ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.3, 887–900
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.3.14

ON THE BILINEAR TIME SERIES MODELS PROVIDED BY GARCH WHITE NOISE: ESTIMATION AND SIMULATION

NABIL LAICHE AND HALIM ZEGHDOUDI[1]

ABSTRACT. This work proposes the estimation of a sample of bilinear time series models mixed by a GARCH white noise, where GARCH model was followed by time varying coefficients. The study allows demonstrating some properties and remarks depending on the behavior of the estimators. Moreover, this work will be validated by a simulations study and digital illustrations using the Matlab software.

1. INTRODUCTION

The applications of nonlinear time series models have undergone several developments over the past thirty years, and is not about economics and finance but also other fields like biology, medicine, chemistry and metrology, etc. Among these models we will suffice here to mention only two because they will be used in this paper. The first is bilinear time series models. These models have appeared in the statistical literature because their applications of modelling of microeconomics and finance are unlimited. Bilinear models were created by Granger and Andersen 1978 see [6], and Subba Rao developed some theory of these models in (1981), and the works of Subba Rao and Gabr (1981) give some general discussions on stationary methods of estimation see [9]. Tong (1981) studied the ergodicity of special bilinear models

[1] *corresponding author*

2010 *Mathematics Subject Classification.* 62M10, 37M10.
Key words and phrases. Bilinear models, GARCH models, the least squares approach, MLE, time varying coefficients.

see [10]. Quinn (1982) and Hannan (1982) discuss properties of some cases of these models. In addition, Bibi and Oyet have investigated this model with time varying coefficients in many works see references [1], [2].

The second GARCH models were introduced by Bollerslev in 1968. These models are extension for ARCH models of Engle (1982). GARCH models have been widely use in financial time series and particularly in analyzing the risks and their applications in forecasting are not limited see references [4], [5].

The paper will be organized as follows, in Section 2 we will give some preliminaries for the bilinear models followed by some theory especially, where we amuse the value of Nilsen-Klimko theorem see [7], and we will present the necessary stability conditions for the proposed model. The main goal of this study is in Section 3 where we will estimate the coefficients of bilinear model, with GARCH white noise driven by time varying coefficients with MLE approach. We will finalize the work in Section 4, with numerical illustration and simulations inlaid with graphs for using the Matlab tool, where this simulation allows deducing some remarks and building comments.

2. Definitions and properties

We begin this section with some definitions and preliminaries identified by a little theoretical study.

Definition 2.1. *We define bilinear time series model of order in a space* (Ω, F, P) *by a following stochastic equation*

$$x_t = \sum_{i=1}^{p} a_i x_{t-i} + \sum_{j=1}^{q} c_j \varepsilon_{t-j} + \sum_{i=1}^{r} \sum_{j=1}^{s} b_{ij} x_{t-i} \varepsilon_{t-j} + \varepsilon_t \,.$$

We denote $BL(p, q, r, s)$, the sequences $\{a_i, 1 \leq i \leq p\}$, $\{c_j, 1 \leq j \leq q\}$.
Let $\{b_{ij}, 1 \leq i \leq r, 1 \leq j \leq s\}$ be the constant coefficients of the model. Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ is white noise part not necessarily identically distributed in the general cases, habitually takes mean zero and variance σ_t^2 . There are situations where we see that white noise is written in the form $\varepsilon_t = \eta_t \sigma_t$, such as the quantities σ_t are strictly positive numbers, and η_t are independent and identically distributed random variables. To illustrate this expression we can give an example of the ARCH and GARCH models which we will present in this paper, where we will

suggest here in our paper that white noise will follow GARCH model, so we have the following definition.

Definition 2.2. *The process* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *is called generalized autoregressive conditionally heteroscedasticity with time varying coefficients of order* (p_0, q_0), *and denoted by* $GARCH(p_0, q_0)$ *each model defined by the following stochastic equation*

$$\varepsilon_t = z_t h_t \,,$$

where

$$h_t^2 = \gamma_0 + \sum_{j=1}^{q_0} \beta_{j,t}(\beta)\varepsilon_{t-j} + \sum_{i=1}^{p_0} \alpha_{i,t}(\alpha)h_{t-i} \,.$$

