

# On the non-negative first-order exponential bilinear time series model

I. Pereira and M. G. Scotto \*

Department of Mathematics

University of Aveiro

Portugal

## Abstract

In this paper the bilinear model  $BL(1, 0, 1, 1)$  driven by exponential distributed innovations is studied in some detail. Conditions under which the model is strictly stationary as well as some properties of the stationary distribution are discussed. Moreover parameter estimation is also addressed.

*MSC:* primary 62M10; secondary 62F30; 62F15

*Keywords:* bilinear processes, tail index, Whittle criterion, Bayesian estimation

---

\*M. G. Scotto, Department of Mathematics, University of Aveiro. Campo Universitário de Santiago, 3810-193 Aveiro, Portugal. E-mail: mscotto@mat.ua.pt. Contract/grant sponsors: POSI/CPS/42069/2001, FCT plurianual funding.

# 1 Introduction

In the analysis of stationary time series the class of linear models with finite variance, which includes ARMA models, plays a central role. However, such models are unlikely to provide a sufficiently broad class capable of accurately capturing features often exhibited by data sets such as sudden burst of large positive and negative values, almost no correlation in data, volatility changes in time, and high threshold exceedances appearing in clusters. Since there is no unifying theory that is applicable to all nonlinear systems the study of such systems has to be restricted to special classes of nonlinear models. Various non-linear models have been introduced addressing these issues. Among the more successful non-linear models we mention the bilinear (BL) models, first proposed and developed by Granger and Andersen (1978). A time series  $\{X_t\}$  is called a bilinear process of order  $(p, q, P, Q)$ , denoted by  $\text{BL}(p, q, P, Q)$ , for some integers  $p, q, P, Q \geq 0$ , if it satisfies the recurrence equation

$$X_t = \sum_{j=1}^p a_j X_{t-j} + \sum_{j=0}^q c_j \epsilon_{t-j} + \sum_{j=1}^P \sum_{k=0}^Q b_{jk} X_{t-j} \epsilon_{t-k},$$

where  $\{\epsilon_t\}$  is a sequence of independent and identically distributed (i.i.d.) random variables and the  $a_j$ ,  $c_j$ , and  $b_{jk}$  are real constants with  $c_0 = 1$ . In terms of potential applications, bilinear models are suitable for modelling seismological data such as records of explosions and earthquakes. One step towards the application of bilinear models to real data sets is the estimation of parameters. Most of the work in parameter estimation in the literature, is focused in the time-domain approach. For instance the least squared method has been considered among others by Pham and Tran (1981), Subba Rao and Gabr (1984) and Guegan and Phan (1989). Kim and Billard (1990) have obtained moment estimators for the first order bilinear model and derived their asymptotic distribution. In contrast, Bayesian analysis of bilinear time series has not received much attention in the literature; see Chen (1992a, b) for details.

A different approach is by considering frequency-domain methods and in particular the *Whittle criterion*. This method was originally proposed to estimate the parameters

of Gaussian ARMA processes (i.e., linear processes with finite variance). The applicability of the Whittle criterion for non-Gaussian and non-linear processes was discussed in detail by Dzhaparidze and Yaglom (1983) who introduced the class of *non-Gaussian mixing processes*. Dzhaparidze and Yaglom showed that for this class of processes the Whittle estimator is weakly consistent and asymptotically normal. Nevertheless the Whittle estimator can also be used when estimating the parameters of ARMA processes with i.i.d infinite variance innovations; see Mikosch *et al.* (1995). Sesay and Subba Rao (1992) used Dzhaparidze and Yaglom's results to estimate the parameters of the bilinear model  $BL(p, 0, p, 1)$ .

When dealing with bilinear models it is common to assume that the innovations are normally distributed. Recently, however, there has been considerable interest in non-negative time-series models. The motivation to include such models comes from the need to account for the non-negative nature of certain data sets such as CPU time to complete a job, call holding times, interarrival times between packets in a network and lengths of on/off cycles; see Adler *et al.* (1997) and references therein.

In this paper we consider the bilinear models  $BL(1, 0, 1, 1)$  driven by exponential distributed innovations. The purpose of this work is two-folded: first we give conditions under which the first-order non-negative bilinear model is strictly stationary and some properties of the stationary distribution, such as moments, are discussed. Moreover, we also consider parameter estimation. In particular we cast the  $BL(1, 0, 1, 1)$  in a Bayesian framework and make inferences by using the Gibbs sampler. Furthermore we also consider the Whittle criterion.

