UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
CADERNOS DE MATEMÁTICA E ESTATÍSTICA
SÉRIE A: TRABALHO DE PESQUISA

ESTIMATION OF THE LONG MEMORY PARAMETER IN NONSTATIONARY TIME SERIES USING SEMI-PARAMETRIC AND PARAMETRIC METHODS

BÁRBARA PATRICIA OLBERMANN
SÍLVIA REGINA COSTA LOPES
VALDÉRIO ANSELMO REISEN

SÉRIE A, Nº 53
PORTO ALEGRE, NOVEMBRO DE 1999

ESTIMATION OF THE LONG MEMORY PARAMETER IN NONSTATIONARY TIME SERIES USING SEMI-PARAMETRIC AND PARAMETRIC METHODS

Olbermann, B.P.[a] **Lopes, S.**[a1] and **Reisen, V.A.**[b]

[a] Instituto de Matemática - UFRGS, Porto Alegre, RS, Brazil
[b] Departamento de Estatística - UFES, Vitória, ES, Brazil

Abstract:
Recently, the study of time series has been focused on time series having the *long memory* property, that is, series in which the dependence between distant observations is not negligible. One model that shows properties of long memory is the $ARFIMA(p, d, q)$ when the degree of differencing d is in the interval (0.0,0.5), range where the process is stationary. In this work, we analyze the estimation of the degree d^* in $ARFIMA(0, d^*, 0)$ processes when $d^* > 0.5$, that is, when the processes are nonstationary, but still have the property of long memory. We present a study, through simulations, for the estimators of d^* with different semiparametric and parametric methods for nonstationary processes when d^* belongs to the intervals (0.5,1.0) and (1.0,1.5).

Keywords: Nonstationary time series, Long memory, Estimation.

[1] Corresponding author. E-mail: slopes@mat.ufrgs.br

1. INTRODUCTION

One characteristic that distinguishes a time series from other set of observations is the fact that the values of a time series, in different time t, are correlated; in other words, the random variable X_t is correlated with the random variables X_s, for all $s \neq t$.

A key problem in time series analysis is determining the degree of correlation between values of a series in different time t with the goal in constructing a model for obtaining good forecasting values.

Many researchers are studying time series with long memory properties. Hurst (1951) was the one who first studied long memory time series. Persistence, or long memory, in a time series is the presence of significative dependence between observations apart for a long period of time. This characteristic has been observed in time series in different fields such as meteorology, astronomy, hidrology and economics (more details can be found in Beran (1994) and Hosking (1981 and 1984)).

We can characterize *persistence* in two different ways:

a) in time domain, the autocorrelation function ρ_k decays hyperbolically to zero, that is, $\rho_k \simeq k^{2d-1}$ when $k \to \infty$.

b) in frequency domain, the spectral density function $f_X(\cdot)$ is unbounded when the frequency is near zero, that is, $f_X(w) \simeq w^{-2d}$ when $w \to 0$.

In this paper we study the $ARFIMA(0, d, 0)$ models, which are a class of processes with long memory property, through the estimation of the differencing parameter d. The contribution due to Geweke and Porter-Hudak (1983), introducing the $\hat{d}_p$ estimator, was very important giving rise to several other works, and presenting a proof for the asymptotic properties only when $d \in (-0.5, 0.0)$. Reisen (1994) proposed a modified form of the regression method based on a smoothed version of the periodogram function obtaining the $\hat{d}_{sp}$ estimator. Robinson (1994, 1995), making use of mild modifications on $\hat{d}_p$, deals simultaneously with $d \in (-0.5, 0.0)$ and $d \in (0.0, 0.5)$ proving the asymptotic properties for this new estimator, denoted here by $\hat{d}_{pr}$. Hurvich and Deo (1998) and also Robinson (1994), addressed the problem of selecting the number of frequencies that must be used in the linear regression model for estimating the differencing parameter in the stationary case. Fox and Taqqu (1986) have considered an approximated maximum likelihood procedure to estimate the parameter, denoted here by $\hat{d}_w$. They adapted the approach of Whittle (1951), introduced for weakly dependent random variables. Fox and Taqqu (1986) and Dahlhaus (1989) have shown that the maximum likelihood estimates of the $ARFIMA(p, d, q)$ model are asymptotically unbiased. The methods based on the use of the log-periodogram regression equation, such as the ones in Geweke and Porter-Hudak (1983), Reisen (1994) and Robinson (1994, 1995), are usually called *semiparametric methods* while the approximated maximum likelihood approach presented by Fox and Taqqu (1986) which uses the spectral density function is usually called *parametric method.*

The main goal of this paper is to evaluate estimators of the degree of differencing d in the case where the stochastic processes are nonstationary, that is, when $d > 0.5$. Hurvich and Ray (1995) consider the asymptotic characteristics of the periodogram ordinates for cases when $d \geq 0.5$ and $d \leq -0.5$. They find that the periodogram of a nonstationary or noninvertible fractionally integrated process at the j-th Fourier frequency $w_j = \frac{2\pi j}{n}$, where n is the sample size, has an asymptotic relative bias depending on j. They examine the impact of periodogram bias only on the $\hat{d}_p$ estimator in finite samples. Here we also consider the impact of periodogram bias on the smoothed regression estimator $\hat{d}_{sp}$ and also on the approximated maximum likelihood estimator $\hat{d}_{\mathrm{W}}$. We analyze the performance of those estimators of d based on their bias and their mean squared errors.