Here z_t are independent identically distributed random variables with zero mean and variance equals 1. The parties $\{\beta_{j,t}(\beta), 1 \le j \le q_0\}$ and $\{\alpha_{i,t}(\alpha), 1 \le i \le p_0\}$ are time varying coefficients take their values in $\mathbb{R}^+$, where consider in this paper that γ_0 is constant, and α and β are vectors of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively in general, but in our study here, we will shorten on space $\mathbb{R}^2$ where $\alpha{=}(\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$. In this paper, we will project the study on a special case of bilinear time series models oriented by the GARCH white noise with time varying coefficients. And let F_t the $\sigma-$field generated by the set of N observations $\{x_t, t = 1, ..., N\}$, so the proposed bilinear model that we are going to estimate its coefficients

$$x_t = ax_{t-s} + bx_{t-s}\varepsilon_{t-1} + \varepsilon_t \tag{2.1}$$

The order of this model is $BL(s, 0, s, 1)$ where $s \ge 1$, and its white noise follows a $GARCH(1, 1)$ model defined with its equation

$$\begin{cases} \varepsilon_t = \eta_t h_t \\ h_t^2 = \gamma_0 + \beta_t(\beta)\varepsilon_{t-1}^2 + \alpha_t(\alpha)h_{t-1}^2 \,. \end{cases} \tag{2.2}$$

If we consider $E(\varepsilon_t^2) = \sigma_t^2$ then, the necessary condition of stability for the model (2.1) will be $a^2{+}\sigma_t^2 b^2 < 1$, see [2], and it is quite clear that the condition of stability is equivalent to the condition $|a| + |b|\, \sigma_t < 1$, we can write the model under following recurrent expression

$$x_t = \sum_{j=1}^{\left[\frac{t}{s}\right]} \left\{ \prod_{i=0}^{j-1} (a + b\varepsilon_{t-si-1}) \right\} \varepsilon_{t-sj} + \varepsilon_t \,. \tag{2.3}$$

We denote by $\left[\frac{t}{s}\right]$ for the integer value of $\frac{t}{s}$, so overall the parameter that we will estimate $\theta = (a, b, \gamma_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$, and we assume that the white noise follows the law $N(0, E(\varepsilon_t^2))$. First to build the principle of the estimation approach we have

$$E(\varepsilon_t^2) = E(\eta_t^2)E(h_t^2) = E(h_t^2)\,.$$

Because previously we have $E(\eta_t^2) = 0$, and in another way we find that

$$\begin{aligned} E(\varepsilon_t^2) &= E\left\{\gamma_0 + \beta_t(\beta)\varepsilon_{t-1}^2 + \alpha_t(\alpha)h_{t-1}^2\right\} \\ &= \gamma_0 + \beta_t(\beta)E(\varepsilon_{t-1}^2) + \alpha_t(\alpha)E(h_{t-1}^2)\,. \end{aligned}$$

As $E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2)$, we get the following expression

$$E(\varepsilon_t^2) = \frac{\gamma_0}{1 - \beta_t(\beta) - \alpha_t(\alpha)}\,.$$

In this situation the condition of stability of the model will be:

$$|a| + \frac{|b|\,\gamma_0}{1 - \beta_t(\beta) - \alpha_t(\alpha)} < 1\,,$$

where $\beta_t(\beta) + \alpha_t(\alpha) < 1$ and according to the following theorem which ensures the stability of white noise GARCH (1,1) .

Theorem 2.1. *"Strict stationarity of the GARCH(1,1) process"*

The condition $C = E\log\{\alpha_t(\alpha)h_t^2 + \beta_t(\beta)\} \in [-\infty, 0[$ *ensures the existence of the unique and stationary solution for GARCH (1,1) model (the condition* C *equivalent to condition* $-\infty \leq E\log\{\alpha_t(\alpha)h_t^2 + \beta_t(\beta)\} < 0$*).*

Proof. See [3]. □

Proposition 2.1. *The model generated by its expression* (2.1) *is convergent through its stability condition.*

Proof. It suffices to demonstrate that $\rho = E\{|x_t|\} < \infty$, and also it suffices to prove the proposition to take case $s = 1$. With the application of Schwartz inequality we will find that:

$$\begin{aligned} \rho &= E\left|\sum_{j=1}^{\infty}\left\{\prod_{i=0}^{j-1}(a + b\varepsilon_{t-i-1})\right\}\varepsilon_{t-j} + \varepsilon_t\right| \\ &\leq \sum_{j=1}^{\infty}\left\{E(\varepsilon_{t-j}^2)\right\}^{0.5}\prod_{i=0}^{j-1}\left\{E(a + b\varepsilon_{t-i-1})^2\right\}^{0.5}\,. \end{aligned}$$

