Afrika Statistika

Afrika Statistika
Vol. 13 (3), 2018, pages 1717 –1732.
DOI: http://dx.doi.org/10.16929/as/1717.130

ISSN 2316-090X

Estimation of a stationary multivariate ARFIMA process

Kévin Stanislas MBEKE [1], **Ouagnina HILI**[1,*]

[1] Laboratory of Mathematics and New Technologies of Information, National Polytechnic Institute Félix HOUPHOUET-BOIGNY Yamoussoukro, P.O. Box 1093, Ivory Coast

Received on May 15, 2018; Accepted on July, 2018

Abstract. In this note, we consider an m-dimensional stationary multivariate long memory ARFIMA (AutoRegressive Fractionally Integrated Moving Average) process, which is defined as : $A(L)D(L)\left(y_1(t),\ldots,y_m(t)\right)^{'} = B(L)\left(\varepsilon_1(t),\ldots,\varepsilon_m(t)\right)^{'}$, where $M^{'}$ denotes the transpose of the matrix M. We determine the minimum Hellinger distance estimator (MHDE) of the parameters of a stationary multivariate long memory ARFIMA. This method is based on the minimization of the Hellinger distance between the random function of $f_n(.)$ and a theoretical probability density $f_\theta(.)$. We establish, under some assumptions, the almost sure convergence of the estimator and its asymptotic normality.

Résumé. Dans cette note, nous considérons un processus ARFIMA (AutoRegressive Fractionally Integrated Moving Average) stationnaire multivarié à longue mémoire défini par : $A(L)D(L)\left(y_1(t),\ldots,y_m(t)\right)^{'} = B(L)\left(\varepsilon_1(t),\ldots,\varepsilon_m(t)\right)^{'}$, où $M^{'}$ représente la transposée de la matrice M. Nous déterminons le minimum de distance de Hellinger d'un estimateur (MHDE) de paramètres d'un processus ARFIMA stationnaire multivarié à longue mémoire. Cette méthode consiste à minimiser la distance de Hellinger entre la densité de probabilité théorique $f_\theta(.)$ et une fonction aléatoire $f_n(.)$. Sous quelques hypothèses, nous établissons la convergence presque sûre de l'estimateur et sa normalité asymptotique.

Key words: Stationary Multivariate ARFIMA process; Estimation; Long memory; Minimum Hellinger distance.
AMS 2010 Mathematics Subject Classification : 62F12, 62H12.

*Ouagnina HILI: o_hili@yahoo.fr
Kévin Stanislas MBEKE : mbekegabrielromain@yahoo.com

1. Introduction

The last decades of macroeconomic and financial economic research have resulted in a vast array of important contributions in the area of long memory modelling. From a theoretical perspective, much effort has focussed on issues of testing and estimation. Time series with long-range dependence appear in many contexts. This paper studies parameter estimation for the ARFIMA model. These processes take into account, in a modeling, the presence of long memory component in the study of a time series. The long memory phenomenon has provided econometricians with processes that take low frequencies into account. Such processes allow to get forecasts to far horizon. This motivates our study. Multivariate processes with long-range dependent properties are found in a large number of applications including finance, macroeconomic and neuroscience. Statistical analysis of such data is challenging because multivariate time series present phase phenomenons.

In this note, the objective is to determine the minimum Hellinger distance estimator (MHDE) of the parameters of a stationary multivariate long memory ARFIMA. Granger and Joyeux (1980) and Hosking (1981) have proposed the ARFIMA (p, d, q) model to define a time series, which presents a character of short or long memory following d. For $\frac{-1}{2} < d < 0$ the process is short memory. For $0 < d < \frac{1}{2}$ the process is long memory. This long memory property is characterized by a slow decay of the autocorrelation function or the sum of unfinished autocorrelations. The process is non-stationary for $d > \frac{1}{2}$ and stationary for $d < \frac{1}{2}$. In spite of ARFIMA process, the notion of long memory has been widely discussed by the authors such as Bitty and Hili (2010) for linear processes with long memory, N'dri and Hili (2013) for Strongly Dependent Multi-Dimensional Gaussian Processes. Mayoral (2007) proposed by Minimum Distance a new method for estimating the parameters of stationary and non-stationary ARFIMA (p, d_0, q) process for $d_0 > -0, 75$.

Our study is essentially based on the long range dependence process as in Kamagaté and Hili (2012). Kamagaté and Hili (2012) and Kamagaté and Hili (2013) estimated by the Minimum Hellinger Distance method a stationary univariate ARFIMA process and by the quasi maximum likelihood approach a non-stationary multivariate ARFIMA process.

In this paper, we generalize the results of Kamagaté and Hili (2012) to the multivariate case. We consider an m-dimensional ARFIMA stationary process $(y_1(t), \dots, y_m(t))$ following $d < \frac{1}{2}$ which is generated by

$$A(L)D(L)\left(y_1(t), \dots, y_m(t)\right)^{'} = B(L)\left(\varepsilon_1(t), \dots, \varepsilon_m(t)\right)^{'}$$

where $M^{'}$ denotes the transpose of the matrix M. After the invertibility of the above process, we establish the consistence and asymptotic normality by using the Minimum Hellinger Distance method. The reasons for choosing this estimation technic lie in the fact that these estimators obtained are efficient and robust (cf.

Beran (1977)).

The two main results of this work are the almost sure convergence of $\widehat{\theta}_n$ to θ and its asymptotic normality. This convergence is a consequence of Lemma 1 and Lemma 3.1 in Hili (1995). Lemma 1 is demonstrated in three steps by applying Prakasa-Rao's (1983) inequality and Borel-Cantelli's inequality. Asymptotic normality is a consequence of Lemma 2 and Lemma 3 proved by Beran (1977) and Wu and Mielniczuk (2002) respectively.

The paper is organized as following. After some notes about the estimator, in section 2, we present a multivariate ARFIMA model. Section 3 is devoted to the estimation of parameters including the consistency of the estimator and its asymptotic normality. In section 4 we establish the main results of this work.

