Long-range Dependence in Mean and Volatility: Models, Estimation and Forecasting

Sílvia R. C. Lopes

Abstract. In this paper we consider the estimation and forecasting future values of some stochastic processes exhibiting *long-range dependence*, both in mean and in volatility. We summarize basic definitions, properties and some results considering ARFIMA and SARFIMA processes, which exhibit *long memory in mean*. We proceed in the same manner considering FIGARCH and Fractionally Integrated Stochastic Volatility (FISV) processes where one can find *long memory in volatility*.

Estimation methods in parametric and semiparametric classes are presented for estimating the fractional parameter based on the classical *Ordinary Least Squares* and two robust methodologies (the *Least Trimmed Squares* and the *MM-estimation*).

An application of the SARFIMA methodology, based on the Nile River monthly flows data, is presented.

Mathematics Subject Classification (2000). Primary 60G10, 62G05, 62G35, 62M10, 62M15; Secondary 62M20.

Keywords. Long-range dependence, long memory, ARFIMA models, SARFIMA models, FIGARCH models, stochastic volatility models, semiparametric and parametric estimation, robust estimation, forecasting.

1. Introduction

Models for *long-range dependence*, or *long memory*, *in mean* were first introduced by Mandelbrot and Wallis [32], Mandelbrot and Taqqu [33], Granger and Joyeux [17] and Hosking [18], following the seminal work of Hurst [21].

We refer the reader to a recent collection of papers in Doukhan et al. [12] which reviews long-range dependence from many different angles, both theoretically and in the applied sense. We also refer the book by Palma [38] for general results on long-range dependence.

Persistence or *long-range dependence* property has been observed in time series in different areas of the science such as meteorology, astronomy, hydrology, and economics, as reported in Beran [4]. The *long-range dependence* can be characterized by two different but equivalent (see Bary [3]) forms given below, where $0.0 < d < 0.5$ is a constant:

- in time domain, the autocorrelation function $\rho_X(\cdot)$ decays hyperbolically to zero, that is, $\rho_X(k) \simeq k^{2d-1}$, when $k \to \infty$;
- in frequency domain, the spectral density function $f_X(\cdot)$ is unbounded when the frequency is near zero, that is, $f_X(w) \simeq w^{-2d}$, when $w \to 0$.

One of the models that exhibits the long-range dependence is the Autoregressive Fractionally Integrated Moving Average (ARFIMA) process. While in the ARFIMA process, the autocorrelation function shows a hyperbolic decay rate, in an ARMA process this function decays in an exponential rate.

However, in an ARFIMA process one can not capture the periodicity frequently present in some real data sets, even though still the long memory feature occurs in these data. The so-called Seasonal Autoregressive Fractionally Integrated Moving Average (SARFIMA) processes are a natural extension of the ARFIMA process. This model takes into account the seasonality inherent to such data.

The ARFIMA framework was also naturally extended towards *volatility models*. The Fractionally Integrated Generalized Autoregressive Conditionally Heteroskedastic (FIGARCH) models were introduced by Baillie, Bollerslev and Mikkelsen [2] and Bollerslev and Mikkelsen [7], motivated by the fact that autocorrelation function of the squared, log-squared, or the absolute value series of an asset return decays slowly, even when the return series has no serial correlation. In order to model long memory in the second moment, Breidt et al. [9] introduced the Fractionally Integrated Stochastic Volatility (FISV) model.

In this paper we will analyze *long-range dependence in mean* and *volatility*. We shall consider estimation and forecasting for different models.

To describe a method that generates SARFIMA processes, it is convenient to have a closed formula for the Durbin-Levinson Algorithm. This algorithm is given by recurrence relations allowing, for the partial autocorrelation function of the process, to go from lag k to lag $(k+1)$. This algorithm relates autocorrelation and partial autocorrelation functions of a process and Brietzke et al. [10] give its closed formula for SARFIMA processes.

