

**NON-STATIONARY GAUSSIAN ARFIMA PROCESSES:
ESTIMATION AND APPLICATION**

Lopes, S.R.C.^{a1}, Olbermann, B.P.^a and Reisen, V.A.^b

^a Instituto de Matemática - UFRGS, Porto Alegre, RS, Brazil

^b Departamento de Estatística - UFES, Vitória, ES, Brazil

September 30, 2002

Abstract:

Recently, the study of time series turned the attention to the ones having *long memory* property. The $ARFIMA(p, d, q)$ model shows this property 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 $ARFIMA(p, d^*, q)$ processes when $d^* > 0.5$, that is, when the processes are non-stationary but still have the property of long memory. We present a simulation study for the estimators of d^* with semiparametric and parametric methods and different sample sizes. The methodology is applied to the experimental data series of UK long interest gilts.

Keywords: Non-stationary Processes, Long Memory, Estimation.

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

1. INTRODUCTION

Persistency, or long memory, in a time series is the presence of significant 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 (see Beran (1994) and Hosking (1981 and 1984) for more details).

Persistency may be characterized,

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

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

The contribution due to Geweke and Porter-Hudak (1983), introducing the GPH 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 SPR estimator. Robinson (1994, 1995), making use of mild modifications on GPH, 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 GPHt. Hurvich and Deo (1999) 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 FT. They adapted the approach by 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. Yajima (1985) discusses estimation associated with the long memory time series considering the maximum likelihood and least squares estimators. Cheung (1993), using Monte Carlo methods, examines the small samples properties of tests for the memory parameter. The papers by Yajima (1985) and Cheung (1993) discussed the estimation and tests for the memory parameter in the stationary case.

Liu (1998) studies the asymptotic theory of non-stationary *ARFIMA* process with special attention to the unit root KPSS test. Velasco (1999) showed that it is possible to estimate consistently the memory of non-stationary and non-invertible processes using methods of log-periodogram designed for stationary cases.

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 non-stationary or non-invertible 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 the periodogram bias only on the GPH estimator in finite samples.

The main goal of this paper is to evaluate estimators of the degree of differencing d in the case where the stochastic processes are non-stationary, that is, when $d > 0.5$.

In Section 2 we summarize the properties of the $ARFIMA(p, d, q)$ process, for $d \in (-0.5, 0.5)$. In Section 3 the estimation methods are presented and Section 4 gives the non-stationary processes. Section 5 presents simulation results and the application is in Section 6. Conclusions and future work are given in Section 7.

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

DEFINITION 1: Let $\{\epsilon_t\}_{t \in \mathbb{Z}}$ be the white noise process with zero mean 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

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

and

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

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

$$\Phi(\mathcal{B})(1 - \mathcal{B})^d X_t = \Theta(\mathcal{B})\epsilon_t, \quad t \in \mathbb{Z}, \quad (2.1)$$

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

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

given by

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

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

If $d \in (-0.5, 0.5)$ then the process $\{X_t\}_{t \in \mathbb{Z}}$ is stationary and invertible and its spectral density function is given by

$$f_X(w) = f_U(w) \left[2 \sin\left(\frac{w}{2}\right) \right]^{-2d}, \quad \text{for } 0 < w \leq \pi, \quad (2.2)$$

where $f_U(\cdot)$ is the spectral density function of the $ARMA(p, q)$ process. One observes that $f_X(w) \simeq w^{-2d}$ when $w \rightarrow 0$.

The term $(1 - \mathcal{B})^d$, in the expression (2.1), is the binomial expansion

$$(1 - \mathcal{B})^d = \sum_{k=0}^{\infty} \binom{d}{k} (-\mathcal{B})^k = 1 - d\mathcal{B} - \frac{d}{2!}(1-d)\mathcal{B}^2 \cdots, \text{ for } d \in \mathbb{R}. \quad (2.3)$$

When $p = 0 = q$, in the expression (2.1), one has the $ARFIMA(0, d, 0)$ process.

The $ARFIMA(p, d, q)$ process exhibits the property of long memory when $d \in (0.0, 0.5)$, of intermediate memory when $d \in (-0.5, 0.0)$ and of short memory when $d = 0$. If $d \geq 0.5$ the $ARFIMA$ process is non-stationary 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 non-invertible. The reader can find more properties of the stationary $ARFIMA(p, d, q)$ process in Hosking (1981).

3. ESTIMATES OF THE DIFFERENCING PARAMETER

We now summarize some methods for the estimation of d : the regression methods using the periodogram function and the smoothed version of the periodogram function; the regression method using the modifications suggested by Robinson (1995) and the approximated maximum likelihood estimator of the differencing parameter using the approach suggested by Whittle (1951).

ESTIMATOR GPH

Consider the set of harmonic frequencies $w_j = \frac{2\pi j}{n}$, $j = 0, 1, \dots, [n/2]$, where n is the sample size and $[x]$ means the integer part of x . By taking the logarithm of the spectral density function $f_X(\cdot)$ and adding $\ln f_U(0)$ and $\ln I(w_j)$ to both sides of the expression (2.2) we have

$$\ln I(w_j) = \ln f_U(0) - d \ln \left[2 \sin\left(\frac{w_j}{2}\right) \right]^2 + \ln \left\{ \frac{f_U(w_j)}{f_U(0)} \right\} + \ln \left\{ \frac{I(w_j)}{f_X(w_j)} \right\}, \quad (3.1)$$

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

The estimator of d is given by

$$\text{GPH} = - \frac{\sum_{j=1}^{g(n)} (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{g(n)} (x_j - \bar{x})^2}, \quad (3.2)$$

where $g(n) = n^\alpha$, $0 < \alpha < 1$ (see Geweke and Porter-Hudak (1983)), $y_j = \ln I(w_j)$, $x_j = \ln\{2 \sin(w_j/2)\}^2$ and $\bar{x} = \frac{1}{g(n)} \sum_{j=1}^{g(n)} x_j$.