We denote by θ the vector of parameters of interest composed of $(d_1, \dots, d_m)$ and matrix coefficients. The Minimum Hellinger Distance estimator of θ is defined by

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta} H_2(f_n, f_\theta),$$

where $H_2(f_n, f_\theta)$ is the Hellinger Distance defined by

$$H_2(f_n, f_\theta) = \left\{ \int_{\mathbb{R}^m} |f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)|^2 dx \right\}^{\frac{1}{2}}. \tag{1}$$

The Minimum Hellinger Distance minimizes the Hellinger Distance between f_n and f_θ. $f_\theta(.)$ is a theoretical probability density, $f_n(.)$ is a random function of $\widehat{\underline{\varepsilon}}(t)$ and $\tilde{f}_n(.)$ a non-parametric kernel density estimator of $\underline{\varepsilon}(t)$ defined by :

$$f_n(x) = \frac{1}{nh_n^m} \sum_{t=1}^{n} K\left(\frac{x - \widehat{\underline{\varepsilon}}(t)}{h_n}\right) \qquad x \in \mathbb{R}^m \tag{2}$$

$$\tilde{f}_n(x) = \frac{1}{nh_n^m} \sum_{t=1}^{n} K\left(\frac{x - \underline{\varepsilon}(t)}{h_n}\right) \qquad x \in \mathbb{R}^m, \tag{3}$$

where $K : \mathbb{R}^m \longrightarrow \mathbb{R}+$ is a kernel function and (h_n) is a sequence of bandwidths and $f : \mathbb{R}^m \longrightarrow \mathbb{R}+$.

2. Multivariate ARFIMA model

The multivariate ARFIMA model was introduced by Sowell (1989). We consider an m-dimensional ARFIMA stationary process $(y_1(t), \dots, y_m(t))$ following $d < \frac{1}{2}$ which is generated by

$$A(L)D(L)\left(y_1(t), \dots, y_m(t)\right)^{'} = B(L)\left(\varepsilon_1(t), \dots, \varepsilon_m(t)\right)^{'} \tag{4}$$

where $M^{'}$ denotes the transpose of the matrix M. L is the backward shift operator which, to any element of a time series, associates the previous observation as

$$L^j X_t = X_{t-j}, \qquad j \in \mathbb{N},$$

$\{\varepsilon_1(t), \ldots, \varepsilon_m(t)\}$ are white noise processes that follow the normal law of mean zero and covariance

$$r\left\{\varepsilon_i(t), \varepsilon_j(s)\right\} = \delta\left(t, s\right) k_{ij}, \qquad i, j = 1, \ldots, m.$$

Denote by $\mathbf{K} = (k_{ij})$ the positive definite covariance matrix. The expression $D(L)$ defined in (4) represents a diagonal $(m \times m)$-matrix of the fractional difference operators of backward shift defined by

$$D(L) = \begin{pmatrix} (1-L)^{d_1} & 0 & \cdots \\ 0 & \ddots & 0 \\ \vdots & 0 & (1-L)^{d_m} \end{pmatrix},$$

with $(1-L)^d = 1 + \sum_{k=1}^{+\infty} \frac{\Gamma(k-d)}{\Gamma(-d)k!} L^k$ and $d_1, \ldots, d_m \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. $\Gamma(.)$ is the gamma function such as $\Gamma\left(j+1\right) = j!$..

Let $A(.)$ and $B(.)$ be matrix polynomials in L of degrees p and q respectively defined as hereinafter by :

$$A(L) = I - A_1 L - \ldots - A_p L^p,$$

$$B(L) = I + B_1 L + \ldots + B_q L^q,$$

where I represents the $(m \times m)-$identity matrix. The $det\, A(L) \neq 0$ and $det\, B(L) \neq 0$ are, respectively, the characteristic polynomial of the matrix polynomials $A(.)$ and $B(.)$. We assume that the roots of characteristic polynomial are all outside the unit disk.

Odaki (1993) and Hosking (1981) showed that the process is invertible for $d > -1$ and stationary for $d < \frac{1}{2}$. Taking into account the conditions on the polynomials, the process (4) is invertible and causal and admits a representation of an autoregressive process of infinite order as following :

$$B(L) \begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = A(L) D\left(L\right) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix},$$

$$\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = (B(L))^{-1} A(L) D(L) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}. \tag{5}$$

Let $\left(B(L)\right)^{-1} A(L) = C(L)$, the equality (5) can be written as follows

$$\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = \begin{pmatrix} c_{11}(L,\theta) & \cdots & c_{1m}(L,\theta) \\ \vdots & \cdots & \vdots \\ c_{m1}(L,\theta) & \cdots & c_{mm}(L,\theta) \end{pmatrix} D(L) \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix}$$

and

$$\begin{pmatrix} \varepsilon_1(t) \\ \vdots \\ \varepsilon_m(t) \end{pmatrix} = \sum_{j=0}^{\infty} \Psi_j(\theta) \begin{pmatrix} y_1(t-j) \\ \vdots \\ y_m(t-j) \end{pmatrix}$$

where $\theta = (A_i, d_l, B_j) \in \Theta \subset \mathbb{R}^{\mathbf{n}}$ is the vector parameters of interest. For $0 \leq i \leq p$, $1 \leq l \leq m$ and $0 \leq j \leq q$. Θ is a compact set. $\mathbf{n} = m\,[m(p+q)+1]$. $\{\Psi_j(\theta)\}_{j=0}^{\infty}$ are $(m \times m)$-matrix associated with the entire series development of the matrix polynomial $C(L)D(L)$ in power of L such as $\sum_{j\in\mathbb{Z}} |\psi_{r,s}(j)| < \infty$ for $1 \leq r, s \leq m$.

