

On BL-GARCH(1,2) Models

Abdelhalim BOUCHEMELLA

University of 08 Mai 1945 Guelma - Algeria

abdelhalingbs@yahoo.fr

Abstract

In this study, we focus on the class of BL-GARCH models, which is initially introduced by Storti & Vitale [11] in order to handle a leverage effects and volatility clustering. First we illustrate some properties of BL-GARCH(1,2) model, like the positivity, stationarity and marginal distribution, then we'll study the statistical inference, applying the composite likelihood on panel of BL-GARCH(1,2) model.

Keywords: Random coefficient Autoregressive model; BL-GARCH models; composite likelihood.

1 Introduction

Classical modelling of time series is not usually appropriate for financial data, such as *ARMA* models do not allow the variability in volatility over time, are not able to capture asymmetries in the conditional variance of financial time series, and fail to generate the squared autocorrelations. In front of these monetary and financial problems, Engle [1982] proposed a new class of Autoregressive conditionally heteroscedastic models (*ARCH*), followed by generalized *ARCH* or *GARCH* suggested by Bollerslev [1986]. Storti and Vitale [2003] proposed an innovative approach to modelling leverage effects in financial time series based on the Bilinear *GARCH* noted by *BL – GARCH* models which are considered as a generalization of *GARCH* models.

In this present paper we study the *BL – GARCH* models, specifically *BL – GARCH* (1, 2) that is widely used and proved its performance for the volatility analysis of financial time series. We focus on studies of Storti & Vitale [2003] and Diongue & Guégan and Wolff [2009], which they have well discussed and treated this class of models.

In recent years, several authors have been interested in composite maximum likelihood methods that are widely used in parametric statistical inference because of the good asymptotic properties of the estimators. The aim of composite likelihood is to reduce and simplify the computational complexity to cope with large datasets and presence of complex interdependencies.

The term pseudo-likelihood was originally proposed by Besag [1974]. Lindsay [1988] used the term composite likelihood for justify his choice to describe the method of construction considered. There are many research and studies in various fields, have applied this method, for example in statistical genetics (Larribe and Fearnhead 2011), in time series (Richard, Davis and Chun Yip Yau 2011) and (Pakel, Shephard and Sheppard 2011), in longitudinal data (Molenberghs and Verbeke 2005).

We organize this work as follows. In Section 2 we study the important properties of $BL - GARCH(1, 2)$ concerning conditions for the positivity of conditional variance, conditions of stationarity and we conclude this section by the properties of marginal distribution. In Section 3 we introduce the $BL - GARCH(1, 2)$ panel model, then we illustrate the good performance of estimators of composite likelihood applied to this model.

2 Properties of $BL - GARCH(1, 2)$

we consider the asset log-returns y_t at time t , assuming that

$$y_t = \mu_t + u_t \quad \text{where } \mu_t = E(y_t / \Omega_{t-1}) \quad (2-1)$$

$$u_t = h_t \varepsilon_t \quad (2-2)$$

$$h_t^2 = a_0 + a_1 u_{t-1}^2 + b_1 h_{t-1}^2 + b_2 h_{t-2}^2 + c_1 u_{t-1} h_{t-1} \quad (2-3)$$

where Ω_{t-1} is the historical information set up to time $t - 1$.

2.1 Positivity of Conditional variance

We can write the model (2-3) in matrix form as :

$$h_t^2 = [1, u_{t-1}, h_{t-1}, h_{t-2}] \begin{bmatrix} a_0 & 0 & 0 & 0 \\ 0 & a_1 & \frac{1}{2}c_1 & 0 \\ 0 & \frac{1}{2}c_1 & b_1 & 0 \\ 0 & 0 & 0 & b_2 \end{bmatrix} \begin{bmatrix} 1 \\ u_{t-1} \\ h_{t-1} \\ h_{t-2} \end{bmatrix} \quad (2-4)$$

$$= K_t' A K_t$$

Proposition 1 *A sufficient set of conditions for positivity of conditional variance h_t^2 is :*

$$a_0 > 0 ; a_1 > 0 ; b_1 > 0 ; b_2 > 0 ; c_1^2 > 4a_1b_1 \quad (2-5)$$

Proof. We note that $h_t^2 > 0$ if and only if A is a positive definite matrix and this implies that all eigenvalues of A are strictly positive.

