



Asian Journal of Mathematics & Statistics

ISSN 1994-5418

On the Estimation and Performance of One-dimensional Autoregressive Integrated Moving Average Bilinear Time Series Models

J.F. Ojo
 University of Ibadan, Nigeria

Abstract: In this study, full and subset one-dimensional autoregressive integrated moving average bilinear models which are capable of achieving stationarity for all non linear series are proposed and were compared to determine which of them perform better. The parameters of the proposed models were estimated using Newton-Raphson iterative method and an algorithm is proposed to eliminate redundant parameters from the full models to have subset models. Akaike Information Criterion (AIC) was used to determine the order of the model. To determine the best model, the residual variance attached to the proposed full and subset models were studied. In the fitted models different sample sizes were used and the statistical properties of the derived estimates are investigated. It was found that the residual variance attached to the full bilinear model was smaller than the subset model and this was so because of the introduction of the d factor in our new models which has made us to capture trend and seasonality in the data, which in turn helps arrive at stationarity easily for any time series data set and at the same time made the full model a better model.

Key words: Parameters, newton-raphson, residual variance, algorithm and stationarity

INTRODUCTION

The bilinear time series models have attracted considerable attention during the last years. They have found a variety of applications including those in economy, biology, ecology, software interfailure, signal processing etc.

An overview of various models and their application can be found by Granger and Anderson (1978), Pham and Tran (1981), Gabr and Rao (1981), Rao *et al.* (1983), Liu (1992), Gonclaves *et al.* (2000), Shangodoyin and Ojo (2003), Wang and Wei (2004), Boonchai and Eivind (2005), Bibi (2006), Doukhan *et al.* (2006), Drost *et al.* (2007), Usoro and Omekara (2008). The bilinear modes studied by the above researchers could not achieve stationarity for all nonlinear series. Rao *et al.* (1983) gave a set of sufficient conditions for the existence of a strictly stationary stochastic process conforming to the following bilinear model:

$$X_t = \sum_{i=1}^p a_i X_{t-i} - \sum_{h=1}^q c_h + \left(\sum_{k=1}^r b_{kl} X_{t-k} \right) e_{t-1} + e_t, \text{ denoted as BL } (p, q, r, 1)$$

where, p is the order of the autoregressive component, q is the order of the moving average component, r is the upper limit of the lag difference for the observed variable X in the bilinear part of the model, l is the limit of the lag difference for the error part of the bilinear model,

a_1, a_2, \dots, a_p are the parameters of the autoregressive component, c_1, c_2, \dots, c_p are the parameters of the moving average component and b_1, \dots, b_{r1} are the parameters of the nonlinear component and e_t are independently and identically distributed as $N(0, \sigma_e^2)$.

In this study, we extend the work of Rao *et al.* (1983) to the proposed one-dimensional autoregressive integrated moving average bilinear and subset bilinear models, which are capable of achieving stationarity for all nonlinear series; this is an important improvement over other bilinear time series models.

In addition, bilinear time series are characterized by too many parameters, some of which are close to zero. In the proposed models, we address this problem by employing the concept of subsetting. Subsetting helps remove these redundant parameters, thereby leading to so-called subset bilinear models. Gabr and Rao (1981) worked on subset bilinear models and tested all the subsets of the best order of the full bilinear model before selecting the best subset. In this study, subsetting concept is introduced to the proposed one-dimensional autoregressive integrated moving average bilinear model to determine its usefulness in achieving a better model.

PROPOSED ONE-DIMENSIONAL BILINEAR TIME SERIES MODELS

We define one-dimensional autoregressive integrated moving average bilinear and subset bilinear time series models as follows:

Model 1 (M1)

$$\psi(B)X_t = \phi(B)\nabla^d X_t + \theta(B)e_t + \left(\sum_{k=1}^r b_{kl} X_{t-k} \right) e_{t-1}, \text{ denoted as BL}(p, d, q, r, 1)$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ and

$$X_t = \psi_1 X_{t-1} + \dots + \psi_{p+d} X_{t-p-d} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} + \left(\sum_{k=1}^r b_{kl} X_{t-k} \right) e_{t-1} \quad (1)$$

ϕ_1, \dots, ϕ_p are the parameters of the autoregressive component; $\theta_1, \dots, \theta_q$ are the parameters of the associated error process; b_1, \dots, b_{r1} are the parameters of the non-linear component and $\theta(B)$ is the moving average operator. p is the order of the autoregressive component; q is the order of the moving average process; $r1$ is the order of the nonlinear component and $\psi(B) = \nabla^d(B)$ is the generalized autoregressive operator. ∇^d is the differencing operator and d is the degree of consecutive differencing required to achieve stationarity.