ESTIMATOR SPR

The regression estimator SPR is obtained by replacing the spectral density function in the expression (3.1) by the smoothed periodogram function, $f_s(\cdot)$, with the Parzen lag window. Reisen (1994) shows that SPR is obtained by the same expression as in (3.2), where now $y_j = \ln f_s(w_j)$, for $j = 1, \dots, g(n)$. The value of $g(n)$ is chosen as in the GPH method. The effect of different truncation points in the Parzen window is also investigated here for this estimator.

ESTIMATOR GPHt

This estimator regresses $\ln\{I(w_j)\}$ on $\ln\{2 \sin(w_j/2)\}^2$, for $j = l, l + 1, \dots, g(n)$, where l is the lower truncation point. In the asymptotic theory both l and $g(n)$ tend to infinity with n , but more slowly, while $\frac{l}{g(n)}$ goes to zero. Details of this method may be found in Robinson (1995).

ESTIMATOR FT

This estimator involves the function

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

where $f_X(\cdot; \eta)$ is the spectral density function of the $\{X_t\}_{t \in \mathbb{Z}}$, and η denote the vector of unknown parameters. The FT estimator is the value of η which minimizes the function $Q(\cdot)$. When we are dealing with $p = 0 = q$, η is given only by the parameter d . For computational purposes, the function

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

is an approximation of $Q(\cdot)$, where w_j are the Fourier frequencies, for $j = 1, \dots, n - 1$. More details of this estimator can be found in Fox and Taquq (1986). For general $ARFIMA(p, d, q)$ Gaussian processes, Dahlhaus (1989) has shown that the maximum likelihood estimator of d is strongly consistent, asymptotically normally distributed and asymptotically efficient in the Fisher sense.

4. ESTIMATION OF THE DEGREE OF DIFFERENCING IN NON-STATIONARY $ARFIMA(p, d^*, q)$ PROCESS

Let $\{\tilde{X}_t\}_{t \in \mathbb{Z}}$ be an $ARFIMA(0, d^*, 0)$ process where $d^* = d + r$, with $d \in (0.0, 0.5)$ and $r > 0$ such that $d^* \in (0.5, 1.5)$. Hence, we have

$$Y_t = (1 - \mathcal{B})^r \tilde{X}_t, \quad t \in \mathbb{Z},$$

so that $\{Y_t\}_{t \in \mathbb{Z}}$ is an $ARFIMA(0, d, 0)$ process. The time series $\{Y_t\}_{t=1}^n$ was simulated as suggested by Hosking (1981) with $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The processes $\{\tilde{X}_t\}_{t \in \mathbb{Z}}$ were obtained through the algebraic form $\tilde{X}_t = (1 - \mathcal{B})^{-r} Y_t$, for $t \in \mathbb{N} - \{0\}$, with $\tilde{X}_1 = Y_1$. We use Fortran (IMSL) subroutines for simulations and estimations results.

The $ARFIMA(p, d^*, q)$ processes were simulated including the autoregressive and moving average components in the $\{\tilde{X}_t\}_{t \in \mathbb{Z}}$ processes.

Observe that the process $\{\tilde{X}_t\}_{t \in \mathbb{Z}}$ doesn't have spectral density function, since it is not a stationary one. Nevertheless, we shall see that in the non-stationary case, the function $f_{\tilde{X}}(\cdot)$, given by a similar expression as (2.2), 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

In this simulation exercise we consider $ARFIMA(p, d^*, q)$ processes for both situations when $p = 0 = q$ and when $p \neq 0$ or $q \neq 0$. The results were obtained by time series with $n = 2^m$ observations, where $m = 8, 9, 10$, over 2,000 replications. In Tables 5.1 to 5.4, one finds the values of the mean, the standard deviation and the mean squared error for GPH, SPR, GPHt and FT estimators. The smallest mean squared error value among all estimation methods is in bolded face character.

The results were obtained considering $\beta = 0.9$ in the truncation point for the smoothed periodogram function in the SPR estimator, as suggested by Reisen (1994). For GPH and SPR estimators we consider $g(n) = n^\alpha = n^{0.5}$. For the GPHt estimator we consider two different limits for $g(n)$: for GPHt(1), we consider $\alpha_1 = 0.6$ and for GPHt(2), $\alpha_2 = 0.7$. In both cases, the lower truncation point was taken $l = 2$.

Estimation in $ARFIMA(0, d^*, 0)$

In Tables 5.1 and 5.2 we consider the estimation of $d^* \in \{0.6, 0.8, 1.0, 1.45\}$. Other values of d^* were also considered but are not presented here to conceive space and the results can be requested.

