Seasonal Fractional ARIMA model with BL-GARCH type innovations

Mor Ndongo, Abdou Kâ Diongue, Simplice Dossou-Gbété

Seasonal Fractional ARIMA model with BL-GARCH type innovations

Mor Ndongo ♯, Abdou Kâ Diongue ♯[1], Simplice Dossou-Gbété §

Abstract

In this paper, we introduce the class of seasonal ARFIMA models with bilinear GARCH (BL-GARCH) type innovations that are capable of capturing simultaneously four key properties of non-linear time series: long range dependence, seasonality, volatility clustering and leverage effects. Stationarity and invertibility conditions are derived and conditional sum of squares (CSS) estimation of the model is also considered. Under some assumptions, we show that the resulting estimators are consistent and asymptotically normally distributed. Monte carlo simulation results are presented to evaluate the small-sample performance of the CSS method for various models.

Keywords: Long memory; Seasonality; Volatility clustering; Leverage effects; CSS method; Monte Carlo experiments.

1. Introduction

When dealing with empirical time series arising from diverse fields of applications, we are confronted with the phenomenon of long memory or long range dependence. A popular way to analyze a long memory time series is to use fractionally integrated autoregressive moving average (FARIMA) processes introduced by Granger and Joyeux [14] and Hosking [16]. Furthermore, most of time series in real life may have a persistent periodic behavior, in addition to long term structure. Unfortunately, the FARIMA model does not allow to take into account a periodic or cyclical behavior. Thus, models dedicated to take seasonal or cyclical components with long memory have been developed. Recent contributions related to the seasonal FARIMA model are Porter-Hudak [19], Hassler [15], Arteche and Robinson [1] and Reisen *et al.*[22]. All of the aforementioned works assume that the conditional variance of time series is a constant over time. However, non-constant variance in non-linear time series is a challenging modelling exercise, considered among many other things by Tong [26]. In particular, the stylized fact that the volatility of financial time series is non-constant has been long recognized in the literature (see e.g. Bollerslev [4], Bollerslev, Engle and Woodridge [5], and Weiss [27]). Thus, the methodology for modeling time series with long memory behavior has been extended to long memory time series with a time-varying conditional variance. See for instance, Ling and Li [17] who developed the ARFIMA model with GARCH type innovations, and recently Reisen *et al* [21] examine the daily average PM_{10} concentrations using a seasonal ARFIMA model with GARCH errors. Given that, these models cannot allow to capture asymmetries or non-linearities (Black [3]), we introduce in this paper a new class of seasonal ARFIMA with BL-GARCH type innovations. This approach allows to model simultaneously: long memory, seasonality, volatility clustering and leverage effects, often observed in financial or economics time series. The stationarity and invertibility of the proposed model are analyzed and, the conditional sum of squares (CSS) estimation of the model parameters is considered, and we study its asymptotic properties. Also, the small-sample performance of the CSS approach

[1]Corresponding author.Tel: +221 77 443 63 81, Fax: +221 33 961 53 38

is illustrated, using Monte Carlo simulations.

The article is organized as follows. In Section 2, the class of SARFIMA-BL-GARCH model is introduced with some important properties concerning the positivity of the conditional variance, as well as, for stationarity and invertibility. In Section 3, the conditional sum of squares estimation is presented, and we derive its asymptotic properties. Finally, Section 4 calibrates the performance of the estimation procedure through Monte Carlo simulations, while Section **??** provides concluding remarks.

2. Model and probabilistic properties

In this section, we introduce the model, we will work with, and we give conditions of existence and invertibility.

2.1. Model

Let B be the back shift operator satisfying $BY_t = Y_{t-1}$ for any process $(Y_t)_{t\in\mathbb{Z}}$, and $s \in \mathbb{N}^*$ the seasonal period, then the polynomial of non-seasonal orders p and q, seasonal orders P and Q are respectively defined by:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \qquad \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$

$$\Phi(B^s) = 1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \cdots - \Phi_{Ps} B^{Ps} \qquad \Theta(B^s) = 1 + \Theta_s B^s + \Theta_{2s} B^{2s} + \cdots + \Theta_{Qs} B^{Qs}.$$

It is assumed that these polynomials have no common zeros and satisfy the conditions $\Phi(z^s)\phi(z) \neq 0$ and $\Theta(z^s)\theta(z) \neq 0$ for $|z| = 1$. Furthermore, in the above equations, $(\Phi_i)_{1\leq i\leq P}$, $(\phi_j)_{1\leq j\leq p}$, $(\Theta_k)_{1\leq k\leq Q}$ and $(\theta_l)_{1\leq l\leq q}$ are unknown parameters.

A zero-mean process $(X_t)_{t\in\mathbb{Z}}$ is said a seasonal ARFIMA model with bilinear GARCH type errors, denoted hereafter as SARFIMA$(p,\ d,\ q) \times (P,\ D,\ Q)_s$-BL-GARCH$(r,\ m)$, if the following equation is satisfied

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D X_t = \theta(B)\Theta(B^s)\varepsilon_t, \tag{2.1}$$

where the long-memory parameters d and D are fractional parameters at the zero (or long-run) and seasonal frequencies, respectively . The process $(\varepsilon_t)_{t\in\mathbb{Z}}$ in equation (2.1) is a BL-GARCH$(r,\ m)$ model defined by:

$$\varepsilon_t = h_t Z_t \tag{2.2}$$

$$h_t^2 = a_0 + \sum_{i=1}^{r} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{m} b_j h_{t-j}^2 + \sum_{k=1}^{r^*} c_k \varepsilon_{t-k} h_{t-k}, \tag{2.3}$$

where a_0, a_i, b_j and c_k are constant parameters, r, m and r^* are non-negative integers with $r^* = \min(r,\ m)$, h_t^2 is the conditional variance of the process $(\varepsilon_t)_{t\in\mathbb{Z}}$ given the σ-fields I_{t-1} generated by the past information $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$. In above, the process $(Z_t)_{t\in\mathbb{Z}}$ is a sequence of independent identically normally distributed random variables with mean 0 and variance 1. This model has the advantage of being characterized by a more flexible parametric structure. In this model, leverage effects are explained by the interactions between past observations and volatilities. More precisely, for $c_k < 0$ a positive quantity is added to h_t^2 if $\varepsilon_{t-k} < 0$ while the same quantity is subtracted if $\varepsilon_{t-k} > 0$. If $c_k = 0$ for all k, the model in (2.1)-(2.3) is the seasonal fractional ARIMA model with generalized autoregressive conditional heteroscedasticity disturbances (SARFIMA$(p,\ d,\ q) \times (P,\ D,\ Q)_s$-GARCH$(r,\ m)$) introduced by Reisen *et al* [21]. In addition, if $D = P = Q = 0$, we obtain the fractionally integrated process with GARCH erros (ARFIMA$(p,\ d,\ q)$-GARCH$(r,\ m)$) proposed by Baillie *et al* [2] and Ling and Li [17].