Model 2 (M2)

$$X_t = \sum_{i=1}^l \psi_{p_i+d} X_{t-p_i-d} - \sum_{h=1}^z \theta_{q_h} e_{t-q_h} + \sum_{j=1}^m b_{rjl} X_{t-r_j} e_{t-1} + e_t, \text{ denoted as SBL}(p, d, q, r, l) \quad (2)$$

where, p_i is the order of subset autoregressive component; q_h is the order of subset moving average process and r_j is the order of subset nonlinear component.

In the models above, e_t are independently and identically distributed as $N(0, \sigma_e^2)$.

The Vector form of BL (p, d, q, r, 1)

It is convenient to study the properties of a process when the model is in the state space form because of the Markovian nature of the model Akaike (1974).

Let

$$\Psi = \begin{pmatrix} -\psi_1 & -\psi_2 & -\psi_3 & \dots & -\psi_{p+d-1} & -\psi_{p+d} \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

$$B_j = \begin{pmatrix} b_{1j} & b_{2j} & b_{3j} & \dots & b_{rj} \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} j=1$$

$$\Theta = \begin{pmatrix} \theta_1 & \theta_2 & \dots & \theta_{q-1} & \theta_q \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

and vectors

$$C^T = (1, 0, 0, \dots, 0)$$

and let

$$X^T = (X_t, X_{t-1}, \dots, X_{t-p+d})$$

(Here T stands for the transpose of a matrix) $t = \dots -1, 0, 1, \dots$. With this notation, we can write the model (1) in the vector form as:

$$X_{t+1} = \Psi X_t - \Theta e_t + B X_{t-1} e_t + C e_t \quad (3)$$

STATIONARITY AND CONVERGENCE OF BL (p, d, q, r, 1)

Here, we give a sufficient condition for the existence of strictly stationary process and convergence conforming to the bilinear model (1). This we do through the following theorem.

Theorem 1

Let $\{e_t, t \in Z\}$ be a sequence of i.i.d. random variables defined on a probability space c_1, c_2 such that $Ee_t = 0$ and

$$Ee_t^2 = \sigma_e^2 < \infty$$

Let C be any column vector with components c_1, c_2, \dots, c_p Ψ and B be two matrices of order $p \times p$ and Θ be matrix of order $q \times q$ such that:

$$\rho(\Psi \otimes \Psi + \sigma_e^2((\Theta \otimes \Theta + B \otimes B) - 2\Theta \otimes B)) = \lambda < 1$$

The series of random vectors

$$\sum_{r \geq 1} \prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r}$$

converges absolutely almost surely as well as in the mean for every fixed t in Z . Further, if

$$X_t = \mathbf{C} \mathbf{e}_t + \sum_{r \geq 1} \prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r}, \quad t \in Z$$

then for every t in Z , $\{X_t, t \in Z\}$ is a strictly stationary process conforming to the bilinear model:

$$X_t = \Psi X_{t-1} - \Theta \mathbf{e}_{t-1} + B X_{t-1} \mathbf{e}_{t-1} + \mathbf{C} \mathbf{e}_t \quad (4)$$

Conversely, if $\{X_t, t \in Z\}$ is a strictly stationary process satisfying:

$$X_t = \Psi X_{t-1} - \Theta \mathbf{e}_{t-1} + B X_{t-1} \mathbf{e}_{t-1} + \mathbf{C} \mathbf{e}_t$$

for every t in Z for some sequence $\{\mathbf{e}_t, t \in Z\}$ of i.i.d. random variables with $E \mathbf{e}_t = 0$ and

$$E \mathbf{e}_t^2 = \sigma_e^2 < \infty$$

and for some matrices Ψ, Θ, B, C of respective orders $p \times p, q \times q, r \times r$ and $p \times 1$, with

$$\rho(\Psi \otimes \Psi + \sigma_e^2((\Theta \otimes \Theta + B \otimes B) - 2\Theta \otimes B)) = \lambda < 1$$

then

$$X_t = \mathbf{C} \mathbf{e}_t + \sum_{r \geq 1} \prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r}$$

for every t in Z .

