now KF is applied to predict β_t by $\hat{\beta}_{t|t-1}$, which are used to compute the fitted values of Y_t in eq. (4.4). For computing one-step ahead forecast from eq. (4.4), predicted value of β_{T+1} is $\hat{\beta}_{T+1|T}$, which is same as $\hat{\beta}_{T|T}$. Further, forecast error variance of Y_{T+1} is $\mathbf{X}'_{T+1}\,\boldsymbol{\Sigma}_{T+1|T}\mathbf{X}_{T+1} + \hat{R}$. Similarly, for computing two-step ahead forecast, it is observed from eq. (4.5), that $\hat{\beta}_{T+2|T}$ is $\hat{\beta}_{T|T}$ with two-step-ahead prediction error variance given by $\boldsymbol{\Sigma}_{T+2\,|\,T} = \boldsymbol{\Sigma}_{T|T} + 2\hat{\mathbf{Q}}$. Here $\hat{R}$ and $\hat{\mathbf{Q}}$ are estimated values of R and $\mathbf{Q}$ obtained by using *Prediction error decomposition technique*. Finally, two-step ahead prediction error variance of Y_{T+2} is computed as $\mathbf{X}'_{T+2}\boldsymbol{\Sigma}_{T+2|T}\mathbf{X}_{T+2} + \hat{R}$.

5. AN ILLUSTRATION

Annual Mackerel catch time-series data of Karnataka (tonnes) during the period 1961 to 2008, obtained from Central Marine Fisheries Research Institute, Kochi, India are considered for data analysis. First 45 data points corresponding to the period 1961 to 2005 are used for model building and the remaining 3 data points, i.e. from 2006 to 2008 are used for validation purpose.

5.1 Directed Scatter Diagram

Preliminary *Exploratory data analysis* is carried out to justify the choice of EXPAR and SETAR models to describe the data under consideration. A directed scatter diagram is a powerful tool for analyzing a *Nonlinear time-series*. It consists of diagrams of (X_t, X_{t-j}), $j = \pm 1, \pm 2, \ldots, \pm p$, where X_t is the annual Mackerel catch of Karnataka (tonnes) at time t, and p is the number of possible lags. The directed scatter diagrams exhibited in Fig. 1 show asymmetry in the joint distribution of observations, indicating thereby that the joint distributions of (X_t, X_{t-1}) are non-Gaussian, as a two-dimensional normal distribution cannot be asymmetric.

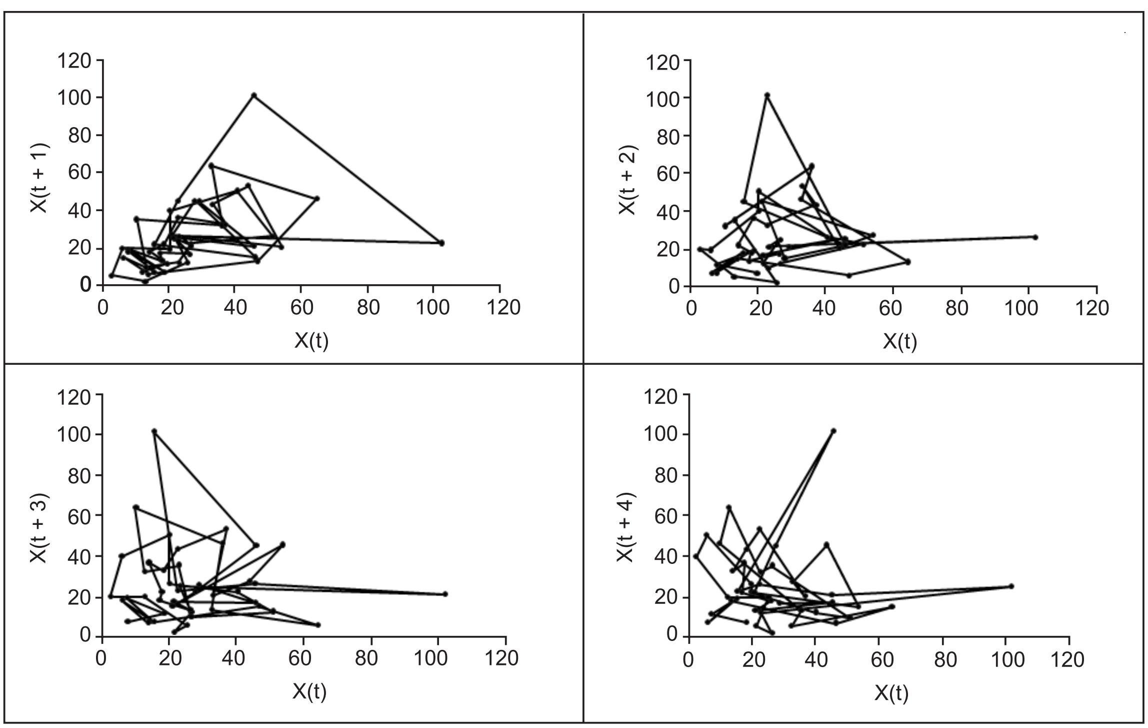

Fig. 1. Directed scatter diagram of annual Mackerel catch data (tonnes) of Karnataka