In Section 2 we present the $ARFIMA(p, d, q)$ process, for $d \in (-0.5, 0.5)$, giving several properties of an $ARFIMA(0, d, 0)$ process. In Section 3 we present methods for estimating the parameter d, in the stationary case and we also extend them for the nonstationary case. In Section 4 we analyze the nonstationary case. Section 5 presents the results for the estimators obtained from the methods described in Section 3 for the parameter of differencing in $ARFIMA(0, d^*, 0)$ nonstationary process but with long memory properties. Conclusions are given in Section 6.

2. THE ARFIMA(p,d,q) PROCESS

DEFINITION 1: Let $\{\epsilon_t\}_{t\in\mathbb{Z}}$ be the white noise process with mean zero and variance $\sigma_\epsilon^2 > 0$, $\mathcal{B}$ the backward-shift operator, that is, $\mathcal{B}X_t = X_{t-1}$, $\Phi(\mathcal{B})$ and $\Theta(\mathcal{B})$ polynomials of orders p and q, respectively, given by

$$\Phi(\mathcal{B}) = 1 - \phi_1(\mathcal{B}) - \cdots - \phi_p(\mathcal{B}^p)$$

and

$$\Theta(\mathcal{B}) = 1 - \theta_1(\mathcal{B}) - \cdots - \theta_q(\mathcal{B}^q),$$

where ϕ_i, $1 \leq i \leq p$, and θ_j, $1 \leq j \leq q$, are real constants. If $\{X_t\}_{t\in\mathbb{Z}}$ is a linear process given by

$$\Phi(\mathcal{B})(1-\mathcal{B})^d X_t = \Theta(\mathcal{B})\epsilon_t, \; t \in \mathbb{Z}, \qquad (1)$$

then $\{X_t\}_{t\in Z}$ is called a *general fractional differenced process* $ARFIMA(p, d, q)$, where d is the *degree* or *parameter of differencing.* The process

$$U_t = (1-\mathcal{B})^d X_t, \; t \in \mathbb{Z},$$

given by

$$\Phi(\mathcal{B})U_t = \Theta(\mathcal{B})\epsilon_t, \; t \in \mathbb{Z},$$

is an *autoregressive moving average process* $ARMA(p,q)$.

For a process to be invertible, that is, there exists a sequence $\{\pi_k\}_{k\geq 0}$ such that $\sum_{k\geq 0}\pi_k < \infty$ and $\epsilon_t = \sum_{k\geq 0}\pi_k X_{t-k}$, for $t \in \mathbb{Z}$, the roots of the equation $\Phi(\mathcal{B}) = 0$ must lie outside the unit circle; for the same process to be stationary the roots of the equation $\Theta(\mathcal{B}) = 0$ must lie outside the unit circle (assuming that $\Phi(\mathcal{B}) = 0$ and $\Theta(\mathcal{B}) = 0$ do not have common roots).

If $d \in (-0.5, 0.5)$, then the process $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and invertible. The term $(1-\mathcal{B})^d$, for $d \in \mathbb{R}$, is defined by the binomial expansion

$$(1-\mathcal{B})^d = \sum_{k=0}^{\infty} \binom{d}{k} (-\mathcal{B})^k = 1 - d\mathcal{B} - \frac{d}{2!}(1-d)\mathcal{B}^2 \cdots. \qquad (2)$$

When $p = q = 0$ in the expression (1), one has the $ARFIMA(0,d,0)$ process.

The $ARFIMA(p,d,q)$ process exhibit the property of long memory when $d \in (0.0, 0.5)$ and of short memory when $d \in (-0.5, 0.0)$. If $d \geq 0.5$ the ARFIMA process is nonstationary although for $d \in [0.5, 1.0)$ it is level- reverting in the sense that there is no long-run impact of an innovation on the value of the process (see Cheung and Lai (1993) and Wu and Crato (1995)). The level-reversion property no longer holds when $d \geq 1$. If $d \leq -0.5$ the ARFIMA process is noninvertible. The reader can find more properties of the $ARFIMA(p,d,q)$ process in Hosking (1981).

Consider now $p = q = 0$ in the expression (1) and let $\{X_t\}_{t\in\mathbb{Z}}$ be an $ARFIMA(0,d,0)$ process. Then the spectral density function of $\{X_t\}_{t\in\mathbb{Z}}$ is given by

$$f_X(w) = \left[2\,\sin(\frac{w}{2})\right]^{-2d}, \text{ for } 0 < w \leq \pi, \qquad (3)$$

as $w \to 0$, $f_X(w) \simeq w^{-2d}$; and the autocorrelation function is given by

$$\rho_k = \frac{(-d)!(k+d-1)!}{(d-1)!(k-d)!}, \text{ for } k \in \mathbb{Z},$$

as $k \to \infty$, $\rho_k \simeq \frac{(-d)!}{(d-1)!}k^{2d-1}$.

