A Class of Semiparametric Estimators for Long-Range Dependent Multivariate Processes

Sílvia R.C. Lopes and Guilherme Pumi

Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil
Corresponding author: S.R.C. Lopes (silvia.lopes@ufrgs.br)

Abstract

In this work we investigate the finite sample performance of a certain class of Gaussian Semiparametric Estimators (GSE) for the memory parameter in long-range dependent multivariate time series. The class of models considered here satisfies simple conditions on the spectral density function, restricted to a small neighborhood of the zero frequency. This includes, but is not limited to, the class of VARFIMA models. We present a simulation study to assess the finite sample properties of the proposed estimator in the context of bivariate VARFIMA$(0, \boldsymbol{d}, 0)$ processes for which the innovation's joint distribution is Gaussian, but the marginals are not. Marginal distributions considered here present heavier tails than the standard Gaussian distribution and include the Student's t, the Logistic and the hyperbolic-Secant distributions.

Keywords: Multivariate processes; Long-range dependence; Semiparametric estimation.

1 Introduction

The estimation of the memory parameter in long-range dependent multivariate time series has a long history. On the parametric case, Sowell (1989) introduced the maximum likelihood approach for the class of VARFIMA processes, later rigorously developed by Hosoya (1997) in a more general context. The method is computationally intensive, especially considering the limited power of the computers in late 1980's and early 1990's. Later Luceño (1996) and, more recently, Tsay (2010) considered computationally cheaper approximations to the exact maximum likelihood. Although the parametric approach does present good properties like $n^{1/2}$-consistency and asymptotic normality, the theory heavily relies on the Gaussianity assumption on the process, requires strong distributional and regularity conditions and it is very sensible regarding misspecifications on the parametric structure of the process.

The class of the so-called Gaussian Semiparametric Estimators (abbreviated GSE) of the memory parameter in long-range dependent processes comprehend estimators based on approximations of the spectral density in the vicinities of the zero frequency. The class of GSE estimators present several advantages over the parametric approach including less distributional assumptions and being more efficient and robust regarding to the short term dependency structure of the process. Although the name, it is important to notice that Gaussianity is nowhere required in the asymptotic theory. It was first introduced in Künsch (1987), but the rigorous asymptotic theory of a GSE type estimator was developed several years later in the seminal papers Robinson (1995a) and Robinson (1995b).

In Lobato (1999) the author introduces a two-step GSE which is $n^{1/2}$-consistent and asymptotically normally distributed under some mild conditions. Closely related to Lobato's two-step GSE is the estimator introduced in Shimotsu (2007) where the author introduces a GSE based on a different approximation for the spectral density function for a large class of processes. The GSE of Shimotsu is also $n^{1/2}$-consistent and asymptotically normally distributed. In Nielsen (2011) the work of Shimotsu is extended to cover non-stationary time series and recently, Pumi and Lopes (2012) and Pumi and Lopes (2013) introduce a generalization of Lobato's and Shimotsu's GSE, respectively.

In this paper we analyze the finite sample performance of the class of estimators introduced in Pumi and Lopes (2012) in the context of bivariate VARFIMA$(0, \boldsymbol{d}, 0)$ processes. We consider the case where the bivariate distribution of the innovation process is multivariate Gaussian, but its marginals present heavier tails than the Gaussian distribution. In order to obtain such prescribed structure, we apply some tools from the theory of copulas as in Lopes et al. (2013). The paper is divided as follows. In the next section we introduce the necessary tools and framework for the work. In section 3 detailed description of the Monte Carlo study proposed in this paper and the data generating process is presented, as well as the simulation results. Section 4 concludes the paper.

2 Framework and Preliminaries

Let us briefly review the class of GSE introduced in Pumi and Lopes (2012). Let $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ be a weakly stationary q-dimensional process and let f denote the spectral density matrix function of $\boldsymbol{X}_t$. Assume that f satisfies the following local approximation

$$f(\lambda) \sim \mathrm{diag}\{\lambda^{-d_i}\mathrm{e}^{\mathrm{i}(\pi-\lambda)d_i/2}\}\, G_0\, \mathrm{diag}\{\lambda^{-d_i}\mathrm{e}^{\mathrm{i}(\pi-\lambda)d_i/2}\}, \quad \text{as } \lambda \to 0^+, \tag{2.1}$$

where $d_i \in (-0.5, 0.5)$, $i = 1, \cdots, q$ and G_0 is a symmetric positive definite real matrix. Processes satisfying (2.1) include several fractionally integrated processes, such as the class of VARFIMA processes. Each coordinate of a process $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$ satisfying (2.1) with $d_i > 0$ exhibit long-range dependence in the sense that the respective marginal spectral density function satisfies $f(\lambda) \sim M\lambda^{-2d_i}$, as $\lambda \to 0^+$, for some constant $M > 0$ and $i \in \{1, \cdots, q\}$. Hence, the parameter $\boldsymbol{d} := (d_1, \cdots, d_q)$ ultimately determines the long run dependence structure of the process.

