Computational Statistics
DOI 10.1007/s00180-006-0005-0

ORIGINAL PAPER

Invariance of the first difference in ARFIMA models

Barbara P. Olbermann · Sílvia R.C. Lopes · Valdério A. Reisen

Abstract A desirable property for an estimator of the fractional ARFIMA parameter is to be first difference invariant. This paper investigates the effects on the fractional parameter estimator in nonstationary ARFIMA(p, d, q) processes before and after applying a first difference. We consider semiparametric and parametric approaches for estimating d. The study is based on a Monte Carlo simulation for different sample sizes. The Brazilian exchange rate series is given as an application of the methodology.

Keywords Non-stationary processes · Long memory · Estimation · Invariance

1 Introduction

The ARFIMA(p, d, q) process was first introduced by Granger and Joyeux (1980), and Hosking (1981). The most useful feature for this process is the long memory. This property is reflected by the hyperbolic decay of the autocorrelation function or by the unboundedness of the spectral density function of the

B. P. Olbermann
Faculdade de Matemática, PUCRS, Av. Ipiranga,
6681, 90619-900 Porto Alegre, RS, Brazil
e-mail: barbarap@pucrs.br

S. R. C. Lopes (✉)
Instituto de Matemática, UFRGS, Av. Bento Gonçalves,
9500, 91509-900 Porto Alegre, RS, Brazil
e-mail: slopes@mat.ufrgs.br

V. A. Reisen
Departamento de Estatística, UFES, Av. Fernando Ferrari,
s/n. 29060-900 Vitória, ES, Brazil

process. While in an ARMA structure, the dependency between observations decays at a geometric rate.

Several applications using ARFIMA processes are considered in some fields such as economics, hydrology and finance. Granger and Joyeux (1980) used a fractionally differenced model with no short-term components to model the US monthly index of consumer food prices for the period January 1947 to June 1978. Based on minimizing the 10-step-ahead forecast errors, they estimated d to be approximately 0.35, after first differencing the original time series. Sowell (1992a) applied the ARFIMA process to correctly model the trend behavior of the postwar US real GNP data. This time series has sample size equal to 172 observations and the author compared test of hypothesis using fractional and non-fractional ARIMA processes in modeling the long-run behavior of the series. Geweke and Porter-Hudak (1983) also found ARFIMA models useful for forecasting other leading indicator series. Chen (1987) used a fractional differencing model with fixed coefficient regression terms to incorporate both long-term dependence and short-term periodic effects into a model for gold prices. Lobato and Velasco (2000) analyzed the long-memory properties for the daily trading volume and the return volatility processes of the 30 stocks that compose the Dow Jones Industrial Average index for the period July 1962 to December 1994. These two components of the vector process exhibit different stochastic properties: although return volatility is considered to be stationary, the trading volume is treated as non-stationary. The authors showed that return volatility and trading volume have the same long memory parameter for the most of the stocks although apparently the long memory of these two series cannot be explained by a common long-memory component.

Several estimation procedures for the fractional ARFIMA parameters have been proposed, mainly, by semiparametric and parametric procedures.

In the first class the regression method proposed by Geweke and Porter-Hudak (1983) was the pioneer. This approach was very important giving rise to several other works. The authors presented a proof when $d \in (-0.5, 0.0)$, nonetheless, the method, denoted here by GPH, has been used for a wide range of d. Robinson (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.

Reisen (1994) proposed a modified form of the regression method based on a smoothed version of the periodogram function. Velasco (1999a) also considered a modified version of the GPH method.

Hurvich and Ray (1995) have introduced a cosine-bell function as a spectral window to reduce the bias of the periodogram function. They find that data tapering, and the elimination of the first periodogram ordinate in the regression equation, can reduce the bias of the estimator. However, the reduction of the bias' estimator has the cost of increasing its variance. Velasco (1999b) has also considered this estimator, and proved its consistency, and asymptotic normality for any d including both non-stationary, and non-invertible processes.

In the parametric class, the reader will find the methods based on the maximum likelihood function suggested in Fox and Taqqu (1986) and

Sowell (1992b), among others. A comparison study between these parametric approaches are given by Cheung and Diebold (1994), Hauser (1999), and the references therein. Reisen et al. (2001a) present an extensive simulation study comparing both semiparametric, and parametric approaches for the estimation of d in ARFIMA(p, d, q) processes.

The main goal of this paper is to analyze which estimation method for the fractional parameter is invariant to first-differencing when the model is described by an ARFIMA(p, d, q) process. Hurvich and Ray (1995) addressed the invariance property only for tapered, and non-tapered GPH estimators. In this work we shall consider three methods in the semiparametric, and one in the parametric class. In our investigation we analyze the case of non-stationary processes with, and without level-reversion property, that is, when $d \in [0.5, 1.0)$ and $d \in [1.0, 1.5)$, respectively. The work is organized as follows: Sect. 2 presents the ARFIMA(p, d, q) processes when $d \in (0.5, 1.5)$, and the estimation methods; the simulation results are in Sect. 3; an application of the methodology is in Sect. 4, and the conclusions are in Sect. 5.

2 The model and the estimators

2.1 Stationary and invertible ARFIMA process

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

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

where $\mathcal{B}$ is the backward-shift operator, that is, $\mathcal{B}X_t = X_{t-1}$. The polynomials $\Phi(\mathcal{B}) = \sum_{i=0}^{p}(-\phi_i)\,\mathcal{B}^i$ and $\Theta(\mathcal{B}) = \sum_{j=0}^{q}(-\theta_j)\,\mathcal{B}^j$ are of orders p and q, respectively, with $\phi_0 = -1 = \theta_0$. The process $\{\epsilon_t\}_{t \in Z}$ is white noise with zero mean, and finite variance σ_ϵ^2. The term $(1 - \mathcal{B})^d$ is the binomial power series of $\mathcal{B}$.

The process $\{X_t\}_{t \in Z}$, given by the expression (2.1), is called a *general fractional differenced zero mean process*, where d is the *fractional differencing parameter*.