3. ESTIMATES OF THE DIFFERENCING PARAMETER

We describe now the estimation of the *degree or parameter of differencing* through the regression method based on the periodogram and smoothed periodogram functions. The regression method using the periodogram function was first introduced by Geweke and Porter-Hudak (1983) and has been widely used for many researchers. However, the periodogram function is not a consistent estimator for the spectral density function (see Brockwell and Davis (1987)). Reisen (1994) proposed a modified form of the regression method based on a consistent estimator for the spectral density function (when $d \in (-0.5, 0.0)$), that is, a smoothed version of the periodogram function. We use the notation $\hat{d}_p$ and $\hat{d}_{sp}$ for the estimators of d based respectively on the *periodogram and smoothed periodogram functions.* Robinson (1994) establishes some consistency properties for $\hat{d}_p$ and also provides an asymptotic distribution theory for any value of d under mild conditions. We shall denote the Robinson estimator by $\hat{d}_{pr}$. We also consider an approximated maximum likelihood estimator of the differencing parameter using the approach suggested by Whittle (1953) and we shall denote it in the sequel by $\hat{d}_{\mathrm{W}}$.

THE ESTIMATOR $\hat{d}_p$

Consider $\{X_t\}_{t \in \mathbb{Z}}$ an $ARFIMA(p,d,q)$ process, with $d \in (-0.5, 0.5)$, whose spectral density function is given by

$$f_X(w) = f_U(w)\Big[2\,\sin(\frac{w}{2})\Big]^{-2d}, \text{ for } w \in [-\pi, \pi],$$

where $f_U(\cdot)$ is the spectral density function of the $ARMA(p,q)$ process.

Consider the set of harmonic frequencies $w_j = \frac{2\pi j}{n}$, $j = 0, 1, \cdots, [n/2]$ where n is the sample size and $[\cdot]$ here means the integer part. By taking the logarithm of the spectral density function $f_X(w)$ and adding $\ln f_U(0)$ and $\ln I(w_j)$ to both sides we have

$$\ln I(w_j) = \ln f_U(0) - d \ln\Big[2\,\sin(\frac{w_j}{2})\Big]^2 + \ln\Big\{\frac{f_U(w_j)}{f_U(0)}\Big\} + \ln\Big\{\frac{I(w_j)}{f_X(w_j)}\Big\}, \quad (4)$$

where $I(\cdot)$ is the periodogram function.

The estimator of d given by

$$\hat{d}_p = -\frac{\sum_{j=1}^{g(n)}(x_j - \bar{x})y_j}{\sum_{j=1}^{g(n)}(x_j - \bar{x})^2} \quad (5)$$

is asymptotically normally distributed with $E(\hat{d}_p) = d$ and $Var(\hat{d}_p) = \frac{\pi^2}{6\sum_{j=1}^{g(n)}(x_j - \bar{x})^2}$, where $g(n)$ is a function of n, $x_j = \ln\{2\sin(w_j/2)\}^2$ and $y_j = \ln I(w_j)$.

THE ESTIMATOR $\hat{d}_{sp}$

This regression estimator is obtained by replacing the spectral density function in the expression (4) by the smoothed periodogram function with the Parzen lag window. Reisen(1994) shows that $\hat{d}_{sp}$, given by the same expression as in (5) but now with $y_j = \ln f_s(w_j)$, is asymptotically normally distributed when $d \in (-0.5, 0.5)$ with $E(\hat{d}_{sp}) = d$ and $Var(\hat{d}_{sp}) \approx 0.539285\frac{m}{n\sum_{j=1}^{g(n)}(x_j - \bar{x})^2}$, where m is a function of n and usually referred to as the truncation point in the Parzen lag window ($m = n^{\beta}, 0 < \beta < 1$).

THE ESTIMATOR $\hat{d}_{pr}$

Robinson (1995) suggests that the range of j in the equation given by the expression (4) should be $j = l, l+1, \cdots, m$ with $l > 1$ where m is given by

$$m = \begin{cases} A(d,\alpha)\, n^{\frac{2\alpha}{2\alpha+1}}, & \text{if } d \in (0.0, 0.25), \\ A(d,\alpha)\, n^{\frac{\alpha}{\alpha+1-2d}}, & \text{if } d \in (0.25, 0.5), \end{cases} \tag{6}$$

with $\alpha \in (0, 2]$ and n the sample size. In the simulations presented in Section 5 we consider $A(d, \alpha) = 1$.

Considering the expression (4) with $j \in \{l, l+1, \cdots, m\}$, the estimator based on the least square method is the estimator of d due to Robinson (1995) and we shall denote it by $\hat{d}_{pr}$. In Theorem 1 of Robinson (1995), one can find some results related to the asymptotic properties of the estimator $\hat{d}_{pr}$, and this theorem does not imply that choosing $l > 1$ in the regression equation is essential in order that this estimator achieves good asymptotic properties, but the author suggests that this may be a desirable practical policy. Considering $l = 1$ is based to the fact that the random variables $e_j = \ln\left\{\frac{I(w_j)}{f_X(w_j)}\right\} + I\!E\left\{-\ln\frac{I(w_j)}{f(w_j)}\right\}$, $j = 1, 2, \cdots, g(n)$ are approximately uncorrelated and identically distributed. However, the Theorem 1 in Robinson (1995) consider that, for $d \in (-0.5, 0.5)$, the variables e_j are asymptotically either uncorrelated or identically distributed when $n \to \infty$ but j stays fixed.