Let f_n denote an arbitrary estimator of f based on a set of observations $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ from $\{\boldsymbol{X}_t\}_{t=0}^{\infty}$. Consider the objective function

$$S(\boldsymbol{d}) := \log\big(\det\{\widehat{G}(\boldsymbol{d})\}\big) - 2\sum_{k=1}^{q} d_k \frac{1}{m}\sum_{j=1}^{m}\log(\lambda_j), \tag{2.2}$$

with

$$\widehat{G}(\boldsymbol{d}) := \frac{1}{m}\sum_{j=1}^{m}\mathrm{Re}\Big[\mathrm{diag}\{\lambda_j^{-d_i}\}f_n(\lambda_j)\,\mathrm{diag}\{\lambda_j^{-d_i}\}\Big], \tag{2.3}$$

where $\lambda_j := 2\pi j/n$, for $j = 1, \cdots, m$ and $m = o(n)$. The estimator of $\boldsymbol{d}$ is then defined as

$$\widehat{\boldsymbol{d}} := \underset{\boldsymbol{d}\in\Theta}{\arg\min}\{S(\boldsymbol{d})\}, \tag{2.4}$$

where Θ denotes the space of admissible estimates, a subset of $(-0.5, 0.5)^q$.

In Pumi and Lopes (2012) the authors derive conditions on f_n in order to obtain (2.4) a consistent and asymptotically normally distributed estimator. It is satisfied by general VARFIMA processes. Now let

$$w_n(\lambda) := \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n}\boldsymbol{X}_t\mathrm{e}^{\mathrm{i}t\lambda} \quad \text{and} \quad I_n(\lambda) := w_n(\lambda)\overline{w_n(\lambda)}'$$

be the discrete Fourier transform and the periodogram of $\boldsymbol{X}_t$ at λ, respectively, where $\overline{A}'$ denotes the conjugate transpose of a complex matrix A.

Let $W_n(\cdot) := \big(W_n^{ij}(\cdot)\big)_{i,j=1}^q$ be an array of functions (called weight functions) and $\{\ell(k)\}_{k\in\mathbb{N}}$ be an increasing sequence of positive integers. For a Fourier frequency λ_j, we define the smoothed periodogram of $\boldsymbol{X}_t$ by

$$f_n(\lambda_j) := \sum_{|k|\leq \ell(n)} W_n(k) \odot w_n(\lambda_{j+k})\overline{w_n(\lambda_{j+k})}', \tag{2.5}$$

where $\odot$ denotes the Hadamard product.

For $i \in \{1, \cdots, q\}$, let $h_i : [0,1] \to \mathbb{R}$ be a collection of functions. Consider the vector of functions $L_n(\cdot) := \big(L_n^i(\cdot)\big)_{i=1}^q$ defined as $L_n^i(\lambda) := h_i\big(\lambda/n\big)$ and let

$$S_n(\lambda) := \left(\frac{L_n^i(\lambda)}{\sqrt{\sum_{t=1}^n L_n^i(t)^2}}\right)_{i=1}^q.$$

The tapered periodogram $I_T(\lambda; n)$ of $\{\boldsymbol{X}_t\}_{t=1}^n$ is then defined by setting

$$I_T(\lambda; n) := w_T(\lambda; n)\overline{w_T(\lambda; n)}', \quad \text{where} \quad w_T(\lambda; n) := \frac{1}{\sqrt{2\pi}}\sum_{t=1}^n S_n(t) \odot \boldsymbol{X}_t \mathrm{e}^{-it\lambda}. \tag{2.6}$$

For a complete discussion on the properties of the ordinary, tapered and smoothed periodogram, we refer the reader to Priestley (1981) while a discussion on their properties under long-range dependence can be found in Hurvich and Beltrão (1993).