The process given by the expression (2.1) is both stationary, and invertible if the roots of $\Phi(\mathcal{B})$ and $\Theta(\mathcal{B})$ are outside the unit circle, and $d \in (-0.5, 0.5)$. Its spectral density function, $f_X(\cdot)$, is given by

$$f_X(w) = f_U(w)\left[2\,\sin\left(\frac{w}{2}\right)\right]^{-2d}, \quad w \in [-\pi, \pi],$$

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

The ARFIMA(p, d, q) process 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 process

Now, we define the process (2.1) with the parameter $d^* = d + 1$, where $d \in (0.0, 0.5)$, and the model (2.1) becomes

$$\Phi(\mathcal{B})(1 - \mathcal{B})^{d^*} X_t = \Theta(\mathcal{B})\epsilon_t, \;\; t \in Z. \tag{2.2}$$

The process (2.2) is non-stationary when $d^* \geq 0.5$; however, it is still persistent. 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 Velasco 1999a). The level-reversion property no longer holds when $d^* \geq 1$.

2.3 Estimation in ARFIMA(p, d^*, q) process

To estimate the fractional parameter we consider semiparametric and parametric estimation methods. In the semiparametric class we deal with the estimator proposed by Geweke and Porter-Hudak (1983), denoted in the sequel by GPH, where its asymptotic variance is given by

$$h(x, n) \times \text{Var(GPH)} \approx \frac{\pi^2}{6}, \tag{2.3}$$

with $h(x, n) = \sum_{j=1}^{g(n)} (x_j - \overline{x})^2$, $x_j = \ln\{\sin(\frac{w_j}{2})\}^2$, w_j, for $j = 1, \ldots, g(n)$, the j-th Fourier frequency and $g(n) = n^{\alpha}$, for $0 < \alpha < 1$ (see Geweke and Porter-Hudak 1983).

We also consider the smoothed periodogram regression (SPR), suggested by Reisen (1994), with asymptotic variance given by

$$\frac{n}{m} h(x, n) \times \text{Var(SPR)} \approx 0.53928,$$

where $h(x, n)$ and x_j, for $j \in \{1, \ldots, g(n)\}$, have the same values as in (2.3) and $m = n^{\beta}$, with $\beta = 0.9$, is the truncation point of the Parzen lag window (see Reisen 1994).

The last estimation procedure in the semiparametric class considered here is the tapered method, denoted by GPHTa and proposed by Hurvich and Ray (1995) and Velasco (1999b). This is also a periodogram regression method where one uses the cosine-bell function as a spectral window to reduce the bias of the periodogram function. The spectral window is given by

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

In this case, the modified periodogram function is given by

$$I(w_j) = \frac{1}{2\pi \sum_{t=0}^{n-1} \lambda(t)^2} \left| \sum_{t=0}^{n-1} \lambda(t) X_t \, \mathrm{e}^{-iw_j t} \right|^2 .$$

This estimator regresses $\ln\{I(w_j)\}$ on $\ln\{2 \sin(\frac{w_j}{2})\}$, for $j = l, l+1, \ldots, g(n)$. Its asymptotic variance (see Velasco 1999a) is given by

$$g(n) \times \mathrm{Var}(\mathrm{GPHTa}) \approx \frac{\pi^2}{8}, \tag{2.4}$$

where $g(n) = n^{\alpha}$, for $0 < \alpha < 1$.

In our investigation we also consider the frequency domain FT estimator, proposed by Fox and Taqqu (1986), as a parametric procedure. The estimate obtained from this procedure has asymptotic variance given by $n\,\mathrm{Var}(\mathrm{FT}) = 6/\pi^2$, where n is the sample size (see Beran 1994).

3 Simulation results

In this section we analyze the behavior of the estimators, presented in Sect. 2.3, before, and after applying a first difference to a time series generated from an ARFIMA(p, d^*, q) model, where $d^* \in [0.5, 1.5)$.

Let $\{X_t\}_{t \in Z}$ be an ARFIMA$(0, d^*, 0)$, given by the expression (2.2) with $p = 0 = q$. From expression (2.2),

$$Y_t = (1 - \mathcal{B}) X_t, \qquad t \in Z,$$

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

The stationary ARFIMA processes were generated using the algorithm proposed by Hosking (1984). The processes $\{X_t\}_{t \in Z}$ were obtained through the algebraic form $X_t = (1 - \mathcal{B})^{-1} Y_t$, for $t \in I\!N - \{0\}$, where $X_1 = Y_1$. We used Fortran (IMSL) subroutines for simulation and estimation results. The ARFIMA(p, d^*, q) processes were simulated including the autoregressive and moving average components in the $\{X_t\}_{t \in Z}$ process. Let $\hat{d}^*$ be the estimator of d^*, and $\hat{d}$ be the fractional estimator obtained from the first differenced data. The main goal is to verify the equality $\hat{d}^* = \hat{d} + 1$.

The simulation study is based on 2,000 replications of time series with three different sample sizes ($n \in \{256, 512, 1,024\}$). However, in the present paper we only show the results for the two first sample sizes. The parameter estimates when using $n = 1,024$ were very similar to the cases presented here and are available upon request. In the semiparametric methods the bandwidths were set at $\alpha = 0.5$ and 0.8. These sample sizes were also considered by Velasco (1999a).

All tables give the mean ($\overline{E}(\cdot)$), the mean squared error (mse) values for the estimators before, and after applying the first difference to the time series and the quantity $\nu \equiv 1 + \overline{E}(\hat{d}) - \overline{E}(\hat{d}^*)$. The smallest mean squared error value is given in the tables by a boldfaced character. The bandwidth $g(n) = n^{\alpha}$ in the semiparametric methods was fixed at $\alpha_1 = 0.5$ (this is a commonly used value in the literature) and at $\alpha_2 = 0.8$ (see Lopes et al. 2004; Velasco 1999a; Hurvich and Ray 1995). All the estimates from the semiparametric class are denoted here, respectively, by GPH(i), SPR(i) and GPHTa(i), for $i = 1, 2$. The truncation point in the Parzen lag window in the SPR estimator was taken $m = n^{\beta}$, with $\beta = 0.9$ (see Reisen 1994 for a discussion on the β value). The value of l was set equal to 2 in the GPHTa estimator (see Theorem 3 in Hurvich and Ray 1995).

The Case of ARFIMA$(0, d^*, 0)$

For pure ARFIMA processes we considered several different values of $d^* \in [0.5, 1.0]$. Table 1 shows the simulation results when $d^* = 0.6, 0.8, 1.0$. For d^* in this range it is expected that the first differentiated series is now belonging to the intermediate or short-memory class.