The rest of the article is organized as follows. In Section 2, we summarize some of the basic probabilistic properties of the  $BL(1, 0, 1, 1)$  model. Section 3 gives the general Bayesian setting for this special bilinear model. Section 4 introduces the Whittle criterion. Finally, a simulation study is presented in Section 5.

## 2 Basic properties of the first-order exponential bilinear model

Let  $(\Omega, \mathcal{A}, P)$  be a probability space and  $\mathbf{X} = \{X_t\}_{t \geq 1}$  a process which is defined from  $(\Omega, \mathcal{A}, P)$  to  $(\mathbb{R}, \mathcal{B})$  satisfying

$$X_t = b\epsilon_{t-1}X_{t-1} + \epsilon_t, \quad (1)$$

where  $\{\epsilon_t\}$  are i.i.d. random variables exponentially distributed in the interval  $(0, \infty)$  with parameter  $\lambda > 0$ . First note that the representation given in (1) is not Markovian and the random pair  $(b\epsilon_{t-1}, \epsilon_t)$  forms an 1-dependent, identically distributed pair. However, by setting  $Z_t = b\epsilon_t X_t$  we see that  $X_t$  has a Markovian representation in the form

$$\begin{cases} X_t = Z_{t-1} + \epsilon_t \\ Z_t = b\epsilon_t Z_{t-1} + b\epsilon_t^2 \end{cases},$$

based on  $\mathbf{Z} = \{Z_t\}_{t \geq 1}$ . The Markovian representation above implies that the study of several properties of the bilinear process in (1), related with its probabilistic structure and the existence of moments, can be obtained via the analysis of stochastic difference equations (SDE) of the form  $Y_t = A_t Y_{t-1} + B_t$ , where  $(A_t, B_t)$  are i.i.d.  $\mathbb{R}^2$ -valued random pairs with some joint distribution and  $Y_0$  is independent of these, with some given starting distribution.

We first present a result related with some useful properties of the exponential distribution. For simplicity in notation we define  $\Gamma'(1, \lambda) = \left. \frac{\partial \Gamma(\alpha, \lambda)}{\partial \alpha} \right|_{\alpha=1}$ .

**Lemma 2.1** *For a random variable  $\epsilon$  exponentially distributed with parameter  $\lambda > 0$  and  $b \in (0, \exp\{-\lambda\Gamma'(1, \lambda)\})$ , define*

$$h(b, u, \lambda) = E(b\epsilon)^u, \quad u \geq 0, \quad b > 0.$$

*Then  $E(\ln b\epsilon) < 0$  and the function  $h$  is strictly convex in  $u$ , and there exist a unique solution  $\kappa = \kappa(b, \lambda) > 0$  to the equation  $h(u) = 1$ , that is*

$$\left(\frac{b}{\lambda}\right)^u \Gamma(u+1) = 1.$$

Furthermore define and  $Q_1 = \{\lambda \in (0, 1) : \Gamma'(1, \lambda) > 0\}$ ,  $Q_2 = \{\lambda \in (0, 1] : \Gamma'(1, \lambda) < 0\}$ ,  $Q_3 = \{\lambda \in [1, \infty) : \Gamma'(1, \lambda) > 0\}$ , and  $Q_4 = \{\lambda \in (1, \infty) : \Gamma'(1, \lambda) < 0\}$ , then it follows that

$$\kappa(b, \lambda) = \begin{cases} > 1 & b \in (0, 1), \lambda \in Q_3, \text{ or } b = 1, \lambda > 1, \\ = 1 & b = 1, \lambda = 1 \\ < 1 & b \in (1, \exp\{-\lambda\Gamma'(1, \lambda)\}), \lambda \in Q_2, \text{ or } b = 1, \lambda < 1, \\ > 0 & b \in (0, 1), \lambda \in Q_1 \text{ or } b \in (1, \exp\{-\lambda\Gamma'(1, \lambda)\}), \lambda \in Q_4 \end{cases}$$