At the first step, three marginal models, viz. ARIMA, EXPAR and SETAR are fitted; these are then used in second step to obtain two combined models corresponding to two methods, viz. A.E.S.(C) and A.E.S.(T.V.). By using the minimum NAIC criterion, best ARIMA model is found to be ARIMA (2,0,0) given by

$$X_t = 18.53 + 0.38X_{t-1} - 0.08X_{t-2} + \varepsilon_t \quad (5.1)$$

with var$\{\varepsilon_t\}$ = 310.96. The standard errors of parameter estimates ($\hat{\mu}, \hat{\varphi}_1, \hat{\varphi}_2$) are respectively computed as (5.34, 0.16, 0.15). It may be noted that roots of characteristic equation, viz. $(1 - 0.38L + 0.08L^2) = 0$ are complex as $0.08 > (-0.38)^2/4$, which lends support to existence of a cyclical pattern in annual Mackerel catch data.

For fitting SETAR model, minimization of NAIC given in eq. (2.6) is carried out using RCGA, where crossover and mutation probabilities are taken as 0.90 and 0.10 respectively. The selection operator is applied using "Tournament selection" method. After 1500 iterations, optimal solution is obtained and fitted SETAR model is

$$X_t = \begin{cases} 10.69 + 0.22X_{t-1} + 0.10X_{t-2} - 0.07X_{t-3} + \varepsilon_t^{(1)}, \\ \qquad\qquad\qquad\qquad \text{if } X_{t-2} \le 22.26 \\ 34.73 + 0.49\,X_{t-1} + \varepsilon_t^{(2)}, \qquad \text{if } X_{t-2} > 22.26 \end{cases} \quad (5.2)$$

with var$\{\varepsilon_t^{(1)}\}$ = 230.74 and var$\{\varepsilon_t^{(2)}\}$ = 867.80. To estimate standard errors of parameters, 1000 bootstrap samples are generated and for each sample, parameters are estimated using RCGA. The standard errors for $\left(\hat{\phi}_0^{(1)}, \hat{\phi}_1^{(1)}, \hat{\phi}_2^{(1)}\, \hat{\phi}_3^{(1)}, \hat{\phi}_0^{(2)}, \hat{\phi}_1^{(2)}\right)$ are respectively computed as (0.54, 0.02, 0.05, 0.02, 0.88, 0.03). By using minimum NAIC criterion given in eq. (2.3), fitted EXPAR model is obtained as

$$X_t = \{0.58 + 0.81\exp(-0.0004X_{t-1}^2)\}\; X_{t-1}$$
$$+\{-0.58 + 0.73\exp(-0.0004X_{t-1}^2)\}\; X_{t-2} + \varepsilon_t \quad (5.3)$$

with var$\{\varepsilon_t\}$ = 307.80. The standard errors of parameter estimates ($\hat{\varphi}_1, \hat{\pi}_1, \hat{\varphi}_2, \hat{\pi}_2$) are respectively computed as (0.16, 0.45, 0.17, 0.43). Using eq. (4.2), fitted combined model through A.E.S.(C) is given by

$$X_t = 3.49 + 0.10X_t^a + 0.60X_t^e + 0.16X_t^s + u_t \quad (5.4)$$

with var$\{u_t\}$ = 299.26. The standard errors of parameter estimates ($\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$) are respectively computed as (11.12, 0.69, 0.50, 0.20). It may be noted that the matrix of regressors comprising X_t^a, X_t^e and X_t^s is of full rank, implying thereby that problem of multicollinearity is not present for the data under consideration. The time-varying coefficients in respect of ARIMA, EXPAR and SETAR marginal models by A.E.S.(T.V.) method are depicted in Fig. 2. It may be noted that there is not much variation of regression coefficient in the EXPAR component. However, it may be emphasized that this estimate for all the years is not only highly significant but also plays a dominant role for a number of time-epochs particularly in the initial phase, implying thereby that EXPAR model is a good model for describing underlying nonlinear dynamics of data generating process. Further, SETAR model is dominant over ARIMA at initial stages of the time-series as well as over some time-epochs from 1990-1994 during which there are fluctuations of large magnitudes. ARIMA dominates in a few time-epochs during 1971-1974 and