Let $\underline{y}_i$ for $1 \leq i \leq n$ be the observations. The innovations $(\varepsilon_1(t), \ldots, \varepsilon_m(t))'$ are not observable, they are estimated by

$$\begin{pmatrix} \widehat{\varepsilon}_1(t) \\ \vdots \\ \widehat{\varepsilon}_m(t) \end{pmatrix} = \sum_{j=0}^{n} \Psi_j\,(\theta) \begin{pmatrix} y_1(t-j) \\ \vdots \\ y_m(t-j) \end{pmatrix}.$$

3. Parameter estimation

To establish the consistency and limit law of the parameter, we need the following assumptions :

Assumption $(A1)$

1. $E(|\underline{\varepsilon}_t|^s) < +\infty$ for $s \geq 1$.
   For all $(u,v) \in \mathbb{R}^{2m}$, we have :
2. $\int_{\mathbb{R}^m} K^2(u)du < \infty$, $\int_{\mathbb{R}^m} u_i K(u)du = 0$ for $1 \leq i \leq m$;
3. $\int_{\mathbb{R}^m} u_i u_j K(u)du = 0$, $\int_{\mathbb{R}^m} u_i^2 K(u)du < \infty$ for $1 \leq j \leq m$;
4. There exists $c > 0$ such as $\sup_{u\in\mathbb{R}^m} |K(u+v) - K(u)| \leq c|v|$.

Assumption $(A2)$. For each $\theta \in \Theta$ and each $x \in \mathbb{R}^m$, the functions $x \mapsto f_\theta(x)$ and $\theta \mapsto f_\theta^{\frac{1}{2}}(x)$ are continuously differentiable and $\underline{\varepsilon}_t$'s admits a density absolutely continuous with respect to the Lebesgue measure $\mathbb{R}^m$, positive in a neighborhood of the origin.

Assumption $(A3)$. For each $x \in \mathbb{R}^m$, the functions $\theta \mapsto \frac{\partial}{\partial\theta_j} f_\theta^{\frac{1}{2}}(x)$, for $1 \leq j \leq q$ and $\theta \mapsto \frac{\partial^2}{\partial\theta_j\partial\theta_k} f_\theta^{\frac{1}{2}}(x)$, for $1 \leq j, k \leq q$ are finished, continuous and defined in $L^2(\mathbb{R}^q)$.

Assumption $(A4)$. $h_n = n^\alpha \mathcal{L}(n)$, where $-1 < \alpha < 0$ with $\mathcal{L}$ a slowly varying function.

$$\lim_{n\to+\infty} h_n = 0, \; \lim_{n\to+\infty} nh_n = +\infty, \; \lim_{n\to+\infty} \frac{\mathcal{L}(an)}{\mathcal{L}(n)} = 1, \; a > 0$$

For each $\theta \in \Theta$, $\sup_{x\in\mathbb{R}^m} \mid \frac{\partial^i f_\theta}{\partial x_k^i}(x) \mid < \infty$, $i = 0, 1, 2, \cdots$ *and* $k = 1, \cdots, m$.

Assumption $(A5)$. For $\theta,\, \theta^{'} \in \Theta,\, \theta \neq \theta^{'}$ implies that $\{x \in \mathbb{R}^m / f_\theta(x) \neq f_{\theta'}(x)\}$ is a set of positive Lebesgue measure.

Assumption $(A6)$. We suppose that there is a constant M such as $\sup_{x\in\mathbb{R}^m} f(x) \leq M < \infty$.

We obtain the following almost-sure convergence theorem and asymptotic distribution laws.

Theorem 1. *Supposing that assumptions (A1)-(A6) are satisfied. Then $\widehat{\theta}_n$ converges almost surely to θ for all $x \in \mathbb{R}^m$. We denote by :*

$$g_\theta(x) = f_\theta^{\frac{1}{2}}(x),\ \dot{g}_\theta(x) = \frac{\partial g_\theta(x)}{\partial\theta},\ \ddot{g}_\theta(x) = \frac{\partial^2 g_\theta(x)}{\partial\theta\partial\theta^t},\ S_\theta(x) = \left[\int_{\mathbb{R}^m} \dot{g}_\theta(x)\dot{g}_\theta^t(x)dx\right]^{-1} \dot{g}_\theta(x).$$

when these quantities exist.

Theorem 2. *labeltheoremHA2 Supposing that assumptions (A1)-(A6) are satisfied and that the following conditions:*

Condition C1 *: The components of $\dot{g}_\theta$ and $\ddot{g}_\theta$ are in L_2 and if the norms of these components are continuous functions of θ.*

Condition C2 *: $\int_{\mathbb{R}^m} \ddot{g}_\theta(x)g_\theta(x)dx$ is a non-singular $(\boldsymbol{n} \times \boldsymbol{n})$-matrix,*

hold. Then the limit distribution of $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$ is $N\left(0, \Sigma^2\right)$, where

$$\Sigma^2 = \frac{1}{4}\left[\int_{\mathbb{R}^m} \dot{g}_\theta(x)\dot{g}_\theta^t(x)dx\right]^{-1} \int_{\mathbb{R}^m} K^2(u)du.$$

4. Proof of the theorems

We need the following lemma to prove Theorem 1.

Lemma 1. *Supposing that assumptions (A1) and (A2) are satisfied. Then*

$$f_n(x) - f_\theta(x) \to 0 \qquad a.s \text{ when } n \to +\infty.$$

Proof. By the triangular inequality, we have

$$\sup_{x\in\mathbb{R}^m} |f_n(x) - f_\theta(x)| \leq (a) + (b) + (c)$$

where

$$(a) = \sup_{x\in\mathbb{R}^m} |f_n(x) - \tilde{f}_n(x)|$$

$$(b) = \sup_{x\in\mathbb{R}^m} |\tilde{f}_n(x) - E\tilde{f}_n(x)|$$

$$(c) = \sup_{x\in\mathbb{R}^m} |E\tilde{f}_n(x) - f_\theta(x)|.$$

Let us split the proof in three steps.

Step 1 : The convergence of (a) to zero after inversion of the process (4).

Considering the conditions on the polynomial functions $A(L)$ and $B(L)$, the process (4) is invertible and can be rewritten as a representation of a autoregressive process of infinite order.