*Proof.* It is easy to check that  $h(b, u, \lambda) = \left(\frac{b}{\lambda}\right)^u \Gamma(u + 1)$ ,  $h'_u(b, u, \lambda) = E[(b\epsilon)^u \ln(b\epsilon)]$  and  $h''_u(b, u, \lambda) = E[(b\epsilon)^u (\ln(b\epsilon))^2] > 0$  where  $h'_u(\cdot)$  and  $h''_u(\cdot)$  denote the first and second derivatives of  $h(\cdot)$  with respect to  $u$ , respectively. Hence  $h(\cdot)$  is strictly increasing in  $b$ , for a fixed value of  $\lambda$ , and convex in  $u$ . Note that  $h'_u(b, 0, \lambda) = \ln b + \lambda\Gamma'(1, \lambda) < 0$ , for  $b \in (0, \exp\{-\lambda\Gamma'(1, \lambda)\})$ . Thus  $h(b, 0, \lambda) = 1$  and the convexity of  $h$  implies that there exist a unique  $\kappa > 0$  such that  $h(\kappa) = 1$ . The values of  $\kappa$  are obtained by a monotony argument and the properties of the Gamma function. This concludes the proof.

The value of  $\kappa = \kappa(b, \lambda) > 0$  is crucial for the existence of moments of the non-negative exponential bilinear process. For example, the necessary and sufficient condition for the existence of the  $m$ th moment is that  $\kappa > m$ , or equivalently,  $b \in (0, (\lambda^m/m!)^{1/m})$ . Note that the equation  $h(b, u, \lambda) = 1$  cannot be solved explicitly, but numerical solutions can be found in Table 1 below.

*Table 1 about here*

It is worth to mention here that when  $b/\lambda > 1$  the process has no finite mean where as  $b/\lambda > 0.705$  implies no finite variance. In order to obtain the stationary solution on (1) as well as its moments, we use the fact that  $X_t$  can be embedded as a function of a SDE of the form

$$\begin{cases} X_t = Z_{t-1} + \epsilon_t \\ Z_t = A_t Z_{t-1} + B_t \end{cases},$$

with  $(A_t, B_t) = (b\epsilon_t, b\epsilon_t^2)$ . The main result of this section is given in Theorem 2.1 below.

**Theorem 2.1** *Let  $\mathbf{X}$  be the non-negative exponential bilinear process defined in (1) and  $\{\epsilon_t\}$  be a sequence of i.i.d. random variables, each exponentially distributed in the interval  $(0, \infty)$ , with parameter  $\lambda > 0$ . It is also assumed that  $X_0$  is independent of  $\mathbf{Z}$ . Suppose that  $b \in (0, \exp\{-\lambda\Gamma'(1, \lambda)\})$ , then the processes  $\mathbf{Z}$  and  $\mathbf{X}$  are strictly stationary if*

$$Z_1 \stackrel{d}{=} b \sum_{m=1}^{\infty} \epsilon_m^2 \prod_{j=1}^{m-1} (b\epsilon_j) \quad (2)$$

and

$$X_0 \stackrel{d}{=} \epsilon_0 + b \sum_{m=1}^{\infty} \epsilon_m^2 \prod_{j=1}^{m-1} (b\epsilon_j). \quad (3)$$

*Proof.* Set  $A_t = b\epsilon_t$  and  $B_t = b\epsilon_t^2$  in Proposition 8.4.3 in Embrechts *et al.* (1997, pp. 457-458). By Lemma 2.1,  $E \ln^+ A < \infty$  and  $E \ln B < 0$  when  $b \in (0, \exp\{-\lambda\Gamma'(1, \lambda)\})$ . Hence (2) is proved. Finally, (3) follows by the Markovian representation of  $X_t$ .

Now we are prepared to obtain the moments of the process.