Set of these eigenvalues are :

$$\left\{ a_0 ; b_2 ; \frac{1}{2} \left(a_1 + b_1 - \sqrt{a_1^2 - 2a_1b_1 + b_1^2 + c_1^2} \right) ; \frac{1}{2} \left(a_1 + b_1 + \sqrt{a_1^2 - 2a_1b_1 + b_1^2 + c_1^2} \right) \right\}$$

■

2.2 Stationarity

We can rewrite the *BL – GARCH*(1,2) as

$$\begin{aligned} h_t^2 &= a_0 + (a_1\varepsilon_{t-1}^2 + b_1 + c_1\varepsilon_{t-1})h_{t-1}^2 + b_2h_{t-2}^2 \\ h_t^2 &= g(\varepsilon_{t-1}) + c(\varepsilon_{t-1})h_{t-1}^2 + d(\varepsilon_{t-1})h_{t-2}^2 \end{aligned}$$

which is a random coefficient autoregressive model of second order [*RCAR*(2)]. we put : $h_t^2 = X_t$; $g(\varepsilon_{t-1}) = e_t$; $c(\varepsilon_{t-1}) = \phi_1$ and $d(\varepsilon_{t-1}) = \phi_2$, we have

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + e_t \quad (2-6)$$

$$\begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ X_{t-2} \end{pmatrix} + \begin{pmatrix} e_t \\ 0 \end{pmatrix}$$

$$\begin{aligned} \underline{X}_t &= \Phi \underline{X}_{t-1} + \underline{e}_t \\ V_t &= a_0 [I_2 - \Psi B]^{-1} \end{aligned} \quad (2-7)$$

with $E(\underline{X}_t) = V_t$, $E(\Phi) = \Psi$ and B is backward operator. This implies that its eigenvalues are in the unit circle.

So in order to the process be second order stationarity if and only if all eigenvalues are within unit circle.

We can also rewrite the *BL – GARCH*(1,2) as

$$\underline{Z}_t = \underline{b}_t + A_t \underline{Z}_{t-1} \quad (2-8)$$

$$\begin{bmatrix} u_t^2 \\ h_t^2 \\ h_{t-1}^2 \\ u_t h_t \end{bmatrix} = \begin{bmatrix} a_0 \varepsilon_t \\ a_0 \\ 0 \\ a_0 \end{bmatrix} + \begin{bmatrix} a_1 \varepsilon_t^2 & b_1 \varepsilon_t^2 & b_2 \varepsilon_t^2 & c_1 \varepsilon_t^2 \\ a_1 & b_1 & b_2 & c_1 \\ 0 & 1 & 0 & 0 \\ a_1 \varepsilon_t & b_1 \varepsilon_t & b_2 \varepsilon_t & c_1 \varepsilon_t \end{bmatrix} \begin{bmatrix} u_{t-1}^2 \\ h_{t-1}^2 \\ h_{t-2}^2 \\ u_{t-1} h_{t-1} \end{bmatrix} \quad (2-9)$$

Remark 2 (i) A_t is $(p + q + r) \times (p + q + r)$ matrix and $r = \min(p, q)$ in general case of $BL - GARCH(p, q)$ model.
(ii) Equation (2-8) is random coefficient $VAR(1)$.
(iii) Iosifescu and Grigorescu (1990) proved that $(\underline{Z}_t)_{t \geq 1}$ is a Markov process

Theorem 3 (Strict Stationarity)

In order to exist a strict stationary solution of equation (2-8) it is necessary and sufficient that

$$\gamma < 0 \quad (2-10)$$

where $\gamma = \lim_{t \rightarrow +\infty} \frac{1}{t} \log \|A_t A_{t-1} \dots A_1\|$, is the largest Lyapunov exponent of the model (2-8).

If this solution exists, then it is unique strictly stationary, non anticipative and ergodic.