According to Giraitis and Leipus [13] or Reisen et al. [22], one can easily show that

$$
\begin{aligned}
(I-B)^d(I-B^s)^D &= \prod_{j=0}^{[\frac{s}{2}]} \left[(1-e^{i\lambda_j}B)(1-e^{-i\lambda_j}B)\right]^{d_j} \\
&= \prod_{j=0}^{[\frac{s}{2}]} (1-2\cos(\lambda_j)B+B^2)^{d_j},
\end{aligned}
\tag{2.4}
$$

with $d_0 = \frac{d+D}{2}$, $d_j = D$, for $j = 1, \ldots, [\frac{s}{2}]-1$, $d_{[\frac{s}{2}]} = \frac{D}{2}$, and $\lambda_j = \frac{2\pi j}{s}$, for $j = 0, \ldots, [\frac{s}{2}]$, where $[x]$ is the greatest integer small than or equal to x, and i is the complex number such that $i^2 = -1$. Otherwise, by means of the expansion we have:

$$
\prod_{j=0}^{[\frac{s}{2}]} (1-2cos\lambda_j B+B^2)^{d_j} = \sum_{j=0}^{+\infty} \psi_j(d,\ \nu)B^j,
$$

where the coefficients $\left(\psi_j(d,\ \nu)\right)_{j\geq 0}$ are given by:

$$
\psi_j(d,\ \nu) = \sum_{\substack{0\leq l_0,\cdots,l_{[\frac{s}{2}]}\leq j,\\ l_0+\cdots+l_{[\frac{s}{2}]}=j}} C_{l_0}(d_0,\nu_0)\cdots C_{l_{[\frac{s}{2}]}}\left(d_{[\frac{s}{2}]},\nu_{[\frac{s}{2}]}\right),
\tag{2.5}
$$

and where $d = (d_0, \ldots, d_{[\frac{s}{2}]})$, $\nu = (\nu_0, \ldots, \nu_{[\frac{s}{2}]})$ with $\nu_j = \cos(\lambda_j)$, for $j = 0, \ldots, [\frac{s}{2}]$. The weights $(C_l(d_i,\nu_i))_{l\in\mathbb{Z}}$ are the Gegenbauer polynomials and they can be computed using the following recursion formula:

$$
\begin{cases}
C_0(d_i,\nu_i) = 1 \\
C_1(d_i,\nu_i) = 2d_i\nu_i \\
C_j(d_i,\nu_i) = 2\nu_i\left(\frac{d_i-1}{j}+1\right)C_{j-1}(d_i,\nu_i) - \left(2\frac{d_i-1}{j}+1\right)C_{j-2}(d_i,\nu_i),\ \forall j>1.
\end{cases}
$$

2.2. *Probabilistic properties*

We now specify some probabilistic properties of the model 2.1 - 2.3. First, we give the conditions for which the conditional variance h_t^2 , defined in (2.3), is non-negative. It is important in practice for estimation theory (using quasi-maximum likelihood methods) that a model as in (2.2) and (2.3) does not generate negative conditional variance h_t^2 in sample, since the log quasi-likelihood involves a term in $\log(h_t^2)$, which explodes to $-\infty$ as h_t^2 approaches 0, and is ill-defined for $h_t^2 \leq 0$. The second set of conditions concerns the second-order stationary of the BL-GARCH process and the stationary for a seasonal fractional ARIMA model. The BL-GARCH model with gaussian distributions was recently proposed by Storti and Vitale [24]. This study was extended by Diongue *et al* [9], considering elliptical distributions, and they give new probabilistic results concerning the stationarity of the process and the moments.

For the positivity of the conditional variance, Storti and Vitale [24] show that, for $a_0 > 0$ and $s_i > 0$ $(i = 1, \ldots, \max(r,m) - r^*)$, with $s_i = a_{r^*+i}$ if $r > m$, or $s_i = a_{r^*+i}$ if $r < m$, a sufficient conditions for $h_t^2 > 0$ is given by:

$$
c_i^2 < 4a_ib_i, \qquad \text{for } i = 1, \ldots, r^*.
\tag{2.6}
$$

They give also the condition of second-order stationary for the BL-GARCH model, restricting to the case $r = m$. Here, we extend this result to the more general case in which $r \neq m$. The result is stated in the following corollary.

Corollary 2.1. The BL-GARCH($r,\ m$) process, defined by equations (2.2) and (2.3), is stationary in wide-sense if and only if all the roots B_i of the polynomial

$$\Pi(B) = 1 - \sum_{i=1}^{r^*} \pi_i B^i - \sum_{i=r^*+1}^{\max(r,m)} \alpha_i B^i, \tag{2.7}$$

with B being the Backward operator, $\pi_i = a_i + b_i$, for $i = 1, \ldots, r^*$, lie outside the unit circle. Note that in (2.7), we set $\alpha_i = a_i$ if $r > m$, $\alpha_i = b_i$ if $r < m$. In other words, the condition can be expressed as follows:

$$\Pi(B) = 0 \qquad \text{for } |B_i| > 1,\ i = 1, \ldots, \max(r,\ m). \tag{2.8}$$

Proof. The proof of Corollary 2.1 follows mainly the lines of the proof of Theorem 2 in Storti and Vitale [24]. ■

Concerning the seasonal fractional ARIMA model, Reisen *et al.*[22] (see also Giraitis and Leipus [13]) show that it is stationary and invertible if and only if the following conditions hold:

$$|D + d| < \frac{1}{2} \quad \text{and} \quad |D| < \frac{1}{2}. \tag{2.9}$$

Therefore, using the model properties in Reisen *et al.*[22] and the conditions listed in Storti and Vitale [24], the following theorem is established for the SARFIMA($p,\ d,\ q$) $\times$ ($P,\ D,\ Q$)$_s$−BL-GARCH($r,\ m$) model.

Theorem 2.2. Let $(X_t)_{t\in\mathbb{Z}}$ be generated by equations (2.1), (2.2) and (2.3). We assume that the conditions (2.6) and (2.8) hold. We suppose that the polynomials $\Phi(z^s)\phi(z)$ and $\Theta(z^s)\theta(z)$ have no common zeros, and the long memory parameters d and D satisfy condition (2.9). Then, the following statements hold:

a) if $\Phi(z^s)\phi(z) \neq 0$, for $|z| = 1$, then $(X_t)_{t\in\mathbb{Z}}$ is second-order stationary and has the following representation

$$X_t = \sum_{j=0}^{\infty} \psi_j(d,\ \nu) \frac{\Theta(z^s)\theta(z)}{\Phi(z^s)\phi(z)}\, \varepsilon_{t-j}. \tag{2.10}$$

Hence, (X_t) is strictly stationary and ergodic.

b) If $\Theta(z^s)\theta(z) \neq 0$, for $|z| \leq 1$, then (X_t) is invertible, and that is, (ε_t) can be written as

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j(d,\ \nu) \frac{\Phi(z^s)\phi(z)}{\Theta(z^s)\theta(z)}\, X_{t-j}, \tag{2.11}$$

where the weights $\left(\pi_j(d,\ \nu)\right)_{j\geq 0}$ are such that $\pi_j(d,\ \nu) = \psi_j(-d,\ \nu)$, with the coefficients $\left(\psi_j(d,\ \nu)\right)_{j\geq 0}$ given by equation (2.5).