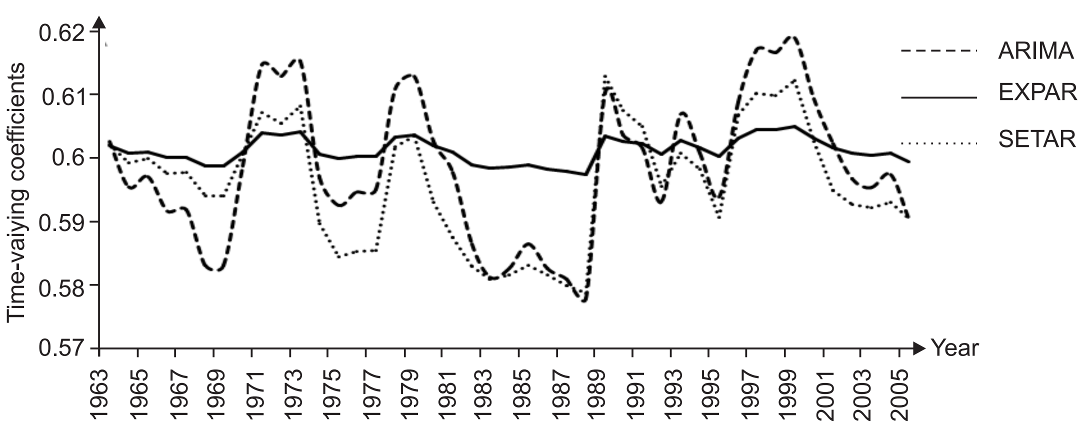

Fig. 2. Time-varying coefficients of the combined model

later parts of the time-series from 1997 up to 2001, where moderate fluctuations in the data are observed. Thus, all the three marginal models are contributing towards improving modelling aspects of the data.

The NAIC, BIC and MSE values of three marginal models as well as two combined models are computed and reported in Table 1. A perusal indicates that combined models improve description of border line nonlinearity present in time-series data. Further, the combined model fitted through A.E.S.(T.V.) has performed better than the one fitted through A.E.S.(C).

Table 1. Goodness of fit of models

Model / Criterion	ARIMA	EXPAR	SETAR	Combined model by A.E.S. (C)	Combined model by A.E.S. (T.V.)
NAIC	6.32	6.10	6.18	5.88	5.29
BIC	258.70	251.20	252.74	248.44	227.87
MSE	289.26	279.17	283.88	271.39	200.22

5.2 Evaluation of Forecasting Performance

In this sub-section, a comparative study is carried out to evaluate the combined models on their ability to forecast. To this end, forecasting performance for 3 data points corresponding to Mackerel catch for years 2006 to 2008 as hold-out data is studied. One-step ahead forecasts are obtained and performance of fitted models is compared on the basis of one-step ahead *Mean square prediction error* (MSPE) given by

$$\text{MSPE} = \frac{1}{3}\sum_{i=0}^{2}\{Y_{T+i+1} - \hat{Y}_{N+i+1}\}^2 \qquad (5.2.1)$$

The forecast error variances of Y_{T+i+1}, $i = 0, 1, 2$ are derived using recursive and theoretical approach. To this end, naive approach is used to obtain one-step and two-step ahead forecasts of fitted ARIMA, EXPAR and SETAR models. As discussed in Section 4, in case of recursive approach, advancing the time-epoch in each step, time update equation of state, $\hat{\boldsymbol{\beta}}_{T+i+1|T+i} = \hat{\boldsymbol{\beta}}_{T+i|T+i}$ is obtained which is used to calculate 3 forecast values of $\hat{Y}_{T+i+1} = \mathbf{X}'_{T+i+1}\ \hat{\boldsymbol{\beta}}_{T+i+1|T+i}$, $i = 0, 1, 2$ for annual Mackerel catch data. The MSPE of forecast values obtained by the five models are computed and reported in Table 2. Since the combined model fitted through A.E.S.(T.V.) has performed best, attempt is made to obtain its theoretical forecast error variances as well. The 3 forecast error variances are calculated as $\mathbf{X}'_{T+i+1}\ \boldsymbol{\Sigma}_{T+i+1\,|\,T+i}\mathbf{X}_{T+i+1} + \hat{R}$, where $\boldsymbol{\Sigma}_{T+i+1\,|\,T+i} = \boldsymbol{\Sigma}_{T+i\,|\,T+i} + \hat{\mathbf{Q}}$. One-step ahead forecast error variance is then summarized by calculating the mean of 3 forecast error variances and is obtained as 98.14. It may be noted from Table 2 that the MSPE for one-step ahead forecast for combined model fitted through A.E.S.(T.V.) method, viz. 91.58 is quite close to the corresponding theoretical forecast error variance, viz. 98.14, indicating thereby that the hold-out data successfully provides a good estimate of true one-step ahead forecast error variance. Similarly, MSPE for two-step ahead forecast is given by