Models for heteroskedastic time series with long memory are of great interest in econometrics and finance, where empirical facts about asset returns have motivated the several extensions of GARCH type models (for a review, see Lopes and Mendes [26]). Many empirical papers have detected the presence of long memory in the volatility of risky assets, market indexes, exchange rates. As the number of models available increases, it becomes of interest a simple, fast, and accurate estimation procedure for the fractional parameter d, independent of the specification of a parametric model. The regression based semiparametric (semiparametric in the sense that a full parametric model is not specified for the spectral density function of the process) estimators seem to be the natural candidates. However, the performance of the semiparametric estimators is greatly affected by their asymptotic statistical properties, besides depending on their definition and estimation method and is also heavily dependent on the number of frequencies $m = g(n)$ used for the regression. Lopes and Mendes [26] considers several long memory models in mean and in volatility presenting some light on the heavy dependency of the frequency number m for the semi-parametric estimation procedures.

The regression method was introduced in the pioneer work of Geweke and Porter-Hudak [16], giving rise to several other proposals. Hurvich and Ray [22] introduced a cosine-bell function as a spectral window, to reduce bias in the periodogram function. They found that data tapering and the elimination of the first periodogram ordinate in the regression equation, could increase the estimator accuracy. However, smaller bias was obtained at the cost of a larger variance. Velasco [48] also considered smoothed versions of the periodogram function while in Velasco [49] the consistency and asymptotic normality of the regression estimators was proved for any d, considering non-stationary and non-invertible processes. Reisen et al. [41] carried out an extensive simulation study comparing both the semiparametric and parametric approaches in ARFIMA processes. Monte Carlo methods were also considered by Lopes et al. [28] for non-stationary ARFIMA processes.

However, not always the ultimate interest is just the estimation of the *fractional* or *seasonal fractional parameter*, respectively, denoted by d and D. Frequently, one can also be interested in forecasting values of the processes. Reisen and Lopes [42] present some simulations and applications of forecasting ARFIMA processes while, more recently, Bisognin and Lopes [5] give an account of the estimation and forecasting issues for SARFIMA processes.

The paper is organized as follows. In Section 2 we define ARFIMA, SARFIMA, FIGARCH and FISV processes presenting their definitions together with some properties and results. Section 3 summarizes a closed formula for the Durbin-Levinson's algorithm relating the partial autocorrelation and the autocorrelation functions for the SARFIMA$(0, D, 0)$ processes. In Section 4 we present one parametric and five semi-parametric estimation procedures and their respective robust versions. In Section 5 some forecasting theory for the models presented here is summarized. In Section 6 we present an application and Section 7 contains a summary of the paper.

2. Long Memory Models

In this section we define both the long memory models in mean and in volatility. We present below the basic definitions, some properties and results. In this section we consider ARFIMA, SARFIMA, FIGARCH and FISV processes.

2.1. ARFIMA(p, d, q) Processes

In this sub-section we define the ARFIMA process, which exhibits the *long memory in mean* characteristic.

Definition 2.1. Let $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ be a white noise process with zero mean and variance $\sigma_\varepsilon^2 > 0$, $\mathcal{B}$ the *backward-shift operator*, that is, $\mathcal{B}^k(X_t) = X_{t-k}$, $\phi(\cdot)$ and $\theta(\cdot)$ polynomials of degrees p and q, respectively, given by

$$\phi(z) = \sum_{j=0}^{p}(-\phi_j)z^j \quad \text{and} \quad \theta(z) = \sum_{k=0}^{q}(-\theta_k)z^k,$$

where ϕ_j, $1 \le j \le p$, and θ_k, $1 \le k \le q$, are real constants with $\phi_0 = -1 = \theta_0$. Let $\{X_t\}_{t\in\mathbb{Z}}$ be a linear process given by

$$\phi(\mathcal{B})(1-\mathcal{B})^d(X_t - \mu) = \theta(\mathcal{B})\varepsilon_t,\ t \in \mathbb{Z}, \tag{2.1}$$

where μ is the *mean* of the process, $d \in (-0.5, 0.5)$ and $\nabla^d \equiv (1-\mathcal{B})^d$ is the *difference operator*, defined as the binomial expansion