We consider two density functions $f_n(.)$ and $\tilde{f}_n(.)$ respectively of $\widehat{\underline{\varepsilon}}_t$ and $\underline{\varepsilon}_t$. By assumption (A1), we have

$$\sup_{x\in\mathbb{R}^m} |f_n(x) - \tilde{f}_n(x)| \leq \frac{c}{nh_n^{m+1}} \sum_{t=1}^{n} \|\underline{\eta}_t\|,$$

where

$$\underline{\eta}_t = \begin{pmatrix} \eta_1(t) \\ \vdots \\ \eta_m(t) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) - \widehat{\varepsilon}_1(t) \\ \vdots \\ \varepsilon_m(t) - \widehat{\varepsilon}_m(t) \end{pmatrix},$$

where $\underline{\eta}_t$ represents the rest when truncating the series from n,

$$\underline{\eta}_t = \sum_{j=n+1}^{\infty} \Psi_j(\theta)\underline{y}_{t-j}.$$

We adapt the notion of invertibility according to Granger and Andersen (1978) as below

$$\lim_{t\to+\infty} E\left(\underline{\varepsilon}_t - \widehat{\underline{\varepsilon}}_t\right)^2 = \Sigma^2 \sum_{j=n+1}^{+\infty} \Psi_{j,n-1}^2 < \infty.$$

Let's examine the following expression

$$E\left(\frac{1}{nh_n^{m+1}} \sum_{t=1}^{n} \|\underline{\eta}_t\|\right)^2.$$

We note by

$$\chi = \frac{1}{nh_n^{m+1}} \sum_{t=1}^{n} \|\underline{\eta}_t\|$$

$$E\left(\chi\right)^2 = E\left(\frac{1}{n^2h_n^{2m+2}} \sum_{t=1}^{n} \|\underline{\eta}_t\|^2\right) + 2E\left(\frac{1}{n^2h_n^{2m+2}} \sum_{t=1}^{n} \sum_{\substack{s=1 \\ t<s}}^{n} \|\underline{\eta}_t\|\|\underline{\eta}_s\|\right).$$

For all reals $u > 0$ and $v > 0$,

$$uv \leq \frac{1}{2}\left(u^2 + v^2\right) \tag{6}$$

then

$$E(\chi)^2 \leq E\left(\frac{1}{n^2 h_n^{2m+2}} \sum_{t=1}^{n} \|\underline{\eta}_t\|^2\right) + \frac{1}{n^2 h_n^{2m+2}} E\left(\sum_{t=1}^{n} \|\underline{\eta}_t\|^2 + \sum_{\substack{s=1 \\ t<s}}^{n} \|\underline{\eta}_s\|^2\right)$$

thus

$$\begin{aligned} E(\chi)^2 &\leq \frac{2}{n^2 h_n^{2m+2}} E\left(\sum_{t=1}^{n} \|\underline{\eta}_t\|^2\right) + \frac{1}{n^2 h_n^{2m+2}} E\left(\sum_{\substack{s=1 \\ t<s}}^{n} \|\underline{\eta}_s\|^2\right) \\ &\leq \frac{2}{n^2 h_n^{2m+2}} E\left(\sum_{t=1}^{n} \|\underline{\eta}_t\|^2\right) + \frac{2}{n^2 h_n^{2m+2}} E\left(\sum_{\substack{s=1 \\ t<s}}^{n} \|\underline{\eta}_s\|^2\right) \\ &\leq \frac{2}{n^2 h_n^{2m+2}} \left(E\left(\sum_{t=1}^{n} \|\underline{\eta}_t\|^2\right) + E\left(\sum_{\substack{s=1 \\ t<s}}^{n} \|\underline{\eta}_s\|^2\right)\right) \\ &\leq \frac{2}{n^2 h_n^{2m+2}} (d + e). \end{aligned}$$

We will now focus on the expressions of (d) and (e).

$$d = E\left(\sum_{t=1}^{n} \| \sum_{j=n+1}^{\infty} \Psi_j(\theta)\underline{y}_{t-j}\|^2\right) \qquad e = E\left(\sum_{s=1}^{n} \| \sum_{j=n+1}^{\infty} \Psi_j(\theta)\underline{y}_{t-j}\|^2\right).$$

Let's consider $\sum_{j=n+1}^{\infty} \Psi_j(\theta)\underline{y}_{t-j}$ a vectorial series in $\mathbb{R}^m$ vectorial space.

$$\sum_{j=n+1}^{\infty} \Psi_j(\theta)\underline{y}_{t-j} = \lim_{N\to\infty} \sum_{j=n+1}^{N} \Psi_j(\theta)\underline{y}_{t-j}.$$

By the triangular inequality, we have

$$d \leq E\left(\sum_{t=1}^{n} \lim_{N\to\infty} \left(\sum_{j=n+1}^{N} \|\Psi_j(\theta)\underline{y}_{t-j}\|\right)^2\right).$$

Therefore

$$d \leq E\left(\sum_{t=1}^{n} \lim_{N\to\infty} \left(\sum_{j=n+1}^{N} \|\Psi_j(\theta)\underline{y}_{t-j}\|^2 + 2 \sum_{J=n+1}^{N} \sum_{\substack{l=n+1 \\ j<l}}^{N} \|\Psi_j(\theta)\underline{y}_{t-j}\| \|\Psi_l(\theta)\underline{y}_{t-l}\|\right)\right).$$

Using inequality (6), we have

$$2\sum_{J=n+1}^{N}\sum_{\substack{l=n+1\\ j<l}}^{N}\|\Psi_j(\theta)\underline{y}_{t-j}\|\|\Psi_l(\theta)\underline{y}_{t-l}\| = \sum_{J=n+1}^{N}\|\Psi_j(\theta)\underline{y}_{t-j}\|^2+\sum_{\substack{l=n+1\\ j<l}}^{N}\|\Psi_l(\theta)\underline{y}_{t-l}\|^2.$$