For the case when $d^* \in (0.5, 1.0)$, see Table 5.1, all estimators present good results, in the sense of minimizing the bias and the mean squared error values, where FT is the best estimator followed by GPHt(2). We believe that the reason for GPHt(2) outperforming the others estimators (among the class of regression estimators) is because more frequencies are involved in the regression equation used to obtain it. The SPR estimator has better performance than GPH in terms of small mean squared error value, which is expected, since SPR uses the smoothed periodogram function to estimate the spectral density function. As n increases the estimators get even better. Except for the SPR estimator, all the methods tend to overestimate the true parameter.

The value $d^* = 1.0$ is exactly the frontier between level and no level reversion properties. The results indicate that all methods perform reasonable well, are very competitive and underestimate d^* (see Table 5.2, $d^* = 1.0$).

For the case $d^* > 1.0$, since the level reversion property does not hold, all estimators perform poorly, are strongly biased with large mean squared error values compared with the case when $d^* < 1.0$ (see Table 5.2, $d^* = 1.45$). To

Table 5.1: Estimation results in the $ARFIMA(0, d^*, 0)$ when $d^* \in \{0.6, 0.8\}$.

n	$Method$	$d^* = 0.6$			$d^* = 0.8$		
		$mean$	sd	mse	$mean$	sd	mse
256	GPH	0.6183	0.2222	0.0497	0.8315	0.2069	0.0437
	SPR	0.5566	0.1773	0.0333	0.7780	0.1779	0.0321
	GPHt(1)	0.6136	0.2004	0.0403	0.8390	0.2046	0.0433
	GPHt(2)	0.6147	0.1321	0.0176	0.8372	0.1384	0.0205
	FT	0.6076	0.0552	0.0031	0.8249	0.0582	0.0040
512	GPH	0.6186	0.1730	0.0303	0.8332	0.1651	0.0283
	SPR	0.5712	0.1443	0.0216	0.7975	0.1419	0.0201
	GPHt(1)	0.6232	0.1508	0.0233	0.8373	0.1510	0.0242
	GPHt(2)	0.6143	0.0958	0.0094	0.8306	0.1005	0.0110
	FT	0.6039	0.0406	0.0017	0.8231	0.0439	0.0025
1024	GPH	0.6155	0.1408	0.0201	0.8328	0.1269	0.0171
	SPR	0.5871	0.1180	0.0141	0.8072	0.0909	0.0091
	GPHt(1)	0.6109	0.1125	0.0128	0.8303	0.1134	0.0137
	GPHt(2)	0.6102	0.0730	0.0054	0.8289	0.0731	0.0062
	FT	0.6057	0.0259	0.0007	0.8190	0.0317	0.0014
2048	GPH	0.6102	0.1163	0.0136	0.8193	0.0977	0.0097
	SPR	0.5886	0.0970	0.0095	0.8110	0.0919	0.0084
	GPHt(1)	0.6081	0.0829	0.0069	0.8195	0.0939	0.0090
	GPHt(2)	0.6042	0.0524	0.0028	0.8276	0.0647	0.0049
	FT	0.6035	0.0188	0.0004	0.8143	0.0225	0.0007

Table 5.2: Estimation results in the $ARFIMA(0, d^*, 0)$ when $d^* \in \{1.0, 1.45\}$.

n	$Method$	$d^* = 1.0$			$d^* = 1.45$		
		$mean$	sd	mse	$mean$	sd	mse
256	GPH	0.9911	0.1843	0.0340	1.0561	0.1482	0.1771
	SPR	0.9739	0.1557	0.0249	1.1313	0.0891	0.1095
	GPHt(1)	0.9933	0.1825	0.0333	1.0304	0.1207	0.1906
	GPHt(2)	0.9947	0.1162	0.0135	1.0239	0.0886	0.1894
	FT	0.9957	0.0469	0.0022	1.0281	0.0894	0.1860
512	GPH	0.9775	0.1522	0.0236	1.0462	0.1239	0.1783
	SPR	0.9761	0.1308	0.0176	1.1254	0.0756	0.1110
	GPHt(1)	0.9936	0.1325	0.0175	1.0283	0.1066	0.1891
	GPHt(2)	0.9964	0.0785	0.0062	1.0230	0.0947	0.1912
	FT	0.9964	0.0314	0.0010	1.0255	0.0886	0.1880
1024	GPH	0.9637	0.1436	0.0219	1.0327	0.1082	0.1857
	SPR	0.9766	0.1355	0.0189	1.1118	0.0550	0.1174
	GPHt(1)	0.9674	0.1336	0.0189	1.0268	0.0970	0.1884
	GPHt(2)	0.9697	0.1111	0.0132	1.0224	0.0888	0.1906
	FT	0.9721	0.0930	0.0094	1.0325	0.1352	0.1924
2048	GPH	0.9932	0.0970	0.0094	1.0339	0.0977	0.1826
	SPR	1.0083	0.0807	0.0066	1.1024	0.0605	0.1244
	GPHt(1)	0.9971	0.599	0.0036	1.0196	0.0741	0.1907
	GPHt(2)	0.9914	0.0401	0.0017	1.0160	0.0646	0.1925
	FT	0.9959	0.0152	0.0002	1.0308	0.1376	0.1946

overcome this, one idea is to apply first difference to the process and then to reestimate d . As investigated by Hurvich and Ray (1995) the GPH estimator is not, in general, invariant under first difference. The estimated \hat{d}^* based on the original data is not, in general, equal to one plus the estimated \hat{d} based on the differenced data. However, this property still remains to be investigated for other fractional parameter estimators and this is the subject of our forthcoming paper.