$$(1-\mathcal{B})^d = \sum_{k=0}^{\infty}\left(\begin{array}{c} d \\ k \end{array}\right)(-\mathcal{B})^k = 1 - d\mathcal{B} - \frac{d}{2!}(1-d)\mathcal{B}^2 - \frac{d}{3!}(1-d)(2-d)\mathcal{B}^3 - \cdots, \tag{2.2}$$

for all $d \in \mathbb{R}$, with

$$\binom{d}{k} = \frac{\Gamma(1+d)}{\Gamma(1+k)\Gamma(1+d-k)},$$

and $\Gamma(\cdot)$ the Gamma function (see Brockwell and Davis [11]). Then, the process $\{X_t\}_{t\in\mathbb{Z}}$ is called a *general fractional differenced* ARFIMA(p, d, q) *process*, where d is the *degree* or *parameter of fractional differencing.*

From the expression (2.1), the process

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

given by

$$\phi(\mathcal{B})U_t = \theta(\mathcal{B})\varepsilon_t,\ t \in \mathbb{Z},$$

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

If $d \in (-0.5, 0.5)$ then the process $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and invertible (see Theorem 2.4) and its *spectral density function* is given by

$$f_X(w) = f_U(w)\left[2\,\sin(\frac{w}{2})\right]^{-2d}, \text{ for } 0 < w \le \pi, \tag{2.3}$$

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 $w \to 0$.

The use of the spectral density function can be seen as a practical device for determining the precise rate of decay of the autocorrelation function.

When $p = 0 = q$, in the expression (2.1), one obtain the so-called *pure* ARFIMA$(0, d, 0)$ *process*.

Its commonly used the following terminology: the ARFIMA(p, d, q) process exhibits the characteristic 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 and in this case the classical estimation procedures presented in Section 4 of this work still hold (see Lopes et al. [28]; Olbermann et al. [36] and Velasco [49]). The level-reversion property no longer holds when $d \geq 1$. If $d \leq -0.5$ the ARFIMA process is non-invertible.

For more details on the properties of the ARFIMA(p, d, q) processes see, for instance, Hosking [18] and Lopes et al. [28]. We refer the reader to Sena Jr et al. [46] for an extensive Monte Carlo simulation study to evaluate the performance of some parametric and semi-parametric estimators for long and short-memory parameters of the ARFIMA(p, d, q) model with conditional heteroskedastic innovation errors (in fact, an ARFIMA-GARCH model).

2.2. SARFIMA$(p, d, q) \times (P, D, Q)_s$ Processes

In many practical situations a time series can exhibit a periodic pattern. This is a common feature in fields such as meteorology, economics, hydrology and astronomy. Sometimes, even in these fields, the periodicity can depend on time, that is, the autocorrelation structure of the data varies from season to season. In our analysis, we consider the seasonality period constant over seasons. However, the periodic pattern of such kind of time series can not be described by an ARFIMA(p, d, q) process.

We shall consider the *autoregressive fractionally integrated moving average with seasonality* processes, denoted hereafter by SARFIMA$(p, d, q) \times (P, D, Q)_s$, which are an extension of the ARFIMA(p, d, q) models, proposed by Granger and Joyeux [17] and Hosking [18]. The SARFIMA processes exhibit *long-range dependence in mean* besides the *seasonality of period* s.

We shall give some definitions and some properties for the SARFIMA(p, d, q) $\times(P, D, Q)_s$ processes. We recall, however, that these properties also hold for the ARFIMA(p, d, q) process when one considers $P = Q = 0$, $D = 0$ and $s = 1$. These properties are still true for the pure version ARFIMA$(0, d, 0)$ process, when $p = 0 = q$.