We give here a little evidence

$$\begin{aligned} E(a+b\varepsilon_{t-i-1})^2 &= E(a^2+b^2\varepsilon_{t-i-1}^2+2ab\varepsilon_{t-i-1}) \\ &= a^2+b^2E(\varepsilon_{t-i-1}^2) \end{aligned}$$

and according to the stability condition $a^2+b^2E(\varepsilon_{t-i-1}^2) \leq \delta < 1$ and we set $M=\max E(\varepsilon_{t-j}^2)$ then

$$\rho \leq M\sum_{j=1}^{\infty}\prod_{i=0}^{j-1}\delta = M\sum_{j=1}^{\infty}\delta^j \leq \frac{M\delta}{1-\delta},$$

which finalize the proof. □

Among the best estimation approaches, the least square method (LS), this method is an explosion for knowledge and its applications are not finished until today, but its realization in bilinear time series subset with GARCH models is very few in the statistical references but with ARCH we can see Weiss (1986) and Pantulla (1988), and the estimation of the GARCH with time varying coefficients almost does not exist.

Definition 2.3. *We define the predictor of a time series and denoted by $x_{t|t-1}$, it represents the orthogonal projection of x_t on the observation up the time $t-1$ the difference*

$$x_{t|t-1} = x_t - \varepsilon_t\,.$$

Definition 2.4. *We say $\hat{\theta}$ an estimator for θ if and only if $\hat{\theta}$ is solution of*

$$\arg\min_{\theta\in\Omega} g_N(\theta)\,.$$

In this paper Ω is an open set included in $\mathbb{R}^7$, and $g_N(\theta)$ the penalty function defined by expression

$$g_N(\theta) = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_t^2(\theta)\,.$$

Least squares method based on the Taylor's formula, where we consider that θ^0 is the true value of θ

$$\begin{aligned} g_N(\theta) &= g_N(\theta^0) + (\theta-\theta^0)\frac{\partial g_N(\theta)}{\partial\theta^T} + (\theta-\theta^0)\frac{\partial^2 g_N(\theta)}{\partial\theta^T\partial\theta}(\theta-\theta^0)^T \\ &\quad + (\theta-\theta^0)\frac{\partial^2 g_N(\tilde{\theta})}{\partial\theta^T\partial\theta}(\theta-\theta^0)^T\,. \end{aligned}$$

T represents transpose of matrix, $\tilde{\theta}$ is an intermediate point between θ and θ^0. In sense of norm we have $\|\theta - \theta^0\| < \nu$, $\nu > 0$, we recall the famous theorem of Klimko and Nilsen which shows the existence of estimators and their asymptotic behaviors, we assume that $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7) = (a, b, \gamma_0, \alpha_1, \alpha_2, \beta_1, \beta_2)$.

Theorem 2.2. *Let be $\{x_t, t \in \mathbb{Z}\}$ a stable process generated by expression of model (2.1), such as $x_{t|t-1}$ almost surely twice continuously differentiable in an open subset $\Omega \subset \mathbb{R}^7$, which contains the true value θ^0 of θ, and as we can say derivatives of order 1 and 2 for $x_{t|t-1}$ are bounded. C_1 and C_2 are two positive constants, so if the following assumptions are verified for all $i, j \in \{1, ..., 7\}$*

H-1: $E_{\theta^0} \left\{ \frac{\partial \varepsilon_t^2}{\partial \theta_i} \right\}^4 \leq C_1, \forall i \in \{1, ..., 7\}$.

H-2: $E_{\theta^0} \left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta_i \partial \theta_j} - E_{\theta^0} \left(\frac{\partial \varepsilon_t^2(\theta)}{\partial \theta_i \partial \theta_j} \mid F_{t-1} \right) \right\}^2 \leq C_2, \forall i, j \in \{1, ..., 7\}$.

H-3: $\frac{1}{2N} \sum_{t=1}^{N} E_{\theta^0} \left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta_i \partial \theta_j} \mid F_{t-1} \right\}$ converges as surely for matrix $O_{ij}(\theta^0)$ of size 7×7, where O_{ij} is strictly positive matrix of constants.

H-4: $\lim_{N \to \infty} \sup_{\omega \to 0} \frac{1}{N\omega} \left| \sum_{t=1}^{N} \left[\left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta_i \partial \theta_j} \right\}_{\theta = \tilde{\theta}} - \left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta_i \partial \theta_j} \right\}_{\theta = \theta^0} \right] \right| < \infty$.

Then, there exists an estimator $\hat{\theta} = (\hat{\theta}_{1,N}, \hat{\theta}_{2,N}, ..., \hat{\theta}_{7,N})$ such as $\hat{\theta} \to \theta^0$ as $N \to \infty$, if the four hypotheses are verified it results in the following hypothesis

H-5: $\frac{1}{N} \sum_{t=1}^{N} \left[E_{\theta^0} \left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta} \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta^T} \mid F_{t-1} \right\} - E_{\theta^0} \left\{ \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta} \frac{\partial \varepsilon_t^2(\theta)}{\partial \theta^T} \mid F_{t-1} \right\} \right] \to 0$ as $N \to \infty$, where $\sqrt{N}(\hat{\theta} - \theta^0)$ follows Gaussian law.