Consequently

$$\begin{aligned} d &\leq 2E\left(\sum_{t=1}^{n}\lim_{N\to\infty}\left(\sum_{j=n+1}^{N}\|\Psi_j(\theta)\underline{y}_{t-j}\|^2+\sum_{\substack{l=n+1\\ j<l}}^{N}\|\Psi_l(\theta)\underline{y}_{t-l}\|^2\right)\right)\\ &\leq 2E\left(\sum_{t=1}^{n}\lim_{N\to\infty}\left(\sum_{j=n+1}^{N}\left(\sum_{i=1}^{m}\psi_{i,tj}^2\right)+\sum_{\substack{l=n+1\\ j<l}}^{N}\left(\sum_{i=1}^{m}\psi_{i,tl}^2\right)\right)\right), \end{aligned} \tag{7}$$

where $\psi_{i,tj}$ and $\psi_{i,tl}$ are the coefficients of the vector $\Psi_j(\theta)\underline{y}_{t-j}$. By using the same argument as in (d), we obtain

$$\begin{aligned} e &= E\left(\sum_{s=1}^{n}\|\sum_{j=n+1}^{\infty}\Psi_j(\theta)\underline{y}_{s-j}\|^2\right)\\ &\leq 2E\left(\sum_{s=1}^{n}\lim_{N\to\infty}\left(\sum_{j=n+1}^{N}\|\Psi_j(\theta)\underline{y}_{s-j}\|^2+\sum_{\substack{l=n+1\\ j<l}}^{N}\|\Psi_l(\theta)\underline{y}_{s-l}\|^2\right)\right)\\ &\leq 2E\left(\sum_{s=1}^{n}\lim_{N\to\infty}\left(\sum_{j=n+1}^{N}\left(\sum_{i=1}^{m}\psi_{i,sj}^2\right)+\sum_{\substack{l=n+1\\ j<l}}^{N}\left(\sum_{i=1}^{m}\psi_{i,sl}^2\right)\right)\right), \end{aligned} \tag{8}$$

where $\psi_{i,sj}$ and $\psi_{i,sl}$ are the coefficients of the vector $\Psi_j(\theta)\underline{y}_{s-j}$. Odaki (1993) characterized the invertibility of the process by a function $f_n(d)$ defined as following

$$f_n(d)=\begin{cases} 1/n & for\ 0<|\,d\,|<\frac{1}{2}\\ \frac{log(n)}{n} & for\ d=-\frac{1}{2}\\ 1/n^{2(1+d)} & for\ d<-\frac{1}{2} \end{cases}.$$

Odaki (1993) also showed that the order of magnitude of the sum of squares of these coefficients is

$$\sum_{i=0}^{+\infty}\psi_{i,tj}^2 = o\left(\frac{1}{n}\right)\ \ if\ d\in(-1/2\,;\,1/2)\,. \tag{9}$$

By (7), (8) and (9), we have

$$\begin{aligned} \frac{c}{n^2h_n^{2m+2}}o\left(\frac{1}{n}\right) &= o\left(\frac{1}{n^3h_n^{2m+2}}\right)\\ &= o\left(\frac{1}{n^{2m\alpha+2\alpha+3}\mathcal{L}^{2m+2}(n)}\right). \end{aligned}$$

Then

$$\sup_{x\in\mathbb{R}^m}|f_n(x)-\tilde{f}_n(x)| = o\left(\frac{1}{n^{2m\alpha+2\alpha+3}\mathcal{L}^{2m+2}(n)}\right) \quad when \quad n\to\infty$$
$$\sup_{x\in\mathbb{R}^m} n^{\frac{1}{4}}|f_{\theta_n}(x)-\tilde{f}_n(x)| = o\left(\frac{1}{n^{2m\alpha+2\alpha+\frac{11}{4}}\mathcal{L}^{2m+2}(n)}\right) \quad when \quad n\to\infty.$$

Hence the convergence of (a) to $o\left(\frac{1}{n^{2m\alpha+2\alpha+\frac{11}{4}}\mathcal{L}^{2m+2}(n)}\right)$ where $2m\alpha+2\alpha+\frac{11}{4}>0$.

Step 2 : We will now prove the almost sure convergence of (b) using the Prakasa-Rao's (1983) inequality

By Prakasa-Rao's (1983)'s inequality , we have

$$\mathbb{P}\left(|\, f_n(x)-Ef_n(x)\,| > \varepsilon\sqrt{\frac{s_n logm}{m}}\right) \leq 2exp\left(-\frac{n\frac{s_n logm}{m}\varepsilon^2}{8c_0M}\right).$$

Let's consider a sequence $(s_n)_{n\in\mathbb{N}^*}$ defined by

$$S_n = nh_n$$

where h_n is a sequence of bandwidths satisfaying Assumption (A4). Let choose (h_n) such that

$$h_n = n^{\alpha}ln(n)\,;\, s_n = n^{1+\alpha}ln(n) \quad with\ -1<\alpha<0.$$

$$\begin{aligned}-1<\alpha<0 \ &\Leftrightarrow\ 0<1+\alpha<1\\ &\Leftrightarrow\ 0<n^{1+\alpha}<n.\end{aligned}$$

Then for $n>1$, the general term sequence s_n is positive. Let's examine the limit of

$$S_n = n^{1+\alpha}ln(n)\ and \quad \frac{S_n ln(n)}{n}.$$

$$0<1+\alpha<1 \ \Rightarrow\ \lim_{n\to+\infty} S_n = +\infty$$

$$0<1+\alpha<1 \ \Rightarrow\ \lim_{n\to+\infty}\frac{S_n ln(n)}{n} = 0.$$

Let be $\delta_m(x,\underline{\varepsilon}_t) = \frac{1}{h_n^m}K\left(\frac{x-\underline{\varepsilon}_t}{h_n}\right)$, there is a positive constant c_0 such as

$$\sup_{x\in\mathbb{R}^m}\frac{1}{h_n^m}K\left(\frac{x-\underline{\varepsilon}_t}{h_n}\right) \leq c_0 s_n \to +\infty.$$