Proof of theorem 1 is given in the Appendix.

Description of Algorithm for Fitting Proposed Full and Subset Bilinear Models

For the sake of simplicity, we will break the algorithm down into the following steps.

Step 1

Fit various order of autoregressive integrated moving average model of the form:

$$X_t = \psi_1 X_{t-1} + \dots + \psi_{p+d} X_{t-p-d} + \mathbf{e}_t - \theta_1 \mathbf{e}_{t-1} - \dots - \theta_q \mathbf{e}_{t-q}$$

Step 2

Choose the model for which Akaike Information Criterion (AIC) is minimum among various order fitted in step 1.

Step 3

Fit possible subsets of chosen model in step 2 using 2^q-1 subsets approach Haggan and Oyetunji (1980).

Step 4

Choose the model for which AIC is minimum among the fitted models in step 3 to have the best subset model and this will form the initial values.

Step 5

Fit various order of the proposed full bilinear model of the form

$$X_t = \psi_1 X_{t-1} + \dots + \psi_{p+d} X_{t-p-d} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} + b_{11} X_{t-1} e_{t-1} + \dots + b_{r1} X_{t-r} e_{t-1} + e_t$$

and choose the model for which AIC is minimum

Step 6

Fit possible subsets of chosen model in step 5 using 2^q-1 subsets approach Shangodoyin and Ojo (2003).

Step 7

The model with the minimum AIC is the best subset proposed bilinear model.

Estimation of parameters of BL (p, d, q, r, 1)

The joint density function of $(e_m, e_{m+1}, \dots, e_n)$, where $m = \max(r, 1)$, is given by:

$$\frac{1}{(2\pi\sigma_e^2)^{(n-m+1)/2}} \exp\left(-\frac{1}{2\sigma_e^2} \sum_{m}^n e_t^2\right) \quad (5)$$

Proceeding as in Rao (1981), the Jacobian of the transformation from $(e_m, e_{m+1}, \dots, e_n)$ to $(X_m, X_{m+1}, \dots, X_n)$ is unity, the likelihood function of $(X_m, X_{m+1}, \dots, X_n)$ is the same as the joint density function of $(e_m, e_{m+1}, \dots, e_n)$. Thus maximising the likelihood function is equivalent to minimizing the function $Q(G)$, which is as follows:

$$Q(G) = \sum_{m}^n e_t^2 \quad (6)$$

with respect to the parameter

$$G = (\psi_1, \dots, \psi_{p+d}; \theta_1, \theta_2, \dots, \theta_q; B_{11}, \dots, B_{r1})$$

For convenience, we shall write $G_1 = \Psi_1$, $G_2 = \Psi_2, \dots, G_R = B_n$, where, $R = p + q + r_1$. Then the partial derivatives of $Q(G)$ are given by

$$\frac{dQ(G)}{dG_i} = 2 \sum_{t=m}^n e_t \frac{de_t}{dG_i} \quad (i = 1, 2, \dots, R) \quad (7)$$

$$\frac{d^2 Q(G)}{dG_i dG_j} = 2 \left(\sum_{t=m}^n e_t \frac{de_t}{dG_i} \frac{de_t}{dG_j} + \sum_{t=m}^n e_t \frac{d^2 e_t}{dG_i dG_j} \right) \quad (8)$$

where, the partial derivatives of e_i satisfy the recursive equations:

$$\begin{cases} \frac{de_i}{d\psi_i} + \sum_{j=1}^s W_j(t) \frac{de_{t-j}}{d\psi_i} = 1, & \text{if } i = 0 \\ = X_{t-i}, & \text{if } i = 1, 2, \dots, p \end{cases} \quad (9)$$