Proof. See [7]. □

Bibi and Oyet in reference [2] found that the white noise of the model (2.1) can be written recurrently as the form

$$\varepsilon_t = x_t - a x_{t-s} + \sum_{i=0}^{t} (-1)^j \left\{ \prod_{i=0}^{j-1} b \right\} \left\{ \prod_{i=0}^{j-1} x_{t-i-s} \right\} (x_{t-j} - a x_{t-j-s}).$$

3. Maximum likelihood estimation

The approach of the maximum likelihood method (MLE) has been applied in certain references where the model will be GARCH but in a situation where their coefficients change over time almost there does not exist. The concept of time varying coefficients GARCH takes several dimensions in applications because there are physical phenomena rebelling on the laws of classical physics, where we note that the coefficients are not constants but take a time varying situation. In finance several models take the situation of variation of the coefficients over time, for example the most well-known models in the form of time-varying coefficients there are models with alternative coefficients. We will estimate the coefficients using MLE for the following model: we assume the observations $\{x_t, t = 1, ..., N\}$. Starting point is the specification of the conditional density of the residuals ε_t, where first we will have the normal distribution assumed. Let we assume the following conditional density form

$$\Phi_\theta(\varepsilon_t \mid F_{t-1}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi y_t(\theta)}} \exp\left\{-\frac{\varepsilon_t^2(\theta)}{y_t(\theta)}\right\} ,$$

and ε_t is GARCH (1,1) white noise defined by its equation (2.2). We assume here that the coefficients take an alternative situation where

$$\beta_t(\beta_1, \beta_2) = \frac{1-(-1)^t}{2}\beta_1 + \frac{1+(-1)^t}{2}\beta_2 , \quad \alpha_t(\alpha_1, \alpha_2) = \frac{1-(-1)^t}{2}\alpha_1 + \frac{1+(-1)^t}{2}\alpha_2 ,$$

and $t \geq 1$. The construction of $y_t(\theta)$ is given when we set $h_{t-1} = h_t$, where this construction makes it possible to estimate the coefficients of model (2.1). So when we put $y_t = h_t^2$ we get:

$$y_t = \frac{\gamma_0 + \beta_t(\beta)\varepsilon_{t-1}^2}{1 - \alpha_t(\alpha)} .$$

For estimating the parameters of the model, we want to get a solution θ which maximize the logarithm likelihood function

$$G(\theta) = \ln\left\{\Phi_\theta(\varepsilon_t \mid F_{t-1})\right\} .$$

Then

$$G(\theta) = \sum_{t=1}^{N} \left\{-\frac{1}{2}\ln(y_t(\theta) - \frac{\varepsilon_t^2(\theta)}{2y_t(\theta)}\right\} - N\ln 2\pi .$$

To illustrate derivation techniques we have

$$\Psi_t(\theta) = -\frac{1}{2}\ln(y_t(\theta) - \frac{\varepsilon_t^2(\theta)}{2y_t(\theta)} .$$

From this situation, we will extract the algorithm that allows calculating the partial derivatives with the coordinates of θ, and we will give the derivatives 1 and 2 of $G(\theta)$ with respect to θ, then:

$$\begin{aligned}\frac{\partial \Psi_t(\theta)}{\partial \theta_1} &= \frac{\partial \Psi_t(\theta)}{\partial a} = \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2(\theta)}{2y_t^2(\theta)}\right\} \frac{\partial y_t(\theta)}{\partial a} - \frac{\partial \varepsilon_t(\theta)}{\partial a}\frac{\varepsilon_t(\theta)}{y_t(\theta)} \\ \frac{\partial \Psi_t(\theta)}{\partial \theta_2} &= \frac{\partial \Psi_t(\theta)}{\partial b} = \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2(\theta)}{2y_t^2(\theta)}\right\} \frac{\partial y_t(\theta)}{\partial b} - \frac{\partial \varepsilon_t(\theta)}{\partial b}\frac{\varepsilon_t(\theta)}{y_t(\theta)}. \\ \frac{\partial \Psi_t(\theta)}{\partial \theta_3} &= \frac{\partial \Psi_t(\theta)}{\partial \gamma_0} = \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2(\theta)}{2y_t^2(\theta)}\right\} \frac{\partial y_t(\theta)}{\partial \gamma_0}\end{aligned}$$

where $\frac{\partial y_t(\theta)}{\partial \gamma_0} = \frac{1}{1-\alpha_t(\alpha)}$.