$$\text{MSPE} = \frac{1}{2}\sum_{i=0}^{1}\{Y_{N+i+2} - \hat{Y}_{N+i+2}\}^2 \qquad (5.2.2)$$

Following the above approach for two-step ahead forecasts of state, $\hat{\boldsymbol{\beta}}_{T+i+2|T+i}$, $i = 0, 1$ are computed, which are used to obtain $\hat{Y}_{T+i+2} = (\mathbf{X}'_{T+i+2} \cdot \hat{\boldsymbol{\beta}}_{T+i+2|T+i})$, $i = 0, 1$ and are reported in Table 3. A perusal of Tables 2 and 3 shows that the combined model fitted through A.E.S.(T.V.) performs better than the one fitted through the A.E.S.(C) for forecasting purposes as well. Finally, from Table 3, for the former model, MSPE for two-step ahead forecasts, viz. 92.95 is found to be quite close to the corresponding theoretical forecast error variance, viz. 101.33 obtained by computing mean of the 2 forecast error variances. This implies that the model is successful for forecasting hold-out data.

Table 2. Performance of one-step ahead forecasts

Model / Criterion	ARIMA	EXPAR	SETAR	Combined model by A.E.S. (C)	Combined model by A.E.S. (T.V.)
MSPE	145.56	142.32	211.79	96.55	91.58

Table 3. Performance of two-step ahead forecasts

Model / Criterion	ARIMA	EXPAR	SETAR	Combined model by A.E.S. (C)	Combined model by A.E.S. (T.V.)
MSPE	189.23	134.99	394.30	166.25	92.95

6. CONCLUSIONS

In this paper, methodology for combining one linear and two nonlinear time-series models capable of describing cyclical data is discussed. To this end, two methods, viz. C*onstant coefficient regression method* and *Time-varying coefficient regression method* are considered. Formulae for optimal one-step and two-step ahead forecasts are derived theoretically for A.E.S.(T.V.). Superiority of this model for modelling as well as forecasting purposes is clearly demonstrated for the data under consideration. As future work, this methodology may be extended to tackle the problem of multicollinearity in the combined models. A systematic study on the value of forecasts of simple models compared with the forecasts of a very flexible nonlinear time-series model may be of considerable interest. In this context, robustness with respect to outliers and/or varying volatility needs to be analyzed. There is also a need to study a more decision-theoretic analysis of the methods, like a Bayesian approach.

ACKNOWLEDGEMENTS

Authors are grateful to the referee for valuable comments. Dr. Himadri Ghosh and Dr. Prajneshu thank the Department of Science and Technology, New Delhi for providing financial assistance during the course of this research work.

REFERENCES

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

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2007). *Time-Series Analysis: Forecasting and Control*. 3rd Ed. Dorling Kindersley, New Delhi.

Deb, K. (2005). *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley and Sons, Singapore.

Durbin, J. and Koopman, S.J. (2001). *Time-Series Analysis by State Space Methods*. Oxford University Press, New York.

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

Ghazal, M.A. and Elhassanein, A. (2009). Dynamics of EXPAR models for high frequency data. *Inter. J. Appl. Math. Statist.,* **14**, 88-96.

Ghosh, H., Sunilkumar, G. and Prajneshu (2006). Self-exciting threshold autoregressive models for describing cyclical data. *Cal. Stat. Assoc. Bull.*, **58**, 15-132.

Granger, C.W.J. and Ramanathan, R. (1984). Improved methods of forecasting. *J. Forecasting*, **3**, 197-204.

Iquebal, M.A., Ghosh, H. and Prajneshu (2010). Application of genetic algorithm for fitting self-exciting threshold autoregressive nonlinear time-series model. *J. Ind. Soc. Agril. Statist.,* **64***,* 391-398.

Terui, N. and van Dijk, H.K. (2002). Combined forecasts from linear and nonlinear time-series models. *Int. J. Forecasting,* **18**, 421-438.

Wu, B. and Chang, C.L. (2002). Using genetic algorithms to parameters (*d*, *r*) estimation for threshold autoregressive models. *Comput. Statist. Data Anal.*, **38**, 315-330.