Structural Time series Modelling and its application in agriculture

S.Ravichandran
Directorate of Rice Research, Hyderabad

1. Introduction

ARIMA time series methodology is widely used for modelling time series data. This methodology can be applied only when either the series under consideration is stationary or it can be made so by differencing, de-trending, or by any other means. Another disadvantage is that this approach is empirical in nature and does not provide any insight into the underlying mechanism. An alternative mechanistic approach, which is quite promising, is the **"Structural time series modelling (Harvey, 1996)".** Here, the basic philosophy is that characteristics of the data dictate the particular type of model to be adopted from the family. Purpose of present lecture is to discuss about STM methodology utilised for modelling time series data in the presence of trend, seasonal and cyclical fluctuations.

2. ARIMA Methodology

ARIMA method consists in developing a suitable ARIMA model, which in turn provides a comprehensive tool kit for time series model identification, parameter estimation forecasting. Box - Jenkins time series models amalgamate three types of processes, viz. autoregressive (AR) of order p, differencing to make a series of degree d and moving average (MA) process of order q and hence is written as ARIMA (p, d, q). This process assumes that the series under consideration is stationary. Main aim of analyzing any time series data is to forecast future values of the series under study. ARIMA model approach deals with short term forecasting rather than long term forecasting.

3. STRUCTURAL TIME SERIES MODELLING

Structural time series models are formulated in such a way that their components are stochastic; in other words, they are regarded as being driven by random disturbances. The key to handling structural time series models is the state space form, with the state of the system representing the various unobserved components such as trends, cycles and seasonals. Once in state space form (SSF), the Kalman filter provides the means of updating the state, as new observations become available. Predictions are made by extrapolating these components into the future. Kalman filter, a smoothing algorithm is used for obtaining the best estimate of the state at any point within the sample. Prediction and smoothing is carried out once the parameters governing the stochastic movements of the state variables have been estimated. Estimation of these parameters, known as **"hyper- parameters"**, is based on Kalman filter. This is because the likelihood function can be expressed in terms of one-step-ahead prediction errors, and these prediction errors emerge as a by-product of the filter. Once a model is estimated, its suitability can be assessed using goodness fit statistics. Ravichandran and Prajneshu (2001) compared the efficiencies of ARIMA and State Space Modelling utilising all-india Marine Products export data.

4. *Kalman filter*

Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least squares method. Kalman filter is a recursive procedure for computing optimal estimator of the state at a particular time, based on information available till that time (Meinhold and Singpurwalla, 1983). The filter is very powerful in several aspects: it supports estimation of past, present and even future states, and it can do so even when the precise nature of the modelled system is unknown.

5. Basic Structural Time Series Model

A **"Structural time series model"** is set up in terms of its various components, like trend, cyclical fluctuations, and seasonal variations, i.e.

$$Y_t = T_t + C_t + S_t + \varepsilon_t \ , \qquad (1)$$

where Y_t is the observed time series at time t, T_t , C_t , S_t , ε_t are the trend, cyclical, seasonal and irregular components.

(i) Local level model (LLM)

In the absence of seasonal and cyclical components, eq. (1) reduces to

$$Y_t = \mu_t + \varepsilon_t \ , \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \ , \quad t = 1, 2, \ldots, T \, . \qquad (2)$$

When the trend component (μ_t) does not show a steady upward or downward movement, it becomes a permanent component called "level". Sometimes, it is assumed to vary according to a random walk, i.e.

$$\mu_t = \mu_{t-1} + \eta_t \ , \quad \eta_t \sim N(0, \sigma_\eta^2) \ . \qquad (3)$$