$$\frac{de_i}{d\theta_i} + \sum_{j=1}^s W_j(t) \frac{de_{t-j}}{d\theta_i} = e_{t-i}, \quad \text{if } i = 1, 2, \dots, q \quad (10)$$

$$\frac{de_i}{dB_{ki}} + \sum_{j=1}^s W_j(t) \frac{de_{t-j}}{dB_{ki}} = -X_{t-k} e_{t-m} \quad (k=1, 2, \dots, r; m_i=1) \quad (11)$$

$$\frac{d^2 e_i}{d\psi_i d\psi_{i'}} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{d\psi_i d\psi_{i'}} = 0 \quad (i, i' = 1, 2, \dots, p) \quad (12)$$

$$\frac{d^2 e_i}{d\theta_i d\theta_{i'}} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{d\theta_i d\theta_{i'}} = 0 \quad (i, i' = 0, 1, 2, \dots, q) \quad (13)$$

$$\frac{d^2 e_i}{d\psi_i dB_{ki}} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{dB_{ki} d\theta_i} + X_{t-k} \frac{d^2 e_{t-m_i}}{d\psi_i} = 0 \quad (i=0, 1, 2, \dots, p; k_i=1, 2, \dots, r, m_i=1) \quad (14)$$

$$\frac{d^2 e_i}{d\theta_i dB_{ki}} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{dB_{ki} d\theta_i} + X_{t-k} \frac{d^2 e_{t-m_i}}{d\theta_i} = 0 \quad (i=1, 2, \dots, q; k_i=1, 2, \dots, r, m_i=1) \quad (15)$$

$$\frac{d^2 e_i}{d\psi_i d\theta_i} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{d\psi_i d\theta_i} = 0 \quad (16)$$

$$\begin{aligned} \frac{d^2 e_i}{dB_{ki} dB_{k'i}} + \sum_{j=1}^s W_j(t) \frac{d^2 e_{t-j}}{dB_{ki} dB_{k'i}} + X'_{t-k} \frac{d^2 e_{t-m_i}}{dB_{k'i}} = -X_{t-k} \frac{de_{t-m}}{dB_{k'i}} \\ (K, k' = 1, 2, \dots, r; m_i m'_i = 1) \end{aligned} \quad (17)$$

$$W_j(t) = \sum_{i=1}^s B_{ij} X_{t-j}$$

We assume that $e_t = 0$ ($t = 1, 2, \dots, m-1$) and

$$\frac{de_i}{dG_i} = 0, \frac{d^2 e_i}{dG_i dG_j} = 0, \quad (i, j = 1, 2, \dots, R; t = 1, 2, \dots, m-1)$$

From these assumptions and Eq. (11) it follows that the second order derivatives with respect to Ψ_i ($i = 0, 1, 2, \dots, p$) and θ_i ($i = 0, 1, 2, \dots, q$) are zero. For a given set of values $\{\Psi_{1i}\}$, $\{\theta_{1i}\}$ and $\{B_{1i}\}$ one can evaluate the first and second order derivatives using the recursive Eq. 9-11 and 17.

Let

$$V(G) = \frac{dQ(G)}{dG_1}, \frac{dQ(G)}{dG_2}, \dots, \frac{dQ(G)}{dG_R}$$

and let $H(G) = [d^2Q(G)/dG_i dG_j]$ be a matrix of second partial derivatives. Expanding $V(G)$, near in a Taylor series, we obtain

$$[V(\hat{G})]_{\hat{G}=G} = 0 = V(G) + H(G)(\hat{G} - G)$$

Rewriting this equation and following Krzanowski (1998), we have

$$\hat{G} - G = H^{-1}(G)V(G)$$

thereby obtaining an iterative equation given by

$$G^{(k+1)} = G^{(k)} - H^{-1}(G^{(k)})V(G^{(k)})$$

where $G^{(k)}$ is the set of estimates obtained at the k th stage of iteration. The estimates obtained by the above iterative equations usually converge. For starting the iteration, we need to have good sets of initial values of the parameters. This can be obtained as follows:

Suppose we wish to fit one-dimensional bilinear model $BL(p, d, q, r, 1)$. We choose the coefficients of the autoregressive integrated moving average models (ARIMA) part of this model equal to the corresponding best subset ARIMA model. These coefficients are used as the initial values for starting the iteration of the Newton-Raphson iterative equation.