Definition 2.2. Let $\{X_t\}_{t \in \mathbb{Z}}$ be a stochastic process given by the expression

$$\phi(\mathcal{B})\Phi(\mathcal{B}^s)\nabla^d\nabla_s^D(X_t - \mu) = \theta(\mathcal{B})\Theta(\mathcal{B}^s)\varepsilon_t, \quad \text{for} \quad t \in \mathbb{Z}, \tag{2.4}$$

where μ is the *mean* of the process, $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a white noise process, $s \in \mathbb{N}$ is the seasonal period, $D, d \in (-0.5, 0.5)$, $\mathcal{B}$ is the *backward-shift operator*, that is, $\mathcal{B}^k(X_t) = X_{t-k}$, and $\mathcal{B}^{sk}(X_t) = X_{t-sk}$, ∇^d and ∇_s^D are, respectively, the *difference* and the *seasonal difference operators*, where ∇_s^D is given by

$$\nabla_s^D \equiv (1-\mathcal{B}^s)^D = \sum_{k\geq 0} \binom{D}{k} (-\mathcal{B}^s)^k = 1 - D\mathcal{B}^s - \frac{D(1-D)}{2!}\mathcal{B}^{2s} - \cdots \tag{2.5}$$

and ∇^d is given by the expression (2.2). The polynomials $\phi(\cdot)$, $\theta(\cdot)$, $\Phi(\cdot)$, and $\Theta(\cdot)$ have degrees p, q, P, and Q, respectively, and are defined by

$$\begin{aligned} \phi(z) &= \sum_{j=0}^{p} (-\phi_j)\, z^j, \quad \theta(z) = \sum_{k=0}^{q} (-\theta_k)\, z^k, \\ \Phi(z) &= \sum_{l=0}^{P} (-\Phi_l)\, z^l, \quad \Theta(z) = \sum_{m=0}^{Q} (-\Theta_m)\, z^m, \end{aligned}$$

where ϕ_j, $1 \leq j \leq p$, θ_k, $1 \leq k \leq q$, Φ_l, $1 \leq l \leq P$, and Θ_m, $1 \leq m \leq Q$ are constants and $\phi_0 = \theta_0 = -1 = \Phi_0 = \Theta_0$. Then, $\{X_t\}_{t\in\mathbb{Z}}$ is a *seasonal fractionally integrated ARIMA*$(p,d,q)\times(P,D,Q)_s$ process with period s, denoted by SARFIMA$(p,d,q)\times(P,D,Q)_s$, where d and D are, respectively, *the degree of fractional differencing and of seasonal fractional differencing* parameters.

Remark 2.3. (a) When $P = Q = 0$, $D = 0$ and $s = 1$ the SARFIMA$(p,d,q)\times(P,D,Q)_s$ process is just the ARFIMA(p,d,q) process (see Beran [4]). In this situation it is already known the behavior of the parameter estimators and also the forecasting properties for these models (see Lopes et al. [28]; Reisen et al. [41]; and Reisen and Lopes [42]).

(b) A particular case of the SARFIMA$(p,d,q)\times(P,D,Q)_s$ process is when $p = q = P = Q = 0$. This process is called the *seasonal fractionally integrated ARIMA model with period* s, denoted by SARFIMA$(0,D,0)_s$, which will be the main goal of our expository paper. It is given by

$$\nabla_s^D (X_t - \mu) \equiv (1-\mathcal{B}^s)^D (X_t - \mu) = \varepsilon_t, \quad t \in \mathbb{Z}. \tag{2.6}$$

In what follows we shall describe some of the properties of the SARFIMA-$(0,D,0)_s$ process. We recall that these properties also hold for the ARFIMA$(0,d,0)$ process, when $D = d$ and $s = 1$ (see, for instance, Hosking [18] and [19]).

Without loss of generality, we shall consider $\mu = 0$ in expressions (2.1), (2.4) and in their pure versions. For the extensions of these properties to the complete SARFIMA process, given by Definition 2.2, we refer the reader to Bisognin [6].

