A COMPARISON OF ESTIMATION METHODS IN NON-STATIONARY ARFIMA PROCESSES

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

October 30, 2002

Abstract:
This paper reports an extensive Monte Carlo simulation study based on six estimators for the long memory fractional parameter when the time series is non-stationary, i.e., $ARFIMA(p, d, q)$ process for $d > 0.5$. Parametric and semiparametric methods are compared. In addition, the effect of the parameter estimation is investigated for small and large sample sizes and non-Gaussian error innovations. The methodology is applied to a well known data set, the so-called UK short interest rates.

Keywords: Non-stationary Time Series, Long Memory, Semiparametric and Parametric Estimations.

[1]Corresponding author's e-mail: slopes@mat.ufrgs.br

1. INTRODUCTION

A time series exhibits long memory when there is significant dependence between observations that are separated by a long period of time. Characteristics of a long memory time series are an autocorrelation function ρ_k that decays hyperbolically to zero and a spectral density function $f_X(\cdot)$ that is unbounded in the neighbourhood of the zero frequency.

The literature on ARFIMA processes has rapidly increased since early contributions by Granger and Joyeux (1980), Hosking (1981) and Geweke and Porter-Hudak (1983). This theory has been widely used in different fields such as meteorology, astronomy, hydrology and economics (further details can be found in Beran (1994) and Hosking (1981) and (1984)).

Geweke and Porter-Hudak (1983) presented a very important work on stationary long memory processes. Their paper gave rise to several other works, and presented a proof for the asymptotic distribution of the long memory parameter $d \in (-0.5, 0.0)$. These authors proposed an estimator of d as the ordinary least squares estimator of the slope parameter in a simple linear regression of the logarithm of the periodogram. Reisen (1994) proposed a modified form of the regression method, based on a smoothed version of the periodogram function. Robinson (1995a), making use of mild modifications on Geweke and Porter-Hudak's estimator, dealt simultaneously with $d \in (-0.5, 0.0)$ and $d \in (0.0, 0.5)$. Hurvich and Deo (1999), among others, addressed the problem of selecting the number of frequencies necessary for estimating the differencing parameter in the stationary case. Fox and Taqqu (1986) considered an approximated method, whereas Sowell (1992) presented the exact maximum likelihood procedure for estimating the fractional parameter. These two papers considered the estimation procedures for the stationary case. Simulation studies comparing estimates of d may be found, for instance, in Smith et al. (1997), Bisaglia and Guégan (1998), Reisen and Lopes (1999), Reisen et al. (2000) and (2001).

Recently, much work has focused on long memory non-stationary stochastic ARFIMA processes. More recent works include Hurvich and Ray (1995), Liu (1998), Velasco (1999a,b) and Velasco and Robinson (2000). Hurvich and Ray (1995) consider the asymptotic characteristics of the periodogram ordinates for both cases, $d \geq 0.5$ and $d \leq -0.5$. They found that the periodogram of a non-stationary or noninvertible fractionally integrated process at the j^{th} Fourier frequency $\lambda_j = \frac{2\pi j}{n}$, where n is the sample size, has an asymptotic relative bias depending on j. For finite sample size, Hurvich and Ray (1995) examined the impact of the periodogram and the tapered periodogram bias on the regression estimator of d proposed by Geweke and Porter-Hudak (1983).

Liu (1998) studies the asymptotic theory of non-stationary ARFIMA process with special attention to the unit root KPSS test. Velasco (1999a) showed that it is possible to estimate consistently the memory of non-stationary process using log-periodogram methods designed for stationary cases. The estimator given by Robinson (1995b) was also extended to non-stationary case in Velasco (1999b). Velasco and Robinson (2000) extended the results of the Whittle's maximum likelihood estimator to include non-stationary ($d \in (0.5, 1.0)$) or intermediate memory ($d \in (-0.5, 0.0)$) observations.