In this work, we consider f_n in (2.3) as the ordinary periodogram (which leads to the single-step version of Lobato, 1999's estimator as considered in Shimotsu, 2007 and Pumi and Lopes, 2012) the smoothed periodogram and the tapered periodogram. Under some mild conditions when f_n is the ordinary periodogram or the tapered periodogram, the estimator (2.4) is $n^{1/2}$-consistent and asymptotically normally distributed (Lobato, 1999, Shimotsu, 2007 and Pumi and Lopes, 2012), while when f_n is the smoothed periodogram, the result is conjectured to be true and there is empirical evidence supporting the claim (see Pumi and Lopes, 2012).

3 Monte Carlo Study

In this section we present the results of a Monte Carlo study conducted to assess the robustness of the estimator (2.4) against heavy tailed marginals. In Pumi and Lopes (2012) the authors present a Monte Carlo study based on bidimensional Gaussian VARFIMA$(0, \boldsymbol{d}, 0)$. In this work, we consider VARFIMA$(0, \boldsymbol{d}, 0)$ process for which the bivariate distribution on the innovation process is bivariate Gaussian, but the marginals are not. We consider the traditional Student's t distribution with 3 and 7 degrees of freedom (denoted by t_3 and t_7, respectively) the Standard Logistic distribution (denoted by Logistic(0,1)) and the hyperbolic-secant distribution (denoted by Hyp(1)) (with densities respectively given by

$$f(x) = \frac{1}{4}\mathrm{sech}\left(\frac{x}{2}\right)^2 \quad \text{and} \quad g(x) = \frac{1}{2}\mathrm{sech}\left(\frac{\pi x}{2}\right),$$

for $x \in \mathbb{R}$) as marginals. These distributions present heavier tails than the standard Normal distribution, with excess of kurtosis ∞ for the t_3, 2 for t_7 and Hyp(1) and 6/5 for the Logistic(0,1). In order to couple the bidimensional and marginal requirements we apply some tools from copulas (as in Lopes et al., 2013). We briefly explain the idea of the method. Let Φ and Φ^{-1} denote the distribution and the quantile function of a Standard Normal random variable, respectively. Also let Φ_ρ denote the distribution function of a bivariate normal distribution with mean $(0,0)'$ and variance-covariance matrix given by $\Omega_\rho := \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. The so-called *Gaussian family of copulas* comprehend the copulas given by

$$C_\rho(u,v) = \Phi_\rho\big(\Phi^{-1}(u), \Phi^{-1}(v)\big),$$

for $\rho \in [-1,1]$ and $(u,v) \in [0,1]^2$. Now suppose we want to generate a sample of size n, say $\{(x_1,y_1),\cdots,(x_n,y_n)\}$, from a bivariate random vector (X,Y) in such a way that the joint distribution function of X and Y is the bivariate Gaussian distribution with variance-covariance matrix Ω_ρ and the marginals of X and Y are two prescribed distributions, say F_1 and G_1, respectively. In order to accomplish the joint and marginal prescription, the following steps are applied:

1. Generate a random sample $\{(u_1,v_1),\cdots,(u_n,v_n)\}$ from C_ρ.
2. The desired sample is obtained by setting $x_i = F_1^{-1}(u_i)$ and $y_i = G_1^{-1}(v_i)$, $i = 1,\cdots,n$.

To obtain another sample from (X,Y) with different marginals F_2 and G_2, one can use the same sample $\{(u_1,v_1),\cdots,(u_n,v_n)\}$ from C_ρ and substitute F_1 and G_1 in step 2 by F_2 and G_2, respectively. The two samples obtained this way have the same joint dependence but different marginal distributions. This method allows one to study how the marginal behavior affects some quantity of interest by keeping the joint behavior of the sample (determined by the copula) fixed and introducing the features of interest directly into the marginals.

The goal of the conducted Monte Carlo study is to analyze the robustness of the estimator (2.4) regarding the marginal specification of the process. The time series are generated from bivariate VARFIMA$(0,\boldsymbol{d},0)$ processes with memory parameter $\boldsymbol{d} \in \{(0.1,0.3),(0.2,0.4),(0.1,0.4)\}$ by using the MA(∞) representation of the process (truncated at 50,000 coefficients). The sample size is held fixed in 1,000 and 1,000 replications of each experiment are performed. The innovation processes are generated by using the methodology above described. The joint distribution of the innovation process is held fixed as bivariate Gaussian with variance-covariance matrix Ω_ρ, for $\rho \in \{0.2, 0.4, 0.7\}$ in each experiment. The marginal distributions are held fixed as t_3, t_7, Logistic$(0,1)$ and Hyp(1) for both components in each experiment.