Estimation of the Parameters of Subset Bilinear (p, d, q, r, 1)

Let us assume that the sets of integers, $\{k_1, k_2, \dots, k_l\}$, $\{q_1, q_2, \dots, q_z\}$ and $\{(r_1, 1), (r_2, 1), \dots, (r_m, 1)\}$ are fixed and known. Following Rao (1981), we can show that maximizing the likelihood function of $(X_{m_1}, X_{m_1+1}, \dots, X_N)$ is equivalent to minimizing the function

$$Q(G) = \sum_{t=m_1}^N e_t^2$$

with respect to the parameters

$$(\psi_{k_1}, \psi_{k_2}, \dots, \psi_{k_l}; \theta_1, \dots, \theta_q; b_{q_1}, \dots, b_{q_m})$$

Let $G^T = (\psi_{k_1}, \psi_{k_2}, \dots, \psi_{k_l}; \theta_1, \dots, \theta_q; b_{q_1}, \dots, b_{q_m})$. For convenience, let us write

$$G_1 = \psi_{k_1}, G_2 = \psi_{k_2}, \dots, G_{l+1} = \psi_{k_l}, G_{l+2} = \theta_1, G_{l+3} = \theta_2, \dots, G_{l+m} = \theta_q, G_{l+m+1} = b_{q_1}, \dots, G_{l+m+R} = b_{q_m}$$

The partial derivatives of $Q(G)$ are

$$W_i = \frac{dQ(G)}{dG_i} = 2 \sum_{t=m_1}^N e_t \frac{de_t}{dG_i}$$

$$h_{ij} = \frac{d^2Q(G)}{dG_i dG_j} = 2 \sum_{t=m_1}^N \left(\frac{de_t}{dG_i} \right) \left(\frac{de_t}{dG_j} \right) + 2 \sum_{t=m_1}^N e_t \frac{d^2e_t}{dG_i dG_j}$$

where the partial derivatives satisfy the recursive equations

$$\frac{de_t}{d\psi_{k_r}} = X_{t-k_r} - \sum_{j=1}^m b_{rj} X_{t-r_j} \frac{de_{t-1}}{d\psi_{k_r}}, (r=1,2,3,\dots,l)$$

$$\frac{de_t}{d\theta_{q_r}} = e_{t-q_r} - \sum_{j=1}^m b_{rj} X_{t-r_j} \frac{de_{t-1}}{d\theta_{q_r}}, (r=1,2,3,\dots,l)$$

$$\frac{de_t}{db_{q_1}} = -X_{t-q_1} e_{t-1} - \sum_{j=1}^m b_{j1} X_{t-r_j} \frac{de_{t-1}}{db_{q_1}}$$

In the calculation of these partial derivatives, we set $e_1 = e_2 = \dots = e_{m_0} = 0$ and

$$\frac{de_i}{dG_i} = \frac{de_2}{dG_i} = \dots = \frac{de_{m_0}}{dG_i} = 0, (i=1,2,\dots,R)$$

Let $W^T(G) = (W_1, W_2, \dots, W_R)$ and $H(G) = (h_{ij})$ to evaluate the second order partial derivatives, let us approximate

$$h_{ij} = 2 \sum_{t=m_1}^N \left(\frac{de_t}{d\theta_i} \right) \left(\frac{de_t}{d\theta_j} \right)$$

as is done in the Marquardt algorithm. Expanding $W(\hat{G})$ near $\hat{\theta} = \theta$ in a Taylor series, we obtain $0 = W(G) + H(G)(\hat{G} - G)$.

Rewriting this equation, we get $(\hat{G} - G) = -H^{-1}(G)W(G)$, thereby obtaining the Newton-Raphson iterative equation:

$$G^{(k+1)} = G^{(k)} - H^{-1}(G^{(k)})W(G^{(k)})$$

$$G^{(k)} = G^{(k+1)} + G^{-1}(G^{(k)})W(G^{(k)}) \quad (18)$$

where $G^{(k)}$ is the set of estimates obtained at the kth iteration.