THE ESTIMATOR $\hat{d}_W$

The estimator of d due to Whittle (1951) is also based on the periodogram. It involves the function

$$Q(\eta) = \int_{-\pi}^{\pi} \frac{I(w)}{f_X(w;\eta)} dw$$

where $I(\cdot)$ is the periodogram and $f_X(w;\eta)$ is the espectral density function of the $\{X_t\}_{t\in\mathbb{Z}}$, and η denote the vector of unknown parameters. The Whittle estimator is the value of η which minimizes the function $Q(\cdot)$. When we are dealing with $ARFIMA(0,d,0)$ models, η is given only by the parameter d. For more details see Fox and Taqqu (1986), Beran (1994) and Dahlhaus (1989). For computational purposes, it is easier to minimize the function

$$\mathcal{L}_n(\eta) = \frac{1}{2n} \sum_{j=1}^{n-1} \{\ln f_X(w_j;\eta) + \frac{I(w_j)}{f_X(w_j;\eta)},$$

where w_j are the Fourier frequencies, for $j = 1, \cdots, n$. Dahlhaus (1989) for general $ARFIMA(0,d,0)$ Gaussian processes, has shown that the maximum likelihood estimator of d is strongly consistent, asymptotically normally distributed and asymptotically efficient in the Fisher sense. This estimator requires more computational time to be obtained.

4. ESTIMATION OF THE DEGREE OF DIFFERENCING IN NONSTATIONARY $ARFIMA(0,d^*,0)$ PROCESS

We have considered in Section 3 stationary stochastic processes. However, in practice, one often deals with nonstationary processes. In this section we present a picture of one time series obtained from a nonstationary process. For the estimation procedure we consider the estimators described previously.

Let us now consider the class of process $\{X_t\}_{t\in\mathbb{Z}}$ $ARFIMA(0,d^*,0)$, that is, processes given by

$$(1-\mathcal{B})^{d^*} X_t = \epsilon_t, \text{ with } d^* = d + r, \tag{7}$$

where $d \in (0.0, 0.5)$ and $r > 0$ such that $d^* \in (0.5, 1.5)$. This equation can be rewritten as

$$Y_t = (1-\mathcal{B})^r X_t, \; t \in \mathbb{Z}.$$

So

$$(1-\mathcal{B})^d Y_t = \epsilon_t, \; t \in \mathbb{Z}, \tag{8}$$

is an $ARFIMA(0,d,0)$ process.

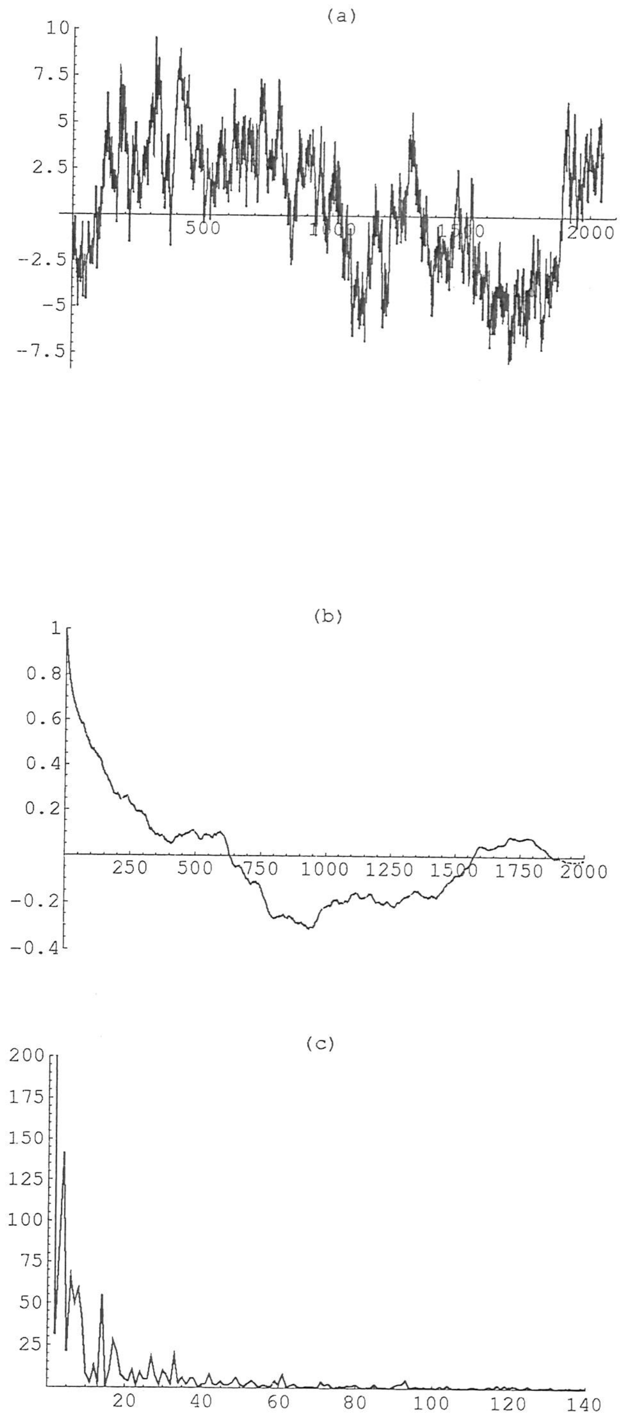

Figure 1: (a) $\{X_t\}_{t=1}^{2048}$ obtained from the process given by the expression (7), when $d^* = 0.8$, with $d = 0.3$, $r = 0.5$ and $\sigma_\epsilon^2 = 1.0$; (b) the sample autocorrelation of the process and (c) the periodogram of the process.