Here, by a Monte Carlo study, we compare the semiparametric methods given in Geweke and Porter-Hudak (1983) and Reisen (1994); two modified forms of the Geweke and Porter-Hudak's estimator; the cosine-bell tapered data and the parametric Whittle method (see Fox and Taqqu, 1986). These estimation methods are investigated in the situation where $d \in (0.5, 1.5)$. The model and the estimation procedures are summarized in Section 2. The simulation study is presented in Section 3. The methodology is applied to a real data set in Section 4. Conclusions are given in Section 5.

2. THE MODEL AND THE ESTIMATORS

2.1. Stationary and Invertible ARFIMA Processes

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a zero mean ARFIMA(p, d, q) process given by

$$\Phi(\mathcal{B})(1-\mathcal{B})^d X_t = \Theta(\mathcal{B})\epsilon_t, \;\; d \in \mathbb{R}, \tag{2.1}$$

where $\mathcal{B}$ is the lag operator. The polynomials $\Phi(\mathcal{B}) = \sum_{i=0}^{p}(-\phi_i)\,\mathcal{B}^i$ and $\Theta(\mathcal{B}) = \sum_{i=0}^{q}\theta_i\,\mathcal{B}^i$ are of orders p and q, respectively, with $\phi_0 = -1$ and $\theta_0 = 1$. The process $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is white noise process with zero mean and finite variance σ_ϵ^2.

The process $\{X_t\}_{t\in\mathbb{Z}}$ is both stationary and invertible if the polynomials $\Phi(\mathcal{B})$ and $\Theta(\mathcal{B})$ have roots outside of the unit circle and $|d| < 0.5$. Its spectral density function, $f_X(\cdot)$, is given by

$$f_X(\lambda) = f_U(\lambda)\left[2\,\sin\left(\frac{\lambda}{2}\right)\right]^{-2d}, \; \lambda \in [-\pi, \pi], \tag{2.2}$$

where $f_U(\cdot)$ is the spectral density function of an ARMA(p, q) process. Hosking (1981), Beran (1994) and Reisen (1994) describe ARFIMA models in detail.

The process (2.1) exhibits *long memory* when $d \in (0.0, 0.5)$, *intermediate memory* when $d \in (-0.5, 0.0)$ and *short memory* when $d = 0$.

2.2. Non-stationary ARFIMA Processes

Now, replacing the parameter d in (2.1) with $d^* = d + r$, where $d \in$

$(0.0, 0.5)$, $r > 0$, $r \in \mathbb{R}$ the resulting model is given by

$$\Phi(\mathcal{B})(1-\mathcal{B})^{d^*} X_t = \Theta(\mathcal{B})\epsilon_t. \tag{2.3}$$

The above equation can be rewritten as

$$Y_t = (1-\mathcal{B})^r X_t$$

such that

$$\Phi(\mathcal{B})(1-\mathcal{B})^{d} Y_t = \Theta(\mathcal{B})\epsilon_t \tag{2.4}$$

is an ARFIMA(p, d, q) process.

For $d^* \geq 0.5$, the process (2.3) is non-stationary, and level-reverting for $d^* \in [0.5, 1.0)$, considering there is no long-run impact of an innovation on the value of the process (see Velasco, 1999a).

2.3. Estimates of the Parameters in ARFIMA(p, d^*, q) Processes

We deal with some well known estimation methods of d to estimate d^*. The first five estimators are semiparametric methods based on an approximated regression equation obtained from the logarithm of the spectral density function of the process. The first method is the one proposed by Geweke and Porter-Hudak (1983), denoted in the following by GPH. The second estimator is the smoothed periodogram regression (SPR), suggested by Reisen (1994).

As a third method we consider the GPH, based on the trimming l and bandwidth m, denoted hereafter by GPHtr, suggested by Robinson (1995a). The GPHtr method regresses $\log\{I(\lambda_j)\}$ on $\log\{2\sin(\lambda_j/2)\}^2$, for $j \in \{l, l+1, \cdots, m\}$, where l tends to infinity more slowly than m. One observes that the independent regressors are related to the fractional integration component in the spectral density function given by (2.2).