**Lemma 2.2** *Assume that  $\mathbf{Z}$  is strictly stationary and write  $Z = Z_1$ . Let  $\kappa$  be the unique positive solution of the equation  $h(u) = 1$ . Then  $E(Z)^u < \infty$  for  $0 \leq u < \kappa$ . Denote by  $p$  the largest integer strictly less than  $\kappa$ . Then for  $r = 1, \dots, p$*

$$EZ^r = \left(\frac{b}{\lambda^2}\right)^r \sum_{i=0}^r \binom{r}{i} \frac{(2r-i)!}{\lambda^{-i}} EZ^i, \quad r = 1, 2, \dots, p.$$

Furthermore, let  $\mathbf{X}$  be the non-negative exponential bilinear process defined in (1) and write  $X = X_0$ . Then

$$EX^r = \sum_{k=0}^r \binom{r}{k} E(Z^k \epsilon^{r-k}), \quad r = 1, 2, \dots, p.$$

*Proof.* The convexity of the function  $h$  and the fact that  $h(b, 0, \lambda) = h(b, \kappa, \lambda) = 1$ , for a fixed values  $b, \lambda > 0$ , implies that  $h(b, \kappa, \lambda) < 1$  for  $0 < u < \kappa$ . According with

Proposition 8.4.3 in Embrechts *et al.* (1997, pp. 457-458),  $E(Z)^u < \infty$  for  $u < \kappa$  and  $Z \stackrel{d}{=} AZ + B$ , providing the first statement. Moreover, the second statement follows by considering the Markovian representation of  $X_t$  and the fact that all moments of the exponential distribution are finite.

**Corollary 2.1** *Let  $\mathbf{X}$  be the strictly stationary non-negative process defined in (1). Suppose that  $0 < 2b^2\alpha^2 < 1$  with  $\alpha = \lambda^{-1} = E(\epsilon_t)$ . Then  $\mathbf{X}$  is second-order stationary, and*

1.  $E(X) = \frac{b\alpha^2 + \alpha}{1 - b\alpha}$ ,  $E(Z) = \frac{2b\alpha^2}{1 - b\alpha}$ ,  $E(A) = b\alpha$ ,  $E(B) = 2b\alpha^2$ ;
2.  $V(X) = V(Z) = \frac{\alpha^2}{q_1}(20b^2\alpha^2 + q_2^2 + 8b\alpha q_2 + q_1)$ , with  $q_1 = 1 - 2b^2\alpha^2$  and  $q_2 = \frac{2b^2\alpha^2}{1 - b\alpha}$ ;
3.  $\gamma(k) = Cov(X_0, X_k) = [E(A)]^{k-1}\gamma(1)$ ,  $k \geq 1$ , with

$$\gamma(1) = E(A)R'(0) + 2E(B)E(Z) + E(B\epsilon) - E(B)[E(Z) + E(\epsilon)](1 - E(A))^{-1}, \quad (4)$$

$$\text{with } E(B\epsilon) = 6b\alpha^3 \text{ and } R'(0) = \frac{\alpha^2}{q_1}(20b^2\alpha^2 + q_2^2 + 8b\alpha q_2) + [E(Z)]^2;$$

4. *The spectral density function is given by*

$$f(\omega) = \frac{1}{2\pi} \left\{ \gamma(0) + 2\gamma(1) \frac{\cos(\omega) - E(A)}{1 + (E(A))^2 - 2E(A)\cos(\omega)} \right\}, \quad -\pi \leq \omega \leq \pi. \quad (5)$$

*Proof.* The statements follows by Lemma 2.2 and after tedious calculations.

For a time series model to be useful for forecasting purposes, it is necessary that it should be invertible. A sufficient condition for the invertibility of the bilinear model is given below.

**Proposition 2.1** *Let  $\mathbf{X}$  be the strictly stationary non-negative exponential bilinear process defined in (1). If  $0 < b < \frac{\lambda}{1 + \sqrt{2}}$  then the model is invertible.*

*Proof.* Following Pham and Tran (1981),  $X_t$  is invertible if  $|b| < \exp\{-E(\log |X_t|)\}$ . By Jensen inequality we have that  $bE(X_t) < 1$ . The result follows by the first statement in Corollary 2.1.

Next result shows that  $\mathbf{X}$  is strongly mixing.

**Lemma 2.3** *Let  $\mathbf{X}$  be the strictly stationary non-negative exponential bilinear process defined in (1). Then  $\mathbf{X}$  is strongly mixing with geometric rate.*

*Proof.* The proof follows as that of Basrak *et al.* (1999, pp. 7-8) with some minor changes. We omit the details.