The $ARFIMA(0,d,0)$ processes $\{Y_t\}_{t \in \mathbb{Z}}$ were simulated as suggested by Hosking (1981) with $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The process $\{X_t\}_{t \in \mathbb{Z}}$ were obtained through the algebraic form $X_t = (1-\mathcal{B})^{-r} Y_t$, for $t \in \mathbb{N} - \{0\}$ with $X_1 = Y_1$.

The Figure 1 shows a time series $\{X_t\}_{t=1}^{2048}$, the sample autocorrelation and periodogram obtained from the process given by the expression (7), when $d^* = 0.8$, $d = 0.3$, $r = 1$ and $\sigma_\epsilon^2 = 1.0$. One can see the nonstationary property from it.

Observe that the process $\{X_t\}_{t \in \mathbb{Z}}$ given by the expression (7) doesn't have spectral density function, since it is not a stationary process. Nevertheless, we shall see that in the nonstationary case, the function $f_X(\cdot)$ given before, plays the role usually played by the spectral density function in determining some of the statistical properties of the periodogram function (more details can be found in Hurvich and Ray (1995)).

5. SIMULATIONS

We compare the performance of estimating d^* for the $ARFIMA(0, d^*, 0)$ process. The results were obtained by considering $n = 2^p$ observations, where $p = 8, 9, 10, 11$ over 300 replications. In the tables we have the sample size n, the mean of the estimators $\overline{d_i^*}$, the mean of the standard deviation $\overline{sd(d_i^*)}$ and the mean of the mean squared error $\overline{mse(d_i^*)}$ for $i = p$, sp, pr and w.

The results on Tables 1 to 5, were obtained considering $\beta = 0.9$ in the truncation point for the smoothed periodogram function, as suggested by Reisen (1994), and for the $\hat{d}_{pr}^*$ estimator we consider $l = 2$, $\alpha = 1$ and $A(d, \alpha) = 1$ in the expression (6).

For the case when $d^* \in (0.5, 1.0)$ the estimators present good results where d_{W}^* is the best followed by d_{pr}^*. This may be explained by the fact that more frequencies are involved in the regression equation used to obtain the Robinson estimator which improved the estimation procedure. The d_{sp}^* has better performance than d_p^* in terms of mean squared error, as it was expected since d_{sp}^* uses the smoothed periodogram to estimate the spectral density function. As n increases the estimators get even better.

In the case when $d^* > 1.0$, the estimators work poorly since now the level reversion property does not hold. None of the methods used in this simulation study seem to be apropriated when $d^* > 1$. One option is to apply first difference and hence estimating d^*. However this procedure must be investigated and it will be a topic for future research.

The box-plots shown in Figures 2, 3 and 4 illustrate results of the estimators for $d^* = 0.6$, 0.8, 1.3 respectively with $n = 2048$. From Figures 2 and 3 it appears that all the estimators have small bias. From Figure 4 we can see that d_{W}^* has less variance than the other estimators.

Table 1: Estimators of the parameter d^* when $d^* = 0.6$.

n	i	$\overline{d_i^*}$	$\overline{sd(d_i^*)}$	$\overline{mse(d_i^*)}$
256	p	0.6074	0.1984	0.0393
	sp	0.5517	0.1714	0.0316
	pr	0.6121	0.1564	0.0245
	w	0.6011	0.0576	0.0033
512	p	0.6207	0.1674	0.0284
	sp	0.5772	0.1435	0.0210
	pr	0.6172	0.1112	0.0126
	w	0.6099	0.0399	0.0017
1024	p	0.6079	0.1282	0.0164
	sp	0.5797	0.1175	0.0142
	pr	0.6077	0.0820	0.0068
	w	0.6040	0.0253	0.0006
2048	p	0.6102	0.1163	0.0136
	sp	0.5896	0.1005	0.0102
	pr	0.6081	0.0654	0.0043
	w	0.6042	0.0181	0.0003

Table 2: Estimators of the parameter d^* when $d^* = 0.8$.

n	i	$\overline{d_i^*}$	$\overline{sd(d_i^*)}$	$\overline{mse(d_i^*)}$
256	p	0.8397	0.2008	0.0418
	sp	0.7831	0.1716	0.0296
	pr	0.8272	0.1284	0.0172
	w	0.8224	0.0598	0.0040
512	p	0.8343	0.1592	0.0264
	sp	0.7918	0.1463	0.0214
	pr	0.8358	0.0859	0.0086
	w	0.8194	0.0430	0.0022
1024	p	0.8352	0.1409	0.0210
	sp	0.8130	0.1195	0.0144
	pr	0.8272	0.0663	0.0052
	w	0.8200	0.035	0.0016
2048	p	0.8237	0.1123	0.0131
	sp	0.8063	0.0963	0.0093
	pr	0.8269	0.0539	0.0036
	w	0.8160	0.0253	0.0009

Table 3: Estimators of the parameter d^* when $d^* = 1.1$.