Proof. The proof of this theorem is given in the appendix. ■

Under stationarity and invertibility conditions, MA(∞) and AR(∞) representations are respectively:

$$X_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} \qquad \text{and} \qquad \varepsilon_t = \sum_{j=0}^{\infty} \tilde{c}_j X_{t-j}, \tag{2.12}$$

where the coefficients $(c_j)_{j\geq 0}$ and $(\tilde{c}_j)_{j\geq 0}$ are determined respectively by:

$$\Theta(z^s)\theta(z)\sum_{j=0}^{\infty}\psi_j(d,\ \nu)z^j = \Phi(z^s)\phi(z)\sum_{j=0}^{\infty}c_j z^j, \qquad |z| \leq 1 \tag{2.13}$$

and

$$\Phi(z^s)\phi(z)\sum_{j=0}^{\infty}\pi_j(d,\ \nu)z^j = \Theta(z^s)\theta(z)\sum_{j=0}^{\infty}\tilde{c}_j z^j, \qquad |z| \leq 1. \tag{2.14}$$

In the particular case where $P = Q = 0$, it is easy to verify that the coefficients $(c_j)_{j\geq 0}$ and $(\tilde{c}_j)_{j\geq 0}$ can be computed using the following recursion formula:

$$c_0 = 1 \quad \text{and} \quad c_j = \psi_j(d, \nu) + \sum_{i=1}^{\min(j,p)} \phi_i c_{j-i} - \sum_{i=1}^{\min(j,q)} \theta_i \psi_{j-i}(d, \nu), \qquad \forall j \geq 1 \tag{2.15}$$

and

$$\tilde{c}_0 = 1 \quad \text{and} \quad \tilde{c}_j = \pi_j(d, \nu) - \sum_{i=1}^{\min(j,p)} \phi_i \pi_{j-i}(d, \nu) + \sum_{i=1}^{\min(j,q)} \theta_i \tilde{c}_{j-i}, \qquad \forall j \geq 1. \tag{2.16}$$

3. Conditional Sum of Squares estimation

Chung [8] proposed a method based on maximization of the CSS function. In this section, we define the CSS method in the estimation of the SARFIMA$(p,\ d,\ q) \times (P,\ D,\ Q)_s$-BL-GARCH$(r,\ m)$ process defined by equations (2.1), (2.2) and (2.3), and we establish the asymptotic properties of the CSS estimator.

3.1. Definition of the estimator

Suppose that $X_1, \ldots, X_n$ are generated by the model (2.1), (2.2) and (2.3). Denote by $\gamma = (d,\ D,\ \phi_1, \ldots, \phi_p,\ \theta_1, \ldots, \theta_q,\ \Phi_1, \ldots, \Phi_P,\ \Theta_1, \ldots, \Theta_Q)^T$, $\delta = (a_0,\ a_1, \ldots,\ a_r,\ b_1, \ldots, b_m,\ c_1, \ldots, c_{r^*})^T$, and $\omega = (\gamma^T,\ \delta^T)^T$ the parameter vector to be estimated. We assume that the parameter ω satisfies the stationary conditions given in Theorem 2.2, and $\omega_0 = (\gamma_0^T,\ \delta_0^T)^T$ is the true value of ω and is in the interior of the compact set $\Lambda \subseteq \mathbb{R}^{p+q+P+Q+r+m+r^*+3}$. Under the assumption of normality of the standardized innovation $(Z_t)_{t\in\mathbb{Z}}$, the likelihood function is equal to

$$\mathcal{L}(\omega) = \mathcal{L}(\omega \mid Z_1, \ldots, Z_n) = \prod_{t=1}^{n} g(Z_t, \omega),$$

where $g(Z_t, \omega)$ is the conditional Gaussian distribution function. Thus, the conditional sum of squares estimator $\hat{\omega}_n$ of ω in Λ maximizes the conditional log-likelihood on I_0, (ignoring the constant)

$$L(\omega) = \frac{1}{n}\sum_{t=1}^{n} l_t, \qquad l_t = -\frac{1}{2}\log(h_t^2) - \frac{\varepsilon_t^2}{2h_t^2}. \tag{3.17}$$

3.2. *Asymptotic properties*

The following theorem provides some results related to the asymptotic properties of the SARFIMA-BL-GARCH process estimators, obtained by the CSS approach.

Theorem 3.1. Suppose that $(X_t)_{t\in\mathbb{Z}}$ are generated by equations (2.1), (2.2) and (2.3). Assume that the hypothesis of Theorem 2.2 holds. Then, the conditional sum of squares estimator $\hat{\omega}_n$ of ω, obtained by maximizing the conditional log-likelihood function (3.17), has the following properties

(a) The CSS estimator $\hat{\omega}_n$ exists and satisfies:

$$\frac{\partial L(\omega)}{\partial \omega} = 0 \quad \text{and} \quad \hat{\omega}_n \xrightarrow{\mathbb{P}} \omega_0, \quad \text{as } n \longrightarrow +\infty.$$

(b) For such a sequence,

$$\sqrt{n}\,(\hat{\omega}_n - \omega_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ \Omega_0^{-1}), \quad \text{as } n \longrightarrow +\infty,$$

where $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{\mathcal{D}}$ denote respectively convergence in probability and in distribution. Furthermore, $\Omega_0 = \operatorname{diag}(\Omega_{\gamma_0},\ \Omega_{\delta_0})$, Ω_{γ_0} and Ω_{δ_0} are values of Ω_γ and Ω_δ at $\omega = \omega_0$, with

$$\Omega_\gamma = \mathbb{E}\left[\frac{1}{h_t^2}\frac{\partial \varepsilon_t}{\partial \gamma}\frac{\partial \varepsilon_t}{\partial \gamma^T} + \frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial \gamma}\frac{\partial h_t^2}{\partial \gamma^T}\right] \quad \text{and} \quad \Omega_\delta = \mathbb{E}\left[\frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial \delta}\frac{\partial h_t^2}{\partial \delta^T}\right].$$

(c) Further, the information matrix Ω_γ and Ω_δ can be estimated consistently by:

$$\hat{\Omega}_\gamma = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{1}{h_t^2}\frac{\partial \varepsilon_t}{\partial \gamma}\frac{\partial \varepsilon_t}{\partial \gamma^T} + \frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial \gamma}\frac{\partial h_t^2}{\partial \gamma^T}\right] \quad \text{and} \quad \hat{\Omega}_\delta = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial \delta}\frac{\partial h_t^2}{\partial \delta^T}\right].$$

Proof. The proof is given in Appendix ■

It is important to not that ε_t, h_t all depend on the theoretically infinite past history of (X_t) or (ε_t). We choose the presample estimates of h_t and ε_t^2 to be $\sum_{t=1}^{n} \varepsilon_t^2/n$. As noted in Ling and Li [17], this will not affect asymptotic efficiency and other asymptotic properties (see also Bollerslev [4] and Weiss [27]).

4. Monte Carlo Simulation

In this section, we study the finite sample performance of the CSS method described previously, to estimate the parameters of SARFIMA$(p,\ d,\ q) \times (P,\ D,\ Q)_s$-BL-GARCH$(r,\ m)$ process. The simulation results give the average values, the root mean square error (RMSE) and the mean absolute error (MAE) of the estimation procedure based on 1000 replications. All calculations were carried out using an R programming environment (see [20]) on a Pentium (R) Dual-Core CPU, 2.60 GHz (2 CPUs) computer.

Because there is no known technique for generating an exact Seasonal ARFIMA model with BL-GARCH type innovations, we adapt the method developed by Stoev and Taqqu [23] for ARFIMA time series with stable innovations. Thus, we approximate the path X_t, $t = 1, \ldots, n$ by the truncation moving average

$$X_t = \sum_{j=0}^{M} c_j \varepsilon_{t-j}, \tag{4.18}$$

where M is the truncation parameter and fixed in this study to 10 000, and the non random constants $(c_j)_{j\geq 0}$ are determined by equation (2.13). The sum in (4.18) is computed using the Fast Fourier Transform (FFT) algorithm, as suggested by Stoev and Taqqu [23]. The BL-GARCH innovations (ε_t) are generated using the method developed in Diongue *et al.*[9]. The simulation algorithm generates $n + 500$ observations for each series, saving only the last n. This operation is performed in order to avoid dependence on initial values.