For parameters θ_4, θ_5 which means $\alpha_i, i = 1, 2$.

$$\frac{\partial \Psi_t(\theta)}{\partial \alpha_i} = \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2(\theta)}{2y_t^2(\theta)}\right\} \frac{\partial y_t(\theta)}{\partial \alpha_i},$$

such as $\frac{\partial y_t(\theta)}{\partial \alpha_i} = \frac{\gamma_0 + \beta_t(\beta)\varepsilon_{t-1}^2(\theta)}{(1-\alpha_t(\alpha))^2}$.

For parameters θ_6, θ_7 which means $\beta_i, i = 1, 2$.

$$\frac{\partial \Psi_t(\theta)}{\partial \beta_i} = \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2(\theta)}{2y_t^2(\theta)}\right\} \frac{\partial y_t(\theta)}{\partial \beta_i},$$

such as $\frac{\partial y_t(\theta)}{\partial \beta_i} = \frac{\varepsilon_{t-1}^2(\theta)}{1-\alpha_t(\alpha)}$.

For indication using the formula (2.3) we can calculate the two derivatives $\frac{\partial \Psi_t(\theta)}{\partial a}, \frac{\partial \Psi_t(\theta)}{\partial b}$ where

$$\frac{\partial \varepsilon_t(\theta)}{\partial a} = -x_{t-s} - \sum_{i=0}^{t} (-1)^j b^j \left\{\prod_{i=0}^{j-1} x_{t-i-s}\right\} (x_{t-j-s}),$$

and

$$\frac{\partial \varepsilon_t(\theta)}{\partial b} = \sum_{i=0}^{t} (-1)^j (j b^{j-1}) \left\{\prod_{i=0}^{j-1} x_{t-i-s}\right\} (x_{t-j} - a x_{t-j-s}).$$

Now we determine the second derivatives, so we have:

$$\begin{aligned}\frac{\partial^2 \Psi_t(\theta)}{\partial a \partial a} &= \left\{\frac{1}{2y_t^2} - \frac{\varepsilon_t^2}{y_t^3}\right\} \left(\frac{\partial y_t}{\partial a}\right)^2 + \left\{-\frac{1}{2y_t} + \frac{\varepsilon_t^2}{2y_t^2}\right\} \frac{\partial^2 y_t}{\partial a \partial a} \\ &\quad + \frac{\partial^2 \varepsilon_t}{\partial a \partial a}\frac{\varepsilon_t}{y_t} + (\frac{\varepsilon_t^2}{y_t^2}\frac{\partial y_t}{\partial a} + \frac{1}{y_t}\frac{\partial \varepsilon_t}{\partial a} - \frac{\varepsilon_t}{y_t}\frac{\partial y_t}{\partial a})\frac{\partial \varepsilon_t}{\partial a},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2\Psi_t(\theta)}{\partial b\partial b} &= \left\{\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right\}\left(\frac{\partial y_t}{\partial b}\right)^2+\left\{-\frac{1}{2y_t}+\frac{\varepsilon_t^2}{2y_t^2}\right\}\frac{\partial^2 y_t}{\partial b\partial b}\\ &+\frac{\partial^2\varepsilon_t}{\partial b\partial b}\frac{\varepsilon_t}{y_t}+(\frac{\varepsilon_t^2}{y_t^2}\frac{\partial y_t}{\partial b}+\frac{1}{y_t}\frac{\partial\varepsilon_t}{\partial b}-\frac{\varepsilon_t}{y_t}\frac{\partial y_t}{\partial b})\frac{\partial\varepsilon_t}{\partial b}\,.\end{aligned}$$

We can generalize for all $i=1,2$.

$$\begin{aligned}\frac{\partial^2\Psi_t(\theta)}{\partial\theta_i\partial\theta_i} &= \left(\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right)\left(\frac{\partial y_t}{\partial\theta_i}\right)^2+\left(-\frac{1}{2y_t}+\frac{\varepsilon_t^2}{2y_t^2}\right)\frac{\partial^2 y_t}{\partial\theta_i\partial\theta_i}\\ &+ \frac{\partial^2\varepsilon_t}{\partial\theta_i\partial\theta_i}\frac{\varepsilon_t}{y_t}+(\frac{\varepsilon_t^2}{y_t^2}\frac{\partial y_t}{\partial\theta_i}+\frac{1}{y_t}\frac{\partial\varepsilon_t}{\partial\theta_i}-\frac{\varepsilon_t}{y_t}\frac{\partial y_t}{\partial\theta_i})\frac{\partial\varepsilon_t}{\partial\theta_i}\,.\end{aligned}$$