For the SPR estimator we also investigated the effect of the truncation point in the Parzen lag window, $m = n^\beta$, when $\beta = 0.8$. The estimation of d^* generates smaller mean squared error values, 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. An extensive simulation study is in Olbermann (1998) and the results are available upon request.

For the GPHt estimator we observed that, for fixed $g(n)$, as l increases the standard deviation and mean squared error values also increase. However, as n gets large, the use of $l = 1$ or $l = 2$ does not make significant changes in terms of the mean squared error values. We also consider $l = 3, 4$ in the

GPHT estimator but the results are worse compared with $l = 1, 2$ (these results are available upon request). This may be explained by the fact that when $l = 3, 4$ one is using the initial frequency for this estimator away from the zero frequency where the long memory property holds.

Estimation in $ARFIMA(p, d^*, q)$

In the case where $p \neq 0$ or $q \neq 0$ we have conducted several simulations for different sizes of n , for different orders p and q and different values of d^* , but we present here only some cases. Other results are available upon request. In Table 5.3 we present the results for $ARFIMA(1, 0.8, 0)$, where $\phi \in \{-0.6, -0.2, 0.2, 0.6\}$ and $n = 256, 1024$. In Table 5.4 we present the results for $ARFIMA(0, 0.8, 1)$, where $\theta \in \{-0.6, -0.2, 0.2, 0.6\}$ and the same values for n . In the FT method, the parameters of the processes are estimated simultaneously by using BCONF subroutine in the IMSL library. For the semiparametric methods, the short-run parameters are estimated using the NSLSE subroutine after the series being differentiated by the estimate of d^* . The estimates present large impact when short-run autoregressive or moving average parameters are included in the model, mainly for the FT estimator. The FT method loses its superiority showing both the largest bias and the largest mean squared error values, mainly when $\phi > 0$ or $\theta < 0$. The bias of d^* is predominantly positive and it is negative for the AR and MA parameters. Also, increasing n the FT estimates get worse. These are not surprising since the presence of AR or MA components in the model makes larger contribution to the spectral density function of the process at certain frequencies. In all methods the bias of d^* is compensated by the bias of the AR or MA parameter. Taking into account the contamination of the short-run parameter the semiparametric methods work reasonably well being the SPR method, in general, more accurate showing smaller bias and mean squared error values. In addition, this estimator shows to have lesser impact when the short-memory parameters are in the model than the other estimators. The bias and the mean squared error values for this estimator increase but not substantially when compared to the other methods. The method GPHT(2) presented the poorest performance to estimate (d, ϕ) or (d, θ) specially for ϕ and θ positives.

The biases in the estimation of the short-run and long-run parameters are much larger for the $ARFIMA(0, d^*, 1)$ than for the $ARFIMA(1, d^*, 0)$ model (compare Tables 5.3 and 5.4).

Table 5.3: Estimation results in the ARFIMA(1, 0.8, 0) model when $\phi \in \{-0.6, -0.2, 0.2, 0.6\}$ and $n = 256, 1024$.