Thus, for fixed value of the seasonal period ($s = 4$), we carry out an experiment of 1000 samples for the four processes summarized in Table 1. For each model, we use four different sample sizes ($n = 100$, $n = 300$, $n = 500$ and $n = 1000$). The simulation results are presented in Tables 2-5.

Table 1: Data generating processes (DGPs).

DGPs	d	D	ϕ	θ	a_0	a_1	b_1	c_1
Model 1	0.10	0.30	–	–	0.01	0.09	0.9	0.15
Model 2	0.10	0.30	0.50	–	0.01	0.09	0.9	0.15
Model 3	0.10	0.30	–	0.30	0.01	0.09	0.9	0.15
Model 4	0.10	0.30	0.50	0.30	0.01	0.09	0.9	0.15

In these tables, the sample sizes are given in the first columns. The estimations of the parameters are given in the next columns, the root mean square error (RMSE) is given in the row below (in parentheses) and the mean absolute error (MAE) is given under the row of the RMSE (in brackets).

- In the first experiment, a SARFIMA$(0, d, 0)\times(0, D, 0)$-BL-GARCH$(1, 1)$ model is considered (i.e. the model without short memory parameters). The simulation results are summarized in Table 2. Results reveal that parameter estimates are satisfactory, even for small sample sizes ($n = 100$, $n = 300$), in the sense that the RMSE and also the MAE are small. We can also remark that the impact of the sample size n on the estimation method. Indeed, when the sample size increases ($n = 1000$), the results improve significantly too. Figure 1 displays the corresponding boxplots[2] and shows the relative scatter of the 1000 estimates. The vertical axis in the figure indicates the deviation from the nominal value of the parameters. This figure confirms the previous results since we observe that confidence intervals are small, and shows also that the impact of the sample sizes on this dispersion.
- In the second experiment, we consider the preceding model (Model 1) with short memory part. Consequently, this experiment is designed to examine the relative performances of the estimators when there are long-memory and short memory components simultaneously. The results are presented in Tables 3 - 5. We observe that, when short memory components are introduced in the model, the estimation of long memory parameters is disturbed. Indeed, the RMSE and MAE obtained in Tables 3, 4 and 5 are larger than those presented in Table 2. This phenomenon is already observed in the literature (e.g. Boutahar *et al.*[6], Diongue and Guégan[11]). However, we observe a significant improvement when the sample size becomes large ($n = 500$, $n = 1000$). We remark also that the estimators of the parameters for the BL-GARCH errors seem not to be affected by the presence of the AR and MA components, particularly for larger sample sizes ($n = 500$, $n = 1000$).

[2]Boxplot of estimates in Model 2, Model 3 and Model 4 are omitted in the paper but are available to authors upon request.

Table 2: CSS estimation for Model 1.

sizes	$\widehat{d}$	$\widehat{D}$	$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{c}_1$
100	0.0980	0.3137	0.0617	0.0825	0.8156	0.1844
	(0.0732)	(0.0893)	(0.1022)	(0.0945)	(0.2286)	(0.1187)
	[0.0605]	[0.0727]	[0.0557]	[0.0750]	[0.1433]	[0.0903]
300	0.1013	0.3158	0.0190	0.0857	0.8883	0.1552
	(0.0503)	(0.0520)	(0.0246)	(0.0490)	(0.0721)	(0.0529)
	[0.0408]	[0.0419]	[0.0119]	[0.0378]	[0.0462]	[0.0415]
500	0.1019	0.3125	0.0140	0.0897	0.8926	0.1541
	(0.0387)	(0.0413)	(0.0102)	(0.0335)	(0.0383)	(0.0409)
	[0.0305]	[0.0326]	[0.0060]	[0.0261]	[0.0289]	[0.0320]
1000	0.0996	0.3097	0.01148	0.0896	0.8972	0.1499
	(0.0264)	(0.0280)	(0.0042)	(0.0198)	(0.0211)	(0.0251)
	[0.0213]	[0.0223]	[0.0030]	[0.0158]	[0.0164]	[0.0202]

Table 3: CSS estimation for Model 2.

sizes	$\widehat{d}$	$\widehat{D}$	$\widehat{\phi}$	$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{c}_1$
100	0.1005	0.3116	0.4907	0.0763	0.0763	0.8114	0.1789
	(0.1068)	(0.0991)	(0.1414)	(0.2127)	(0.0964)	(0.2406)	(0.1214)
	[0.0906]	[0.0789]	[0.1100]	[0.0694]	[0.0763]	[0.1476]	[0.0917]
300	0.0942	0.3083	0.4984	0.0211	0.0837	0.8825	0.1500
	(0.0870)	(0.0548)	(0.1002)	(0.0316)	(0.0467)	(0.0932)	(0.0531)
	[0.0767]	[0.0431]	[0.0842]	[0.0134]	[0.0368]	[0.0519]	[0.0409]
500	0.0957	0.3074	0.4989	0.0135	0.0876	0.8957	0.1511
	(0.0771)	(0.0413)	(0.0856)	(0.0096)	(0.0331)	(0.0381)	(0.0375)
	[0.0671]	[0.0326]	[0.0719]	(0.0056)	[0.0256]	[0.0278]	[0.0296]
1000	0.1004	0.3058	0.4973	0.0113	0.0901	0.8978	0.1529
	(0.0634)	(0.0284)	(0.07138)	(0.0043)	(0.0205)	(0.0216)	(0.0261)
	[0.0526]	[0.0229]	[0.0584]	[0.0031]	[0.0159]	[0.0166]	[0.0207]

Table 4: CSS estimation for Model 3.

sizes	$\widehat{d}$	$\widehat{D}$	$\widehat{\theta}$	$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{c}_1$
100	0.1265	0.3116	0.2945	0.0760	0.0778	0.8061	0.1748
	(0.1050)	(0.0973)	(0.1366)	(0.2014)	(0.0987)	(0.2481)	(0.1191)
	[0.0863]	[0.0779]	[0.1090]	[0.0692]	[0.0773]	[0.1514]	[0.0886]
300	0.1268	0.3064	0.3011	0.0199	0.0826	0.8882	0.1508
	(0.0792)	(0.0518)	(0.0844)	(0.0271)	(0.0466)	(0.0778)	(0.0522)
	[0.0658]	[0.0413]	[0.0686]	[0.0123]	[0.0363]	[0.0470]	[0.0400]
500	0.1350	0.3061	0.3118	0.0135	0.0871	0.8961	0.1481
	(0.0738)	(0.0409)	(0.0758)	(0.0086)	(0.0325)	(0.0357)	(0.0366)
	[0.0612]	[0.0323]	[0.0612]	[0.0056]	[0.0257]	[0.0270]	[0.0289]
1000	0.1381	0.3000	0.3181	0.0110	0.0896	0.8972	0.1446
	(0.0615)	(0.0292)	(0.0564)	(0.0044)	(0.0199)	(0.0221)	(0.0258)
	[0.0509]	[0.0228]	[0.0458]	[0.0030]	[0.0158]	[0.0170]	[0.0207]

Table 5: CSS estimation for Model 4.