Theorem 2.4. *Let* $\{X_t\}_{t\in\mathbb{Z}}$ *be the SARFIMA*$(0,D,0)_s$ *process given by the expression* (2.6)*, with zero mean and* $s \in \mathbb{N}$ *as the seasonal period. Then,* (i). *when $D > -0.5$, $\{X_t\}_{t\in\mathbb{Z}}$ is an invertible process with infinite autoregressive representation given by*

$$\Pi(\mathcal{B}^s)X_t = \sum_{k\geq 0} \pi_k \mathcal{B}^{sk}(X_t) = \sum_{k\geq 0} \pi_k X_{t-sk} = \varepsilon_t,$$

where

$$\pi_k = \frac{-D(1-D)\cdots(k-D-1)}{k!} = \frac{(k-D-1)!}{k!(-D-1)!} = \frac{\Gamma(k-D)}{\Gamma(k+1)\Gamma(-D)}\cdot \tag{2.7}$$

When $k \to \infty$, $\pi_k \sim \frac{1}{\Gamma(-D)} k^{-D-1}$.

(ii). *when $D < 0.5$, $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary process with an infinite moving average representation given by*

$$X_t = \Psi(\mathcal{B}^s)\varepsilon_t = \sum_{k\geq 0} \psi_k \mathcal{B}^{sk}(\varepsilon_t) = \sum_{k\geq 0} \psi_k \varepsilon_{t-sk},$$

where

$$\psi_k = \frac{D(1+D)\cdots(k+D-1)}{k!} = \frac{(k+D-1)!}{k!(D-1)!} = \frac{\Gamma(k+D)}{\Gamma(k+1)\Gamma(D)}\cdot \tag{2.8}$$

When $k \to \infty$, $\psi_k \sim \frac{1}{\Gamma(D)} k^{D-1}$.

In the following, we assume that $D \in (-0.5, 0.5)$.

(iii). *The process $\{X_t\}_{t\in\mathbb{Z}}$ has spectral density function given by*

$$f_X(w) = \frac{\sigma_\varepsilon^2}{2\pi}\left[2\sin\left(\frac{sw}{2}\right)\right]^{-2D}, \quad 0 < w \leq \pi. \tag{2.9}$$

At the seasonal frequencies, for $\nu = 0, 1, \cdots, \lceil s/2 \rceil$, where $\lceil x \rceil$ means the integer part of x, it behaves as

$$f_X\left(\frac{2\pi\nu}{s} + w\right) \sim f_\varepsilon\left(\frac{2\pi\nu}{s}\right)(sw)^{-2D}, \quad \textit{when} \quad w \to 0\cdot$$

In the following, let A be the set $\{1, 2, \cdots, s-1\}$, and $\mathbb{Z}_\geq$ be the set $\{k \in \mathbb{Z} | k \geq 0\}$.

(iv). *The process $\{X_t\}_{t\in\mathbb{Z}}$ has autocovariance and autocorrelation functions of order k, $k \in \mathbb{Z}_\geq$, given, respectively, by*

$$\gamma_X(sk+\xi) = \begin{cases} \frac{(-1)^k\Gamma(1-2D)}{\Gamma(1+k-D)\Gamma(1-k-D)}\sigma_\varepsilon^2 = \gamma_X(k), & \textit{if } \xi = 0 \\ 0, & \textit{if } \xi \in A, \end{cases} \tag{2.10}$$

and

$$\rho_X(sk+\xi) = \begin{cases} \frac{\Gamma(k+D)\Gamma(1-D)}{\Gamma(1+k-D)\Gamma(D)} = \rho_X(k), & \textit{if } \xi = 0 \\ 0, & \textit{if } \xi \in A. \end{cases} \tag{2.11}$$

When $k \to \infty$, $\rho_X(sk) \sim \frac{\Gamma(1-D)}{\Gamma(D)} k^{2D-1}$.

(v). *The process* $\{X_t\}_{t\in\mathbb{Z}}$ *has partial autocorrelation function given by*

$$\phi_X(sk+\xi, sl+\eta) = \begin{cases} -\binom{k}{l} \frac{\Gamma(l-D)\Gamma(k-l+1-D)}{\Gamma(-D)\Gamma(1+k-D)} = \phi_X(k,l), & \text{if } \eta = 0 \\ 0, & \text{if } \eta \in A, \end{cases} \quad (2.12)$$

for any $k, l \in \mathbb{Z}_{\geq}$, *and* $\xi \in A \cup \{0\}$.