n	ϕ	Property	GPH	SPR	GPHt(1)	GPHt(2)	FT
256	-0.6	mean(\hat{d}^*)	0.8380	0.7788	0.8035	0.7735	0.8474
		sd(\hat{d}^*)	0.2058	0.1770	0.2127	0.1470	0.0979
		mse(\hat{d}^*)	0.0438	0.0318	0.0452	0.0223	0.0118
		mean($\hat{\phi}$)	-0.5906	-0.5670	-0.5664	-0.5711	-0.5993
		sd($\hat{\phi}$)	0.1337	0.1229	0.1629	0.1060	0.0722
		mse($\hat{\phi}$)	0.0179	0.0162	0.0276	0.0120	0.0052
	-0.2	mean(\hat{d}^*)	0.8418	0.7780	0.8271	0.8007	1.0134
		sd(\hat{d}^*)	0.2023	0.1691	0.2094	0.1399	0.4265
		mse(\hat{d}^*)	0.0426	0.0291	0.0446	0.0196	0.2273
		mean($\hat{\phi}$)	-0.2068	-0.1625	-0.1919	-0.1898	-0.1971
		sd($\hat{\phi}$)	0.1828	0.1656	0.1931	0.1330	0.0938
		mse($\hat{\phi}$)	0.0334	0.0288	0.0373	0.0178	0.0088
	0.2	mean(\hat{d}^*)	0.8458	0.7846	0.8618	0.8884	1.2413
		sd(\hat{d}^*)	0.2046	0.1715	0.2102	0.1416	0.5591
		mse(\hat{d}^*)	0.0439	0.0296	0.0480	0.0278	0.5070
		mean($\hat{\phi}$)	0.1653	0.2186	0.1521	0.1162	0.2110
		sd($\hat{\phi}$)	0.2009	0.1789	0.2010	0.1354	0.1451
		mse($\hat{\phi}$)	0.0415	0.0323	0.0427	0.0253	0.0211
	0.6	mean(\hat{d}^*)	0.9122	0.8488	1.0172	1.1131	1.3009
		sd(\hat{d}^*)	0.1999	0.1670	0.2078	0.1397	0.5554
		mse(\hat{d}^*)	0.0525	0.0302	0.0903	0.1176	0.5590
		mean($\hat{\phi}$)	0.4729	0.5315	0.3793	0.2903	0.5333
		sd($\hat{\phi}$)	0.1841	0.1571	0.1891	0.1306	0.1726
		mse($\hat{\phi}$)	0.0500	0.0293	0.0844	0.1129	0.0342
1024	-0.6	mean(\hat{d}^*)	0.8292	0.8071	0.8241	0.8082	0.9047
		sd(\hat{d}^*)	0.1372	0.1159	0.1119	0.0789	0.1131
		mse(\hat{d}^*)	0.0197	0.0135	0.0131	0.0063	0.0237
		mean($\hat{\phi}$)	-0.5994	-0.5930	-0.6021	-0.5990	-0.6000
		sd($\hat{\phi}$)	0.0816	0.0688	0.0634	0.0482	0.0279
		mse($\hat{\phi}$)	0.0066	0.0047	0.0040	0.0023	0.0007
	-0.2	mean(\hat{d}^*)	0.8283	0.8066	0.8201	0.8086	1.4551
		sd(\hat{d}^*)	0.1358	0.1171	0.1140	0.0786	0.5409
		mse(\hat{d}^*)	0.0192	0.0137	0.0134	0.0062	0.7214
		mean($\hat{\phi}$)	-0.2064	-0.1926	-0.2048	-0.2019	-0.1971
		sd($\hat{\phi}$)	0.1255	0.1116	0.1063	0.0747	0.0441
		mse($\hat{\phi}$)	0.0158	0.0125	0.0113	0.0055	0.0019
	0.2	mean(\hat{d}^*)	0.8335	0.8156	0.8291	0.8428	1.4793
		sd(\hat{d}^*)	0.1331	0.1232	0.1073	0.0782	0.5301
		mse(\hat{d}^*)	0.0188	0.0154	0.0123	0.0079	0.7418
		mean($\hat{\phi}$)	0.1707	0.1871	0.1717	0.1544	0.2000
		sd($\hat{\phi}$)	0.1382	0.1285	0.1141	0.0818	0.0642
		mse($\hat{\phi}$)	0.0199	0.0166	0.0138	0.0087	0.0041
	0.6	mean(\hat{d}^*)	0.8526	0.8318	0.9052	0.9931	1.5173
		sd(\hat{d}^*)	0.1393	0.1196	0.1128	0.0724	0.5147
		mse(\hat{d}^*)	0.0222	0.0153	0.0238	0.0425	0.7792
		mean($\hat{\phi}$)	0.5408	0.5611	0.4912	0.4065	0.5898
		sd($\hat{\phi}$)	0.1337	0.1162	0.1119	0.0743	0.0972
		mse($\hat{\phi}$)	0.0213	0.0150	0.0243	0.0429	0.0095

Table 5.4: Estimation results in the ARFIMA(0, 0.8, 1) model when $\theta \in \{-0.6, -0.2, 0.2, 0.6\}$ and $n = 256, 1024$.