### 3 Bayesian Inference

Let  $\underline{x} = (x_1, \dots, x_n)$  be a sample generated by the non-negative exponential bilinear process defined in (1) being  $\theta = (b, \lambda)$  the vector of unknown parameters. We assume that one has observed the first observation before formulating the model and let  $\epsilon_1 = 0$ . Since  $\epsilon_2, \dots, \epsilon_n$  are i.i.d. exponentially distributed random variables, the conditional likelihood may be approximated by

$$L(\theta|\underline{x}) \propto \lambda^{n-1} \exp\{-\lambda \sum_{t=2}^n \epsilon_t\},$$

with  $\epsilon_t = x_t - bx_{t-1}\epsilon_{t-1}$ . Since we have no information on the parameters we use a non-informative prior for  $\theta$  of the form  $p(\theta) \propto \lambda^{-1} I_{\Theta}(\theta)$  where  $\Theta = \{(b, \lambda) : b, \lambda > 0\}$ . For this prior the posterior distribution of  $\theta$  is

$$p(\theta|\underline{x}) \propto \lambda^{n-2} \exp\{-\lambda \sum_{t=2}^n \epsilon_t\}, \quad (6)$$

with support  $\theta \in \Theta^* = \{(b, \lambda) \in \mathbb{R}^+ : x_t - bx_{t-1}\epsilon_{t-1} > 0, t = 2, 3, \dots\}$ . It is convenient to notice that we have a constrained parameter problem since the support of the posterior distribution depends on the data. Furthermore the support also depends on the sequence of innovations  $\{\epsilon_t\}$ , which are unknown. Moreover, since the inference for the vector of unknown parameters will be done through the Gibbs sampler, we have to derive the set of full conditional posterior densities. The results are summarized below.

**Proposition 3.1** *Given the approximate posterior function (6), the full conditional posterior densities are*

$$p(\lambda|b, \underline{x}) \propto \lambda^{n-2} \exp\{-\lambda(\sum_{t=2}^n x_t - b \sum_{t=2}^n x_{t-1}\epsilon_{t-1})\}, b(1 + \sqrt{2}) < \lambda < \infty \quad (7)$$

and

$$p(b|\lambda, \underline{x}) \propto \exp\{b\lambda \sum_{t=2}^n x_{t-1}\epsilon_{t-1}\}, \quad 0 < b < \min\{1, \gamma_1, \gamma_2\}, \quad (8)$$

with  $\gamma_1 = \min_{3 \leq t \leq n} \frac{x_t}{x_{t-1}\epsilon_{t-1}}$ , and  $\gamma_2 = \lambda(\sqrt{2} - 1)$ .

Note that the restriction on the support of the full conditional posterior densities is imposed in order to ensure the invertibility of the model. Moreover, by (7) and (8) the steps required to implement the Gibbs sampler are the following: (a) from an arbitrary set of initial values  $\theta^0 = (\lambda^{(0)}, b^{(0)})$  obtain an estimation of  $\epsilon_t$ , say  $\epsilon_t^0$ , by  $\epsilon_1^0 = 0$  and

$$\epsilon_t^0 = \begin{cases} x_1 & t = 2 \\ x_t + \sum_{i=1}^{t-2} (-b^{(0)})^i x_{t-i} \prod_{j=t-i}^{t-1} x_j & t = 3, 4, \dots, n \end{cases};$$

(b) draw  $\lambda^{(1)}$  from  $p(\lambda|b^{(0)}, \underline{x})$  with  $\epsilon_t = \epsilon_t^0$ ; (c) draw  $b^{(1)}$  from  $p(b|\lambda^{(1)}, \underline{x})$  with  $\epsilon_t = \epsilon_t^0$ ; (d) recalculate the residuals  $\epsilon_t$  by using the value  $b^{(1)}$ ; (e) generate  $N$  sets of random numbers by repeating stages (b), (c), and (d) from the full conditional distributions (7) and (8), obtaining a sequence of Gibbs for  $\theta$  given by  $(\lambda^{(1)}, b^{(1)}), \dots, (\lambda^{(N)}, b^{(N)})$ . Since the estimates  $(\lambda^{(k)}, b^{(k)})$ ,  $k \geq 1$  generated by the above procedure are correlated, we will only include in the final sample the observations

$$(\lambda^{(l)}, b^{(l)}), (\lambda^{(l+k)}, b^{(l+k)}), \dots, (\lambda^{(l+km)}, b^{(l+km)}),$$

to obtain an approximate i.i.d sample.