n	i	$\overline{d_i^*}$	$\overline{sd(d_i^*)}$	$\overline{mse(d_i^*)}$
256	p	1.0627	0.1789	0.0333
	sp	1.0423	0.1346	0.0214
	pr	1.0547	0.1234	0.0173
	w	1.0371	0.0559	0.0071
512	p	1.0538	0.1516	0.0250
	sp	1.0567	0.1189	0.0160
	pr	1.0474	0.0932	0.0115
	w	1.0454	0.0418	0.0047
1024	p	1.0453	0.1154	0.0163
	sp	1.0725	0.0944	0.0096
	pr	1.0353	0.0703	0.0091
	w	1.0397	0.0330	0.0047
2048	p	1.0533	0.0991	0.0120
	sp	1.0763	0.0727	0.0058
	pr	1.0401	0.0559	0.0067
	w	1.0422	0.0320	0.0043

Table 4: Estimators of the parameter d^* when $d^* = 1.3$.

n	i	$\overline{d_i^*}$	$\overline{sd(d_i^*)}$	$\overline{mse(d_i^*)}$
256	p	1.0771	0.1807	0.0822
	sp	1.1262	0.1222	0.0451
	pr	1.0577	0.1100	0.0715
	w	1.0623	0.0867	0.0640
512	p	1.0781	0.1568	0.0738
	sp	1.1235	0.1070	0.0426
	pr	1.0469	0.0959	0.0732
	w	1.0625	0.0891	0.0643
1024	p	1.0616	0.1208	0.0714
	sp	1.1243	0.0816	0.0375
	pr	1.0529	0.0938	0.0698
	w	1.0509	0.0861	0.0694
2048	p	1.0617	0.1166	0.0704
	sp	1.1228	0.0788	0.0376
	pr	1.0403	0.0872	0.0750
	w	1.0535	0.1098	0.0727

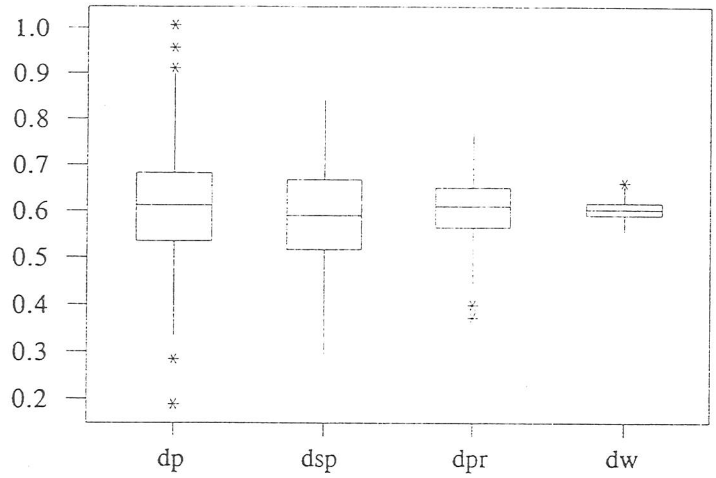

Figure 2: Estimators of d^*, when $d^* = 0.6$ and sample size $n = 2048$.

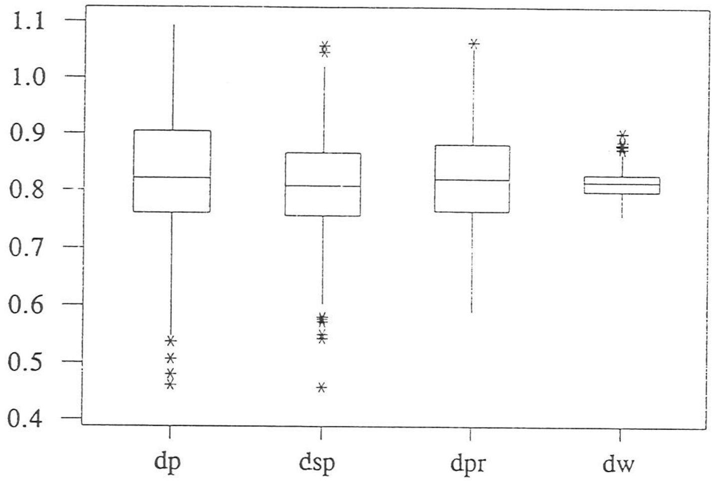

Figure 3: Estimators of d^*, when $d^* = 0.8$ and sample size $n = 2048$.

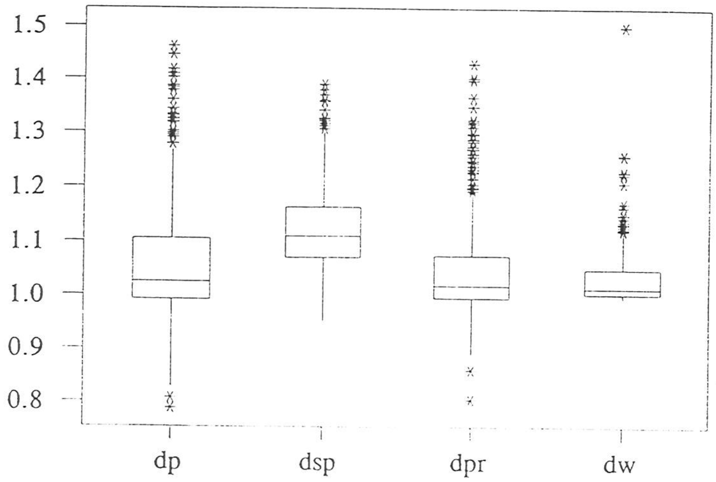

Figure 4: Estimators of d^*, when $d^* = 1.3$ and sample size $n = 2048$.