n	θ	Property	GPH	SPR	GPHt(1)	GPHt(2)	FT
256	-0.6	$\text{mean}(\hat{d}^*)$	0.8306	0.7776	0.8544	0.8830	1.8474
		$\text{sd}(\hat{d}^*)$	0.2112	0.1718	0.2170	0.1419	0.2403
		$\text{mse}(\hat{d}^*)$	0.0455	0.0300	0.0500	0.0270	1.1547
		$\text{mean}(\hat{\theta})$	-0.5725	-0.6031	-0.5575	-0.5495	-0.5956
		$\text{sd}(\hat{\theta})$	0.1280	0.1013	0.1414	0.1013	0.0634
		$\text{mse}(\hat{\theta})$	0.0174	0.0102	0.0218	0.0128	0.0076
	-0.2	$\text{mean}(\hat{d}^*)$	0.8248	0.7700	0.8424	0.8613	1.6226
		$\text{sd}(\hat{d}^*)$	0.2109	0.1778	0.2087	0.1408	0.4838
		$\text{mse}(\hat{d}^*)$	0.0451	0.0325	0.0453	0.0236	0.9171
		$\text{mean}(\hat{\theta})$	-0.1635	-0.2134	-0.1489	-0.1372	-0.1940
		$\text{sd}(\hat{\theta})$	0.1939	0.1633	0.2011	0.1445	0.0967
		$\text{mse}(\hat{\theta})$	0.0389	0.0268	0.0430	0.0248	0.0094
	0.2	$\text{mean}(\hat{d}^*)$	0.8260	0.7663	0.8100	0.7677	1.0172
		$\text{sd}(\hat{d}^*)$	0.2145	0.1734	0.2104	0.1413	0.3746
		$\text{mse}(\hat{d}^*)$	0.0466	0.0312	0.0443	0.0210	0.1874
		$\text{mean}(\hat{\theta})$	0.2373	0.2195	0.2151	0.1693	0.2269
$\text{sd}(\hat{\theta})$		0.2274	0.1867	0.2235	0.1516	0.1677	
$\text{mse}(\hat{\theta})$		0.0530	0.0354	0.0501	0.0239	0.0288	
0.6	$\text{mean}(\hat{d}^*)$	0.7640	0.6992	0.6195	0.4793	0.8214	
	$\text{sd}(\hat{d}^*)$	0.2065	0.1720	0.2184	0.1520	0.1857	
	$\text{mse}(\hat{d}^*)$	0.0439	0.0397	0.0803	0.1259	0.0349	
	$\text{mean}(\hat{\theta})$	0.5570	0.5007	0.4038	0.2562	0.5841	
	$\text{sd}(\hat{\theta})$	0.1955	0.1723	0.2372	0.1779	0.1824	
	$\text{mse}(\hat{\theta})$	0.0400	0.0395	0.0947	0.1497	0.0335	
1024	-0.6	$\text{mean}(\hat{d}^*)$	0.8325	0.8127	0.8337	0.8471	1.9000
		$\text{sd}(\hat{d}^*)$	0.1319	0.1147	0.1151	0.0743	0.0000
		$\text{mse}(\hat{d}^*)$	0.0184	0.0133	0.0144	0.0077	1.2100
		$\text{mean}(\hat{\theta})$	-0.5825	-0.5933	-0.5823	-0.5763	-0.5998
		$\text{sd}(\hat{\theta})$	0.0719	0.0632	0.0651	0.0483	0.0287
		$\text{mse}(\hat{\theta})$	0.0054	0.0040	0.0045	0.0029	0.0008
	-0.2	$\text{mean}(\hat{d}^*)$	0.8322	0.8077	0.8342	0.8378	1.8976
		$\text{sd}(\hat{d}^*)$	0.1352	0.1160	0.1143	0.0753	0.0074
		$\text{mse}(\hat{d}^*)$	0.0193	0.0135	0.0142	0.0071	1.2047
		$\text{mean}(\hat{\theta})$	-0.1677	-0.1899	-0.1666	-0.1643	-0.1978
		$\text{sd}(\hat{\theta})$	0.1250	0.1068	0.1084	0.0762	0.0446
		$\text{mse}(\hat{\theta})$	0.0166	0.0114	0.0128	0.0077	0.0019
	0.2	$\text{mean}(\hat{d}^*)$	0.8183	0.7951	0.8046	0.7884	1.5340
		$\text{sd}(\hat{d}^*)$	0.1313	0.1136	0.1088	0.0690	0.3119
		$\text{mse}(\hat{d}^*)$	0.0175	0.0129	0.0118	0.0049	0.6359
		$\text{mean}(\hat{\theta})$	0.2185	0.1941	0.2040	0.1865	0.1957
$\text{sd}(\hat{\theta})$		0.1406	0.1224	0.1164	0.0759	0.0573	
$\text{mse}(\hat{\theta})$		0.0200	0.0150	0.0135	0.0059	0.0033	
0.6	$\text{mean}(\hat{d}^*)$	0.7584	0.6992	0.6290	0.4876	0.8178	
	$\text{sd}(\hat{d}^*)$	0.2109	0.1785	0.2137	0.1523	0.1721	
	$\text{mse}(\hat{d}^*)$	0.0459	0.0418	0.0746	0.1206	0.0298	
	$\text{mean}(\hat{\theta})$	0.5526	0.5006	0.4187	0.2699	0.5728	
	$\text{sd}(\hat{\theta})$	0.2046	0.1816	0.2349	0.1785	0.1961	
	$\text{mse}(\hat{\theta})$	0.0438	0.0426	0.0877	0.1406	0.0389	

Table 5.5: Estimation results in the ARFIMA(1, 0.8, 1) model when $\phi \in \{-0.6, 0.2\}$, $\theta \in \{-0.2, 0.2, 0.6\}$ and $n = 256, 1024$.