We can rewrite the Prakasa-Rao's (1983) inequality as following

$$\mathbb{P}\left(\mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > \varepsilon\sqrt{\frac{s_n ln(n)}{n}} \right) \leq 2exp\left(-\frac{s_n ln(n)\varepsilon^2}{8c_0 M \frac{m}{n}} \right),$$

$$\tilde{f}_n(x) = \frac{1}{nh_n^m} \sum_{t=1}^{n} K\left(\frac{x - \underline{\varepsilon}_t}{h_n} \right) \qquad x \in \mathbb{R}^m.$$

By Assumption $(A6)$, we have

$$Q = 8c_0 M \frac{m}{n} < \infty,$$

and it follows,

$$\mathbb{P}\left(\mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{\alpha}{2}} \varepsilon ln(n) \right) \leq 2exp\left(-\frac{\varepsilon^2 n^{\alpha+1} ln^2(n)}{Q} \right)$$

$$\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{2\alpha+1}{4}} \varepsilon ln(n) \right) \leq 2exp\left(-\frac{\varepsilon^2 n^{\frac{2\alpha+3}{2}} ln^2(n)}{Q} \right). \tag{10}$$

We will now dominate the next expression

$$\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{2\alpha+1}{4}} \varepsilon ln(n) \right).$$

By Assumption $(A4)$, we obtain

$$\begin{aligned} -1 < \alpha < 0 \;&\Leftrightarrow\; 1 < 2\alpha + 3 < 3 \\ &\Leftrightarrow\; 1 < n^{\frac{2\alpha+3}{2}} < \infty \\ &\Leftrightarrow\; 0 < \frac{\varepsilon^2 n^{\mu} ln^2(n)}{Q} < \infty \end{aligned}$$

where $\mu = \frac{2\alpha+3}{2} > 0$. There exists a sequence (V_n) such as $V_n = \beta ln(n)$ with $\beta \geq 2$ for a certain rank

$$\frac{\varepsilon^2 n^{\mu} ln^2(n)}{Q} > \beta ln(n)$$

under the constraints $\begin{cases} \varepsilon > 1 \\ n > 1 \\ Q > 0 \\ \mu > 0 \end{cases}$.

Then

$$exp\left(\frac{\varepsilon^2 n^{\mu} ln^2(n)}{Q} \right) > n^{\beta}.$$

Therefore

$$exp\left(-\frac{\varepsilon^2 n^{\mu} ln^2(n)}{Q}\right) < \frac{1}{n^{\beta}}.$$

By Inequality (10), we have

$$\begin{aligned}
\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{2\alpha+1}{4}} \varepsilon ln(n)\right) &\leq \frac{2}{n^{\beta}} \\
\sum_{n\geq 1}\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{2\alpha+1}{4}} \varepsilon ln(n)\right) &\leq \sum_{n\geq 1}\frac{2}{n^{\beta}} \\
\sum_{n\geq 1}\mathbb{P}\left(n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid > n^{\frac{2\alpha+1}{4}} \varepsilon ln(n)\right) &< +\infty.
\end{aligned}$$

By using the Borel-Cantelli's Lemma, we have

$$n^{\frac{1}{4}} \sup_{x\in\mathbb{R}^m} \mid \tilde{f}_n(x) - E\tilde{f}_n(x) \mid = 0 \, a.s \quad when \, n \to \infty.$$

Hence the almost sure convergence of $\left(\tilde{f}_n(x) - E\tilde{f}_n(x)\right)$ to zero.

Step 3 : Let us show the convergence of the bias (c). By (3) we have

$$\begin{aligned}
E\left(\tilde{f}_n(x)\right) &= \frac{1}{nh_n^m} E\left(\sum_{t=1}^{n} K\left(\frac{x-\varepsilon_1}{h_n}\right)\right) \\
&= \frac{1}{h_n^m} E\left(K\left(\frac{x-\varepsilon_1}{h_n}\right)\right) \\
&= \frac{1}{h_n^m}\int_{\mathbb{R}^m} K\left(\frac{x-z}{h_n}\right) f_\theta(z)dx \\
&= \int_{\mathbb{R}^m} K(v) f_\theta\left(x - vh_n\right) dv.
\end{aligned}$$

By using the Taylor's expansion in a neighbourhood of x and under Point (3) of (A1), we obtain

$$f_\theta\left(x - vh_n\right) = f_\theta(x) + \sum_{k=1}^{m} \frac{\partial f_\theta}{\partial x_k}(x)\left(-h_n\right) v_k + \frac{h_n^2}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} \frac{\partial^2 f_\theta}{\partial x_j \partial x_k}(x) v_j v_k + o\left(h_n^2\right)$$

$$\begin{aligned}
E\left(\tilde{f}_n(x)\right) - f_\theta(x) &= \int_{\mathbb{R}^m} K(v) f_\theta\left(x - vh_n\right) dv - f_\theta(x)\int_{\mathbb{R}^m} K(v)dv \\
&= \int_{\mathbb{R}^m} K(v)\left[f_\theta\left(x - vh_n\right) - f_\theta(x)\right] dv \\
\sup_{x\in\mathbb{R}^m} |E\tilde{f}_n(x) - f_\theta(x)| &\leq \int_{\mathbb{R}^m} K(v)\left[\frac{h_n^2}{2}\sum_{k=1}^{m} |\frac{\partial^2 f_\theta}{\partial x_k^2}(x)| v_k^2 + o\left(h_n^2\right)\right] dv \\
&\leq \Delta h_n^2 \int_{\mathbb{R}^m} K(v)\left[v_k^2 + o(1)\right] dv,
\end{aligned}$$

where

$$\Delta = \frac{1}{2}\sum_{k=1}^{m}\frac{\partial^2 f_\theta}{\partial x_k^2}(x).$$

Under (A1) and (A4), $\sup_{x\in\mathbb{R}^m}|E\tilde{f}_n(x) - f_\theta(x)| \to 0\ a.s$ when $n \to \infty$