## 4 Whittle estimation

For estimating the parameters of a stationary Gaussian linear process, say  $\hat{\mathbf{X}} = \{\hat{X}_t\}_{t \geq 1}$ , with strictly positive and continuous spectral density  $f(\omega, \cdot)$ , Whittle suggested a procedure which is based on the periodogram. In his setup, the Gaussian log-likelihood function of  $\hat{\mathbf{X}}$  is approximated by

$$L(\hat{\mathbf{X}}; \tilde{\theta}_n) = \frac{n}{4\pi} \int_{-\pi}^{\pi} \left\{ \log f(\omega, \theta_n) + \frac{I_n(\omega)}{f(\omega, \theta_n)} \right\} d\omega, \quad (9)$$

where  $I_n(\omega)$  is the periodogram of the observations. The Whittle estimate  $\tilde{\theta}_n$  of the set of parameters  $\theta$  is obtained by minimizing (9). For non-Gaussian and non-linear processes Dzhaparidze and Yaglom (1983) have introduced a class of non-Gaussian mixing processes, for which the Whittle's criterion remains valid. It is worth to mention here that an important features of SDE is that, in virtue of Proposition 8.4.3 in Embrechts *et al.* (1997, pp. 457-458) not all moments exist. This is due to the fact that SDE with light-tailed input are, in general, heavy-tailed. It is not clear, indeed, from Dzhaparidze and Yaglom's work the assumption of finite absolute moments of all orders since  $\mathbf{W}(\theta)$  only depends on the cumulants up to fourth order. This is, indeed, the argument used by Sesay and Subba Rao (1992) to justify the use of the Whittle criterion to estimate the parameters of stationary BL( $p, 0, p, 1$ ). In Section 5 we will analyze the impact of the moments on the estimates obtained through the Whittle criterion.

## 5 Simulation results

The object of this section is to compare the performance of Bayes and Whittle estimators presented in the previous sections. Through the simulation study we want to highlight the following issues: a) how the results depend on the underlying bilinear parameter  $b$  and the parameter  $\lambda$ ; b) what is the impact of sample size on the simulation results and c) what is the influence of the tail parameter  $\kappa$ . We consider four distinct bilinear models with exponential innovations leading to values of the tail parameter  $\kappa$  between 9.78 and 156.41. Note than all models have finite variance and are invertible. For each of the four models we simulate time series of length  $n = 100, 500, 1000, 10000$  with 500 independent replicates. A closer look at the tables reveals that both estimators of  $b$  and  $\lambda$  tend to be positively skewed. The skewness increases when the sample size becomes large (up to size 1000) and reduces for very large sample size. A comparison of the standard deviations for the two estimators shows the superiority of the Whittle estimator for both small and large sample sizes. The simulation study also indicates that what is important for estimation purposes is the parameter  $\lambda$ . For example, when the parameter  $\lambda$  is small the Whittle criterion provides better estimates

that the Bayesian estimator regardless the values of the parameter  $b$  and the tail parameter  $\kappa$ . In contrast, large values of  $\lambda$  implies a better performance of the Bayes estimates. Another observation concerns the speed of convergence. Note that large values of  $\lambda$  the speed of convergence towards the true values is relatively slow for the Whittle estimator. Moreover, it is worthwhile to mention that the actual computation of the Whittle estimates is simpler and faster (in cpu time) than the computation of the Bayes estimates. In summary, we can conclude that the overall performance of the Whittle estimates seems to be better than the Bayes estimates.