Eqs. (2) and (3) together form the LLM. It may be noted that these equations are already in state space form. However, level μ_t is not directly observable but can be estimated. Further, forecast values are weighted average of the data points. When $\sigma_\varepsilon^2 = 0$, forecast is just the last observation and when $\sigma_\eta^2 = 0$, level is constant and the best forecast is sample mean. Level of time series vary over time depending on signal to noise ratio $q = \sigma_\eta^2 / \sigma_\varepsilon^2$. Estimation of μ_t , conditional on σ_ε^2 and σ_η^2 , is done recursively using Kalman filter and smoother **(Harvey, 1996)**. The parameters σ_ε^2 and σ_η^2 are unknown and are treated as hyperparameters. Likelihood function can be evaluated by Kalman filtering via prediction error decomposition **(Shumway and Stoffer, 2000)**. Once σ_ε^2 and σ_η^2 are known, one-step-ahead prediction of level, i.e. estimator of μ_{t+1} given $Y_t = \{ Y_1 , Y_2 , \ldots , Y_t \}$, viz.

$$a_{t+1} = E(\mu_{t+1} \mid Y_t), \qquad (4)$$

is evaluated recursively by Kalman filter. Prediction error variance

$$P_{t+1} = Var(\mu_{t+1} | Y_t) = Var(a_{t+1}) \quad (5)$$

is also obtained recursively. Reduced form of LLM is ARIMA (0, 1, 1) model.

(ii) Local linear trend model (LLTM)

As described by Harvey (1996), LLTM is given by eq. (10) along with the following two equations:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \ , \quad (6)$$

$$\beta_t = \beta_{t-1} + \xi_t \ , \ t = \ldots, -1, 0, 1, \ldots \ , \quad (7)$$

where $\eta_t \sim N(0, \sigma_\eta^2)$ and $\xi_t \sim N(0, \sigma_\xi^2)$. It may be mentioned that η_t , ξ_t and ε_t are independent of one another. If $\sigma_\eta^2 = \sigma_\xi^2 = 0$, eqs. (6) and (7) collapse to

$$\mu_t = \mu_{t-1} + \beta \ , \quad t = 1, 2, ..., T, \quad (8)$$

which can equivalently be written as

$$\mu_t = \alpha + \beta t \ , \quad t = 1, 2, ..., T, \quad (9)$$

showing that the deterministic linear trend is a limiting case.

LLTM is in state space form with state vector $\alpha_t = (\mu_t , \beta_t)$. Updating and prediction are carried out using Kalman filter by assuming that σ_ε^2 , σ_η^2 and σ_ξ^2 are known. Otherwise these can be estimated using maximum likelihood method for state space models (de Jong, 1988 ; Koopman and Shephard, 1992). Reduced form of a LLTM is ARIMA (0, 2, 2) model. Andrews (1994) showed that forecast function of LLTM performs better than that of corresponding ARIMA model. **Ravichandran and Muthuraman (2006)** utilised STM model trend for modelling and forecasting India's rice production. **Ravichandran and Prajneshu (2002)** compared the efficiencies of STM model for trend and Bayesian Analysis of Time Series (BATS) model while forecasting India's foodgrain production.

(iii) Local linear trend model with intervention effect (LLTMI)

Intervention analysis is concerned with making inference about effects of known events. These effects are measured by including intervention, or dummy variables in a dynamic model (Harvey and Durbin, 1986). LLTMI is described by the following equations:

$$Y_t = \mu_t + \lambda w_t + \varepsilon_t \qquad (10)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \lambda w_t + \eta_t \qquad (11)$$

$$\beta_t = \beta_{t-1} + \lambda w_t + \xi_t, \qquad (12)$$

where w_t is intervention variable and λ is its coefficient. The quantity w_t depends on the form which intervention is assumed to take. As for LLTM, estimation of state vector $\alpha_t = (\mu_t , \beta_t)$ for LLTMI is similarly carried out by putting the model in state space form and applying Kalman filter recursively by treating w_t as an explanatory variable. **Ravichandran and Prajneshu (2002)** utilised this model for sunflower yield forecasting.