The fourth method is a modified form of the GPH method, denoted hereafter by MGPH, obtained by replacing in the regression equation the quantity $2\ \sin(\frac{\lambda_j}{2})$ by j (see Velasco, 1999b, p. 101).

The cosine-bell tapered data method, denoted in the following by GPHTa, is the fifth approach considered here. In this method the modified periodogram function is given by

$$I(\lambda_j) = \frac{1}{2\pi\sum_{t=0}^{n-1} g(t)^2}\left|\sum_{t=0}^{n-1} g(t) X_t e^{-i\lambda_j t}\right|^2,$$

where the tapered data is obtained from the cosine-bell function

$$g(t) = \frac{1}{2}\left[1 - \cos\left(\frac{2\pi(t+0.5)}{n}\right)\right].$$

This estimator was also used in the works by Hurvich and Ray (1995) and Velasco (1999a).

The sixth estimator is the parametric technique proposed by Fox and Taqqu (1986), denoted in the following by FT. For computational purposes this estimator is obtained by minimizing a finite and discrete form of a function depending on the vector of unknown parameters.

3. MONTE CARLO SIMULATION STUDY

The processes $\{Y_t\}_{t\in\mathbb{Z}}$ in equation (2.4) were simulated as suggested by Hosking (1984), where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is a Gaussian white noise process with zero mean and variance $\sigma_\epsilon^2 = 1.0$. The error process was generated using the RNNOR subroutine in the IMSL library. The process $\{X_t\}_{t\in\mathbb{Z}}$ is 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 estimation results are obtained for time series with small and large sample sizes, i.e., for $n = 100$ and 300 respectively, and in both cases are based upon 1,000 replications of the error process. We calculate the empirical values of the mean, the bias, the standard deviation (*sd*) and the mean squared error (*mse*). These quantities are given in the tables. The largest bias and mean squared error, in absolute values, are given in bold-face. The truncation point in the Parzen lag window, for the SPR's method, is $\nu = n^\beta$, where we consider $\beta = 0.9$ (see Reisen, 1994, for a discussion on the value of β). The bandwidth m in the semiparametric methods is a function of n, that is, $m = n^\alpha$. In the GPH, SPR and GPHTa methods, we fixed $\alpha = 0.5$ (as it is widely used). For both GPHtr and MGPH methods we use two different values of the bandwidth m: we consider $\alpha_1 = 0.6$ and $\alpha_2 = 0.7$ and we denote the estimators respectively by GPHtr(i) and MGPH(i), for $i = 1, 2$. Also, the trimming number l in the regression methods is $l = 1$ except for GPHtr and GPHTa, where we consider $l = 2$. The reason for taking $l = 2$ in the GPHTa estimator is given by Theorem 3 in Hurvich and Ray (1995).

First, we deal with ARFIMA$(0, d^*, 0)$ models and the results are in Tables 1 and 2. Table 1 presents the cases where $d^* \in \{0.6, 1.0\}$. Other values of d^* were also simulated and the results are available upon request. For $d^* < 1.0$ we can see the estimates are positively biased, except in the SPR method. The FT method seems, in general, to be more accurate (smaller bias and mean squared error values) than the other methods which also give good results. Note that increasing the bandwidth m in the GPHtr and MGPH methods results in a decrease in the bias and a substantial decrease in the *mse* of the estimates (one observes that the *mse* values decrease almost by half). The SPR estimator has better performance than GPH estimator in the sense of minimizing the *mse* values, as it was expected since SPR uses the smoothed periodogram to estimate the spectral density function. The tapering method (GPHTa) gives estimates with the largest *mse* value due to the large value of its *sd*.