The convergence of (a),(b) and (c) implies Lemma 1. □

Beran (1977) and Hili (1995) consider $\mathbb{F}$ the set of all densities with respect Lebesgue measure on $\mathbb{R}$. We define the functional $T : \mathbb{F} \to \Theta$ as following Let be $g \in \mathbb{F}$ we set

$$A(g) = \left\{\theta \in \Theta : H_2\left(g, f_\theta\right) = \min_{\theta\in\Theta} H_2\left(g, f_\theta\right)\right\}$$

where H_2 is the Hellinger Distance. If $A(g)$ is unique, then $T(g)$ is defined as the value of this element. Elsewhere, they choose an arbitrary but unique element of $A(g)$ and call it $T(g)$.

proof of Theorem 1. Theorem 1 is a consequence of Lemma 1 and Lemma 3.1 in Hili (1995). We have

$$\sup_{x\in\mathbb{R}^m}|f_n(x)-f_\theta(x)| \le \sup_{x\in\mathbb{R}^m}|f_n(x)-\tilde{f}_n(x)|+\sup_{x\in\mathbb{R}^m}|\tilde{f}_n(x)-E\tilde{f}_n(x)|+\sup_{x\in\mathbb{R}^m}|E\tilde{f}_n(x)-f_n(x)|.$$

By Lemma 1

$$|f_n(x) - f_\theta(x)| \to 0 \quad a.s \quad when\, n \to +\infty.$$

Hence

$$\mathbb{P}\left\{\lim_{n\to+\infty} f_n^{\frac{1}{2}}(x) = f_\theta^{\frac{1}{2}}(x) \quad for\ all\ x\right\} = 1.$$

Since

$$\int_{\mathbb{R}^m} f_n(x)dx = \int_{\mathbb{R}^m} f_\theta(x)dx = 1.$$

Consequently

$$H_2\left(f_n, f_\theta\right) = \left\{\int_{\mathbb{R}^m} |f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)|^2 dx\right\}^{\frac{1}{2}} \to 0\, a.s\, when\, n \to \infty.$$

By Lemma 3.1 in Hili (1995), $T\left(f_\theta\right) = \theta$ uniquely on Θ, then the functional T is continuous at f_θ in the Hellinger topology. Therefore

$$\widehat{\theta}_n = T\left(f_n(x)\right) \to T\left(f_\theta(x)\right) = \theta$$

almost surely when $n \to \infty$.
This achieves the proof of Theorem 1. □

Proof of Theorem 2. Lemma 2 and Lemma 3 below were respectively proved by Beran (1977) and by Wu and Mielniczuk (2002).

Lemma 2. *Let's suppose that assumptions (A2) and (A5) and the conditions C1 and C2 of Theorem 2 are satisfied and that* θ *lies in interior of* Θ. *So for any density sequence* $\{f_n\}$ *convergent to* f_θ *in the Hellinger metric, we have*

$$T\left(f_n(x)\right) = \theta + \int_{\mathbb{R}^m} S_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx + R_n \int_{\mathbb{R}^m} \dot{g}_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx.$$

where $R_{\mathbf{n}}$ *is a non-singular* $m\left[m(p+q)+1\right]-matrix$ *whose components of* $\sqrt{n}R_{\mathbf{n}}$ *tend to zero when* $n \longrightarrow +\infty$.

Lemma 3. *Let's suppose that assumptions (A1), (A2) and (A4) are satisfied, then* $N\left(0, f_\theta(x)\int_{\mathbb{R}^m} K^2(u)du\right)$ *is the limit distribution of* $\sqrt{nh_n}\left[f_n(x) - f_\theta(x)\right]$.

Proof. Let us focus now on the proof Theorem 2, referring to the above lemmas. By Lemma 2, we have

$$T\left(f_n(x)\right) = \theta + \int_{\mathbb{R}^m} S_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx + R_n \int_{\mathbb{R}^m} \dot{g}_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx.$$

Since $T\left(f_n(x)\right) = \widehat{\theta}_n$ and by multiplying the equation above by $\sqrt{n}$, we have

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) = \sqrt{n}\int_{\mathbb{R}^m} S_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx + \sqrt{n}R_n \int_{\mathbb{R}^m} \dot{g}_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx.$$

The components of $\sqrt{n}R_{\mathbf{n}} \to 0, when\ n \to \infty$. Then

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) = L_n + o_p(1),$$

where

$$L_n = \sqrt{n}\int_{\mathbb{R}^m} S_\theta(x)\left[f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right]dx.$$

Let's examine the limit law of L_n to deduce the limit law of $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$, where $S_\theta(x) \in L_2$ and $S_\theta \perp f_\theta^{\frac{1}{2}}$ where $\perp$ means orthogonality in L_2. By assumption $(A2)$, $f_\theta^{\frac{1}{2}}(x) > 0$ and the following algebraic equality we can rewrite L_n.

$$f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x) = \frac{f_n(x) - f_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)} - \frac{\left(f_n(x) - f_\theta(x)\right)^2}{2f_\theta^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x) + f_\theta^{\frac{1}{2}}(x)\right)^2}.$$

$$L_n = \sqrt{n}\int_{\mathbb{R}^m} S_\theta(x)\left[\frac{f_n(x) - f_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)} - \frac{\left(\widehat{f}_n(x) - f_\theta(x)\right)^2}{2f_\theta^{\frac{1}{2}}(x)\left(\widehat{f}_n^{\frac{1}{2}}(x) + f_\theta^{\frac{1}{2}}(x)\right)^2}\right]dx, \tag{11}$$

by distributivity in (11), we have

$$L_n = \sqrt{n}\int_{\mathbb{R}^m} S_\theta(x)\frac{f_n(x) - f_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)}dx + E_n, \tag{12}$$

where

$$E_n = -\sqrt{n}\int_{\mathbb{R}^m} \frac{S_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_\theta^{\frac{1}{2}}(x)\right)^2}\left(f_n(x)-f_\theta(x)\right)^2 dx.$$