We consider 3 different estimators of the memory parameter $\boldsymbol{d}$, all of them are variations of (2.4). We denote by $\widehat{\boldsymbol{d}}_{\text{op}}$ the estimator (2.4) with f_n taken as the ordinary periodogram. The estimator based on the tapered periodogram is denoted by $\widehat{\boldsymbol{d}}_{\text{tp}}$, while the estimator based on the smoothed periodogram is denoted by $\widehat{\boldsymbol{d}}_{\text{sp}}$. For the tapered periodogram we apply the so-called cosine-bell tapering function and for the smoothed periodogram, we apply the so-called Bartlett's weights (see Hurvich and Beltrão, 1993, Priestley, 1981 and Pumi and Lopes, 2012). The smoothed periodogram is truncated at $\lfloor n^{0.9} \rfloor = 501$ while the cut-off point of the estimator (2.4) is $m = \lfloor n^{0.85} \rfloor = 354$. These are the values suggested in Pumi and Lopes (2012).

In Table 3.1 the mean estimated values along with the mean square error ($\times 100$, given in parenthesis) values are presented. We observe that the overall performance of all estimators is very good. From the table we observe that in most cases the estimator based on the smoothed periodogram $\widehat{\boldsymbol{d}}_{\text{sp}}$ presents the smallest bias and the smallest mean square error among all estimators. The estimator based on the tapered periodogram presents uniformly higher mean square error which is coherent with the results in Pumi and Lopes (2012). Judging by the results obtained, it is clear that the excess of kurtosis in the marginal distributions considered in this study did not affect significantly the performance of the estimators. We observe, nevertheless, that the estimation results for the case where the marginals are t_7, Logistic$(0,1)$

Table 3.1: Simulation results. Presented are the mean estimated and the mean square error ($\times 100$, given in parenthesis) values.