Table 1: Estimates of the parameter d^* for the ARFIMA$(0, d^*, 0)$ model when $d^* \in \{0.6, 1.0\}$.

d^*	n	*Methods*	m	*mean*	*bias*	*sd*	*mse*
0.6	100	GPH	10	0.6134	0.0134	0.2807	0.0789
		SPR	10	0.5109	**-0.0891**	0.2261	0.0590
		GPHtr(1)	16	0.6100	0.0100	0.3392	0.1150
		GPHtr(2)	25	0.6181	0.0181	0.2267	0.0517
		MGPH(1)	16	0.6067	0.0067	0.2056	0.0423
		MGPH(2)	25	0.6009	0.0009	0.1546	0.0239
		GPHTa	10	0.6169	0.0169	0.4862	**0.2364**
		FT	-	0.6058	0.0058	0.0905	0.0082
	300	GPH	17	0.6182	0.0182	0.2154	0.0467
		SPR	17	0.5618	**-0.0382**	0.1755	0.0322
		GPHtr(1)	31	0.6073	0.0073	0.1912	0.0366
		GPHtr(2)	54	0.6099	0.0099	0.1250	0.0157
		MGPH(1)	31	0.6122	0.0122	0.1475	0.0219
		MGPH(2)	54	0.6056	0.0056	0.1062	0.0113
		GPHTa	17	0.6208	0.0208	0.3291	**0.1086**
		FT	-	0.6041	0.0041	0.0519	0.0027
1.0	100	GPH	10	0.9692	-0.0308	0.2381	0.0576
		SPR	10	0.9330	-0.0670	0.2022	0.0453
		GPHtr(1)	16	0.9856	-0.0144	0.2885	0.0834
		GPHtr(2)	25	0.9970	-0.0030	0.1895	0.0359
		MGPH(1)	16	0.9679	-0.0321	0.1801	0.0334
		MGPH(2)	25	0.9604	-0.0396	0.1320	0.0190
		GPHTa	10	1.1153	**0.1153**	0.4967	**0.2597**
		FT	-	0.9970	-0.0030	0.1895	0.0359
	300	GPH	17	0.9835	-0.0165	0.1734	0.0303
		SPR	17	0.9789	-0.0211	0.1482	0.0224
		GPHtr(1)	31	0.9982	-0.0012	0.1553	0.0241
		GPHtr(2)	54	0.9971	-0.0029	0.1035	0.0107
		MGPH(1)	31	0.9856	-0.0144	0.1182	0.0142
		MGPH(2)	54	0.9783	-0.0217	0.0861	0.0079
		GPHTa	17	1.0648	**0.0648**	0.3120	**0.1014**
		FT	-	0.9948	-0.0052	0.0433	0.0019

When the process is a random walk ($d^* = 1.0$) all procedures perform well and are very competitive. The mean of the estimates underestimates the true parameter for all methods, except for GPHTa. This estimator again presented the largest value for both *sd* and *mse*. The empirical results in Table 1 also indicate that the *sd* and *mse* values decrease when n increases.

Table 2 presents results for $d^* = 1.2$ where the level-reversion property does not hold. We have observed that when $d^* > 1.0$ the non-tapered estimators give an average value of approximately 1 for their estimates, independently of the value of d. The estimates are negatively biased while tapering gives estimates which are positively biased (see Figure 1). This study reveals that, in practical situations, if one of these non-tapered methods give an estimate close to one, this does not necessarily indicate that the series follows a random walk. One solution is to take first differences of the series and then to estimate d^*. However, as pointed out by Hurvich and Ray (1995), the GPH estimator is not, in general, invariant under first differences. The estimated d^* based on the original data is not, in general, equal to one plus the estimated d^* based on the differenced data (this is the subject of the paper Olbermann et al. (2002) in http://athena.mat.ufrgs.br/ slopes).