Table 5: Estimators of the parameter d^* when $d^* = 1.45$.

n	i	$\overline{d_i^*}$	$\overline{sd(d_i^*)}$	$\overline{mse(d_i^*)}$
256	p	1.0424	0.1327	0.1837
	sp	1.1236	0.0786	0.01127
	pr	1.0149	0.0625	0.1932
	w	1.0211	0.0803	0.1903
512	p	1.0454	0.1294	0.1804
	sp	1.1243	0.0775	0.1121
	pr	1.0182	0.0667	0.1909
	w	1.0269	0.0831	0.1858
1024	p	1.0305	0.1011	0.1861
	sp	1.1095	0.0644	0.1201
	pr	1.0095	0.0398	0.1957
	w	1.0183	0.0686	0.1910
2048	p	1.0341	0.1026	0.1835
	sp	1.1068	0.0671	0.1223
	pr	1.0070	0.0294	0.1971
	w	1.0284	0.0996	0.1875

The estimation of d^*, considering $\beta = 0.8$, generates smaller mean squared error, but greater standard deviation compared with the results obtained when $\beta = 0.9$. We also considered $\beta = 0.7$ but the results obtained were not better than those when $\beta \in \{0.8, 0.9\}$. For the sake of brevity we do not present here these simulations. For other simulations with results and respective analysis we refer the reader to Olbermann (1998).

Now, we present the results obtained with the modifications suggested by Robinson (1994) to estimate the degree of differencing through the method of regression using the periodogram function with more values of l and m in the expression (6).

In the following two tables we consider the same notation as before. Observe that the results for the estimator $\hat{d}_p^*$ appear in bold face on the first line for every sample size n in Tables 6 and 7. This makes easier to compare the estimators $\hat{d}_p^*$ and $\hat{d}_{pr}^*$.

In Table 6 we consider $l \in \{1, 2\}$ and $\alpha \in \{0.5, 1, 2\}$ in the expression (6). When $\alpha = 0.5$ we have $m = g(n) = n^{0.5}$, which is the value of $g(n)$ used before in Tables 1 to 5 for $\hat{d}_p^*$.

Table 6: Estimators of the parameter d^* when $d^* = 0.7$. The first line for every different value of n, given in bold face characters, presents the estimator d_p^*.

n	m	l	$\overline{d_{pr}^*}$	$\overline{sd(d_{pr}^*)}$	$\overline{mse(d_{pr}^*)}$
256	**16**	**1**	**0.7143**	**0.0668**	**0.0447**
		2	0.7345	0.2817	0.04176
	40	1	0.7380	0.1162	0.0149
		2	0.7314	0.1353	0.0197
	84	1	0.7142	0.0848	0.0074
		2	0.7097	0.0937	0.0088
512	**22**	**1**	**0.7316**	**0.1670**	**0.0288**
		2	0.7130	0.225	0.0506
	64	1	0.7141	0.0920	0.0086
		2	0.7194	0.1044	0.0113
	147	1	0.7141	0.0620	0.0040
		2	0.7212	0.0732	0.0058
1024	**32**	**1**	**0.7283**	**0.1459**	**0.0220**
		2	0.7286	0.1552	0.0248
	101	1	0.7218	0.0748	0.0060
		2	0.7237	0.0854	0.0078
	256	1	0.7130	0.0488	0.0025
		2	0.7088	0.00511	0.0027
2048	**45**	**1**	**0.7183**	**0.1191**	**0.0145**
		2	0.7236	0.01296	0.0173
	161	1	0.7148	0.0585	0.0036
		2	0.7169	0.0630	0.0042
	445	1	0.7148	0.0348	0.0014
		2	0.7138	0.0366	0.0015

In Table 7 we consider $l \in \{1, 2\}$ and $\alpha \in \{0.4, 0.5, 1, 2\}$ in the expression (6). When $\alpha = 0.4$ we get $g(n) = n^{0.5}$, the same value of $g(n)$ used in Tables 1 to 5.

Table 7: Estimators of the parameter d^* when $d^* = 1.3$. The first line for every different value of n, given in bold face characters, presents the estimator d^*_p.