Proof. See [3]. ■

Example 4 For the $BL-GARCH(2,1)$ we consider the matrix

$$A_t = \begin{bmatrix} a_1 \varepsilon_t^2 & b_1 \varepsilon_t^2 & b_2 \varepsilon_t^2 & c_1 \varepsilon_t^2 \\ a_1 & b_1 & b_2 & c_1 \\ 0 & 1 & 0 & 0 \\ a_1 \varepsilon_t & b_1 \varepsilon_t & b_2 \varepsilon_t & c_1 \varepsilon_t \end{bmatrix}$$

for some values attributed to the coefficients a_1, b_1, b_2, c_1 we can simulate $\hat{\gamma}$ and $\varepsilon_t \sim N(0, 1)$ or $\varepsilon_t \sim t_n$.

Estimation of γ from 1000 simulations of size $t = 1000$.

This simulation does know us the region of stationarity of $BL - GARCH(1, 2)$

If $a_1 + b_1 + b_2 = 1$, there is no stationary solution (strict or 2nd order), while there is a strictly stationary solution of an IGARCH model under general conditions.

2.3 Marginal distribution

From (2-8) and by recursive, we have

$$\underline{Z}_t = \underline{b}_t + \sum_{k=1}^{+\infty} A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k}$$

we put, for $k > 0$,

$$A_{t,k} = A_t A_{t-1} \dots A_{t-k+1} \text{ and } \underline{Z}_{t,k} = A_{t,k} \underline{b}_{t-k}$$

we denote \otimes the Kronecker product and $\|\cdot\|$ matrix norm, then

$$\begin{aligned} E \|\underline{Z}_{t,k}\|^m &= E \|A_{t,k}^{\otimes m} \underline{b}_{t-k}^{\otimes m}\| = E \|A_t^{\otimes m} A_{t-1}^{\otimes m} \dots A_{t-k+1}^{\otimes m} \underline{b}_{t-k}^{\otimes m}\| \\ &= \|A^{(m)} \underline{b}^{(m)}\| \end{aligned}$$

using product matrices independence $A_t A_{t-1} \dots A_{t-k+1} \underline{b}_{t-k}$ (because (ε_t) are *iid*), we have

$$\begin{aligned} \|\underline{Z}_t\|_m &= [E \|\underline{Z}_t\|^m]^{1/m} \\ &= \sum_{k=0}^{\infty} \|\underline{Z}_{t,k}\|_m \\ &\leq \left\{ \sum_{k=0}^{\infty} \|(A^{(m)})^k\|^{1/m} \right\} \|\underline{b}^{(m)}\|^{1/m} \end{aligned}$$

If the spectral radius $\rho(A^{(m)}) < 1$ of the matrix $A^{(m)}$, then $\|(A^{(m)})^k\| \rightarrow 0$ as $k \rightarrow +\infty$, thus $\|u_t\|_m \leq \|\underline{Z}_t\|_m$. So $\rho(A^{(m)}) < 1$ is sufficient condition for the existence of $E(u_t^{2m})$. For more details, it is recommended to refer to [3].

Theorem 5 *Suppose that $E(\varepsilon_t^{2m})$ and $\rho(A^{(m)}) < 1$, then for each $t \in \mathbb{Z}$, $(\underline{Z}_t)_t$ defined by (2-8) converge in L^m and the process $(u_t^2)_t$ defined as the first component of \underline{Z}_t is $2m$ -order strictly stationary solution.*

we make a simulation given some values of coefficients and a distribution of ε_t in order to calculate $\rho(A^{(m)})$, $E(\varepsilon_t^{2m})$, and $E(u_t^{2m})$

3 The BL-GARCH Panel

We assumed that we have panel of asset returns with T observations and N assets. The return on asset i at time t is y_{it} where $i = 1, \dots, N$ and $t = 1, \dots, T$, given by

$$\begin{aligned} y_{it} &= \mu_{it} + u_{it} \\ u_{it} &= h_{it} \varepsilon_{it} \\ h_{it}^2 &= a_{0i} (1 - a_1 - b_1 - b_2) + a_1 u_{it-1}^2 + b_1 h_{it-1}^2 + b_2 h_{it-2}^2 + c_1 u_{it-1} h_{it-1} \end{aligned} \quad (3-1)$$

where $a_{0i} > 0$, $a_1 + b_1 + b_2 < 1$, $a_1, b_1, b_2 \in [0, 1[$ and $c_1 > 2a_1 b_1$.