Table 2: Estimates of the parameter d^* for the ARFIMA$(0, 1.2, 0)$ model.

n	*Methods*	m	*mean*	*bias*	*sd*	*mse*
100	GPH	10	1.0829	-0.1171	0.2354	0.0691
	SPR	10	1.0582	-0.1418	0.1788	0.0521
	GPHtr(1)	16	1.0760	-0.1240	0.2606	0.0832
	GPHtr(2)	25	1.0698	-0.1302	0.1843	0.0509
	MGPH(1)	16	1.0705	-0.1295	0.1814	0.0496
	MGPH(2)	25	1.0463	**-0.1537**	0.1378	0.0426
	GPHTa	10	1.3437	0.1437	0.4771	**0.2480**
	FT	-	1.0628	-0.1372	0.0939	0.0276
300	GPH	17	1.0690	-0.1310	0.1708	0.0463
	SPR	17	1.0889	-0.1111	0.1283	0.0288
	GPHtr(1)	31	1.0588	-0.1412	0.1589	0.0452
	GPHtr(2)	54	1.0594	-0.1406	0.1199	0.0341
	MGPH(1)	31	1.0587	-0.1413	0.1229	0.0351
	MGPH(2)	54	1.0471	**-0.1525**	0.1021	0.0337
	GPHTa	17	1.2780	0.0780	0.3001	**0.0960**
	FT	-	1.0651	-0.1349	0.0747	0.0238

Now, we deal with ARFIMA(p, d^*, q) models and in Table 3 we present the case where $p = 1$, $q = 0$, $d^* = 0.6$ and $n = 300$. Simulations for different sizes of n and also different cases when $p = 0$ and $q = 1$ are available upon request. Besides revealing the estimates of d^*, this table also shows the mean of the estimates for the short-run parameter and their corresponding standard deviation and mean squared error values. In the FT method, the parameters of the process 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^*. Table 3 reveals the impact of the

Table 3: Estimates of the parameters in the ARFIMA$(1, 0.6, 0)$ model when $n = 300$.