Where from the expression (2,3) we can give the derivatives $\frac{\partial^2\varepsilon_t}{\partial\theta_i\partial\theta_i}$.

$$\frac{\partial^2\Psi_t(\theta)}{\partial\gamma_0\partial\gamma_0}=\left(\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right)\left(\frac{\partial y_t}{\partial\gamma_0}\right)^2=\left(\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right)\left(\frac{1}{1-\alpha_t(\alpha)}\right)^2,$$

and for all $i=1,2$.

$$\frac{\partial^2\Psi_t(\theta)}{\partial\beta_i\partial\beta_j}=\frac{\partial^2 y_t}{\partial\beta_i\partial\beta_j}\left(-\frac{1}{2y_t}+\frac{\varepsilon_t^2}{2y_t^2}\right)+\left(\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right)\frac{\partial y_t}{\partial\beta_i}\frac{\partial y_t}{\partial\beta_j},$$

and the same thing for all $i=1,2$.

$$\frac{\partial^2\Psi_t(\theta)}{\partial\alpha_i\partial\alpha_j}=\frac{\partial^2 y_t}{\partial\alpha_i\partial\alpha_j}\left(-\frac{1}{2y_t}+\frac{\varepsilon_t^2}{2y_t^2}\right)+\left(\frac{1}{2y_t^2}-\frac{\varepsilon_t^2}{y_t^3}\right)\frac{\partial y_t}{\partial\alpha_i}\frac{\partial y_t}{\partial\alpha_j}\,.$$

With the same generalization, we give all 49 derivatives for each $i,j\in\{1,2,...,7\}$. Then, it is easy to obtain the partial derivatives matrices:

$$L(\theta)=\left\{\frac{\partial G(\theta)}{\partial\theta_i}\right\}_{i=1,\dots,7}^T,\ O(\theta)=\left\{\frac{\partial^2 G(\theta)}{\partial\theta_i\partial\theta_j}\right\}_{i,j=1,\dots,7}$$

The value of this program can also tell the true value θ^0, then we look for the solution of the following equation where we can find the estimated value $\hat\theta$. The construction of an algorithm is based to give a better approximation for the estimated value $\hat\theta$, through the proposed true value θ^0. This algorithm is known in numerical analysis by the Newton-Raphson approximation method see [8], so we have the function

$$L(\hat\theta)=L(\theta)+O(\theta)(\hat\theta-\theta)=0\,.$$

To finding the approximate solution for the solution, we will ask this recursive expression

$$\hat{\theta} = \theta - O^{-1}(\theta)L(\theta)\,.$$

We apply now Newton-Raphson iterative

$$\begin{aligned}
\theta^1 &= \theta^0 - O^{-1}(\theta^0)L(\theta^0)\\
\theta^2 &= \theta^1 - O^{-1}(\theta^1)L(\theta^1)\\
&\vdots\\
\theta^m &= \theta^{m-1} - O^{-1}(\theta^{m-1})L(\theta^{m-1})\,,
\end{aligned}$$

where the repetition of the iterative ones each time can give a better approximation and then if m tends to infinity then θ^m will converge to the estimated value $\hat{\theta}$.

4. Numerical illustrations

This section is devoted to a simulation of a particular case of a bilinear model (2.1), with time varying GARCH white noise determined by its formula (2.2).

$$\begin{cases} x_t = ax_{t-s} + bx_{t-s}\varepsilon_{t-1} + \varepsilon_t \\ \varepsilon_t = \eta_t h_t,\ h_t^2 = \gamma_0 + \beta_t(\beta)\varepsilon_{t-1}^2 + \alpha_t(\alpha)h_{t-1}^2\,. \end{cases}$$

This model is very successful in financial and economic applications, and has been discussed in detail by gaussian white noise reference [2] using the least squares method (LS), but here the simulation will be according to MLE. Where the model oriented by GARCH white noise with their time varying coefficients in alternative situation as defined above. The tool that we have applied here is Matlab 2013. We will give some model simulations from changing number of simulations (NS) and sample size (N). We have the following notations the true value θ^0, estimated value $\hat{\theta}$, we use here in the simulations some concepts like kurtosis (ku), skewness (sk) and estimated variance (var), the simulation presents some properties and remarks illustrated by the following tables, where each table gives remarks or observations related to the theory of bilinear time series models. We take real value θ^0 =(a,b,γ_0, α_1, α_2, β_1, β_2)=(0.15,0.45, 0.05, 0.01,0.2,0.3,0.4).