We consider a_{0i} as a nuisance parameters and $\theta = (a_1, b_1, b_2, c_1)$ as a vector of interest parameters. Using the covariance tracking suggested by Engle and Mezrich (1996), we have

$$E(y_{it}^2) = a_{0i} \quad (3-2)$$

then we can use the method of moment to estimate the nuisance parameter. According to Barndorff-Nielsen (1996), Lancaster (2000) and Sartori (2003) in their papers, they assumed a stochastic independence over i and t . Then the maximum likelihood estimator of θ is typically inconsistent for finite T and $N \rightarrow \infty$. In order to overcome this problem and get a consistent estimator Engle, Shephard and Sheppard (2008) allowed that T to be large and N relates to T and reduce the rate of convergence to \sqrt{T} not \sqrt{NT} , noted in [2], followed by the same study and consideration of Pakel, Shephard and Sheppard (2011).

3.1 Composite maximum likelihood

In this subsection we apply composite maximum likelihood method, that is widely used in time series in place of full likelihood when for example we want to reduce the computational complexity, or make inference about parameters of interest without making assumptions on the whole joint distribution of the data.

Given the data $y = (y_1, y_2, \dots, y_T)$ where $y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})$ and let $f(y_{it}/\Omega_{i,t-1})$ be the conditional density for y_{it} , we put $\phi_i = (a_{i0}, \theta)$ and $\phi_{(N)} = (\phi_1, \phi_2, \dots, \phi_N)$. Our estimation procedure focuses on two-step. We begin by application of moment method to estimate nuisance parameters using (*), then we apply composite likelihood to estimate θ which is defined by

$$CL(y, \phi_{(N)}) = \frac{1}{T} \sum_{t=1}^T \left[\frac{1}{N} \sum_{i=1}^N \log f(y_{it}/\Omega_{i,t-1}; \phi_i) \right] \quad (3-3)$$

In our situation we use the variation-free as Engle, Shephard and Sheppard (2008) and Engle, Hendry and Richard (1983), then we obtain the composite maximum likelihood estimator by solving

$$\hat{\theta}_{CL} = \arg \max_{\theta} \frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \log f(y_{it}/\Omega_{i,t-1}; \hat{a}_{i0}, \theta) \quad (3-4)$$

where \hat{a}_{i0} for each i is obtained by solving

$$\sum_{t=1}^T Q_{it}(\theta, \hat{a}_{i0}) = 0$$

From (3-2) we have

$$Q_{it}(\theta, a_{i0}) = y_{it}^2 - a_{0i} ; E(Q_{it}(\theta, a_{i0}^*)) = 0 \quad (3-5)$$

$$\frac{1}{T} \sum_{t=1}^T Q_{it}(\theta, \hat{a}_{i0}) = 0 \quad (3-6)$$

where a_{i0}^* is the true value of a_{i0} for each i . Stacking (3-5) for $i = 1, \dots, N$, we have

$$Q_{(N)}(y_t, \phi_{(N)}) = \begin{bmatrix} y_{1t}^2 - a_{01} \\ \vdots \\ y_{Nt}^2 - a_{0N} \end{bmatrix} \implies E(Q_{(N)}(y_t, \phi_{(N)}^*)) = 0 \quad (3-7)$$

On the other hand, for the interest parameter θ , we use composite likelihood, considering the following three typical distributions :

* The score function for the normal density composite likelihood function is

$$W_1(y_t, \theta, \phi_{(N)}) = \frac{\partial}{\partial \theta} \frac{1}{N} \left(-\frac{1}{2} \sum_{i=1}^N \log h_{it}^2 - \frac{1}{2} \sum_{i=1}^N \frac{u_{it}^2}{h_{it}^2} \right) \quad (3-8)$$

* The score function for the cauchy density composite likelihood function is

$$W_2(y_t, \theta, \phi_{(N)}) = \frac{\partial}{\partial \theta} \left(-N \log \pi + \sum_{i=1}^N \log h_{it} - \sum_{i=1}^N \log (h_{it}^2 + u_{it}^2) \right) \quad (3-9)$$