One observes that the FT method gives estimated mean value very close to the true parameter d^* with very small mean squared error values. The semiparametric methods also performed reasonably well but the GPHTa procedure has the largest mean squared error values. The variance of the semiparametric estimates decreases with the increase of the bandwidth. This also reflects in the decrease of the mean squared error value of the estimates. An investigation in this direction can also be found in Hurvich and Ray (1995), and Lopes et al. (2002).

By looking at the ν quantity, one observes its closeness to the zero value, indicating that the relation $E(\hat{d}^*) \simeq 1 + E(\hat{d})$ holds for all considered methods. Therefore, the estimators are, in mean, nearly invariant. The study reveals that these methods can be used to model non-stationary time series with the level reversion property. Also, the estimate of d ($\hat{d}$) after first difference plus 1 ($\hat{d}+1$) is close to the true parameter. This empirical study visions that $E(\hat{d}) + 1 = d^*$. This can be useful in practical situations where the practitioner, after first differentiating an ARFIMA$(0, d^*, 0)$ time series will get an estimate of $d \in (-0.5, 0.0)$ with small bias. One observes that when the sample size n increases the bias, and the mean squared error values decrease.

For $d^* > 1.0$ all estimators, except the GPHTa, are very much biased and the invariance property for the first difference does not hold anymore (see Table 2). The picture of the empirical study, for this case, changes compared with the case when $d^* \leq 1.0$. Now all estimators, besides GPHTa, underestimate d^*. The bias increases dramatically as the value of d^* increases. The estimates are always close to one no matter what is the value of d^*. The conclusion is that $E(\hat{d}^*) \simeq 1$ for GPH, SPR, and FT estimators. This may cause a problem in practical situations since obtaining a value close to one does not necessarily indicates that the series is an ARFIMA$(0, d^*, 0)$ model with $d^* = 1.0$. The property $E(\hat{d}^*) \simeq 1 + E(\hat{d})$ seems to hold only for GPHTa. However, it is interesting

Table 1 Estimation results from the ARFIMA$(0, d^*, 0)$ model, when $d^* \in [0.5, 1.0]$, for different sample sizes

n	d^*	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH(1)	0.6188	0.0438	−0.3822	0.0459	−0.0010
		GPH(2)	0.6133	0.0066	−0.3910	0.0062	−0.0043
		SPR(1)	0.5549	0.0312	−0.3877	0.0229	0.0574
	0.6	SPR(2)	0.6003	0.0043	−0.3947	0.0037	0.0050
		GPHTa(1)	0.6119	0.1607	−0.4018	0.1621	−0.0137
		GPHTa(2)	0.6135	0.0136	−0.3902	0.0141	−0.0040
		FT	0.6070	**0.0028**	−0.3972	**0.0027**	−0.0042
		GPH(1)	0.8365	0.0460	−0.1844	0.0440	−0.0209
		GPH(2)	0.8308	0.0084	−0.2002	0.0063	−0.0310
		SPR(1)	0.7801	0.0314	−0.2148	0.0243	0.0052
256	0.8	SPR(2)	0.8194	0.0054	−0.2065	0.0041	−0.0258
		GPHTa(1)	0.8213	0.4576	−0.1995	0.3429	−0.0208
		GPHTa(2)	0.8304	0.0145	−0.1984	0.0029	−0.0250
		FT	0.8252	**0.0039**	−0.1946	**0.0028**	−0.0199
		GPH(1)	0.9994	0.0369	0.0014	0.0430	0.0020
		GPH(2)	1.0000	0.0051	−0.0027	0.0062	−0.0027
		SPR(1)	0.9670	0.0273	−0.0437	0.0256	−0.0107
	1.0	SPR(2)	1.0036	0.0035	−0.0141	0.0042	−0.0176
		GPHTa(1)	0.9784	0.4756	−0.0136	0.3569	0.0080
		GPHTa(2)	1.0089	0.0125	0.0140	0.0015	−0.0074
		FT	0.9934	**0.0023**	−0.0037	**0.0028**	0.0029
		GPH(1)	0.6273	0.0307	−0.3781	0.0305	−0.0054
		GPH(2)	0.6122	0.0037	−0.3936	0.0034	−0.0059
		SPR(1)	0.5810	0.0208	−0.3808	0.0154	0.0382
	0.6	SPR(2)	0.6036	0.0025	−0.3959	0.0021	0.0005
		GPHTa(1)	0.6178	0.0949	−0.3930	0.0944	−0.0108
		GPHTa(2)	0.6120	0.0078	−0.3939	0.0072	−0.0044
		FT	0.6066	**0.0015**	−0.3978	**0.0014**	−0.0044
		GPH(1)	0.8365	0.0301	−0.1859	0.0284	−0.0224
		GPH(2)	0.8242	0.0049	−0.1992	0.0033	−0.0234
		SPR(1)	0.8003	0.0215	−0.2063	0.0159	−0.0066
512	0.8	SPR(2)	0.8193	0.0032	−0.2031	0.0021	−0.0223
		GPHTa(1)	0.8205	0.1000	−0.1971	0.1005	−0.0176
		GPHTa(2)	0.8252	0.0089	−0.1942	0.0079	−0.0200
		FT	0.8242	**0.0025**	−0.1957	**0.0013**	−0.0199
		GPH(1)	0.9879	0.0216	0.0178	0.0312	0.0299
		GPH(2)	0.9998	0.0027	−0.0015	0.0034	−0.0014
		SPR(1)	0.9915	0.0151	−0.0175	0.0185	−0.0090
	1.0	SPR(2)	1.0051	0.0018	−0.0088	0.0023	−0.0014
		GPHTa(1)	1.0256	0.0984	−0.0008	0.1005	−0.0264
		GPHTa(2)	1.0077	0.0073	0.0021	0.0077	0.0007
		FT	0.9960	**0.0010**	0.0113	**0.0015**	0.0153

to note that for all methods the relation $1+\overline{E}(\hat{d}) \simeq d^*$ holds. This is a very useful information for practical purposes. If the time series is an ARFIMA$(0, d^*, 0)$ model, with $1.0 < d^* < 1.5$, the first difference fractional estimate will be in the range $(0.0, 0.5)$ with small bias.