From expression (2.12), *when* $k = l$, *the partial autocorrelation function of order* k *is given by*

$$\phi_X(sk, sk) = \frac{D}{k-D} = \phi_X(k,k), \quad \text{for all} \quad k \in \mathbb{Z}_{\geq}. \quad (2.13)$$

Proof. For a proof see Brietzke et al. [10]. □

Remark 2.5. (a) The spectral density function of the SARFIMA$(0, D, 0)_s$ process in the seasonal frequencies is unbounded when $0.0 < D < 0.5$, and it has zeros when D is negative.

(b) Among seasonal frequencies the SARFIMA process has similar behavior to the ARFIMA process.

(c) The SARFIMA$(p, d, q) \times (P, D, Q)_s$ process is stationary when $d + D$ and D are less than 0.5 and the polynomials $\phi(\mathcal{B}) \cdot \Phi(\mathcal{B}^s) = 0$ and $\theta(\mathcal{B}) \cdot \Theta(\mathcal{B}^s) = 0$ have no roots in common and the roots of $\phi(\mathcal{B}) \cdot \Phi(\mathcal{B}^s) = 0$ are outside of the unit circle. When we consider all the above assumptions and also $d + D, D > 0$, then the process has *seasonal long memory*.

(d) If $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary stochastic SARFIMA$(p, d, q) \times (P, D, Q)_s$ process (see expression (2.4)), with $d, D \in (-0.5, 0.5)$, its spectral density function is given by

$$f_X(w) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{|\theta(e^{-iw})|^2}{|\phi(e^{-iw})|^2} \frac{|\Theta(e^{-isw})|^2}{|\Phi(e^{-isw})|^2} \left[2\sin\left(\frac{w}{2}\right)\right]^{-2d} \left[2\sin\left(\frac{sw}{2}\right)\right]^{-2D},$$

for all $0 < w \leq \pi$, where σ_ε^2 is the variance of the white noise $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ process.

The following theorem shows that the stochastic process $\{X_t\}_{t\in\mathbb{Z}}$, given by expression (2.6), with seasonality $s \in \mathbb{N}$ and $D < 0.5$, is ergodic.

Theorem 2.6. *Let* $\{X_t\}_{t\in\mathbb{Z}}$ *be a SARFIMA*$(0, D, 0)_s$ *process given by expression* (2.6), *with zero mean, seasonal period* $s \in \mathbb{N}$ *and* $D < 0.5$. *Then,* $\{X_t\}_{t\in\mathbb{Z}}$ *is an ergodic process.*

Proof. Let $\{X_t\}_{t\in\mathbb{Z}}$ be a SARFIMA$(0, D, 0)_s$ process, given by expression (2.6), with zero mean and seasonal period $s \in \mathbb{N}$. Let $D < 0.5$ be the seasonal fractional differencing parameter. From item (i) in Theorem 2.4, the process $\{X_t\}_{t\in\mathbb{Z}}$ has an infinite moving average representation given by

$$X_t = \Psi(\mathcal{B}^s)\varepsilon_t = \sum_{k\geq 0} \psi_k \mathcal{B}^{sk}(\varepsilon_t) = \sum_{k\geq 0} \psi_k \varepsilon_{t-sk},$$

where the coefficients $\{\psi_k\}_{k\geq 0}$ are given by the expression (2.8) and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a white noise process. From item (ii) in Theorem 2.4, for $D < 0.5$, this is a stationary process. So, one has

$$\sigma_\varepsilon^2 \sum_{k\geq 0} \psi_k^2 = \gamma_X(0) = \mathbb{E}({X_t}^2) < \infty. \tag{2.14}$$

In fact,

$$\begin{aligned} \mathbb{E}({X_t}^2) = \quad & \mathbb{E}\left[(\sum_{k\geq 0} \psi_k \varepsilon_{k-t})(\sum_{j\geq 0} \psi_j \varepsilon_{j-t})\right] = \mathbb{E}\left[\sum_{k\geq 0} \psi_k^2 \varepsilon_{k-t}^2 + \sum_{k,j\geq 0, k\neq j} \psi_k \psi_j \varepsilon_{k-t}\varepsilon_{j-t}\right] \\ = \quad & \sum_{k\geq 0} \psi_k^2 \mathbb{E}(\varepsilon_{k-t}^2) = \sigma_\varepsilon^2 \sum_{k\geq 0} \psi_k^2. \end{aligned}$$