6. Structural time series model for cyclical fluctuations

In the absence of seasonal components, STM model reduces to

$$Y_t = \mu_t + \psi_t + \varepsilon_t \ . \qquad (13)$$

For modelling cyclical fluctuations, the following three models are discussed:

i. Cycle Plus Noise Model (CNM)

Here, the trend μ_t is assumed to be constant. Thus, the functional form of CNM **(Harvey, 1996)** is

$$Y_t = \mu + \psi_t + \varepsilon_t \ , \ t = 1, 2, ..., T \ . \qquad (14)$$

The cyclical fluctuations ψ_t can be expressed as a mixture of sine and cosine terms:

$$\psi_t = \alpha \cos (\lambda_c t) + \beta \sin (\lambda_c t) \ , \qquad (15)$$

where λ_c, $(\alpha^2 + \beta^2)^{1/2}$, $\tan^{-1} (\beta / \alpha)$ represent respectively the frequency, amplitude, and phase. The cycles in eq. (4) need to be made stochastic by allowing the parameters α and β to evolve over time. Following Harvey (1996), the final form of eq. (4) can be written as

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix}, \qquad (16)$$

where the correlation coefficient $\rho \in [0\ ,\ 1]$ is a damping factor, and k_t & k_t^* are uncorrelated white-noise disturbance terms. Further, $\psi_0 = \alpha$, $\psi_0^* = \beta$. The new parameters are ψ_t, the current value of the cycle, and ψ_t^*, which appears by construction in order to form ψ_t. Eq. (16) is a vector AR (1) process. Let L denote the lag operator, i.e.

$$L\,\psi_t = \psi_{t-1}\,. \qquad (17)$$

Then eq. (5) can be written as

$$\begin{bmatrix} 1 - L\rho\cos\lambda_c & -L\rho\sin\lambda_c \\ L\rho\sin\lambda_c & 1 - L\rho\cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} k_t \\ k_t^* \end{bmatrix}. \qquad (18)$$

Substituting for ψ_t in eq. (3) yields

$$Y_t = \mu + \frac{(1 - L\rho\cos\lambda_c)\,k_t + (L\rho\sin\lambda_c)\,k_t^*}{(1 - 2L\rho\cos\lambda_c + L^2\rho^2)} + \varepsilon_t\,,\ t = 1, 2, ..., T, \qquad (19)$$

where ε_t is assumed to be uncorrelated with k_t and k_t^*.

For estimation of parameters, eq. (19) has to be put in state space form **(Meinhold and Singpurwalla, 1983)** and then Kalman filter, prediction and smoothing **(Koopman *et al.*, 1999)** are applied. In eq. (19), unobservable state, conditional on variances σ_ε^2 and σ_k^2, is done recursively using Kalman filter and smoother. In general, these are unknown and are treated as hyperparameters. Likelihood function can be evaluated by Kalman filter via prediction error decomposition **(Shumway and Stoffer, 2000)**. Maximizing likelihood function with respect to hyperparameters, using quasi-Newton optimization procedure, is referred to as “hyperparameter estimation”. After estimation of parameters, prediction and smoothing are performed. The reduced form from ARIMA family corresponding to CNM is **“Constant + ARIMA (2,2)”.**

ii. Trend Plus Cycle Model (TCM)

As described by **Harvey (1996)**, TCM is given by the following equations:

$$Y_t = \mu_t + \psi_t + \varepsilon_t \ , \ t = 1, 2, \ldots, T \ , \tag{20}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \ , \tag{21}$$

$$\beta_t = \beta_{t-1} + \xi_t \ , t = \ldots, -1, 0, 1, \ldots \tag{22}$$

In this, ε_t, η_t and ξ_t are the disturbance terms which follow Gaussian distributions with means 0 and variances σ_ε^2 , σ_η^2 and σ_ξ^2, called hyper-parameters of the model. As mentioned in CNM model, estimation of state vector and hyper-parameters is carried out by putting the model in state space form and subsequently Kalman filter is applied with proper initial values. The state space formulation of TCM is as follows:

The measurement equation is

$$Y_t = [1 \quad 0 \quad 1 \quad 0]\, \alpha_t + \varepsilon_t \ , \ t = 1, 2, \ldots, T, \tag{23}$$

where as the transition equation is

$$\alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \cdots \\ \psi_t \\ \psi_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & . & & \\ 0 & 1 & . & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & . & \rho\cos\lambda_c & \rho\sin\lambda_c \\ & & . & -\rho\sin\lambda_c & \rho\cos\lambda_c \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \cdots \\ \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_t \\ \xi_t \\ \cdots \\ k_t \\ k_t^* \end{bmatrix} \tag{24}$$