Table 2 Estimation results from the ARFIMA$(0, d^*, 0)$ model, when $d^* \in (1.0, 1.5)$, for different sample sizes

n	d^*	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH(1)	1.0522	0.0371	0.1067	0.0438	0.0545
		GPH(2)	1.0442	0.0090	0.0990	0.0059	0.0548
		SPR(1)	1.0402	0.0270	−0.0450	0.0285	−0.0852
	1.1	SPR(2)	1.0615	0.0045	0.0871	0.0044	0.0256
		GPHTa(1)	1.1450	0.1689	0.1001	0.1766	−0.0449
		GPHTa(2)	1.0817	0.0121	0.1055	0.0148	0.0237
256		FT	1.0447	**0.0061**	0.0983	**0.0029**	0.0536
		GPH(1)	1.0445	0.1844	0.4598	0.0462	0.4152
		GPH(2)	1.0281	0.1857	0.4581	0.0067	0.4300
		SPR(1)	1.1227	**0.1147**	0.3981	0.0323	0.2754
	1.45	SPR(2)	1.0699	0.1478	0.4439	0.0048	0.3740
		GPHTa(1)	1.1939	0.1735	0.4334	0.1642	0.2395
		GPHTa(2)	1.3127	0.0311	0.4641	0.0150	0.1515
		FT	1.0271	0.1875	0.4928	**0.0048**	0.4657
		GPH(1)	1.0516	0.0249	0.1008	0.0294	0.0492
		GPH(2)	1.0422	0.0069	0.1008	0.0033	0.0586
		SPR(1)	1.0587	0.0164	0.0643	0.0180	0.0056
	1.1	SPR(2)	1.0615	0.0031	0.0933	0.0021	0.0318
		GPHTa(1)	1.1357	0.0952	0.1057	0.0963	−0.0300
		GPHTa(2)	1.0768	0.0069	0.1054	0.0071	0.0286
		FT	1.0455	**0.0048**	0.0984	**0.0013**	0.0529
512		GPH(1)	1.0391	0.1823	0.4632	0.0305	0.4241
		GPH(2)	1.0194	0.1900	0.4575	0.0034	0.4380
		SPR(1)	1.1210	0.1135	0.4171	0.0212	0.2961
	1.45	SPR(2)	1.0522	0.1602	0.4489	0.0023	0.3967
		GPHTa(1)	1.5500	**0.0965**	0.4566	0.0917	−0.0934
		GPHTa(2)	1.2793	0.0357	0.4620	0.0073	0.1827
		FT	1.0219	0.1889	0.4731	**0.0019**	0.4512

Lopes and Pinheiro (2006) have showed that an estimator of d^* based on the wavelet theory is very competitive for both situations with, and without level-reversion property. We believe that the invariance property will hold for this estimator even when $d^* > 1.0$ (this is a topic for a forthcoming paper). An extensive simulation study related to the bias of d^* for a nonstationary ARFIMA process is presented by Lopes et al. (2002, 2004).

The Case of ARFIMA(p, d^*, q)

We now analyze the invariance property for the first difference in ARFIMA $(1, d^*, 0)$ model, when $d^* + 0.8$, and the results are presented in Table 3. From this table one can see that the semiparametric estimators follow the same behavior observed in Table 1. For $\phi = -0.6$ the estimators do not suffer large impact with the autoregressive component presented in the model. For positive ϕ (for instance, $\phi = 0.6$), the SPR method dominates the study by showing smaller bias, and mean squared error values. As it is known, to

Table 3 Estimation results from the ARFIMA$(1, d^*, 0)$ model, when $d^* = 0.8$ and $\phi \in \{-0.6, 0.6\}$, for different sample sizes

ϕ	n	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH(1)	0.8257	0.0424	−0.1984	0.0470	−0.0241
		GPH(2)	0.6957	0.0194	−0.3401	0.0262	−0.0359
		SPR(1)	0.7689	0.0316	−0.2242	0.0266	0.0069
	256	SPR(2)	0.6846	0.0192	−0.3480	0.0262	−0.0326
		GPHTa(1)	0.8044	0.1781	−0.2212	0.2015	−0.0256
		GPHTa(2)	0.7044	0.0232	−0.3309	0.0317	−0.0353
		FT	0.8457	**0.0112**	−0.1956	**0.0040**	−0.0413
−0.6		GPH(1)	0.8242	0.0299	−0.1977	0.0295	−0.0219
		GPH(2)	0.7331	0.0097	−0.2976	0.0132	−0.0308
		SPR(1)	0.7909	0.0209	−0.2148	0.0162	−0.0057
	512	SPR(2)	0.7249	0.0090	−0.3029	0.0129	−0.0279
		GPHTa(1)	0.8097	0.0957	−0.2066	0.0944	−0.0163
		GPHTa(2)	0.7337	**0.0125**	−0.1977	**0.0019**	−0.0269
		FT	0.8714	0.0150	−0.1964	0.0019	−0.0678
		GPH(1)	0.9042	0.0529	−0.1176	0.0562	−0.0218
		GPH(2)	1.1310	0.1204	0.2165	0.1796	0.0854
		SPR(1)	0.8479	**0.0319**	−0.1442	**0.0308**	0.0079
	256	SPR(2)	1.1393	0.1199	0.2086	0.1713	0.0693
		GPHTa(1)	0.9609	0.1878	−0.0726	0.2016	−0.0335
		GPHTa(2)	1.1512	0.1387	0.2171	0.1929	0.0659
		FT	1.2830	0.5433	−0.1456	0.0336	−0.4286
0.6		GPH(1)	0.8714	0.0333	−0.1521	0.0308	−0.0235
		GPH(2)	1.1147	0.1047	0.1661	0.1373	0.0514
		SPR(1)	0.8335	**0.0208**	−0.1759	**0.0168**	−0.0094
	512	SPR(2)	1.1220	0.1063	0.1621	0.1332	0.0400
		GPHTa(1)	0.8625	0.0987	−0.1537	0.0969	−0.0162
		GPHTa(2)	1.1238	0.1135	0.1664	0.1437	0.0426
		FT	1.3994	0.6515	−0.1857	0.0181	−0.5851

compute the FT estimates it is necessary to make use of a numerical method to achieve the maximum value of a function which depends on the periodogram and the spectral density functions. This procedure depends on the parameter range values. Here, we have decided to work with a very large parameter width to avoid any estimate to attain the boundary constraints. This can dramatically changes the picture of the FT estimates and it may explain that FT method presented estimates with very large biases. It seems that the method has difficulties for distinguishing between positive correlation caused by d and positive correlation caused by ϕ. This affects substantially the estimates in this method, by presenting large values of both bias, and mean squared error. The estimator GPHTa has the largest standard deviation value for $\hat{d}^*$, and $\hat{d}$ even when the sample size increases. One study case of the previously reported results is also shown graphically by the boxplot figures (see Figs 1, 2).