$\boldsymbol{d}$	$\widehat{\boldsymbol{d}}$	t_3	t_7	Logistic$(0,1)$	Hyp(1)
$\rho = 0.2$					
(0.1, 0.3)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.0962 (0.072) 0.2838 (0.098)	0.0959 (0.075) 0.2841 (0.103)	0.0958 (0.075) 0.2842 (0.103)	0.0959 (0.075) 0.2841 (0.103)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.1008 (0.211) 0.2919 (0.182)	0.0938 (0.178) 0.2849 (0.197)	0.0937 (0.187) 0.2883 (0.185)	0.0945 (0.228) 0.2857 (0.197)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.0973 (0.068) 0.2918 (0.079)	0.0971 (0.071) 0.2930 (0.083)	0.0971 (0.071) 0.2931 (0.083)	0.0971 (0.071) 0.2928 (0.083)
(0.2, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.1917 (0.077) 0.3801 (0.113)	0.1913 (0.081) 0.3804 (0.117)	0.1913 (0.081) 0.3804 (0.118)	0.1913 (0.081) 0.3803 (0.117)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.1932 (0.234) 0.3900 (0.198)	0.1909 (0.190) 0.3869 (0.191)	0.1925 (0.201) 0.3873 (0.192)	0.1921 (0.209) 0.3907 (0.207)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.1945 (0.070) 0.3969 (0.091)	0.1944 (0.073) 0.3985 (0.095)	0.1944 (0.073) 0.3986 (0.095)	0.1944 (0.074) 0.3983 (0.095)
(0.1, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.0965 (0.072) 0.3804 (0.112)	0.0962 (0.075) 0.3806 (0.116)	0.0962 (0.075) 0.3807 (0.117)	0.0962 (0.075) 0.3806 (0.117)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.1011 (0.244) 0.3939 (0.190)	0.0907 (0.213) 0.3853 (0.193)	0.0961 (0.160) 0.3912 (0.184)	0.0925 (0.268) 0.3885 (0.203)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.0975 (0.068) 0.3972 (0.091)	0.0974 (0.071) 0.3989 (0.095)	0.0974 (0.071) 0.3991 (0.096)	0.0974 (0.071) 0.3987 (0.095)
$\rho = 0.4$					
(0.1, 0.3)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.0969 (0.068) 0.2841 (0.096)	0.0967 (0.070) 0.2848 (0.098)	0.0967 (0.070) 0.2849 (0.098)	0.0967 (0.070) 0.2848 (0.098)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.0919 (0.204) 0.2868 (0.179)	0.0936 (0.221) 0.2857 (0.209)	0.0920 (0.188) 0.2847 (0.213)	0.0998 (0.240) 0.2935 (0.204)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.0979 (0.064) 0.2929 (0.077)	0.0979 (0.066) 0.2946 (0.078)	0.0979 (0.066) 0.2948 (0.078)	0.0979 (0.066) 0.2944 (0.079)
(0.2, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.1924 (0.073) 0.3804 (0.111)	0.1922 (0.075) 0.3811 (0.112)	0.1922 (0.075) 0.3811 (0.112)	0.1922 (0.075) 0.3810 (0.113)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.1916 (0.246) 0.3895 (0.193)	0.1936 (0.205) 0.3887 (0.212)	0.1935 (0.175) 0.3892 (0.204)	0.1919 (0.240) 0.3911 (0.216)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.1952 (0.066) 0.3985 (0.092)	0.1953 (0.068) 0.4006 (0.094)	0.1953 (0.068) 0.4008 (0.094)	0.1952 (0.068) 0.4004 (0.094)
(0.1, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.0979 (0.068) 0.3814 (0.108)	0.0981 (0.070) 0.3823 (0.108)	0.0981 (0.070) 0.3824 (0.109)	0.0981 (0.070) 0.3822 (0.109)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.0979 (0.209) 0.3924 (0.176)	0.0918 (0.178) 0.3841 (0.190)	0.0972 (0.159) 0.3918 (0.184)	0.0974 (0.194) 0.3911 (0.204)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.0991 (0.064) 0.3997 (0.093)	0.0995 (0.066) 0.4022 (0.096)	0.0996 (0.066) 0.4024 (0.096)	0.0995 (0.066) 0.4019 (0.096)
$\rho = 0.7$					
(0.1, 0.3)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.1001 (0.058) 0.2878 (0.080)	0.1008 (0.058) 0.2895 (0.077)	0.1008 (0.058) 0.2896 (0.077)	0.1007 (0.059) 0.2893 (0.078)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.0934 (0.204) 0.2863 (0.199)	0.0923 (0.220) 0.2826 (0.224)	0.0947 (0.165) 0.2862 (0.189)	0.0967 (0.230) 0.2886 (0.226)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.1018 (0.055) 0.2978 (0.069)	0.1029 (0.057) 0.3007 (0.068)	0.1030 (0.057) 0.3009 (0.068)	0.1028 (0.057) 0.3004 (0.068)
(0.2, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.1955 (0.060) 0.3842 (0.092)	0.1963 (0.060) 0.3857 (0.088)	0.1963 (0.060) 0.3858 (0.088)	0.1962 (0.060) 0.3856 (0.088)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.1889 (0.195) 0.3878 (0.193)	0.1883 (0.213) 0.3883 (0.191)	0.1917 (0.192) 0.3898 (0.191)	0.1921 (0.214) 0.3897 (0.231)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.1996 (0.055) 0.4042 (0.092)	0.2010 (0.056) 0.4075 (0.095)	0.2011 (0.056) 0.4077 (0.095)	0.2009 (0.056) 0.4072 (0.095)
(0.1, 0.4)	$\widehat{\boldsymbol{d}}_{\text{op}}$	0.1050 (0.063) 0.3889 (0.082)	0.1069 (0.066) 0.3916 (0.077)	0.1070 (0.066) 0.3918 (0.077)	0.1068 (0.065) 0.3914 (0.078)
	$\widehat{\boldsymbol{d}}_{\text{tp}}$	0.0938 (0.180) 0.3890 (0.193)	0.0988 (0.249) 0.3907 (0.232)	0.0930 (0.187) 0.3855 (0.186)	0.1007 (0.251) 0.3940 (0.210)
	$\widehat{\boldsymbol{d}}_{\text{sp}}$	0.1080 (0.065) 0.4093 (0.104)	0.1107 (0.070) 0.4140 (0.114)	0.1108 (0.070) 0.4143 (0.115)	0.1105 (0.070) 0.4136 (0.113)

and Hyp(1) (whose excesses of kurtosis are relatively close) are more alike compared to each other than when compared to t_3 (whose excess of kurtosis is ∞). Still, in most cases this difference is only on the third decimal place. Also, we observe that there seems to be an influence of the correlation (ρ) between the components on the innovation processes on the estimated values. The higher the correlation, the smaller bias and the mean square error tend to be, except in some cases for $\widehat{\boldsymbol{d}}_{\text{tp}}$.