sizes	$\widehat{d}$	$\widehat{D}$	$\widehat{\phi}$	$\widehat{\theta}$	$\widehat{a}_0$	$\widehat{a}_1$	$\widehat{b}_1$	$\widehat{c}_1$
100	0.0904	0.3137	0.4753	0.2626	0.0790	0.0789	0.7974	0.1830
	(0.1081)	(0.1003)	(0.3194)	(0.3437)	(0.2122)	(0.1002)	(0.2625)	(0.1277)
	[0.0931]	[0.0805]	[0.2336]	[0.2599]	[0.0721]	[0.0785]	[0.1603]	[0.0947]
300	0.0841	0.3107	0.4650	0.2453	0.0199	0.0810	0.8901	0.1555
	(0.0840)	(0.0527)	(0.2680)	(0.2717)	(0.0251)	(0.0465)	(0.0734)	(0.0527)
	[0.0735]	[0.0415]	[0.1867]	[0.1935]	[0.0122]	[0.0365]	[0.0460]	[0.0411]
500	0.0880	0.3116	0.4728	0.2588	0.0139	0.0862	0.8963	0.1526
	(0.0775)	(0.0421)	(0.2332)	(0.2302)	(0.0090)	(0.0320)	(0.0355)	(0.0378)
	[0.0672]	[0.0334]	[0.1603]	[0.1602]	[0.0058]	[0.0255]	[0.0269]	[0.0297]
1000	0.0924	0.3053	0.4560	0.2493	0.0116	0.0896	0.8964	0.1485
	(0.0679)	(0.0303)	(0.2216)	(0.2016)	(0.0047)	(0.0197)	(0.0224)	(0.0262)
	[0.0582]	[0.0238]	[0.1459]	[0.1347]	[0.0032]	[0.0155]	[0.0170]	[0.0211]

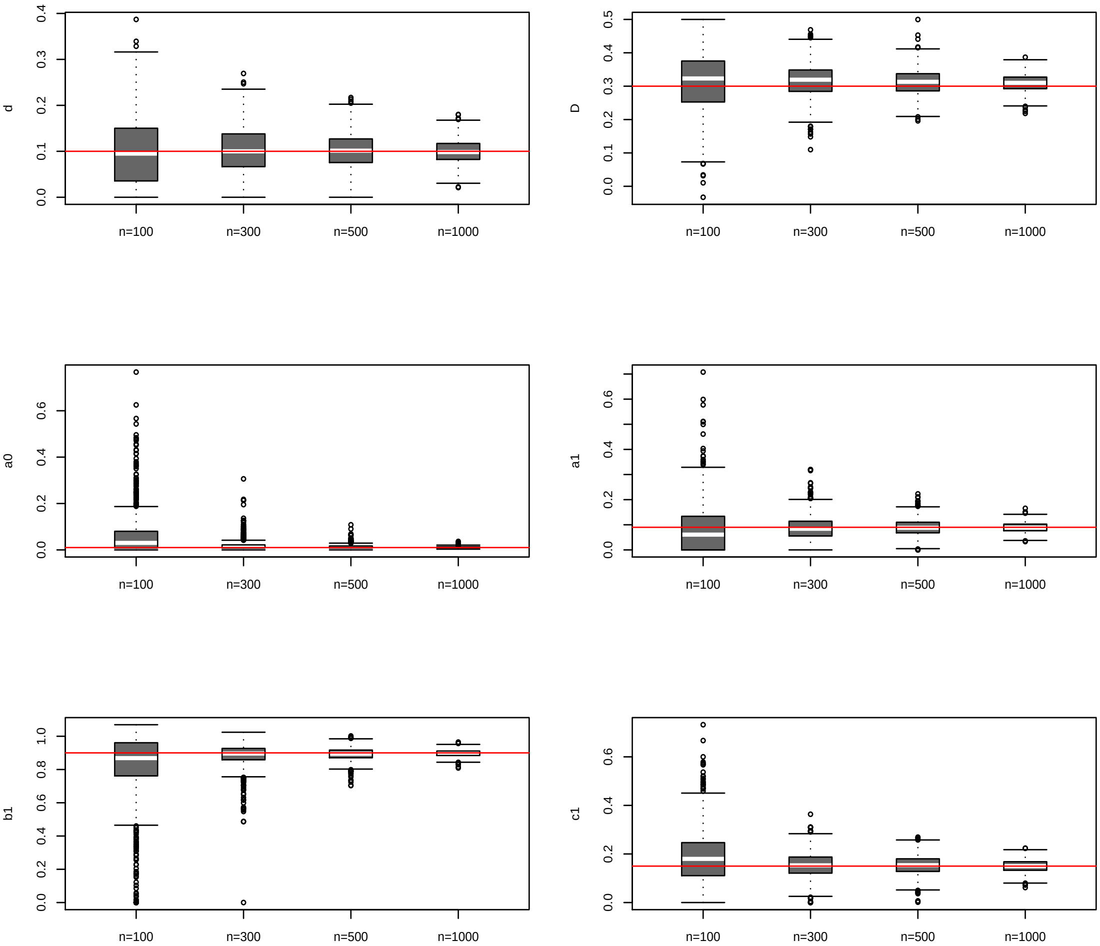

Figure 1: Boxplot of estimates of $d = 0.1$, $D = 0.3$, $a_0 = 0.01$, $a_1 = 0.09$, $b_1 = 0.9$ and $c_1 = 0.15$ in Model 1, with sample size $n = 100$, $n = 300$, $n = 500$ and $n = 1000$, based on 1000 replications.

5. Empirical study

The daily returns of the kenyan stock market index are used for the empirical study in this article to illustrate both the usefulness and the applicability of the proposed model and estimation method.
The raw series has a sample size of 6181 observations, measured form March 1st of 1991 to January 31st of 2008. Table **??** gives the summary statistics of the returns of kenyan stock index for the full sample.

For modeling purpose, the time series is divided into two parts; learning and prediction sets. The 5816 observations from March 1st of 1991 until January 31st of 2007 are considered as learning set and the remaining 365 observations are considered for the prediction study.

6. Conclusion

In this article, we have developed the family of seasonal ARFIMA process with BL-GARCH type errors, which should prove useful in many fields of time series analysis. It is much more flexible in the simultaneous modeling of long-memory behaviour, seasonal components, volatility clustering and leverage effects; often encountered in financial data. Under some assumptions, the model is shown to be stationary and invertible. The parameter estimation problem is addressed to the conditional sum of squares procedure proposed by Chung [8], and under some regularly conditions we establish the asymptotic properties of the resulting estimates. Finite sample behaviours of this method were studied using Monte Carlo simulations. It indicate that the approach can yield asymptotic efficient estimates. Consequently, it is reasonable method to deal with seasonal long-range dependent data containing volatility clustering and leverage effects.

Since the results from the model presentation and the estimation methodology are encouraging, it will be interesting to examine the empirical application of the SARFIMA-BL-GARCH model in financial data. We raise also the question of the "best" estimation procedure. This issue appears to be further investigated.