ϕ	*Methods*	$mean(\hat{d}^*)$	$bias(\hat{d}^*)$	$sd(\hat{d}^*)$	$mse(\hat{d}^*)$	$mean(\hat{\phi})$	$sd(\hat{\phi})$	$mse(\hat{\phi})$
	GPH	0.6200	0.0200	0.2030	0.0416	-0.6827	0.1162	0.0138
	SPR	0.5556	**-0.0444**	0.1698	0.0308	-0.6621	0.1030	0.0120
	GPHtr(1)	0.5952	-0.0048	0.1976	0.0390	-0.6696	0.1262	0.0168
	GPHtr(2)	0.5633	-0.0367	0.1238	0.0166	-0.6737	0.0789	0.0069
-0.7	MGPH(1)	0.6033	0.0033	0.1506	0.0227	-0.6875	0.0815	0.0068
	MGPH(2)	0.5701	-0.0299	0.1037	0.0116	-0.6814	0.0628	0.0043
	GPHTa	0.6281	0.0281	0.3321	**0.1110**	-0.6237	0.2545	**0.0705**
	FT	0.6365	0.0365	0.0687	0.0060	-0.8772	0.1467	0.0529
	GPH	0.6166	0.0166	0.2026	0.0413	-0.4790	0.1524	0.0236
	SPR	0.5544	**-0.0456**	0.1700	0.0309	-0.4486	0.1337	0.0205
	GPHtr(1)	0.5977	-0.0023	0.1991	0.0396	-0.4654	0.1600	0.0267
	GPHtr(2)	0.5620	-0.0380	0.1258	0.0173	-0.4643	0.1005	0.0113
-0.5	MGPH(1)	0.6023	0.0023	0.1467	0.0215	-0.4864	0.1060	0.0114
	MGPH(2)	0.5675	-0.0325	0.1033	0.0117	-0.4732	0.0838	0.0077
	GPHTa	0.5620	-0.0380	0.1258	**0.0895**	-0.4155	0.2810	**0.0860**
	FT	0.6102	0.0102	0.0768	0.0060	-0.5421	0.1618	0.0279
	GPH	0.6175	0.0175	0.2040	0.0419	-0.1855	0.1899	0.0362
	SPR	0.5609	**-0.0391**	0.1665	0.0292	-0.1473	0.1641	0.0297
	GPHtr(1)	0.6072	0.0072	0.1987	0.0395	-0.1772	0.1923	0.0374
	GPHtr(2)	0.5813	-0.0187	0.1296	0.0171	-0.1744	0.1289	0.0173
-0.2	MGPH(1)	0.6113	0.0113	0.1435	0.0207	-0.1972	0.1343	0.0180
	MGPH(2)	0.5836	-0.0164	0.1012	0.0105	-0.1830	0.1019	0.0106
	GPHTa	0.6224	0.0224	0.3211	**0.1035**	-0.1396	0.3078	**0.0983**
	FT	0.6034	0.0034	0.0756	0.0057	-0.1962	0.0834	0.0070
	GPH	0.6225	0.0225	0.2129	0.0458	0.2001	0.0563	0.0032
	SPR	0.5627	-0.0373	0.1777	0.0329	0.2433	0.1827	0.0352
	GPHtr(1)	0.6396	0.0396	0.1963	0.0400	0.1734	0.1952	0.0387
	GPHtr(2)	0.6692	0.0692	0.1310	0.0219	0.1352	0.1311	0.0213
0.2	MGPH(1)	0.6267	0.0267	0.1482	0.0227	0.1771	0.1517	0.0235
	MGPH(2)	0.6461	0.0461	0.1057	0.0133	0.1538	0.1104	0.0143
	GPHTa	0.6404	0.0404	0.3211	0.1046	0.2008	0.3015	**0.0908**
	FT	0.7854	**0.1854**	0.5382	**0.3237**	0.2236	0.1501	0.0231
	GPH	0.6698	0.0698	0.2046	0.0467	0.4168	0.1930	0.0441
	SPR	0.6074	0.0074	0.1682	0.0283	0.4747	0.1638	0.0274
	GPHtr(1)	0.7333	0.1333	0.1977	0.0568	0.3581	0.1860	0.0547
	GPHtr(2)	0.8404	0.2404	0.1247	0.0733	0.2575	0.1206	0.0733
0.5	MGPH(1)	0.7039	0.1039	0.1406	0.0305	0.3836	0.1402	0.0332
	MGPH(2)	0.7854	0.1854	0.0954	0.0435	0.3066	0.1002	0.0474
	GPHTa	0.6855	0.0855	0.3195	0.1093	0.4048	0.2838	**0.0895**
	FT	1.2428	**0.6428**	0.6796	**0.8746**	0.5033	0.1692	0.0286
	GPH	0.7276	0.1276	0.2061	0.0587	0.5579	0.1838	0.0539
	SPR	0.6774	0.0774	0.1673	0.0339	0.6048	0.1504	0.0316
	GPHtr(1)	0.9121	0.3121	0.1895	0.1333	0.3930	0.1770	0.1255
	GPHtr(2)	1.0600	0.4600	0.1287	0.2245	0.2562	0.1244	**0.2124**
0.7	MGPH(1)	0.8303	0.2303	0.1430	0.0735	0.4699	0.1403	0.0726
	MGPH(2)	0.9596	0.3596	0.1025	0.1384	0.3491	0.1074	0.1346
	GPHTa	0.7923	0.1923	0.3217	0.4129	0.4912	0.2753	0.1192
	FT	1.2839	**0.6839**	0.6417	**0.8789**	0.6052	0.1793	0.0411

estimates when the short-run AR parameter is included in the model. For negative values of ϕ, the methods still work well and are very competitive.