4 Conclusions

In this work we compare the finite sample performance among 3 estimators of the memory parameter in the context of multivariate long-range dependent processes. These estimators belong to a certain general class of Gaussian Semiparametric Estimators introduced in Pumi and Lopes (2012) generalizing previous work of Lobato (1999) and Shimotsu (2007). These estimators are based on a certain objective function which depends on the choice of a spectral density matrix estimator. In this work we consider 3 estimators based on different spectral density estimators, namely the ordinary, the tapered and the smoothed periodogram.

A Monte Carlo simulation study based on bivariate VARFIMA$(0, \boldsymbol{d}, 0)$ processes is conducted to analyze the finite sample performance of the considered estimators. The innovation process is taken to be jointly Gaussian but coupled with heavy tailed marginals. Student's t, Logistic and Hyperbolic-Secant distributions are considered. To accomplish the prescribed distributional structure, some tools from the theory of copulas are applied to generate the time series necessary to the study.

The overall performance of the considered estimators are satisfactory, with a considerable advantage of the estimator based on the smoothed periodogram over the others. The simulation results also show that the tail thickness in the innovation processes' marginals do not considerably affect the performance of the considered estimators in this study. Our findings also suggest an increase on the performance of the estimators as the correlation on the innovation process increases.

Acknowledgements

G. Pumi's research was supported by a REUNI Pos-Doctoral fellowship from CAPES-Brazil via the REUNI program. S.R.C. Lopes' research was partially supported by CNPq-Brazil, by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1 and also by INCT *em Matemática*. The authors are grateful to the (Brazilian) National Center of Super Computing (CESUP-UFRGS) for the computational resources.

References

[1] Hosoya, Y. (1997). "A Limit Theory for Long-Range Dependence and Statistical Inference on Related Models". *Annals of Statistics*, **25**, 105-137.

[2] Hurvich, C.M. and Beltrão, K.I. (1993). "Asymptotics for the Low-Frequency Ordinates of the Periodogram of a Long-Memory Time Series". *Journal of Time Series Analysis*, **14**(5), 455-472.

[3] Künsch, H. (1987). "Statistical Aspects of Self-Similar Processes". In Prokhorov, Yu. and Sazanov, V.V. (Eds.), *Proceedings of the First World Congress of the Bernoulli Society*. Utrecht: VNU Science Press, 67-74.

[4] Lobato, I.N. (1999). "A Semiparametric Two-Step Estimator in a Multivariate Long Memory Model". *Journal of Econometrics*, **90**, 129-153.

[5] Lopes, S.R.C.; Pumi, G. and Zaniol, K. (2013). "Mallows Distance in VARFIMA$(0, \boldsymbol{d}, 0)$ Processes". *Communications in Statistics: Computations and Simulations*, **42**(1), 24-51.

[6] Luceño, A. (1996). "A Fast Likelihood Approximation for Vector General Linear Processes with Long Series: Application to Fractional Differencing". *Biometrika*, **83**(3), 603-614.

[7] Nielsen, F.S. (2011). "Local Whittle Estimation of Multi-Variate Fractionally Integrated Processes". *Journal of Time Series Analysis*, **32**(3), 317-335.

[8] Priestley, M.B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

[9] Pumi, G. and Lopes, S.R.C. (2012). "A Semiparametric Estimator for Long-Range Dependent Multivariate Processes". Submitted.

[10] Pumi, G. and Lopes, S.R.C. (2013). "A Generalization of a Gaussian Semiparametric Estimatior on Multivariate Long-Range Dependent Processes". Submitted.

[11] Robinson, P.M. (1995a). "Log-Periodogram Regression of Time Series with Long Range Dependence". *Annals of Statistics*, **23**(3), 1048-1072.

[12] Robinson, P.M. (1995b). "Gaussian Semiparametric Estimation of Long Range Dependence". *Annals of Statistics*, **23**(5), 1630-1661.

[13] Shimotsu, K. (2007). "Gaussian Semiparametric Estimation of Multivariate Fractionally Integrated Processes". *Journal of Econometrics*, **137**, 277-310.

[14] Sowell, F. (1989). "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models". Working Paper, Carnegie-Mellon University.

[15] Tsay, W-J. (2010). "Maximum Likelihood Estimation of Stationary Multivariate ARFIMA Processes". *Journal of Statistical Computation and Simulation*, **80**(7-8), 729-745.