References

[1] J., Arteche and P. M., Robinson, Semiparametric inference in seasonal and cyclical long memory processes. Journal of Time Series Analysis, 21(1), (2000), 1-25.
[2] R. T. Baillie, C. F. Chung, M. A. Tieslau. Analyzing industrialized countries inflation by the fractionally integrated ARFIMA-GARCH model. Journal of Applied Econometrics 11, (1996), 23-40.
[3] F. Black, Studies in stock price volatility changes. Proceedings of the American Statistical Association Annual Meetings, Business and Economics Section,Washington, D.C., 1976, pp. 177181.
[4] T., Bollerslev, Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31, (1986) 307-327.
[5] T., Bollerslev, R. F. Engle and J. M. Woodridge. A Capital As- set Pricing Model With Time-Varying Covariance. Journal of Political Economy, 96, (1988), 116-131.
[6] M. Boutahar, V. Marimoutou, L. Nouira, Estimation Methods of the Long Memory Parameter: Monte Carlo Analysis and Application. Journal of Applied Statistics, 34(3), (2007), 261-301.
[7] I. V. Basawa, P. D. Feigin and C. C. Heyde. Asymptotic Properties of Maximum Likelihood Estimators for Stochastic Processes. Sankhya, Série A, 38 (1976), 259-270.
[8] C. F. Chung, Estimating a generalized long memory process. Journal of Econometrics, 73, (1996a), 237-259.
[9] A. K. Diongue, D. Guégan and R. C. Wolff, BL-GARCH models with elliptical distributed innovations. Journal of Statistical Computation and Simulation, 0, (2009) 1-17.
[10] Z. Ding, C. W. Granger, and R. F. Engle, A long memory property of stock market returns and a new model, J. Empirical Finance 1, (1993), 83106.
[11] A. K. Diongue and D. Guégan. Estimation of k-Factor GIGARCH Process: A Monte Carlo Study, Communications in Statistics - Simulation and Computation, 37(10), (2008), 2037-2049
[12] L. Ferrara and D. Guégan. Comparison of parameter estimation methods in cyclical long memory time series. In: Dunis, C., Timmerman, J., eds. Developments in Forecasts Combination and Portfolio Choice. Chap. 8, New York: John Wiley, (2001).
[13] L. Giraitis and R. Leipus, A generalized fractionally differencing approach in long memory modelling. Lithuanian Mathematical Journal, 35, (1995) 65-81.
[14] C.W.J. Granger, R. Joyeux, An intoduction to long memory time series and fractional differencing, Journal of Time Series Analysis 1 (1980) 15-30.
[15] U., Hassler, (Mis)specification of long memory in seasonal time series. Journal of Time Series Analysis, 15(1), (1994), 19-30.
[16] , J. R. M. Hosking, Fractional Differencing. Biometrika, 68(1), (1981), 165-176.
[17] S. Ling and W. K. Li. On Fractionally Integrated Autoregressive Moving-Average Time series Models with Conditional Heteroscedasticity. Journal of the American Statistical Association, 92, (1997), 1184-1194.
[18] D. B. Nelson, Conditional heteroskedasticity in asset returns:A new approach, Econometrica 59 (1991), pp. 347370.
[19] S. Porter-Hudak, An application of the seasonal fractionally differenced model to the monetary aggegrates, Journal of the American Statistical Association 85 (410) (1990) 338-344.
[20] R Development Core Team. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0. URL http://www.R-project.org.
[21] V. A. Reisen, A. J. Q. Sarnaglia, N. C. Reis Jr., C. Lévy-Leduc and J. M. Santos. Modeling and forecasting daily average PM_{10} concentrations by a seasonal ARFIMA model with volatility. arXiv:1206.2425v1, [stat.AP] 12 Jun 2012.
[22] V. A. Reisen, A. L. Rodrigues and W. Palma, Estimation of seasonal fractionally integrated processes. Computational Statistics & Data Analysis, 50, (2006) 568-582.

[23] S., Stoev and M., Taqqu. Simulation methods for linear fractional stable motion and FARIMA using the Fast Fourier Transform. Fractals, **12**, Nb 1, (2004) 95-121.
[24] G. Storti and C. Vitale, BL-GARCH models and asymmetries in volatility, Statist. Methods Appl. 12 (2003), 19-40.
[25] W. F. Stout, Almost Sure Convergence, New York: Academic Press, (1974).
[26] H. Tong, Non-linear Time Series: A Dynamical System Approach, Oxford Publications, Oxford University Press, 1990.
[27] A. A. Weiss, ARMA Models With ARCH Errors. Journal of Time Series Analysis, 5, (1984), 129-143.

Appendix : Proofs of Theorems

Proof of Theorem 2.2

The proof of this theorem follows mainly the proof of Theorem 2.3 in Ling and Li [17].
a) The process in (2.1) can be written as

$$X_t = \frac{\Theta(B^s)\theta(B)}{\Phi(B^s)\phi(B)}(1-B)^{-d}(1-B^s)^{-D}\varepsilon_t = \Psi_1(B)\varepsilon_t,$$

where $\Psi_1(z) = \frac{\Theta(z^s)\theta(z)}{\Phi(z^s)\phi(z)}(1-z)^{-d}(1-z^s)^{-D}$. Now the power series expansion $(1-z)^{-d}(1-z^s)^{-D}$ converges for all $|z| \le 1$, when $|D+d| < 1/2$ and $|D| < 1/2$ (see Reisen *et al.*[22] for more details). The series $\Theta(z^s)\theta(z)$ converge also for all z, since it is a polynomial, and that $1/\Phi(z^s)\phi(z)$ converges for all $|z| \le 1$ when all the roots of the equation $\Phi(z^s)\phi(z) = 0$ lie outside the circle $|z| = 1$. Thus, the power series expansion of $\Psi_1(z)$ converges for all $|z| \le 1$, and so $(X_t)_t$ exists with representation (2.10). From Corollary 2.1, the process $(\varepsilon_t)_t$ is second-order stationary and hence $\mathbb{E}(\varepsilon_t^2) < \infty$. Similar to the proof of Theorem 2.2 in Ling and Li [17], we can show that $(X_t)_t$ is also second-order stationary. Indeed, we have

$$\begin{aligned}\mathbb{E}\left(X_t^2\right) &= \mathbb{E}\left[\left(\sum_{k=0}^{\infty} c_k\varepsilon_{t-k}\right)^2\right] && (6.19)\\ &= \sum_{k=0}^{\infty} c_k^2\mathbb{E}\left(\varepsilon_{t-k}^2\right) + \sum_{k_1,k_2=0,\ k_1<k_2} c_{k_1}c_{k_2}\mathbb{E}\left(\varepsilon_{t-k_1}\varepsilon_{t-k_2}\right), && (6.20)\end{aligned}$$

where the coefficients $(c_k)_{k\ge 0}$ are determined by equation (2.13). By representation (2.2), we have $(\varepsilon_t)_t$ is a measurable function of i.i.d. random variables Z_t's with mean zero and variance 1. Hence, the second term of equation (6.20) is equal to zero. Thus, we have:

$$\mathbb{E}\left(X_t^2\right) = \sum_{k=0}^{\infty} c_k^2\mathbb{E}\left(\varepsilon_{t-k}^2\right) = \left(\sum_{k=0}^{\infty} c_k^2\right)^2 \mathbb{E}\left(\varepsilon_t^2\right). \tag{6.21}$$

Under the conditions (2.9), we have $\sum c_k^2 < \infty$, and we know that the right side of equation (6.21) is finite, and so $(X_t)_t$ is second-order stationary. If $(\varepsilon_t)_t$ is a function of i.i.d random variables, hence so is $(X_t)_t$. Consequently, $(X_t)_t$ is strictly stationary and ergodic.

b) The proof is similar to a) except that the conditions are required on the convergence of

$$\Psi_2(z) = \frac{\Phi(z^s)\phi(z)}{\Theta(z^s)\theta(z)}(1-z)^{d}(1-z^s)^{D}, \text{ for all } |z| \le 1.$$

A similar argument as in a) can be used to show that the power series expansion of $\Psi_2(z)$ is convergent for all $|z| \le 1$, and so $(X_t)_t$ is invertible and we know that (2.11) holds. ∎

Proof of Theorem 3.1

The proof of Theorem 3.1 uses the following lemma.