TABLE 1

$NS = 250$		$s = 1$
N		$\hat{\theta} = (\hat{a}, \hat{b}, \hat{\gamma}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$
100		$(0.1704, 0.3702, 0.0536, -0.0037, 0.1990, 0.3281, 0.3299)$
250	θ^0	$(0.1755, 0.3634, 0.0625, 0.0010, 0.1978, 0.3225, 0.3430)$
500		$(0.1769, 0.3788, 0.0698, 0.0065, 0.2004, 0.3275, 0.3571)$
1000		$(0.1723, 0.3755, 0.0613, 0.0073, 0.1997, 0.3269, 0.3580)$
2000		$(0.1715, 0.3728, 0.0515, -0.0012, 0.1979, 0.3282, 0.3570)$

TABLE 2

$NS = 500$		$s = 1$
N		$\hat{\theta} = (\hat{a}, \hat{b}, \hat{\gamma}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$
100		$(0.1521, 0.3509, 0.0488, -0.0046, 0.1775, 0.2952, 0.3409)$
250	θ^0	$(0.1675, 0.3631, 0.0532, 0.0011, 0.1841, 0.3202, 0.3522)$
500		$(0.1748, 0.3694, 0.0581, 0.0001, 0.1942, 0.3285, 0.3514)$
1000		$(0.1768, 0.3674, 0.0541, 0.0026, 0.1958, 0.3336, 0.3570)$
2000		$(0.1749, 0.3731, 0.0523, 0.0017, 0.1977, 0.3330, 0.3577)$

TABLE 3

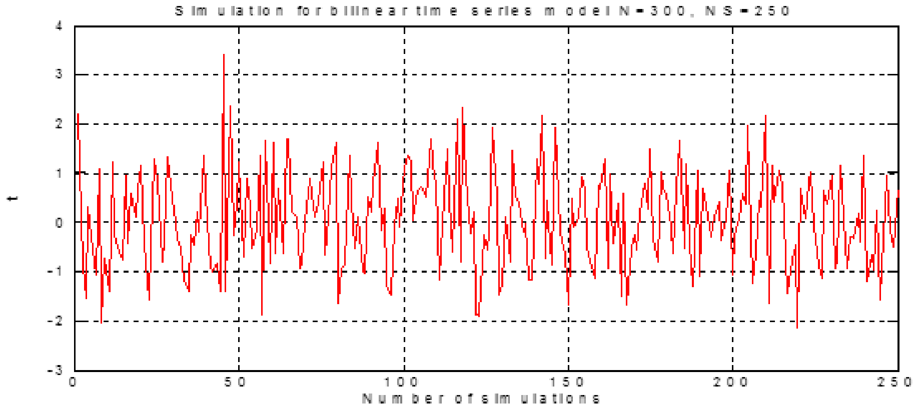
$NS = 1000$		$s = 1$
N		$\hat{\theta} = (\hat{a}, \hat{b}, \hat{\gamma}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$
100		$(0.1737, 0.3664, 0.0418, -0.0087, 0.1821, 0.3131, 0.3363)$
250	θ^0	$(0.1695, 0.3587, 0.0543, -0.0102, 0.1908, 0.3188, 0.3423)$
500		$(0.1755, 0.3669, 0.0571, -0.0035, 0.1949, 0.3277, 0.3512)$
1000		$(0.1746, 0.3739, 0.0530, -0.0021, 0.1962, 0.3290, 0.3564)$
2000		$(0.1725, 0.3763, 0.0512, -0.0010, 0.1986, 0.3293, 0.3584)$

TABLE 4

N	NS	order of the model s	coefficients	$(\hat{a}, \hat{b})$
		$s = 1$		$(0.0197, 0.2966)$
250	250	$s = 2$	(a, b)	$(-0.0064, 0.2401)$
		$s = 3$		$(0.0005, 0.3277)$

TABLE 5

N	NS	(a, b)	$var(\hat{a}, \hat{b})$	$ku\ (\hat{a}, \hat{b})$	$sk\ (\hat{a}, \hat{b})$
120			$(0.0444, 0.0404)$	$(3.4780, 3.1526)$	$(0.2555, 0.5649)$
480	250	$(0.2, 0.58)$	$(0.0133, 0.0139)$	$(3.2573, 3.2430)$	$(0.2979, 0.3894)$
900			$(0.0063, 0.0066)$	$(2.9518, 2.9279)$	$(0.0730, 0.2837)$