Analyzing the invariance property we can see the difference $(\overline{E}(\hat{d}) + 1) - \overline{E}(d^*)$ is not very large, and it is comparable with the case in Table 1, except for the FT method when $\phi > 0$ (see Figs. 1, 2).

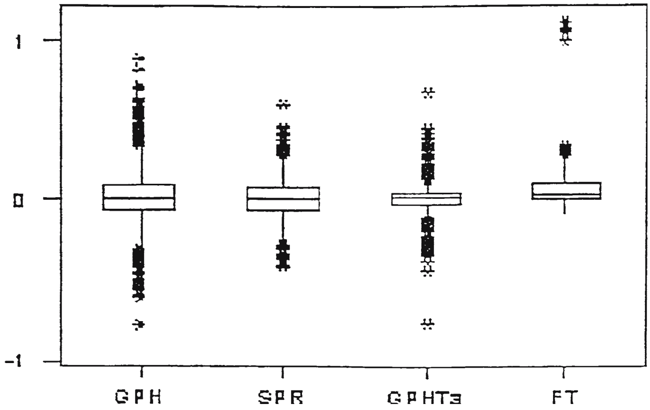

Fig. 1 Box-plot of $\hat{d}^* - (1 + \hat{d})$, for the ARFIMA$(1, d^*, 0)$ model, when $d^* = 0.8$, $\phi = -0.6$ and $n = 512$

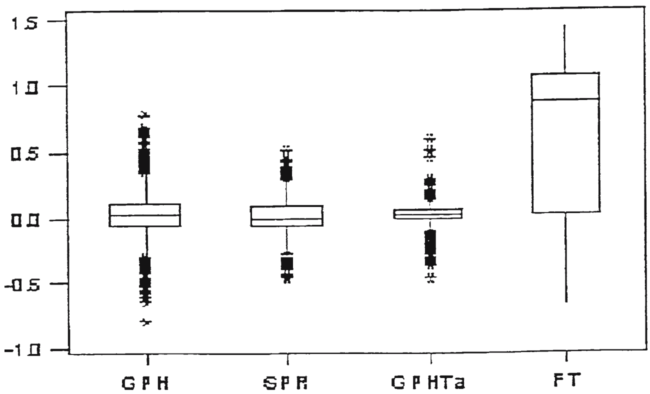

Fig. 2 Box-plot of $\hat{d}^* - (1 + \hat{d})$, for the ARFIMA$(1, d^*, 0)$ model, when $d^* = 0.8$, $\phi = +0.6$ and $n = 512$

The picture of the estimates from the ARFIMA$(0, d^*, 1)$ model (see Table 4) is similar to the ARFIMA$(1, d^*, 0)$ case but in the opposite way. Now, FT method presents large bias for negative value of θ. Again the semiparametric methods are not too much affected by the correlation structure caused by θ, and the value of ν is relatively small, comparable with the ARFIMA$(0, d^*, 0)$ process.

We now consider the case when $p = 1 = q$. Table 5 presents the simulation results when $\phi = -0.6$ and $\theta = 0.2$, and when $\phi = 0.2$ and $\theta = 0.6$, both for $d^* = 0.8$, and sample size equal to $n = 256$. In this situation, the SPR is the estimator that suffers less impact, compared with the others, from the short-memory parameter combinations considered here. The method GPHTa has the largest mean squared error value for both $\hat{d}^*$, and $\hat{d}$ in this case.

Table 4 Estimation results from the ARFIMA$(0, d^*, 1)$ model, when $d^* = 0.8$ and $\theta \in \{-0.6, 0.6\}$, for different sample sizes

θ	n	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH(1)	0.8410	0.0419	−0.1937	0.0454	−0.0347
		GPH(2)	0.9526	0.0297	−0.0515	0.0287	−0.0041
		SPR(1)	0.7818	0.0295	−0.2235	0.0256	−0.0053
	256	SPR(2)	0.9469	**0.0257**	−0.0585	0.0242	−0.0054
		GPHTa(1)	0.8198	0.1612	−0.2060	0.1875	−0.0258
		GPHTa(2)	0.9524	0.0368	−0.0486	0.0394	−0.0010
		FT	1.8599	1.1664	−0.1946	**0.0040**	−1.0545
−0.6		GPH(1)	0.8405	0.0321	−0.1929	0.0306	−0.0334
		GPH(2)	0.9181	0.0175	−0.0962	0.0142	−0.0143
		SPR(1)	0.8005	0.0217	−0.2128	0.0173	−0.0133
	512	SPR(2)	0.9145	**0.0156**	−0.1006	0.0121	−0.0151
		GPHTa(1)	0.8216	0.0962	−0.1955	0.0966	−0.0171
		GPHTa(2)	0.9126	0.0206	−0.0922	0.0202	−0.0048
		FT	1.8996	1.2095	−0.1979	**0.0018**	−1.0975
		GPH(1)	0.7395	0.0493	−0.2729	0.0536	−0.0124
		GPH(2)	0.4332	0.1443	−0.6030	0.1695	−0.0361
		SPR(1)	0.6841	0.0457	−0.2886	**0.0304**	0.0273
	256	SPR(2)	0.4178	0.1525	−0.6072	0.1700	−0.0249
		GPHTa(1)	0.6515	0.1493	−0.3553	0.1764	−0.0068
		GPHTa(2)	0.4539	0.1352	−0.5887	0.1668	−0.0426
		FT	0.8128	**0.0373**	−0.1977	0.0389	−0.0105
0.6		GPH(1)	0.7779	0.0284	−0.2423	0.0303	−0.0202
		GPH(2)	0.4693	0.1146	−0.5578	0.1316	−0.0271
		SPR(1)	0.7386	0.0237	−0.2573	**0.0177**	0.0041
	512	SPR(2)	0.4609	0.1184	−0.5622	0.1334	−0.0231
		GPHTa(1)	0.7092	0.0848	−0.4622	0.0778	−0.1714
		GPHTa(2)	0.4822	0.1093	−0.5504	0.1308	−0.0326
		FT	0.8351	**0.0220**	−0.1893	0.0212	−0.0244

Table 5 Estimation results from the ARFIMA$(1, d^*, 1)$ model, when $d^* = 0.8$, for different values of ϕ and θ and sample size $n = 256$