Lemma 6.1. Suppose that $(X_t)_{t\in\mathbb{Z}}$ are generated by equations (2.1), (2.2) and (2.3), and the assumptions of Theorem 3.1 hold. Then

$$-\frac{1}{n}\sum_{t=1}^{n}\begin{pmatrix} \dfrac{\partial^2 l_t}{\partial\gamma\partial\gamma^T} & \dfrac{\partial^2 l_t}{\partial\delta\partial\gamma^T} \\ \dfrac{\partial^2 l_t}{\partial\gamma\partial\delta^T} & \dfrac{\partial^2 l_t}{\partial\delta\partial\delta^T} \end{pmatrix} \xrightarrow{\text{a.s}} \begin{pmatrix} \Omega_\gamma & 0 \\ 0 & \Omega_\delta \end{pmatrix}, \quad \text{as } n \longrightarrow \infty,$$

where Ω_γ and Ω_δ are positive matrices, with

$$\Omega_\gamma = \mathbb{E}\left[\frac{1}{h_t^2}\frac{\partial\varepsilon_t}{\partial\gamma}\frac{\partial\varepsilon_t}{\partial\gamma^T} + \frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial\gamma}\frac{\partial h_t^2}{\partial\gamma^T}\right] \qquad \text{and} \qquad \Omega_\delta = \mathbb{E}\left[\frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial\delta}\frac{\partial h_t^2}{\partial\delta^T}\right],$$

where $\xrightarrow{\text{a.s}}$ denotes convergence almost surely.

Proof. The proof of Lemma 6.1 follows mainly the lines of the proof of Theorem 3.1. in Ling and Li [17]. ■

To prove the consistency and asymptotic normality of the CSS estimator, we will check the conditions listed by Basawa et al.[7]. They have analyzed the asymptotic properties of ML estimator in process with dependent observations, by given a set of sufficient conditions. Thus to prove condition (a), we need to check the following three conditions:

(1) $-n^{-1}\sum \dfrac{\partial l_t(\omega_0)}{\partial\omega} \xrightarrow{\mathbb{P}} 0;$

(2) There exists a nonrandom matrix $M(\omega_0) > 0$ such that for all $\varepsilon > 0$,

$$\mathbb{P}\left\{-n^{-1}\sum \frac{\partial^2 l_t(\omega_0)}{\partial\omega\partial\omega^T} \geq M(\omega_0)\right\} > 1-\varepsilon, \quad \text{for all} \quad n \geq n_1(\varepsilon);$$

(3) There exists a constant $M < \infty$ such that

$$\mathbb{E}\left|\frac{\partial^3 l_t(\omega)}{\partial\omega_i\partial\omega_j\partial\omega_k}\right| < M, \quad \text{for all} \quad \omega \in \Lambda,$$

where ω_i is the ith component of ω.

Consider the first-order derivatives of the function l_t with respect to the parameters. Thus, we have for each t:

$$\frac{\partial l_t}{\partial\gamma} = \frac{1}{2h_t^2}\left(\frac{\varepsilon_t^2}{h_t^2}-1\right)\frac{\partial h_t^2}{\partial\gamma} - \frac{\varepsilon_t}{h_t^2}\frac{\partial\varepsilon_t}{\partial\gamma} \quad \text{and} \quad \frac{\partial l_t}{\partial\delta} = \frac{1}{2h_t^2}\left(\frac{\varepsilon_t^2}{h_t^2}-1\right)\frac{\partial h_t^2}{\partial\delta}. \tag{6.22}$$

These derivatives involve that of ε_t and h_t^2, which we now specify as follows:

$$\frac{\partial\varepsilon_t}{\partial d} = -\sum_{k=1}^{\infty}\frac{1}{k}\,\varepsilon_{t-k}; \qquad \frac{\partial\varepsilon_t}{\partial D} = -\sum_{k=1}^{\infty}\frac{1}{k}\,\varepsilon_{t-ks};$$

$$\frac{\partial\varepsilon_t}{\partial\phi_j} = -\phi^{-1}(B)\varepsilon_{t-j},\ j=1,\ldots,p; \qquad \frac{\partial\varepsilon_t}{\partial\theta_j} = -\theta^{-1}(B)\varepsilon_{t-j},\ j=1,\ldots,q;$$

$$\frac{\partial\varepsilon_t}{\partial\Phi_j} = -\Phi^{-1}(B^s)\varepsilon_{t-js},\ j=1,\ldots,P; \qquad \frac{\partial\varepsilon_t}{\partial\Theta_j} = -\Theta^{-1}(B^s)\varepsilon_{t-js},\ j=1,\ldots,Q;$$

$$\frac{\partial h_t^2}{\partial\gamma} = 2\sum_{i=1}^{r} a_i\varepsilon_{t-i}\frac{\partial\varepsilon_{t-i}}{\partial\gamma} + \sum_{j=1}^{m} b_j\frac{\partial h_{t-j}^2}{\partial\gamma} + \sum_{k=1}^{r^*} c_k\varepsilon_{t-k}\frac{\partial h_{t-k}}{\partial\gamma} + \sum_{k=1}^{r^*} c_k h_{t-k}\frac{\partial\varepsilon_{t-k}}{\partial\gamma},$$

and similarly, we have

$$\frac{\partial h_t^2}{\partial\delta} = \tilde{\varepsilon}_t + \sum_{j=1}^{m} b_j\frac{\partial h_{t-j}^2}{\partial\delta} + \sum_{k=1}^{r^*} c_k\varepsilon_{t-k}\frac{\partial h_{t-k}}{\partial\delta},$$

where $\tilde{\varepsilon}_t = (1,\ \varepsilon_{t-1}^2,\ldots,\varepsilon_{t-r}^2,\ h_{t-1}^2,\ldots,h_{t-m}^2,\ \varepsilon_{t-1}^2 h_{t-1}^2,\ldots,\varepsilon_{t-r^*}^2 h_{t-r^*}^2)^T$.

Using the expression in equation (6.22), we have $\mathbb{E}\left[\frac{\partial l_t}{\partial\omega}\right]_{\omega=\omega_0} = 0$. According the assumptions of Theorem 2.2 and Corollary 2.1, the process $(X_t)_{t\in\mathbb{Z}}$ and $(\varepsilon_t)_{t\in\mathbb{Z}}$ are strictly stationary and ergodic. Hence, applying the ergodic theorem, we obtain condition (1). To check condition (2), we need to compute the derivatives of order two as follows:

$$\frac{\partial^2 l_t}{\partial\gamma\partial\gamma^T} = -\frac{1}{h_t^2}\frac{\partial\varepsilon_t}{\partial\gamma}\frac{\partial\varepsilon_t}{\partial\gamma^T} - \frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial\gamma}\frac{\partial h_t^2}{\partial\gamma^T}\left(\frac{\varepsilon_t^2}{h_t^2}\right) + \left(\frac{\varepsilon_t^2}{h_t^2}-1\right)\frac{\partial}{\partial\gamma}\left[\frac{1}{2h_t^2}\frac{\partial h_t^2}{\partial\gamma^T}\right] - \frac{2\varepsilon_t}{h_t^2}\frac{\partial\varepsilon_t}{\partial\gamma}\frac{\partial h_t^2}{\partial\gamma^T} + \frac{\varepsilon_t}{h_t^2}\frac{\partial^2\varepsilon_t}{\partial\gamma\partial\gamma^T},$$

and

$$\frac{\partial^2 l_t}{\partial\delta\partial\delta^T} = -\frac{1}{2h_t^4}\frac{\partial h_t^2}{\partial\delta}\frac{\partial h_t^2}{\partial\delta^T}\left(\frac{\varepsilon_t^2}{h_t^2}\right) + \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t^2}-1\right)\frac{\partial}{\partial\delta}\left[\frac{1}{h_t^2}\frac{\partial h_t^2}{\partial\delta^T}\right].$$