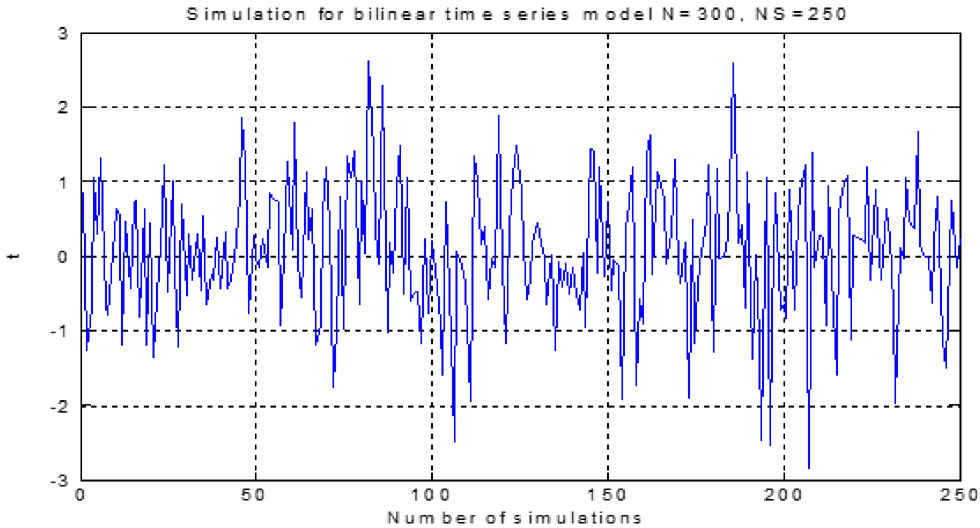


FIGURE 1

The graph red present the model with its true values, and graph blue represent the model replaced by its estimated values, where the true values $a = 0.2$ and $b = 0.45$ with white noise GARCH(1,1).

4.1. **Comments and conclusion.** We are going to simulate the model with the Matlab tool, and observe firstly that the significant criteria of convergence are verified, that is: to say when N tends towards infinity then variance will tend to zero, Kurtosis tends towards 3 and Skewness approximate to zero where these results are illustrated in Table 5. On the other hand, if the sample size increased the estimators approximate towards the true values. Generally, also if the number of the simulations large therefore the estimators give a better approximation towards the true values see tables 1, 2 and 3. Moreover, we also observe among the model orders that the best approximation will be given where $s = 1$, this result is obvious according to the table 4. There are also disturbances in the simulation because the coefficients of GARCH white noise coefficients take alternative situation. We observe that if the coefficients are very small then we will find that the true values and the estimated values are identical, see tables. As a fundamental result the estimators when the model takes a bilinear expression with GARCH white noise, then the estimators will follow the best estimation standards. The asymptotic behavior of estimators according to MLE with GARCH is very effective in the approximation, where there are cases where the true values of their estimators will be almost identical. When we change the model coefficients by their estimators, then the two graphs will almost be identical, they illustrate the asymptotic behavior of the estimators.

REFERENCES

[1] A. BIBI. A. OYET: *A note on the properties of some timevarying bilinear models,* Statistics and probability letters, **12** (2002), 399–411.

[2] A. BIBI, A. J. OYET: *Estimation of some bilinear time series models with time varying coefficients,* Stochastic Analysis and Aplications, **22**(2) (2004), 355–376.

[3] F. CHRISTIAN, J. M. ZAKOIAN: *GARCH models: Structure, statistical inference and financial applications,* Hoboken, NJ: Wiley, 2010.

[4] C. M. HAFNER: *Nonlinear time series analysis with applications to foreign exchange rate volatility*, Springer, 1997.

[5] X. EVDOKIA, D. STAVROS: *ARCH Models for Financial Applications,* John Wiley, 2010.

[6] C. W. J. GRANGER, A. P. ANDERSEN: *Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Göttingen, 1978.

[7] L. A. KLIMKO, P. I. NELSON: *On conditional leastsquares estimation for stochastic processes*, The Annals of Statistics, **3** (1978), 629–642.

[8] J. PAN, G. LI, Z. XIE: *Stationary solution andparametric estimation for bilinear model driven by ARCH models*, Journal of Science in China., **12** (2002), 1523–1537.

[9] T. SUBBA RAO, M. M. GABR: *An Introduction to Bispectral Analysis and Bilinear Time Series Models*, Lecture Notes in Statistics, Springer-Verlag, New York, 1984.

[10] H. TONG: *Threshold models in nonlinear time seriesanalysis*, Lecture Notes in Statistics, Springer-Verlag, 1983.

LAPS LABORATORY
BADJI MOKHTAR UNIVERSITY
ANNABA- ALGERIA
E-mail address: haninabil2012@gmail.com

LAPS LABORATORY
BADJI MOKHTAR UNIVERSITY
ANNABA- ALGERIA
E-mail address: zeghdoudihalim@yahoo.fr