NUMERICAL EXAMPLE: THE WOLFER SUNSPOT DATA

To present the application of the models proposed, we will use a real time series dataset, the Wolfer sunspot, available in Box *et al.* (1994). The scientists track solar cycles by counting sunspots-cool planet-sized areas on the Sun where intense magnetic loops poke through the star's visible surface. It was Rudolf Wolf who devised the basic formula for calculating sunspots in 1848; these sunspot counts are still continued.

As the Wolfer sunspot data set represent a non-stationary series, the bilinear models proposed in this paper may be applied. The Wolfer sunspot data set, available in Box *et al.* (1994), in this study is considered at two different sample size of 150 and 250. For the fitted model below we have used the algorithm and the estimation technique in the previous section.

Table 1: Goodness of fit of one-dimensional full and subset autoregressive integrated bilinear models at sample sizes of 50, 150 and 250. Two models are compared, namely M1: BL (p, 1, q, r, 1), M2: SBL (p, 1, q, r, 1). All models are significant at $p < 0.001$

Sample size Model	Sample size of 150		Sample size of 250	
	Full bilinear	Subset bilinear	Full bilinear	Subset bilinear
Residual variance	198.2	211.4	256.70	258.70
R2	0.60	00.57	0.60	0.59
F (Statistic)	30.03	64.48	58.26	69.40

Fitted Model M1 and M2 at sample size 150

M1

$$X_t = 0.21742X_{t-1} + 0.172224X_{t-3} - 0.518088X_{t-4} - 0.218600X_{t-5} - 0.135334X_{t-6} - 0.269434X_{t-7} + 0.630377et - 1 - 0.119139et - 2 - 0.763971et - 3 - 0.000351X_{t-1}et - 1 + 0.006676X_{t-2}et - 1 - 0.001134X_{t-3}et - 1 - 0.011233X_{t-4}et - 1 - 0.003409X_{t-5}et - 1 + 0.002608X_{t-6}et - 1 - 0.020809X_{t-7}et - 1 + 0.011283X_{t-8}et - 1 + et$$

M2

$$X_t = -0.217421X_{t-1} + 0.172224X_{t-3} - 0.518088X_{t-4} - 0.218600X_{t-5} - 0.135334X_{t-6} - 0.269434X_{t-7} + 0.630377et - 1 - 0.119139et - 2 - 0.763971et - 3 - 0.012484X_{t-4}et - 1 - 0.002564X_{t-5}et - 1 + 0.005071X_{t-6}et - 1 - 0.011200X_{t-7}et - 1 + et$$

Fitted Model M1 and M2 at sample size 250

M1

$$X_t = -0.712478X_{t-1} - 0.153047X_{t-2} + 0.032479X_{t-3} - 0.606080X_{t-4} - 0.351330X_{t-5} - 0.422284X_{t-6} - 0.407042X_{t-7} - 0.311950X_{t-8} + 0.809607et - 1 - 0.048903et - 2 - 0.673588et - 3 - 0.003340X_{t-1}et - 1 - 0.008671X_{t-2}et - 1 - 0.007744X_{t-3}et - 1 - 0.005649X_{t-4}et - 1 - 0.006420X_{t-5}et - 1 - 0.012716X_{t-6}et - 1 - 0.006439X_{t-7}et - 1 + et$$

M2

$$X_t = -0.712478X_{t-1} - 0.153047X_{t-2} + 0.032479X_{t-3} - 0.606080X_{t-4} - 0.351330X_{t-5} - 0.422284X_{t-6} - 0.407042X_{t-7} - 0.311950X_{t-8} + 0.809607et - 1 - 0.048903et - 2 - 0.673588et - 3 - 0.003128X_{t-1}et - 1 - 0.007883X_{t-2}et - 1 - 0.009624X_{t-3}et - 1 - 0.008877X_{t-5}et - 1 - 0.013244X_{t-6}et - 1 - 0.004914X_{t-7}et - 1 + et$$

The fitted models' residual variances, coefficient of determination (R-squared) and F-statistic are given in Table 1.