n	m	l	$\overline{d^*_{pr}}$	$\overline{sd(d^*_{pr})}$	$\overline{mse(d^*_{pr})}$
256	**16**	**1**	**1.0771**	**0.1807**	**0.0822**
		2	1.0760	0.1932	0.0873
	21	1	1.0835	0.1711	0.0758
		2	1.0801	0.1688	0.0767
	52	1	1.0699	0.1149	0.0661
		2	1.0577	0.1100	0.0715
	101	1	1.0563	0.1039	0.0701
		2	1.0552	0.1053	0.0710
512	**22**	**1**	**1.0781**	**0.1568**	**0.0738**
		2	1.0654	0.1562	0.0793
	32	1	1.0756	0.1425	0.0705
		2	1.0694	0.1646	0.0802
	85	1	1.0541	0.1063	0.0718
		2	1.0469	0.0959	0.0732
	181	1	1.0543	0.0889	0.0683
		2	1.0404	0.0883	0.0752
1024	**32**	**1**	**1.0616**	**0.1208**	**0.0714**
		2	1.0764	0.1326	0.0675
	47	1	1.0568	0.1025	0.0696
		2	1.0650	0.0121	0.0711
	141	1	1.0475	0.0859	0.0711
		2	1.0529	0.0938	0.0698
	322	1	1.0322	0.0728	0.0769
		2	1.0357	0.0755	0.0755
2048	**45**	**1**	**1.0617**	**0.1166**	**0.0704**
		2	1.0540	0.1187	0.0745
	69	1	1.0484	0.0994	0.0731
		2	1.0551	0.1044	0.0708
	231	1	1.0425	0.0824	0.0731
		2	1.0403	0.0872	0.0750
	574	1	1.0348	0.0741	0.0758
		2	1.0267	0.0640	0.0787

From Tables 6 and 7 we observe that while the value of m in the expression (6) increases, the value of the variance of the estimators decreases.

We also consider other values for $A(d, \alpha)$, in the expression (6), besides the value $A(d, \alpha) = 1$, but we didn't find significant differences on the results. For this reason we do not present these simulations here.

6. CONCLUSION

When the differencing parameter d^* is in the interval (0.5,1.0) the analyzed estimators behaved better than when $d^* \in (1.0, 1.5)$ as was reported by Hurvich and Ray (1995). In the first case the results obtained for the estimator $\hat{d}^*_W$ were better than any other estimator. The estimator $\hat{d}_{pr}$ is valuable in the sense of the asymptotic properties, even though for practical purposes it is useless since it depends on the value of the parameter. So, one will need iterative procedures when dealing with the estimator $\hat{d}_{pr}$.

In the nonstationary case with no level-reversion property, that is, when $d^* \in (1.0, 1.5)$, we observe that all different estimators for the differencing parameter underestimate it. This was also reported by Hurvich and Ray (1995) for the estimators $\hat{d}_p$ and $\hat{d}_{pr}$ analyzed by them.

ACKNOWLEDGEMENTS

V.A. Reisen was partially supported by CNPq-Brazil. S. Lopes was partially supported by CNPq-Brazil and also by Pronex *Fenômenos Críticos em Probabilidade e Processos Estocásticos* (ConvênioFINEP/MCT/CNPq 41/ 96/0923/00).

REFERENCES

[1] Beran, J., *Statistics for Long-Memory Processes* (Chapman & Hall, New York, 1994).

[2] Brockwell, P.J. and R.A. Davis, *Time Series: Theory and Methods* (Springer-Verlag, New York, 1987).

[3] Cheung,Y.-W. and K.S. Lai, A Fractional Cointegration Analysis of Purchasing Power Parity,*Journal of Business & Economic Statistics*, **11** (1993) 103-112.

[4] Dahlhaus, R, Efficient parameter estimation for self-similar processes, *The Annals of Statistics*, **17** (1989) 1749-1766.

[5] Fox, R. and M. S. Taqqu, Large-sample properties of parameter estimates for strongky dependent stationary Gaussian time series, *The Annals of Statistics*, **14** (1986) 517-532.

[6] Geweke, J. and S. Porter-Hudak, The Estimation and Application of Long Memory Time Series Model, *Journal of Time Series Analysis*, **4**(4) (1983) 221-238.

[7] Hosking, J., Fractional Differencing, *Biometrika*,**68**(1) (1981) 165-167.

[8] Hosking, J., Modelling Persistence in Hydrological Time Series using Fractional Differencing, *Water Resources Research*, **20**(12) (1984) 1898-1908.

[9] Hurst, H.R., Long-term storage in reservoirs, *Trans. Am. Soc. Civil Eng.* **116** (1951) 770-799.

[10] Hurvich, C.M. and R.S. Deo, Plug-in Selection of the Number of Freqüências in Regression Estimates of the Memory Parameter of a Long-Memory Time Series, *Journal of Time Series Analysis*, **20**(3) (1999) 331-341.

[11] Hurvich, C.M. and B.K. Ray, Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes, *Journal of Time Series Analysis*, **16**(1) (1995) 017-042.

[12] Olbermann, B.P., *Estimação do Parâmetro de Diferenciação em Séries Temporais Não Estacionárias com Características de Longa Dependência*, Dissertação de Mestrado. Instituto de Matemática da UFRGS, Porto Alegre (1998).

[13] Reisen, V.A., Estimation of the Fractional Difference Parameter in the ARIMA(p,d,q) model using the Smoothed Periodogram, *Journal of Time Series Analysis*, **15**(3) (1994) 335-350.

[14] Robinson, P.M., Rates of convergence and optimal spectral bandwidth for long range dependence, *Probability Theory and Related Fields*, **99** (1994) 443-473.

[15] Robinson, P.M., Log-periodogram regression of time series with long range dependence, *Annals of Statistics*, **23**(3) (1995) 1630-1661.

[16] Whittle, P.,*Hypothesis Testing in Time Series Analysis* (Hafner, New York, 1951).

[17] Wu, P. and N. Crato, New Tests for Stationarity and Parity Reversion: Evidence on New Zeland Real Exchange Rates, *Empirical Economics*, **20** (1995) 559-613.