* The score function for the student density composite likelihood function is

$$W_3(y_t, \theta, \phi_{(N)}) = \frac{\partial}{\partial \theta} N \left(\log \Gamma \left(\frac{\nu + 1}{2} \right) - \log \Gamma \left(\frac{\nu}{2} \right) - h_{it}^2 - \frac{1}{2} \sum_{i=1}^N \frac{u_{it}^2}{h_{it}^2} \right) \quad (3-10)$$

For $i = 1, 2, 3$ we put

$$W_i(y_t, \theta, \hat{\phi}_{(N)}) = 0 \quad (3-11)$$

where $\hat{\phi}_{(N)}$ is a moment estimator

The sample moment conditions for each of (3-8), (3-9) and (3-10) are given by

$$\frac{1}{T} \sum_{t=1}^T W_i(y_t, \hat{\theta}, \hat{\phi}_{(N)}) = 0 ; \quad \text{for } i = 1, 2, 3 \quad (3-12)$$

We put

$$K_i(y_t, \theta^*, \phi_{(N)}^*) = \begin{bmatrix} Q_{(N)}(y_t, \phi_{(N)}^*) \\ W_i(y_t, \theta^*, \phi_{(N)}^*) \end{bmatrix}$$

Then we imply that

$$E[K_i(y_t, \theta^*, \phi_{(N)}^*)] = 0 \quad \text{and} \quad \frac{1}{T} \sum_{t=1}^T K_i(y_t, \hat{\theta}, \hat{\phi}_{(N)}) = 0 ; \quad \text{for } i = 1, 2, 3$$

(3-6) and (3-11) are the first order condition for the maximisation problem of (3-4).

3.2 Asymptotic behavior

In this subsection we attempt to obtain the asymptotic properties of composite likelihood estimator, based on a reasonable initial moment estimator for nuisance parameters. We show under which initial conditions to have a consistent estimator and asymptotic normality with the standard root- T convergence rate and N can potentially increase with T .

Engle, Shephard and Sheppard (2008) have obtained consistency property and central limit theorem for $\hat{\theta}_{CL}$ under some regularity conditions, and also Billy Wu, Qiwei Yao and Shiwu Zhu (2013). Through the following two fundamental theorems, we will show the consistency and central limit theorem for $\hat{\theta}_{CL}$ when $T \rightarrow \infty$ while N can potentially increase with T .

Theorem 6 *We consider the following assumptions*

- (i) *The condition (3-5) holds.*
- (ii) *We assume that the parameters spaces are compacts.*
- (iii) *Suppose that*

$$\arg \max \frac{1}{TN} \sum_{t=1}^T \sum_{i=1}^N \log f(y_{it}/\Omega_{i,t-1}; a_{i0}^*, \theta) \xrightarrow{p} \theta^*$$

- (iv) *$\log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)$ is continuously differentiable in a_{i0}*
- (v) *Assume that the following sum satisfies a weak law large number as $T \rightarrow \infty$*

$$\frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N \sup_{a_{i0}, \theta} \left| \frac{\partial \log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)}{\partial a_{i0}} \right|$$

- (vi) *Assume that*

$$\sup_{\theta} \max_{i \in \{1, \dots, N\}} |\hat{a}_{i0} - a_{i0}| \xrightarrow{p} 0$$

then there exists a solution of the likelihood equation (3-11), for which

$$\hat{\theta} \xrightarrow{p} \theta^*$$

Proof. See [2]. ■

Theorem 7 For any consistent solution of the likelihood equation (3-11), we assume that