ϕ	θ	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH(1)	0.8223	0.0395	−0.2029	0.0467	−0.0252
		GPH(2)	0.5902	0.0531	−0.4429	0.0653	−0.0331
		SPR(1)	0.7595	**0.0306**	−0.2270	**0.0243**	0.0135
−0.6	0.2	SPR(2)	0.5782	0.0553	−0.4462	0.0644	−0.0245
		GPHTa(1)	0.7781	0.1413	−0.2479	0.1703	−0.0260
		GPHTa(2)	0.6043	0.0525	−0.4321	0.0680	−0.0365
		FT	0.8605	0.0329	−0.1559	0.0296	−0.0164
		GPH(1)	0.7581	0.0450	−0.2635	0.0488	−0.0216
		GPH(2)	0.5339	0.0795	−0.5012	0.0971	−0.0351
		SPR(1)	0.7033	**0.0386**	−0.2810	**0.0288**	0.0157
0.2	0.6	SPR(2)	0.5202	0.0839	−0.5073	0.0982	−0.0275
		GPHTa(1)	0.6883	0.1553	−0.3182	0.1738	−0.0065
		GPHTa(2)	0.5486	0.0776	−0.4933	0.1000	−0.0419
		FT	0.8888	0.0801	−0.1190	0.0802	−0.0078

Table 6 Estimate of the parameter in ARFIMA$(0, d^*, 0)$ when the generated process was an ARFIMA$(1, d^*, 0)$ model with $d^* = 1.1$, $\phi \in \{-0.6, 0.6\}$ and $n \in \{256, 512\}$

ϕ	n	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH	1.0474	0.0369	0.0888	0.0459	0.0415
	256	SPR	1.0326	**0.0271**	0.0377	0.0305	0.0052
		GPHTa	1.0455	0.5357	0.0653	0.4142	0.0198
		FT	1.7763	0.5784	0.0921	**0.0041**	−0.6842
−0.6		GPH	1.0475	0.0275	0.0981	0.0293	0.0506
	512	SPR	1.0548	**0.0171**	0.0618	0.0199	0.0070
		GPHTa	1.0462	0.2461	0.0906	0.2003	0.0444
		FT	1.7880	0.5802	0.0961	**0.0018**	−0.6919
		GPH	1.0833	0.0358	0.1771	0.0525	0.0938
	256	SPR	1.0777	**0.0211**	0.1225	**0.0278**	0.0447
		GPHTa	1.1826	0.5077	0.1965	0.4276	0.0139
		FT	1.8231	0.5761	0.1079	0.0308	−0.7152
0.6		GPH	1.0663	0.0252	0.1388	0.0303	0.0725
	512	SPR	1.0748	**0.0149**	0.1006	**0.0168**	0.0258
		GPHTa	1.1022	0.2508	0.1423	0.2065	0.0400
		FT	1.8223	0.5752	0.0860	0.0180	−0.7363

Table 7 Estimate of the parameter in ARFIMA$(0, d^*, 0)$ when the generated process was the ARFIMA$(1, d^*, 0)$ model with $d^* = 1.45$, $\phi \in \{-0.6, 0.6\}$ and $n \in \{256, 512\}$

ϕ	n	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH	1.0470	0.1819	0.4575	0.0508	0.4105
	256	SPR	1.1245	**0.1142**	0.3914	0.0357	0.2669
		GPHTa	1.1875	1.2843	0.4334	0.4405	0.2459
		FT	1.8050	0.2020	0.5105	**0.0082**	−0.2945
−0.6		GPH	1.0477	0.1787	0.4599	0.0301	0.4122
	512	SPR	1.1248	**0.1120**	0.4153	0.0206	0.2905
		GPHTa	1.2092	0.5180	0.4370	0.2192	0.2279
		FT	1.8047	0.2013	0.4828	**0.0031**	−0.3219
		GPH	1.0541	0.1800	0.5367	0.0563	0.4826
	256	SPR	1.1340	**0.1085**	0.4703	**0.0319**	0.3363
		GPHTa	1.2826	1.2181	0.5678	0.4574	0.2852
		FT	1.8081	0.2000	0.9437	0.4619	0.1356
0.6		GPH	1.0448	0.1807	0.5065	0.0307	0.4617
	512	SPR	1.1265	**0.1118**	0.4572	**0.0178**	0.3307
		GPHTa	1.2360	0.4945	0.5233	0.2361	0.2873
		FT	1.8070	0.2004	1.0540	0.7234	0.2470

Parameters' estimation in model misspecification

We now analyze the simulation experiment results when one specifies the wrong process. Tables 6, 7 and 8 present these results where one estimates the parameter in an ARFIMA$(0, d^*, 0)$ model when the generated process was the ARFIMA$(1, d^*, 0)$, with $d^* \in \{1.1, 1.45\}$ (see Tables 6, 7). We also

Table 8 Estimate of the parameter in ARFIMA$(0, d^*, 0)$ when the generated process was the ARFIMA$(0, d^*, 1)$ model with $d^* = 0.8, \theta \in \{-0.6, 0.6\}$ and $n \in \{256, 512\}$

θ	n	Method	$\overline{E}(\hat{d}^*)$	mse$(\hat{d}^*)$	$\overline{E}(\hat{d})$	mse$(\hat{d})$	ν
		GPH	0.8393	0.0443	−0.1894	0.0440	−0.0287
	256	SPR	0.7882	**0.0286**	−0.2170	0.0254	−0.0053
		GPHTa	0.8264	0.3997	−0.2109	0.3865	−0.0373
		FT	1.8157	1.0909	−0.1962	**0.0039**	−1.0119
−0.6		GPH	0.8389	0.0313	−0.1884	0.0292	−0.0273
	512	SPR	0.7998	**0.0207**	−0.2094	0.0167	−0.0091
		GPHTa	0.8227	0.2204	−0.2089	0.2116	−0.0316
		FT	1.8143	1.0892	−0.1981	**0.0019**	−1.0125
		GPH	0.7481	0.0441	−0.2596	0.0519	−0.0077
	256	SPR	0.6893	**0.0412**	−0.2815	**0.0296**	0.0292
		GPHTa	0.5476	0.3049	−0.3744	0.3783	0.0779
		FT	0.3802	0.1814	−0.1924	0.0358	0.4273
0.6		GPH	0.7786	0.0264	−0.2367	0.0288	−0.0154
	512	SPR	0.7348	**0.0221**	−0.2549	**0.0179**	0.0102
		GPHTa	0.6345	0.1939	−0.3167	0.1905	0.0488
		FT	0.4023	0.1608	−0.1904	0.0203	0.4073