The FT method generally gives better estimates. For positive AR parameters the bias of the estimates are positive and they increase with ϕ. The FT method loses its superiority, showing the largest bias and *mse* values. In this situation, the bias of $\hat{d}^*$ is predominantly positive, and negative for the AR parameter. The SPR method has a bias and *mse* that is increasing more slowly than that of the other estimation methods as ϕ increases. This is even more evident when ϕ increases from 0.5 to 0.7. In the GPHtr and MGPH approaches, one observes that larger values of m produce opposite effect than those in the ARFIMA$(0, d^*, 0)$ model. Now the bias and the mean squared error values tend to increase with m, specifically for positive short-run parameter values. This is not surprising as long as the presence of AR components in the model makes a larger contribution to the spectral density function of the process at certain frequencies (away from the zero frequency, but that are still involved in the regression equation). The GPHTa presents larger *mse* compared with the other semiparametric methods. The *sd* and *mse* values of $\hat{d}^*$ for the semiparametric methods are essentially the same, as it was found for the case ARFIMA$(0, 0.6, 0)$, when $n = 300$, given in Table 1. The ARFIMA$(1, 0.8, 0)$ was also considered and the results are available upon request. It was noticed the same behaviour observed in the case of Table 3 for all methods where the bias and mean squared error values increase significantly for positive values of ϕ. The SPR method, in general, has better behaviour compared with the others. For instance, when $\phi = 0.7$ and $n = 300$, the estimates of d^* and ϕ are respectively 0.8804 and 0.6050 and their corresponding mean squared error values are 0.0335 and 0.0307, respectively.

We have also considered the performance of estimating the parameter d^* in ARFIMA(p, d^*, q) models when the innovation distributions are non Gaussian. We considered innovation processes with the following distributions: uniform over the interval $(-\sqrt{3}, \sqrt{3})$, exponential with rate 1 (with re-centered zero mean), Student's t with 3 degrees of freedom and χ^2 with 1 degree of freedom (with re-centered zero mean). All estimators have the same behaviour as when the distribution of the innovation process is Gaussian. As the results are quite similar we do not report them here, however they are available upon request. Non-Gaussian distributed innovation processes have also been the focus of many works, more recently in Velasco (2000) for stationary time series.

4. APPLICATION

Does the UK short interest rates time series have long memory or a unit root? The UK short interest rate time series is a 91 days of UK treasury bill rates, measured quarterly, from quarter 1 in 1952 to quarter 4 in 1988 (with 148 observations). This series is presented and analyzed by Mills (1997) with special interest in the theory of the unit root tests. From the plots of the time series and its sample autocorrelation given in Mills (1997), it is clear that this data exhibits non-stationary behaviour.

The GPH and GPHTa estimates indicated that the process is a random walk. The other procedures gave $\hat{d}^* < 1.0$. The GPHtr and MGPH methods gave similar estimates where their values are between SPR and FT estimates. In the FT method we tried different values for p and q, and the results suggested an ARFIMA$(0, d^*, 0)$ model. When using the semiparametric methods the ARMA process order was identified after the time series has been differentiated by $\hat{d}^*$. The choices of the model were made by testing the AR and MA estimates and the AIC (Akaike's information) criterion. The estimates of the short-run parameters and other statistics were obtained by using the MINITAB package. The results related to the identification, estimation models and forecasting analyses are given in Table 4.

The residual analyses were performed for the fitted models, and they all indicated that the errors are approximately Gaussian white noise. Based on the modified Box-Pierce test (MBP) the hypothesis of adequated model is not rejected for all cases at 5% of significance level.

Forecasting issues were also considered by calculating the quantities mean squared error (mse) and the mean absolute percentage error (MAPE) (see Wei, 1990, p. 179) at the forecast initial value $t = 135$. This analysis indicates that the series may be a non-stationary long memory one with $d^* < 1.0$. Hence, the use of $d^* = 1.0$ may over-differentiate the series.

Table 4: Identification, estimation, and forecast results for the UK short interest rates.

Estimate	ARFIMA(0,FT,0)	ARFIMA(0,SPR,1)	ARFIMA(0,1,0)
$\hat{d}^*$	0.8514	0.7675	-
$\sigma(\hat{d}^*)$	0.0640	0.1140	-
$\hat{\theta}_1$	-	-0.1936	-
$\sigma(\hat{\theta}_1)$	-	0.0836	-
$\hat{\sigma}_\epsilon^2$	1.4671	1.4553	1.5089
AIC	53.30	56.12	57.18
MBP	0.108	0.107	0.064
mse	1.126	1.041	1.171
MAPE(%)	33.70	32.20	34.60