We could see the performance of the two models above using the residual variance attached to each model. The residual variance of full bilinear model is smaller than that of subset model. The proposed model gave us the best model at full model which is an improvement. The usual convention is that the subset model is always better than the full model. But in this proposed model, testing all subsets of the models is not necessary.

CONCLUSION

This study focused on new bilinear models that could handle all non-linear series. Bilinear models at different levels of sample sizes were considered using the non-linear real series. Full bilinear model emerged as the better model when compared with subset model. And this is an improvement in the model proposed. Moreover, estimation of parameters witnessed a unique, consistent and convergent estimator that has prevented the models from exploding, thereby making stationarity possible. The introduction of the d factor in our new

models has made us to capture trend and seasonality in the data, which in turn helps arrive at stationarity easily for any time series data set and at same time made the full model a better model.

APPENDIX

Proof of Theorem 1

Here we prove the theorem 1 from section 4. For the sake of simplicity, we will break the proof down into the following steps.

Step 1

For almost sure convergence, we show that

$$\sum_{r \geq 1} E \left| \left(\prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) C \mathbf{e}_{t-r} \right) \right| < \infty \quad (\text{A.1})$$

for every $i=1, 2, \dots, p$. This implies that

$$\sum_{r \geq 1} \prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) C \mathbf{e}_{t-r}$$

is absolutely convergent almost surely as well as in the mean

Step 2

We establish (A.1) for $i = 1$. The general case is clear. First, we note that for every t in Z , $r \geq 1$ and $s = 1, 2, \dots, p$

$$E \left| \left((\Psi - \Theta \mathbf{e}_{t-r} + B \mathbf{e}_{t-r}) C \mathbf{e}_{t-r} \right)_s \right| \leq K_0$$

where, K_0 is a constant that depends only on Ψ, B, C and σ^2 .

Step 3

If $r \geq 2$, then

$$E \left| \left(\prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) C \mathbf{e}_{t-r} \right)_1 \right| \leq K_1 \lambda^{(r-1)/2}$$

for some constant $K_1 > 0$

Step 4

Now, for any $s = 1, 2, \dots, p$

$$\begin{aligned} E \left| \left(\prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) C \mathbf{e}_{t-r} \right)_s \right| &\leq K_1 \lambda^{(r-1)/2} \\ &= \left(\left(\prod_{j=1}^{r-1} (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \right) \otimes \left(\prod_{j=1}^{r-1} (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \right) \right)_{1s,1s} \\ &= \left(\prod_{j=1}^{r-1} (\Psi - \Theta \mathbf{e}_{t-j} + B \mathbf{e}_{t-j}) \right) \otimes (\Psi - \Theta \mathbf{e}_{t-r} + B \mathbf{e}_{t-r})_{1s,1s} \end{aligned}$$

Consequently,

$$\begin{aligned} E \left(\left(\prod_{j=1}^{t-1} (\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}) \right) \right)_{1s}^2 &= \prod_{j=1}^{t-1} (E(\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}) \otimes (\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}))_{1s1s} \\ &= (E[(\Psi - \Theta \mathbf{e}_t + \mathbf{B} \mathbf{e}_t) \otimes (\Psi - \Theta \mathbf{e}_t + \mathbf{B} \mathbf{e}_t)])_{1s1s}^{t-1} \\ &= ((E(\Psi \otimes \Psi - \mathbf{e}_t \Theta \otimes \Psi + \mathbf{e}_t \mathbf{B} \otimes \Psi + \mathbf{e}_t^2 \Theta \otimes \Theta - 2\mathbf{e}_t^2 \Theta \otimes \mathbf{B} + \mathbf{e}_t \mathbf{B} \otimes \Psi + \mathbf{e}_t^2 \mathbf{B} \otimes \mathbf{B}))_{1s1s})^{t-1} \\ &= ((\Psi \otimes \Psi + \sigma^2((\Theta \otimes \Theta + \mathbf{B} \otimes \mathbf{B}) - 2\Theta \otimes \mathbf{B}))_{1s1s})^{t-1} \\ &\leq k\lambda^{t-1} \end{aligned}$$

Hence

$$E \left(\left| \prod_{j=1}^t (\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r} \right| \right) \leq K_1 \lambda^{(t-1)/2}$$

for a suitable choice of K_1 .