The covariance matrix of the vector of disturbances in eq. (24) is a diagonal matrix with diagonal elements $\{ \sigma_\eta^2 \ , \ \sigma_\xi^2 \ , \ \sigma_k^2 \ , \ \sigma_{k^*}^2 \}$. Once the parameters are estimated using prediction error decomposition, Kalman filter, prediction and smoothing can be applied.

iii. Cyclical Trend Model (CTM)

Here the cycle is actually incorporated within trend. Thus the model is given by the equations (Harvey, 1996):

$$Y_t = \mu_t + \varepsilon_t \ , \ t = 1, 2, \ldots, T \ , \quad (25)$$

$$\mu_t = \mu_{t-1} + \psi_{t-1} + \beta_{t-1} + \eta_t \ , \quad (26)$$

$$\beta_t = \beta_{t-1} + \xi_t \ . \quad (27)$$

Estimation of parameters and hyper-parameters is carried out by putting the model in state space form and applying Kalman filter. The essential difference between TCM and CTM is that, in the former, the observation Y_t depends on cyclical fluctuations ψ_t explicitly whereas, in the latter, it does so on ψ_{t-1} implicitly through the trend μ_t. ARIMA (2, 2, 4) model is the corresponding analogue of both TCM and CTM. **Ravichandran, Prajneshu and Savita Wadhwa (2002)** utilised the above versions of STM models for cyclical fluctuations in all-India lac forecasting which has prominent cycles since modelling and forecasting using these models found out to be better than traditional ARIMA forecasting model.

7. STM models for describing seasonal fluctuations

A STM is set up in terms of various components of interest, such as trend (μ_t), cyclical fluctuations (ψ_t), seasonal variations (γ_t) and error terms (ε_t), i.e.

$$Y_t = \mu_t + \psi_t + \gamma_t + \varepsilon_t . \quad (28)$$

However, if cyclical fluctuations are not present, eq. (19) reduces to

$$Y_t = \mu_t + \gamma_t + \varepsilon_t \ . \quad (29)$$

In reality, μ_t varies over time and it seems logical that μ_t depends on μ_{t-1}. We, therefore, let

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \ , \quad (30)$$

$$\beta_t = \beta_{t-1} + \xi_t \ , \quad (31)$$

where the error terms ε_t , η_t , ξ_t are assumed to be mutually uncorrelated and follow respectively N (0 , σ_ε^2), N (0 , σ_η^2), and N (0 , σ_ξ^2) distributions.

Eqs. (29), (30) and (31) together are called “Basic Structural Model” (BSM). The seasonal component γ_t in eq. (29) can be considered as either **“Dummy seasonal” or “Trigonometric seasonal” (Harvey, 1996). S.Ravichandran and Prajneshu (2004)** utilised these models for describing seasonal fluctuations in fisheries in Tamil Nadu.

8. Goodness of fit

Goodness of fit statistics is used for assessing overall model fit. Basic measure of goodness of fit in time series models is prediction error variance. Comparison of fit between different models is based on Akaike information criterion (AIC).

$$\text{AIC} = -2 \log L + 2n , \qquad (39)$$

where L is the likelihood function, which is expressed in terms of estimated one - step - ahead prediction errors $\hat{v}_t = Y_t - \hat{Y}_{t|t-1}$. Here n is the number of hyperparameters estimated from the model. Schwartz - Bayesian information criterion (SBC) is also used as a measure of goodness of fit which is given as

$$\text{SBC} = -2 \log L + n \log T , \qquad (40)$$

where T is the total number of observations. Lower the values of these statistics, better is the fitted model.

9. Software for STM modelling

STM models can be fitted to the data using Structural Time Series Analyser, Modeller and Predictor (STAMP) software package **(Koopman *et al.*, 2000)** or by using SsfPack2.2 (Koopman *et al.*, 1999) software package or by SAS (Statistical Analysis System) version 9.2. SSfPack is also available for S-Plus users. STAMP handles a wide variety of models, including basic GARCH and stochastic volatility models.