estimate the parameter in an ARFIMA$(0, d^*, 0)$ when the generated process was the ARFIMA$(0, d^*, 1)$, with $d^* = 0.8$ (see Table 8). From the empirical results in Tables 6, 7 and 8 one observes that the semiparametric methods do not depend on the corrected specification of the spectral density function to obtain an estimate for d^*. Therefore, the bias of the fractional estimator is unaffected by the precise order of the ARMA structure. However, a wrong specification for the model leads to a large bias for the parameter d^* when one uses the FT method. We refer the reader to Reisen et al. (2001b) where an extensive Monte Carlo study was carried out for analyzing misspecification of the model considering other estimation methods besides GPH, SPR and FT. In this later work the authors consider different values for α in $g(n) = n^{\alpha}$, and also consider the bias and the mean squared error values for the estimate of the short memory components when they are present in the model.

4 An application

As an application, we consider the monthly Brazilian exchange rate data from the period January 1979 to March 2002. The original time series, with 279 observations (see Fig. 3), clearly shows that the mean level changes with time. The correlogram (see Fig. 4) declines very slowly indicating a long memory behavior for this series.

The estimates are shown in Table 9 for the original, and first differenced series together with the asymptotic standard deviation value (denoted by σ) for each estimation method. Observe that for each semiparametric method we present two different values for the bandwidth $g(n) = n^{\alpha}$. In Table 9 we consider

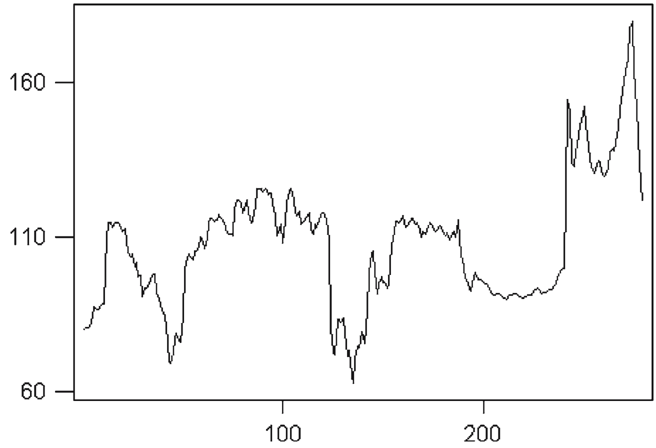

Fig. 3 Brazilian exchange rate data (01/1979 to 03/2002)

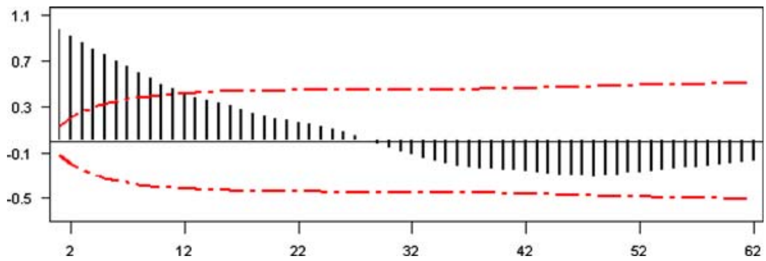

Fig. 4 Sample autocorrelation function of the Brazilian exchange rate data

$\alpha_1 = 0.5$ and $\alpha_2 = 0.8$ and the estimates of the semiparametric class are denoted, respectively, by GPH(i), SPR(i) and GPHTa(i), for $i = 1, 2$. Table 10 presents the upper and lower bounds of the 95% confidence interval for the unit root based on each four estimation procedures. From Tables 9 and 10 we observe that different bandwidths produce totally different estimates. This may indicate that the candidate model probably has short-memory component. Large value of α produces estimates with smaller standard deviation and they are very similar to the FT estimate. Results from the tables also suggest a unit root process.

In Table 11 we give the results for the augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests where the null hypothesis is the presence of a unit root and the alternative hypothesis is the stationarity for the series. From this table we observe that for both tests we can not reject the null hypothesis at 5% of significance level, indicating that the Brazilian exchange rate data possesses a unit root.

The above investigation suggests to perform an analysis of the short-term components for the data. When using the semiparametric methods the ARMA process order was identified after differentiating the original time series by

Table 9 Estimation results for the original and the differenced Brazilian exchange rate series

Method	GPH(1)	GPH(2)	SPR(1)	SPR(2)	GPHTa(1)	GPHTa(2)	FT
Original Data							
Estimate	0.5020	1.0528	0.6390	1.0556	0.9330	1.2741	1.0959
σ	0.2102	0.0770	0.0906	0.0332	0.2777	0.1171	0.0467
Differenced Data							
Estimate	–0.3415	–0.0104	–0.2002	0.0734	–0.2130	0.3616	0.2571
σ	0.2102	0.0770	0.0907	0.0332	0.2777	0.1171	0.0467

Note: The numbers 1 and 2, in parenthesis, for the semiparametric methods, indicate that $g(n) = n^{\alpha_i}$, with $\alpha_1 = 0.5$ and $\alpha_2 = 0.8$. σ means the asymptotic standard deviation value

Table 10 Confidence interval for d^* at 95% confidence level based on the estimation methods

Method	GPH(2)	SPR(2)	GPHTa(2)	FT
Lower Bound	0.9019	0.9905	1.0446	1.0044
Upper Bound	1.2037	1.1207	1.5036	1.1874

Note: The number 2, in parenthesis, for the semiparametric methods, indicates that $g(n) = n^{\alpha_2}$, with $\alpha_2 = 0.8$.