Since, $\lambda < 1$, we have

$$\sum_{r=1}^t E \left(\left| \prod_{j=1}^t (\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r} \right| \right) < \infty$$

for every $i = 1, 2, \dots, p$.

Thus (A.1) is established.

It is obvious that the vector-valued stochastic process $\{X_t, t \in Z\}$ defined by

$$X_t = \mathbf{C} \mathbf{e}_t + \sum_{r=1}^t \prod_{j=1}^r (\Psi - \Theta \mathbf{e}_{t-j} + \mathbf{B} \mathbf{e}_{t-j}) \mathbf{C} \mathbf{e}_{t-r}, t \in Z$$

is strictly stationary. Thus $\rho(\Psi \otimes \Psi + \sigma^2((\Theta \otimes \Theta + \mathbf{B} \otimes \mathbf{B}) - 2\Theta \otimes \mathbf{B})) = \lambda < 1$ is a sufficient condition for strictly stationary of the model 2.4 Hence the proof.

REFERENCES

- Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. Ann. Inst. Statist. Math., 26: 363-387.
- Bibi, A., 2006. Evolutionary transfer functions of bilinear process with time varying coefficients. Comput. Math. Appl., 52: 331-338.
- Boonchai, K.S. and S. Eivind, 2005. Multivariate bilinear time series: A stochastic alternative in population dynamics. Geophysical Research Bstracts, Vol. 7, 02219. <http://www.cosis.net/abstracts/EGU05/02219/EGU05-J-02219.pdf>.
- Box, G.E.P., G.M. Jenkins and G.C. Reinsel, 1994. Time Series Analysis: Forecasting and Control. 3rd Edn., Prentice Hall, New Jersey, ISBN-13: 978-0130607744, pp: 592.
- Doukhan, P., A. Latour and D. Oraichi, 2006. A simple integer-valued bilinear time series model. Adv. Applied Prob., 38: 559-578.

- Drost, F.C., R. van den Akker and B.J.M. Werker, 2007. Note on integer-valued bilinear time series models. Econometrics and Finance Group Center, ISBN 0924-7815, Tilburg University, The Netherlands.
- Gabr, M.M. and T.S. Rao, 1981. The estimation and prediction of subset bilinear time series models with application. *J. Time Series Anal.*, 2: 89-100.
- Gonclaves, E., P. Jacob and N. Mendes-Lopes, 2000. A decision procedure for bilinear time series based on the asymptotic separation. *Statistics*, 33: 333-348.
- Granger, C.W.J. and A.P. Anderson, 1978. Introduction to Bilinear Time Series Models. Vandenhoeck and Ruprecht, Göttingen, Germany, pp: 45-69.
- Haggan, V. and O.B. Oyetunji, 1980. On the selection of subset autoregressive time series models. UMIST Technical Report No. 124 (Dept. of Mathematics).
- Liu, J., 1992. On stationarity and asymptotic inference of bilinear time series models. *Statist. Sinica*, 20: 479-494.
- Pham, T.D. and L.T. Tran, 1981. On the first order bilinear time series model. *J. Applid Prob.*, 18: 617-627.
- Rao, T.S., 1981. On theory of bilinear time series models. *J. R. Statist. Soc. Series B*, 43: 244-255.
- Rao, M.B., T.S. Rao and A.M. Walker, 1983. On the existence of some bilinear time series models. *J. Time Series Anal.*, 4: 60-76.
- Shangodoyin, D.K. and J.F. Ojo, 2003. On the performance of bilinear time series autoregressive moving average models. *J. Niger. Statist. Assoc.*, 16: 1-12.
- Usoro, A.E. and C.O. Omekara, 2008. Lower diagonal bilinear moving average vector models. *Adv. Applied Mathematical Anal.*, 3: 49-54.
- Wang, H.B. and B.C. Wei, 2004. Separable lower triangular bilinear model. *J. Applied Prob.*, 41: 221-235.