10. Further scope of research

Techniques such as Generalized Kalman Filter (GKF) and Extended Kalman Filter (EKF) involving nonlinear state space models may be utilised in their research problems. Software for fitting STM models for non-linear state space models is not readily available and hence has to be developed. More emphasis may be made on **"Structural time series modelling"** approach in favour of ARIMA time -series modelling approach for development of models using time-series data. Modelling and forecasting of data dealing with single-species as well as multi - species is highly desirable. Researchers should apply STM models with multivariate extension in some real-life situations, e.g. in fisheries, entomology, there are interacting species exhibiting cyclical fluctuations and having prey-predator, competition, or symbiotic types of interaction for seasonally fluctuating fish species. Appropriate estimation procedures along with relevant computer programs to handle such situations have to be developed in order to apply **"Multivariate structural time series models"** to real-life data. STM modelling can be utilised for describing seasonal and cyclical fluctuations in India's rainfall data as the rainfall fluctuates due to erratic monsoon influenced by global climate change. Such studies would go a long way in helping policy makers in framing appropriate modelling and forecasting policies.

10. Concluding Remarks

STM modelling and forecasting methodology can be effectively utilized in agriculture where lots of data on various agricultural commodities are collected over time. In fisheries also, data on marine products export, data on catch, effort and other parameters of interest etc. are usually over a long period of time and at regular intervals. Modelling these time series data can be effectively carried out using STM procedure and forecasting can also be carried out with better accuracy. STM methodology can be effectively utilized in fisheries and agriculture where observations are collected over regular intervals and where linear regression modelling procedure is utilized for forecasting these parameters. Since, advanced statistical software packages are available, STM modelling methodology can be efficiently utilized, when there are more number of data points and when there are more number of explanatory variables. In Indian agriculture and fisheries also, STM modelling methodology can be effectively utilized since accurate modelling and forecasting would help policy planners in framing suitable policies.

References

1. **Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., (1994)**. *Time series analysis* : *Forecasting and control*, 3rd edition. Prentice Hall, New Jersey.

2. **de Jong P. (1988)**. The likelihood for a state space model. *Biometrika* **75** : 165-169.

3. **Harvey A C. (1996)**. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge Univ. Press, U.K.

4. **Harvey A C and Durbin J. (1986)**. The effects of seat belt legislation on British road casualties : a case study in structural time series modelling. *Journal of Royal Statistical Society,* A 149 : 187-227.

5. **Koopman S J, Harvey A C, Doornik J A and Shephard N. (2000)**. *STAMP : Structural Time Series Analyser, Modeller and Predictor*. Timberlake Consultants Press.

6. **Meinhold, R. J. and Singpurwalla, N.D., (1983)**. Understanding Kalman filter. *Amer. Statistician,* **37**, 123-127.

7. **S.Ravichandran and Prajneshu (2001)** “State space modelling versus ARIMA Time Series Modelling” in Journal of Indian Soc. Of Ag.Statistics, 54(1), 2001, PP.43-51.

8. **Ravichandran and Prajneshu (2002)** “Dynamical modelling and forecasting of India’s foodgrain production”, Proc. Nat. Acad. Sci. India, Vol. 72, B(I).

9. **S.Ravichandran and P.Muthuraman (2006)** “Structural time-series modelling and forecasting India’s rice production", Agricultural situation in India, pp.773 - 776.

10. **S.Ravichandran , Prajneshu and Savita Wadhwa (2002)** “Structural Time-Series Models for Describing Cyclical Fluctuations”, J.of Ind. Soc. of Agrl.Stat., 55(1), pp. 70-78.

11. **S.Ravichandran and P.Muthuraman (2003) “**Structural time series models for describing sunflower yield”, Indian Journal of Agricultural Sciences.

12. **S.Ravichandran and Prajneshu (2004)** “Structural time-series models for describing seasonal fluctuations in fisheries” , Indian J. of Animal Sciences.

13. **Shumway, R.H. and Stoffer, D.S., (2000).** *Time series analysis and its applications.* Springer Verlag, New York.

****----****