Table 11 Results of the augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) tests for the Brazilian exchange rate time series

Test	ADF	PP
Statistic Value	–3.1216	–15.0253
Lag-order	6	5
p-Value	0.1035	0.2675

Table 12 Estimate results of the short-term components for the Brazilian exchange rate series

Estimate	GPH(2)	SPR(2)	GPHTa(2)	FT
$\hat{\phi}_1$	0.2501	0.2459	0.2514	0.2022
s.e.($\hat{\phi}_1$)	0.1182	0.1149	0.1193	0.0925
$\hat{\theta}_1$	0.6608	0.6691	0.6579	0.7419
s.e.($\hat{\theta}_1$)	0.0915	0.0880	0.0927	0.0632
$\hat{\sigma}^2_\epsilon$	0.9708	0.9714	0.9706	0.9777
MBP	29.8	30.0	29.8	31.5

Notes: MBP – the modified Box-Pierce chi-square statistic with 21 degrees of freedom. The number 2, in parenthesis, for the semiparametric methods, indicates that $g(n) = n^{\alpha_2}$, with $\alpha_2 = 0.8$

the estimate of d^*. Only the estimation method FT is a one-step procedure performing the estimation of the entire vector of parameters. We tried several models with different orders by using the Minitab package. Table 12 shows the estimation results and other statistics related to this analysis and presents the estimates of the AR and MA components and their standard errors (denoted respectively by s.e.($\hat{\phi}_1$) and s.e.($\hat{\theta}_1$)).

The residual analysis was also performed for the fitted model and it indicates that the errors are approximately Gaussian white noise. The hypothesis of adequated model is not rejected at 5% of significance level, for all four estimation methods, based on the modified Box-Pierce (MBP) test.

All methods indicate that the series is a non-stationary ARFIMA$(1, d^*, 1)$ model. Looking at the estimation results for the first differenced data, all estimates approximately present the invariance property.

5 Conclusions

This work presented the performance of four estimation methods, belonging to the parametric and semiparametric classes, for non-stationary ARFIMA models with main interest when $d^* \in [0.5, 1.5)$. The two general approaches to estimate this parameter are: use the original non-stationary time series to estimate d^*; take first difference of the original time series, then estimate d in the transformed stationary time series, and finally add 1 to the estimate of d to obtain an estimate for d^*.

For the ARFIMA$(0, d^*, 0)$ case, when $d^* \in [0.5, 1.0]$, and the corrected order of the process is considered, the parametric FT method is, in mean, the best estimation method with the invariance property for the first difference. However, the semiparametric methods also gave good results. Therefore, when $d^* \leq 1.0$, the simulation study reveals that both approaches perform well with all four estimation methods. On average, the estimate for d^* derived from the original time series equals one plus the estimate for d derived from the transformed series. However, the study also reveals that the second approach, when one uses first differences, performs far better than the first when $d^* > 1.0$. In these cases, the invariance property does not hold.

When $d^* > 1.0$, the tapered estimator indicates to be more nearly invariant to the first difference than the other methods. The procedures GPH, SPR, and FT produced estimates very close to one no matter the value of d^*.

The simulation study was extended to a general class of ARFIMA(p, d^*, q) model. The estimates presented similar performance to the ARFIMA$(0, d^*, 0)$ case except for the FT method. The estimates from this procedure were strongly affected by short-memory structures.

The authors also considered the semiparametric methods proposed in Robinson (1995), and Velasco (1999a). The results were very similar to the GPH method, and are available from the authors.

When misspecification for the model occurs, the estimation methods in the semiparametric class are not affected by the corrected specification of the spectral density function since they are model order independent.

As an application the Brazilian exchange rate series was analyzed and the estimates and tests indicated a non-stationary long memory time series modelled by an ARFIMA$(1, d^*, 1)$ process with $d^* = 1.0$.

Acknowledgments V.A. Reisen was partially supported by CNPq-Brazil. S.R.C. Lopes was partially supported by CNPq-Brazil and by Pronex *Probabilidade e Processos Estocásticos* (Convênio MCT/CNPq/FAPERJ – Edital 2003). S.R.C. Lopes and B.P. Olbermann were partially supported by *Fundação de Amparo à Pesquisa no Estado do Rio Grande do Sul* (FAPERGS Foundation).

The authors would like to thank the editor and the anonymous referee for their valuable comments and suggestions that improved this version of the manuscript.

References

Beran J (1994) Statistics for long-memory processes. Chapman & Hall, New York

Chen SK (1987) Modeling gold price movements and forecasts. Managerial Finance 13 (2):25–28

Cheung Y-W, Diebold FX (1994) On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. J Econom 62:301–316

Fox R, Taqqu MS (1986) Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. Ann Stat 14:517–532

Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series model. J Time Ser Anal 4(4):221–238

Granger CWJ, Joyeux R (1980) An introduction to long memory time series models and fractional differencing. J Time Ser Anal 1(1):15–29

Hauser MA (1999) Maximum likelihood estimators for ARMA and ARFIMA models: a Monte Carlo study. J Stat Plann Inference 80:229–255

Hosking J (1981) Fractional Differencing. Biometrika 68(1):165–167

Hosking J (1984) Modelling persistence in hydrological time series using fractional differencing. Water Resour Res 20(12):1898–1908

Hurvich CM, Ray BK (1995) Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. J Time Ser Anal 16(1):017–042

Lobato IN, Velasco C (2000) Long memory in stock-market trading volume. J Bus Econom Stat 18(4):410–427

Lopes SRC, Olbermann BP, Reisen VA (2004) A comparison of estimation methods in non-stationary ARFIMA processes. J Stat Comput Simul 74(5):339–347

Lopes SRC, Olbermann BP, Reisen VA (2002) Non-stationary Gaussian ARFIMA processes: estimation and application. Braz Rev Econom 22(1):103–126

Lopes SRC, Pinheiro AS (2006) Wavelets for estimating the fractional parameter in non-stationary ARFIMA process (in press)

Reisen VA (1994) Estimation of the fractional difference parameter in the ARIMA(p,d,q) model using the smoothed periodogram. J Time Ser Anal 15(3):335–350

Reisen VA, Abraham B, Lopes SRC (2001a) Estimation of parameters in ARFIMA processes: a simulation study. Commun Stat Simul Comput 30(4):787–803

Reisen VA, Sena Jr MR, Lopes SRC (2001b) Error and model misspecification in ARFIMA processes. Braz Rev Econom 21(1):101–135

Robinson PM (1995) Log-periodogram regression of time series with long range dependence. Ann Stat 23(3):1630–1661

Sowell F (1992a) Modeling long-run behavior with the fractional ARIMA model. J Monet Econ 29:277–302

Sowell F (1992b) Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J Econom 53:165–188

Velasco C (1999a) Gaussian semiparametric estimation of non-stationary time series. J Time Ser Anal 20(1):87–127

Velasco C (1999b) Non-stationary log-periodogram regression. J Econom 91:325–371