MBP- *the modified Box-Pierce chi-square statistic with 11 degrees of freedom.*

5. CONCLUSION

The performance of the semiparametric and parametric methods of estimating the fractional difference parameter in a non-stationary process were investigated. For the ARFIMA$(0, d, 0)$ model, when $d^* \in (0.5, 1.0)$, the non-tapered and tapered estimates perform well and the FT method, in general, gives better estimates. This last estimator is outperformed by the semiparametric methods when positive and large short-run parameters are included in the process. Also, the estimators improve as the sample sizes increase. In the non-stationary case with no level-reversion property we observed that all non-tapered estimators are strongly biased and underestimate the true parameter value. In this situation the bias of GPHTa is positive and is always smaller than the value for the other estimators. The time series UK short interest rates was analyzed as an example and the results indicate that it may belong to a class of non-stationary long memory process with the fractional parameter $d^* \in (0.7, 1.0)$.

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* (Convênio FINEP/MCT/CNPq 41/ 96/ 0923/00) and also by *Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul* (FAPERGS). The authors wish to thank the associate editor (Jane L. Harvill) and an anonymous referee for their valuable comments that improved an earlier version of the paper.

REFERENCES

Beran, J. (1994) *Statistics for Long-Memory Processes*. Chapman and Hall, New York.

Bisaglia, L. and Guégan, D. (1998) A comparison of techniques of estimation in long-memory processes. *Computational Statistics and Data Analysis*, **27**, 61-81.

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

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

Granger, C.W.J. and Joyeux, R. (1980) An Introduction to Long Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, **1**, 15-29.

Hosking, J. (1981) Fractional Differencing. *Biometrika*, **68**, 165-167.

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

Hurvich, C.M. and Deo, R.S. (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**, 331-341.

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

Liu, M. (1998) Asymptotics of Nonstationary Fractional Integrated Series. *Econometric Theory*, **14**, 641-662.

Mills, T.C. (1997) *The Econometric Modelling of Financial Time Series.* Cambridge University Press, London.

Olbermann, B.P., Lopes, S.R.C. and Reisen, V.A. (2002) Invariance of the First Difference in ARFIMA Models. Submitted.

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**, 335-350.

Reisen, V.A. and Lopes, S.R.C. (1999) Some Simulations and Applications of Forecasting Long-Memory Time Series Models. *Journal of Statistical Planning and Inference*, **80**, 269-287.

Reisen, V.A., Abraham, B. and Lopes, S.R.C. (2001) Estimation of the Parameters in ARFIMA Processes: A Simulation Study. *Communications in Statistics: Simulation and Computation*, **30**, 787-803.

Reisen, V.A., Abraham, B. and Toscano, E.M.M. (2000) Parametric and Semiparametric Estimation of Stationary Univariate ARFIMA Models. *Brazilian Journal of Probability and Statistics*, **14**, 185-206.

Robinson, P.M. (1995a) Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**, 1048-1072.

Robinson, P.M. (1995b) Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, **23**, 1630-1661.

Smith, J., Taylor, N. and Yadav, S. (1997) Comparing the Bias and Misspecification in ARFIMA Models. *Journal of Time Series Analysis*, **18**, 507-527.

Sowell, F. (1992) Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. *Journal of Econometrics*, **53**, 165-188.

Velasco, C. (1999a) Non-stationary log-periodogram regression. *Journal of Econometrics*, **91**, 325-371.

Velasco, C. (1999b) Gaussian Semiparametric Estimation of Non-stationary Time Series. *Journal of Time Series Analysis*, **20**, 87-127.

Velasco, C. (2000) Non-Gaussian Log-Periodogram Regression. *Econometric Theory*, **16**, 44-79.

Velasco, C. and Robinson, P.M. (2000) Whittle Pseudo-Maximum Likelihood Estimation for Nonstationary Time Series. Preprint.

Wei, W. S. (1990) *Time Series Analysis Univariate and Multivariate Methods.* Addison Wesley, New York.

Figure 1: Box-Plot of estimators of d in ARFIMA$(0, 1.4, 0)$ model when $n = 300$.