ϕ	n	θ	Property	GPH	SPR	GPHt(1)	GPHt(2)	FT
-0.6	256	-0.2	mean(\hat{d}^*)	0.8403	0.7852	0.8270	0.8090	1.0061
			sd(\hat{d}^*)	0.2089	0.1706	0.2085	0.1386	0.4184
			mse(\hat{d}^*)	0.0452	0.0293	0.0442	0.0193	0.2174
			mean($\hat{\phi}$)	-0.4934	-0.5312	-0.4972	-0.5590	-0.5856
			sd($\hat{\phi}$)	0.2505	0.2366	0.2656	0.1618	0.1289
			mse($\hat{\phi}$)	0.0740	0.0606	0.0810	0.0278	0.0168
			mean($\hat{\theta}$)	-0.0415	-0.1333	-0.0623	-0.1411	-0.1645
			sd($\hat{\theta}$)	0.3209	0.2855	0.3084	0.2301	0.2090
	mse($\hat{\theta}$)	0.1279	0.0859	0.1139	0.0564	0.0449		
	256	0.2	mean(\hat{d}^*)	0.8223	0.7579	0.7947	0.7125	0.8605
			sd(\hat{d}^*)	0.1975	0.1699	0.2066	0.1385	0.1712
			mse(\hat{d}^*)	0.0395	0.0306	0.0427	0.0268	0.0329
			mean($\hat{\phi}$)	-0.5956	-0.6074	-0.6020	-0.6215	-0.5966
			sd($\hat{\phi}$)	0.0841	0.0798	0.1071	0.0759	0.0693
			mse($\hat{\phi}$)	0.0071	0.0064	0.0114	0.0062	0.0048
			mean($\hat{\theta}$)	0.2302	0.1557	0.1921	0.0898	0.2371
			sd($\hat{\theta}$)	0.2473	0.2267	0.2627	0.1918	0.2049
	mse($\hat{\theta}$)	0.0620	0.0533	0.0690	0.0489	0.0433		
	1024	0.2	mean(\hat{d}^*)	0.7355	0.7042	0.7950	0.7625	0.8086
			sd(\hat{d}^*)	0.1216	0.0942	0.1099	0.0696	0.1156
			mse(\hat{d}^*)	0.0189	0.0180	0.0121	0.0062	0.0134
			mean($\hat{\phi}$)	-0.6176	-0.6241	-0.6023	-0.6092	-0.6016
			sd($\hat{\phi}$)	0.0455	0.0416	0.0448	0.0388	0.0369
			mse($\hat{\phi}$)	0.0024	0.0023	0.0020	0.0016	0.0013
mean($\hat{\theta}$)			0.1165	0.0778	0.1934	0.1536	0.1884	
sd($\hat{\theta}$)			0.1618	0.1294	0.1421	0.0949	0.0742	
mse($\hat{\theta}$)	0.0330	0.0316	0.0202	0.0111	0.0056			
0.2	256	-0.6	mean(\hat{d}^*)	0.8439	0.7896	0.8821	0.9359	1.3094
			sd(\hat{d}^*)	0.2130	0.1777	0.2024	0.1319	0.4841
			mse(\hat{d}^*)	0.0473	0.0316	0.0477	0.0358	0.4936
			mean($\hat{\phi}$)	0.1444	0.2014	0.1069	0.0378	0.2144
			sd($\hat{\phi}$)	0.2464	0.2178	0.2375	0.1642	0.1944
			mse($\hat{\phi}$)	0.0638	0.0474	0.0650	0.0532	0.0379
			mean($\hat{\theta}$)	-0.6165	-0.6062	-0.6212	-0.6335	-0.6034
			sd($\hat{\theta}$)	0.0859	0.0817	0.0904	0.0811	0.0762
	mse($\hat{\theta}$)	0.0076	0.0067	0.0086	0.0077	0.0058		
	256	0.6	mean(\hat{d}^*)	0.7581	0.7033	0.6563	0.5548	0.8888
			sd(\hat{d}^*)	0.2080	0.1711	0.2161	0.1457	0.2690
			mse(\hat{d}^*)	0.0450	0.0386	0.0673	0.0913	0.0801
			mean($\hat{\phi}$)	0.1467	0.1357	0.1444	0.0868	0.0484
			sd($\hat{\phi}$)	0.2078	0.2164	0.2766	0.2911	0.1925
			mse($\hat{\phi}$)	0.0459	0.0509	0.0795	0.0974	0.0600
			mean($\hat{\theta}$)	0.5129	0.4498	0.4043	0.2448	0.5278
			sd($\hat{\theta}$)	0.2729	0.2783	0.3042	0.2993	0.3442
	mse($\hat{\theta}$)	0.0819	0.0999	0.1307	0.2155	0.1236		
	1024	0.6	mean(\hat{d}^*)	0.7846	0.7541	0.7643	0.6681	0.8435
			sd(\hat{d}^*)	0.1358	0.1102	0.1142	0.0829	0.1536
			mse(\hat{d}^*)	0.0186	0.0142	0.0143	0.0242	0.0254
			mean($\hat{\phi}$)	0.1755	0.1841	0.1794	0.1578	0.1875
			sd($\hat{\phi}$)	0.1144	0.1007	0.1080	0.1248	0.0766
			mse($\hat{\phi}$)	0.0136	0.0103	0.0120	0.0173	0.0060
mean($\hat{\theta}$)			0.5561	0.5356	0.5411	0.4234	0.5932	
sd($\hat{\theta}$)			0.2044	0.1687	0.1730	0.1644	0.1162	
mse($\hat{\theta}$)	0.0435	0.0325	0.0333	0.0581	0.0135			

For the $ARFIMA(1, 0.8, 1)$ model (see Table 5.5) the bias in the short-run parameters, in some situations, is markedly more severe and it depends on the AR and MA parameter values. The biases in the fractional parameter for the semiparametric methods are relatively small for $\phi = -0.6$ compared with $\phi = 0.2$ for the combination of θ values considered. The method FT always overestimates the fractional parameter while in the semiparametric methods the overestimation and underestimation occur for negative and positive ϕ values, respectively.

6. APPLICATION

The methodology is applied to the data UK long interest gilts: yield on 20 year UK gilts, quarterly, from quarter 1 in 1952 to quarter 4 in 1988 (148 observations). This time series is presented and analyzed by Mills (1997) with special interest in unit root tests. The time series, its sample autocorrelation and periodogram functions are shown in Mills (1997). Clearly, this data exhibits non-stationary behaviour. The unit root test presented in Mills (1997) provides no evidence against the hypothesis that the time series contains a single unit root. However, the sample autocorrelation function shows a slow decay, leading us to suspect non-stationary long memory property.

In Table 6.1 we present the value of the different estimators studied in Section 3, its standard deviation, the number of regressors $g(n)$ for each regression method and the t -statistic test value under the null hypothesis $H_0 : d^* = 1.0$ against the alternative hypothesis $H_1 : d^* < 1.0$. The t -statistic test is obtained from the ratio $\frac{(\hat{d}^* - d^*)}{\sigma(\hat{d}^*)}$, where $\sigma(\hat{d}^*)$ is the asymptotic standard deviation of \hat{d}^* . From the results given in this table, when we compare the t -statistic test value for each method with the normal distribution, we observe that all estimators are very close to 1 and its suggests that the adequated model is $ARFIMA(0, 1, 0)$, in accordance with the unit root test presented in Mills (1997). To end this analysis the forecasting accuracy was performed by computing the mean squared error value of 5-step prediction ($MSE(5)$) of the time series from the origin $n = 135$. The $MSE(5)$ values are not relatively different and this was expected since the estimates of d^* are very similar for all methods considered and the orders of the models are all the same. However, the use of the GPH or SPR methods makes a small forecasting gain compared with the other estimators.