*Tables 2 and 3 about here*

## References

- [1] Adler, R. J., Feldman, R., and Taqqu, M. (eds.) (1997). *A User's Guide to Heavy-Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions and Processes*. Birkhäuser, Boston.
- [2] Basrak, B., Davis, R. A., and Mikosch, T. (1999). The sample ACF of a simple bilinear process. *Stoch. Proc. Appl.*, **83**, 1–14.
- [3] Chen, C. W. S. (1992a). Bayesian inferences and forecasting in bilinear time series models. *Commun. Statist.- Theory Meth.*, **21**, 1725–1743.
- [4] Chen, C. W. S. (1992b). Bayesian analysis of time series models:a Gibbs sampling approach. *Commun. Statist.- Theory Meth.*, **21**, 3407–3425.
- [5] Dzhaparidze, K. O. and Yaglom, A. M. (1983). Spectrum parameter estimation in time series analysis. *Developments in Statistics*, **4**, New York: Academic Press.
- [6] Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag,Berlin.
- [7] Granger, C. W. J. and Andersen, A. P. (1978). *An Introduction to Bilinear Time Series Models*. Göttingen: Vandenhoeck & Ruprecht.
- [8] Guegan, D. and Pham, D. T. (1989). A note on the estimation of the parameters of the diagonal bilinear model by the method of least squares. *Scand. J. Statist.*, **16**, 129–136.
- [9] Kim, W. K. and Billard, L. (1990). Asymptotic properties for the first-order bilinear time series model. *Commun. Statist.- Theory Meth.*, **19**, 1171–1183.
- [10] Mikosch, T. Gadrich, T. Klüppelberg, C. and Adler, R. (1995). Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.*, **23**, 305–326.
- [11] Pham, D. T. and Tran, T. L. (1981). On the first-order bilinear time series model. *J. Appl. Prob.*, **18**, 617–627.
- [12] Sesay, S. A. O. and Subba Rao, T (1992). Frequency-domain estimation of bilinear time series models. *J. Time Series Anal.*, **6**, 521–545.
- [13] Subba Rao, T and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Lecture Notes in Statistics 24, Springer-Verlag.

$\lambda$	2.0	4.0	0.5	4.0	2.0	4.0	0.5
$b$	0.70	0.90	0.10	0.70	0.30	0.50	0.05
$\kappa$	5.65	9.78	11.24	13.13	15.64	19.19	24.53

Table 1: Values of  $\kappa = \kappa(b, \lambda)$  for different combinations of  $b$  and  $\lambda$ .

		<b>Whittle est. <math>b_W</math></b>	<b>Bayes est. <math>b_B</math></b>	<b>Whittle est. <math>\lambda_W</math></b>	<b>Bayes est. <math>\lambda_B</math></b>
Model	$n$	Mean (StD)	Mean (StD)	Mean (StD)	Mean(StD)
1 $\lambda = 0.5$ $b = 0.05$ $\kappa = 24.5$	100	0.55(0.03)	0.03(0.02)	0.53(0.25)	0.50(0.06)
	500	0.05(0.08)	0.03(0.02)	0.51(0.12)	0.45(0.04)
	1000	0.05(0.07)	0.02(0.02)	0.50(0.08)	0.45(0.04)
	10000	0.05(0.01)	0.02(0.02)	0.50(0.02)	0.45(0.04)
2 $\lambda = 15.0$ $b = 0.25$ $\kappa = 156.4$	100	0.08(0.15)	0.25(0.07)	6.34(0.03)	15.3(1.60)
	500	0.10(0.10)	0.25(0.01)	10.82(0.02)	15.04(0.69)
	1000	0.15(0.09)	0.24(0.01)	12.38(0.02)	15.03(0.48)
	10000	0.24(0.02)	0.24(0.02)	14.61(0.09)	15.01(0.14)
3 $\lambda = 30.0$ $b = 0.90$ $\kappa = 87.4$	100	0.14(0.07)	0.89(0.17)	6.65(0.01)	30.55(3.11)
	500	0.06(0.07)	0.88(0.10)	13.83(0.01)	30.10(1.34)
	1000	0.14(0.10)	0.88(0.11)	18.18(0.01)	30.03(1.03)
	10000	0.71(0.11)	0.86(0.15)	27.70(0.09)	29.98(0.47)
4 $\lambda = 4.0$ $b = 0.90$ $\kappa = 9.80$	100	0.81(0.47)	0.18(0.22)	3.798(0.06)	2.90(0.60)
	500	0.80(0.11)	0.05(0.07)	3.93(0.03)	2.63(0.21)
	1000	0.81(0.07)	0.03(0.02)	3.98(0.04)	2.61(0.20)
	10000	0.89(0.02)	0.02(0.01)	3.99(0.08)	2.57(0.18)

Table 2: Estimates of  $b$  and  $\lambda$  via Whittle and Bayes approaches. Standard Deviations (StD) in parenthesis

Model	500	1000	10000
1	1.84	1.90	1.92
2	8.17	6.76	2.77
3	16.0	17.3	7.89
4	2.38	2.61	2.76

Table 3: Estimates of the skewness.