Similarly, we find $(\partial^2 l_t)/(\partial\gamma\partial\delta^T)$ and $(\partial^2 l_t)/(\partial\delta\partial\gamma^T)$.
According to Lemma 6.1, the matrix Ω_0 is positive definite, and hence for any constant vector $C \neq 0$,

$$\frac{1}{n}\sum C^T\left(\frac{\partial^2\ l_t(\omega_0)}{\partial\omega\ \partial\omega^T}\right)C \xrightarrow{\text{a.s}} -C^T\Omega_0 C.$$

For any given C, let $0 < \Delta(C) < C^T\Omega_0 C/2$. Then

$$\forall\ \varepsilon > 0,\ \lim_{n\to+\infty}\mathbb{P}\left(\left|n^{-1}\sum C^T\left(\frac{\partial^2\ l_t(\omega_0)}{\partial\omega\ \partial\omega^T}\right)C + C^T\Omega_0 C\right| > \varepsilon\right) = 0$$

Particularly, we have for $\varepsilon = \Delta$:

$$\lim_{n\to+\infty}\mathbb{P}\left(\left|n^{-1}\sum C^T\left(\frac{\partial^2\ l_t(\omega_0)}{\partial\omega\ \partial\omega^T}\right)C + C^T\Omega_0 C\right| \leq \Delta\right) = 1.$$

In other words, we have $\forall\ \varepsilon > 0$, there exists $n_1 = n_1(\varepsilon)$ such that $\forall\ n \geq n_1$:

$$\left|\ \mathbb{P}\left(\left|n^{-1}\sum C^T\left(\frac{\partial^2\ l_t(\omega_0)}{\partial\omega\ \partial\omega^T}\right)C + C^T\Omega_0 C\right| \leq \Delta\right) - 1\right| < \varepsilon$$

$$\Longrightarrow\ \mathbb{P}\left(n^{-1}\sum C^T\left(\frac{\partial^2\ l_t(\omega_0)}{\partial\omega\ \partial\omega^T}\right)C + C^T\Omega_0 C \leq \Delta\right) > 1-\varepsilon$$

$$\Longrightarrow \mathbb{P}\left(n^{-1}\sum C^T\left(\frac{\partial^2\, l_t(\omega_0)}{\partial\omega\,\partial\omega^T}\right)C + C^T\Omega_0 C \le \frac{C^T\Omega_0\, C}{2}\right) > 1-\varepsilon$$

$$\Longrightarrow \mathbb{P}\left(n^{-1}\sum C^T\left(\frac{\partial^2\, l_t(\omega_0)}{\partial\omega\,\partial\omega^T}\right)C \le -\frac{C^T\Omega_0\, C}{2}\right) > 1-\varepsilon$$

$$\Longrightarrow \mathbb{P}\left(-n^{-1}\sum C^T\left(\frac{\partial^2\, l_t(\omega_0)}{\partial\omega\,\partial\omega^T}\right)C \ge \frac{C^T\Omega_0\, C}{2}\right) > 1-\varepsilon$$

So the condition (2) is satisfied. To prove condition (3), we remark that:

$$\frac{\partial^3\varepsilon_t}{\partial d^3} = \ln^3(1-B)\varepsilon_t = \left(-\sum_{k=1}^{\infty}\frac{1}{k}B^k\right)^3\varepsilon_t.$$

$$\begin{aligned}
\Longrightarrow \mathbb{E}\left(\frac{\partial^3\varepsilon_t}{\partial d^3}\right)^2 &= \mathbb{E}\left(\sum_{k_1,k_2,k_3=1}\frac{1}{k_1k_2k_3}\,\varepsilon_{t-k_1-k_2-k_3}\right)^2\\
&= \sum_{k_1,k_2,k_3=1}\frac{1}{k_1^2\,k_2^2\,k_3^2}\,\mathbb{E}\left(\varepsilon^2_{t-k_1-k_2-k_3}\right)\\
&= \left(\frac{\pi^2}{6}\right)^3\mathbb{E}\left(\varepsilon_t^2\right) < \infty
\end{aligned}$$

Similarly, we can show that $\mathbb{E}\left(\frac{\partial^3\varepsilon_t}{\partial\omega_i^3}\right)^2 < \infty$, where ω_i is the ith component of ω. Thus, differentiating $\partial^2 l_t/\partial\omega\partial\omega^T$ and using the preceding discussion, we obtain condition (3).

Finally, using the first-order Taylor expansion of the expression $\frac{\partial\, L(\omega)}{\partial\, \omega}$ at $\omega = \omega_0$, and the conditions (1), (2) and (3), we obtain the existence and the consistency of the CSS estimator $\hat{\omega}_n$.

To prove condition (b) of Theorem 3.1, we remark at first, by Lemma 6.1, we have:

$$\frac{1}{n}\sum\left(\frac{\partial l_t}{\partial\omega}\frac{\partial l_t}{\partial\omega^T}\right)_{\omega=\omega_0} \xrightarrow{a.s} \Omega_0. \tag{6.23}$$

Let us define S_n as

$$S_n = \sum_{t=1}^{n}\eta_0^T\left(\frac{\partial l_t}{\partial\omega}\right)_{\omega=\omega_0}$$

where η_0 is an arbitrary constant vector with $\eta_0^T\eta \neq 0$. We have

$$\mathbb{E}\left(\eta_0^T\frac{\partial l_t(\omega_0)}{\partial\omega}\right) = \eta_0^T\mathbb{E}\left(\frac{\partial l_t(\omega_0)}{\partial\omega}\right) = 0.$$

Then $\left(\eta_0^T\dfrac{\partial l_t(\omega_0)}{\partial\omega}\right)_t$ are a sequence of martingale differences and S_n is the sum of martingale difference. Hence S_n is a martingale and we have $\frac{1}{n}\mathbb{E}(S_n^2) = \eta_0^T\Omega_0\eta > 0$. From the stationarity and the ergodicity of (X_t) and (ε_t), we have:

$$\left[\frac{1}{n}\mathbb{E}\left(S_n^2\right)\right]^{-1}\left[\frac{1}{n}\mathbb{E}\left(S_n^2 \mid I_n\right)\right] \xrightarrow{a.s} 1. \tag{6.24}$$

Using the central limit theorem of Stout [25], we get:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\frac{\partial l_t}{\partial \omega} \right)_{\omega=\omega_0} \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ \Omega_0), \quad \text{as } n \longrightarrow +\infty. \tag{6.25}$$

Using the first-order Taylor expansion of the expression $\frac{\partial\, L(\omega)}{\partial\, \omega}$ at $\omega = \omega_0$, and the conditions (6.23), (6.24) and (6.25), we have:

$$\sqrt{n}\,(\hat{\omega}_n - \omega_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\ \Omega_0^{-1}), \quad \text{as } n \longrightarrow +\infty.$$

The condition (c) of Theorem 3.1 is directly obtained from Lemma 6.1. ■