Table 6.1: Identification, estimation, statistic test and forecasting results for the UK long interest gilts.

<i>Property</i>	GPH	SPR	GPHt(1)	GPHt(2)	FT
\hat{d}^*	1.1344	1.1361	0.9141	1.0317	0.9592
$\sigma(\hat{d}^*)$	0.2579	0.1140	0.2286	0.1605	0.0640
$g(n)$	12	12	20	33	-
<i>t</i> -statistic*	0.5211	1.1939	-0.3758	0.1975	-0.6375
<i>MSE</i> (5)	0.729	0.728	0.916	0.819	0.862

**t*-test for $H_0 : d^* = 1; H_1 : d^* < 1$.

7. CONCLUSION AND FUTURE WORK

When the differencing parameter d^* is in the interval (0.5,1.0) the analyzed estimators behaved better than when $d^* \in (1.0, 1.5)$.

In the non-stationary 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 GPH and GPH with tapered data.

We also study the impact of the autoregressive and moving average components over the degree of differencing estimators.

The time series UK long interest gilts was analyzed as an example and the results indicate that this time series may belongs to the class of *ARFIMA*(0, 1, 0) processes.

Studies when the innovations are non Gaussian distributed remain to be done and we will also consider others estimators for the long memory parameter.

ACKNOWLEDGEMENTS

V.A. Reisen was partially supported by CNPq-Brazil. S.R.C. Lopes was partially supported by CNPq-Brazil, by Pronex *Fenômenos Críticos em Probabilidade e Processos Estocásticos* (ConvênioFINEP/MCT/CNPq 41/96/0923/00) and by *Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul* (FAPERGS).

This work was presented at the 22th *Econometric Brazilian Meeting*, in December of 2000, and the authors want to thank Professor Francisco Cribari-Neto (UFPE), who was the referee of this work in that meeting, for his valuable comments and suggestions.

The authors also thank the Editor and two anonymous referees whose comments led to a great improvement of the paper.

References

- [1] Beran, J. (1994), *Statistics for Long-Memory Processes*. Chapman & Hall, New York.
- [2] Cheung, Y.-W. (1993), Tests for Fractional Integration: A Monte Carlo Investigation. *Journal of Time Series Analysis*, **14**(4), 331-345.
- [3] Cheung, Y.-W. and K.S. Lai (1993), A Fractional Cointegration Analysis of Purchasing Power Parity. *Journal of Business & Economic Statistics*, **11**, 103-112.
- [4] Dahlhaus, R. (1989), Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, **17**, 1749-1766.
- [5] Fox, R. and M. S. Taqqu (1986), Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**, 517-532.
- [6] Geweke, J. and S. Porter-Hudak (1983), The Estimation and Application of Long Memory Time Series Model. *Journal of Time Series Analysis*, **4**(4), 221-238.
- [7] Hosking, J. (1981), Fractional Differencing. *Biometrika*, **68**(1), 165-167.
- [8] Hosking, J. (1984), Modelling Persistence in Hydrological Time Series using Fractional Differencing. *Water Resources Research*, **20**(12), 1898-1908.
- [9] Hurvich, C.M. and R.S. Deo (1999), Plug-in Selection of the Number of Frequencies in Regression Estimates of the Memory Parameter of a Long-Memory Time Series. *Journal of Time Series Analysis*, **20**(3), 331-341.
- [10] Hurvich, C.M. and B.K. Ray (1995), Estimation of the Memory Parameter for Nonstationary or Noninvertible Fractionally Integrated Processes. *Journal of Time Series Analysis*, **16**(1), 017-042.
- [11] Liu, M. (1998), Asymptotics of Nonstationary Fractional Integrated Series. *Econometric Theory*, **14**, 641-662.
- [12] Mills, T.C. (1997), *The Econometric Modelling of Financial Time Series*. Cambridge University Press, London.
- [13] Olbermann, B.P. (1998), *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.

- [14] Reisen, V.A. (1994), Estimation of the Fractional Difference Parameter in the ARIMA(p,d,q) model using the Smoothed Periodogram. *Journal of Time Series Analysis*, **15**(3), 335-350.
- [15] Robinson, P.M. (1994), Rates of convergence and optimal spectral bandwidth for long range dependence. *Probability Theory and Related Fields*, **99**, 443-473.
- [16] Robinson, P.M. (1995), Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**(3), 1630-1661.
- [17] Velasco, C. (1999), Non-stationary log-periodogram regression. *Journal of Econometrics*, **91**, 325-371.
- [18] Whittle, P. (1951), *Hypothesis Testing in Time Series Analysis*. Hafner, New York.
- [19] Wu, P. and N. Crato (1995), New Tests for Stationarity and Parity Reversion: Evidence on New Zeland Real Exchange Rates. *Empirical Economics*, **20**, 559-613.
- [20] Yajima, Y. (1985), On Estimation of Long Memory Time Series Models. *Austral. J. Statist.* **27**(3), 303-320.