Using inequality

$$2f_\theta^{\frac{1}{2}}\left(f_n^{\frac{1}{2}}+f_\theta^{\frac{1}{2}}\right)^2 = 2f_\theta^{\frac{3}{2}}+\gamma > 2f_\theta^{\frac{3}{2}} \qquad with \quad \gamma>0,$$

and posing $\delta = inf_{x\in\mathbb{R}^m} f(x)$, we can take E_n in absolute value as following

$$|E_n| \leq \sqrt{n}\int_{\mathbb{R}^m} \frac{|S_\theta(x)|}{2f_\theta^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_\theta^{\frac{1}{2}}(x)\right)^2}\left(f_n(x)-f_\theta(x)\right)^2 dx.$$

$$\begin{aligned}|E_n| &\leq \sqrt{n}\int_{\mathbb{R}^m} \frac{|S_\theta(x)|}{2\delta^{\frac{3}{2}}}\left(f_n(x)-f_\theta(x)\right)^2 dx \\ &\leq \frac{1}{2}\delta^{\frac{-3}{2}}\int_{\mathbb{R}^m} |S_\theta(x)|\sqrt{n}\left(f_n(x)-f_\theta(x)\right)^2 dx.\end{aligned}$$

By Lemma 1

$$n^{\frac{1}{4}}\sup_{x\in\mathbb{R}^m} \mid f_n(x)-f_\theta(x)\mid \to 0\, a.s \quad when\; n\to\infty$$

then

$$\sqrt{n}\left(f_n(x)-f_\theta(x)\right)^2 \to 0\; a.s \quad when\, n\to\infty.$$

S_θ is continuous and bounded (for θ fixed). By applying Vitali's Theorem on the sequence

$$W_n(x) = |S_\theta(x)|\sqrt{n}\left(f_n(x)-f_\theta(x)\right)^2$$

$|E_n| \to 0$ in probability when $n\to\infty$.

Let us consider the first term on the right of equation (12)

$$\sqrt{n}\int_{\mathbb{R}^m} S_\theta(x)\frac{f_n(x)-f_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)}dx. \tag{13}$$

Therefore, by Lemma 3, the limit distribution of (13) is $N\left(0,\Sigma^2\right)$, where

$$\begin{aligned}\Sigma^2 &= \int_{\mathbb{R}^m}\left(\frac{S_\theta(x)}{2f_\theta^{\frac{1}{2}}}\right)\left(\frac{S_\theta(x)}{2f_\theta^{\frac{1}{2}}}\right)^t\int_{\mathbb{R}^m} K^2(u)du f_\theta(x)dx \\ &= \frac{1}{4}\int_{\mathbb{R}^m} S_\theta(x)S_\theta^t(x)dx\int_{\mathbb{R}^m} K^2(u)du \\ &= \frac{1}{4}\left[\int_{\mathbb{R}^m}\dot{g}_\theta(x)\dot{g}_\theta^t(x)dx\right]^{-1}\int_{\mathbb{R}^m} K^2(u)du.\end{aligned}$$

Hence, we get the result. □

Acknowledgements We thank the anonymous referee for a careful reading and comments, which helped to improve the quality of the note.

References

Beran, R.(1977). Minimum Hellinger distance estimates for parametric models. *Ann. Statist.*, 5(3), 445-463. Doi : http://dx.doi.org/10.1214/aos/1176343842.

Bitty, A. L. et Hili, O.(2010). Estimateurs du minimum de distance de hellinger des processus linéaires à longue mémoire, *Comptes Rendus Mathématique*, 348(7-8) : 445-448. Doi : https://doi.org/10.1016/j.crma.2010.02.020

Granger, C. W. J. and Andersen, A.(1978), On the inverstibility of time series models, *Stoch. Process. Applic.*, 8, 87-92. Doi : https://doi.org /10.1016/0304-4149(78)90069-8

Granger, C. W. J. and Joyeux, R. (1980), An introduction to long-memory time series and fractional differencing. *Journal of Time Series Analysis*, 1, 15-29. Doi : https://doi.org/10.1111/j.1467-9892.1980.tb00297.x

Hili, O. (1995). On the estimation of nonlinear time series models, Stochastics and Stochastics. *Reports*, 52(3-4), 207-226. Doi :http://dx.doi.org /10.1080/17442509508833972.

Hosking, J. R. M.(1981), Fractional differencing. *Biometrika*, 68, 165-176. Doi :https:/doi.org/10.1093/biomet/68.1.165

Kamagaté, A. et Hili, O. (2012), Estimation par le minimum de distance de Hellinger d'un processus ARFIMA. *C. R. Acad. Sci. Paris, Ser. I*, 350(13-14), 721-725. Doi :http://dx.doi.org/10.1016/j.crma.2012.07.005.

Kamagaté, A. and Hili, O. (2013), The Quasi Maximum Likelihood Approach to Statistical Inference on a Nonstationary Multivariate ARFIMA Process. *Random Oper. and Stoch. Equa.*, 21(3),305-320. Doi: https://doi.org/10.1515/rose-2013-0014.

Mayoral, L. (2007) Minimum distance estimation of stationary and non-stationary ARFIMA processes.*Econome. J.*, 10(1)124-148. Doi: https://doi.org/10.1111/j.1368-423X.2007.00202.x

N'dri A. and Hili 0. (2013), Hellinger distance estimation of strongly dependent Multi-dimensional Gaussian processes. *International Journal of Statistics and Probability*, 2(3),70-84. Doi https://doi.org/10.5539ijsp.v2n3p70.

Odaki, M. (1993), On the invertibility of fractionally differenced ARIMA processes. *Biometrika*, 80, 703-709. Doi https://doi.org/10.1093/biomet/80.3.703

Prakasa, Rao, B. L. S. (1983), Nonparametric Functional Estimation, *New York : Academic Press.*

Sowell, F. (1989), A decomposition of block Toeplitz matrices with applications to vector time series, *GSIA mimograph, Carnegie Mellon University*, November 1989.

Wu, W. B. and Mielniczuk, J. (2002), Kernel density estimation for linear process, *The Annals of Statistics*, 30, 1441-1459.