- (i) $Q_{it}(\theta, a_{i0})$ is once continuously differentiable.
- (ii) (a_{i0}^*, θ^*) is an interior point of $(\Lambda_i \times \Theta)$.
- (iii) we put

$$Y_{t,T} = \frac{1}{N} \sum_{i=1}^N \left[\frac{\partial \log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)}{\partial \theta'} - \left(\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)}{\partial \theta \partial a_{i0}} \right) Q_{it}(\theta, a_{i0}) \right]$$

$$D_{i,\theta\theta,T} = \left[\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)}{\partial \theta \partial \theta'} \right] - \left[\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \log f(y_{it}/\Omega_{i,t-1}; a_{i0}, \theta)}{\partial \theta \partial a_{i0}} \right] \left[\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 Q_{it}(\theta, a_{i0})}{\partial \theta'} \right]$$

$$D_{\theta\theta,T} = \frac{1}{N} \sum_{i=1}^N D_{i,\theta\theta,T}$$

- (iv) we assume that $(Y_{t,T})$ obeys a central limit theorem i.e

$$\frac{1}{T} \sum_{t=1}^T Y_{t,T} \xrightarrow{d} N(0, I_{\theta\theta})$$

where $I_{\theta\theta}$ is assumed that has diagonal elements definite positive.

- (v) That as $T \rightarrow \infty$; $D_{\theta\theta,T} \xrightarrow{d} D_{\theta\theta} > 0$, where $D_{\theta\theta}$ is invertible

Then $\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, D_{\theta\theta}^{-1} I_{\theta\theta} D_{\theta\theta}^{-1})$

Proof. The demonstration is well detailed in [2]. ■

Conclusion 8 Through this work we have tried to study, in the first part the fundamental probabilistic properties of BL-GARCH(1,2), basing on studies of Abdou Kâ Diongue, D. Guégan and R.C. Wolff (2009) and G. Storti & Vitale (2003), that have been made in this class of models. In the second part, we have studied the statistical inference, extending the model on panel data structure, using one of efficient method well called composite likelihood that was introduced by Lindsay (1988), this method has good properties under some general regularity conditions as the consistency property and the asymptotic normality of estimators.

References

- [1] Abdou Kâ Diongue, D. Guégan and R.C. Wolff. BL-GARCH models with elliptical distributed innovations. *Journal of Statistical Computation and Simulation*. Vol. 00, No. 0, Month 2009, pp. 1-17.
- [2] R.F. Engle, N. Shephard and K. Sheppard. Fitting vast dimensional time-varying covariance models. (2008), Working paper.
- [3] C. Francq and J-M. Zakoian. *GARCH Models*. WILEY 2010.
- [4] B. Lindsay. Composite likelihood methods. In *statistical inference from stochastic processes*, (Edited by N. U. Prabhu). American Mathematical Society, Providence, RI. (1988), pp. 221-239.
- [5] D.B. Nelson. Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, (1990), pp. 318-334.
- [6] L. Pace, A. Salvan and N. Sartori. Adjusting composite likelihood ratio statistics. *Statistica Sinica* 21, (2011), pp.129-148.
- [7] C. Pakel, N. Shephard and K. Sheppard. Nuisance parameters, composite likelihoods and panel of GARCH models. *Statistica Sinica* 21 (2011), pp. 307-329.
- [8] P. Posedel. Properties and estimation of GARCH(1,1) model. *Metodološki zvezki*, Vol. 2, (2005), pp. 243-257.
- [9] Richard A. Davis and Chun Yip Yau. Comments on pairwise likelihood in time series models. *Statistica Sinica* 21, (2011), pp. 255-277.
- [10] A. Rudogph. Acentral limit theorem for random coefficient autoregressive models and ARCH/GARCH models. *Adv. Prob.* 30, (1998), pp. 113-121.
- [11] G. Storti and C. Vitale. BL-GARCH models and asymmetries in volatility. *Statist. Methods Appl.* 12 (2003), pp. 19-40.
- [12] C. Tim Ng, H. Joe, D. Karlis and J. Liu. Composite likelihood for time series models with a latent autoregressive process. *Statistica Sinica* 21, (2011), pp. 279-305.
- [13] C. Varin, N. Reid and D. Firth. An overview of composite likelihood methods. *Statistica Sinica* 21, (2011), pp. 5-42.
- [14] B. Wu, Q. Yao and S. Zhu. Estimation in the Presence of Many Nuisance Parameters: composite likelihood and plug-in likelihood. *Stochastic Processes and Their Applications*, 123 (7). (2013), pp. 2877-2896.d in time series models. *Statistica Sinica* 21, (2011),pp.2