Hence, from expression (2.14), $\sum_{k\geq 0} \psi_k^2 < \infty$. Lemma 3.1 in Olbermann [37] proves the ergodicity for moving average processes of finite and infinite order. This lemma requires the coefficients of an infinite moving average representation to be squared absolutely summable. Therefore, one concludes that the process $\{X_t\}_{t\in\mathbb{Z}}$, given by (2.6), is ergodic. □

For general definition and properties of ergodicity in stochastic processes see Durret [13].

For SARFIMA$(0, D, 0)_s$ processes, the next theorem shows that the conditional expectation and conditional variance depend only on the past values distant from multiples of the seasonality s. This theorem is very important when one needs to generate the mentioned processes.

Theorem 2.7. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be the SARFIMA$(0, D, 0)_s$ process given by the expression (2.6), with zero mean, $s \in \mathbb{N}$ as the seasonal period and $D \in (-0.5, 0.5)$. The conditional expectation and the conditional variance of X_t, given X_l, for all $l < t$, denoted respectively by $m_t \equiv \mathbb{E}(X_t|X_l, l < t)$ and $v_t \equiv Var(X_t|X_l, l < t)$, are given by*

$$\begin{cases} m_\zeta = 0, \quad \text{for} \quad \zeta = 1, \cdots, s-1, \\ \\ m_{sk} = \sum_{j=1}^{k} \phi_X(sk, sj)\, X_{sk-sj}, \quad \text{for} \quad k \in \mathbb{N}, \\ \\ m_{sk+\zeta} = \sum_{j=1}^{k} \phi_X(sk+\zeta, sj)\, X_{sk+\zeta-sj}, \end{cases} \tag{2.15}$$

and

$$\begin{cases} v_\zeta = \sigma_\varepsilon^2, \quad for \quad \zeta = 1, \cdots, s-1, \\ \\ v_{sk} = \sigma_\varepsilon^2 \prod_{j=1}^{k} (1 - \phi_X^2(sj, sj)), \quad for \quad k \in \mathbb{N}, \\ \\ v_{sk+\zeta} = v_{sk}, \end{cases} \tag{2.16}$$

where $t = \zeta$ determines the mean and the variance for lags smaller than s, $t = sk$ for multiple lags of s, and $t = sk+\zeta$ for not multiple lags of s, $\phi_X(\cdot,\cdot)$ is the partial autocorrelation function of the process $\{X_t\}_{t\in\mathbb{Z}}$ given by item (v) *in Theorem* 2.4*, and σ_ε^2 is the variance of the white noise process.*

Proof. For a proof see Bisognin and Lopes [5]. □

2.3. FIGARCH(p, d, q) Processes

In this section we shall consider one natural extension of the ARFIMA framework towards volatility models. Models for time series with *long-range dependence in volatility* are of great interest in econometrics and finance. Lopes and Mendes [26] review some extensions of the GARCH class of processes and study the performance of 300 regression type estimators for several long memory models, including FIGARCH processes. The authors show that the performance of the semiparametric estimators are affected by their asymptotic statistical properties besides by their strong dependency on the number of frequencies used for the regression.

Denote by $\mathcal{F}_t$ the σ-field of events generated by $\{X_l; l \leq t\}$ and assume that $\mathbb{E}(X_t|\mathcal{F}_{t-1}) = 0$ *a.s.*. Following Engle [14] and Bollerslev [8] we specify a GARCH(p, q) model by

$$X_t = \sigma_t Z_t, \tag{2.17}$$

where Z_t is an independent identically distributed random variable with zero mean and unit variance such that $X_t|\mathcal{F}_{t-1}$ is an independent random variable with zero mean and variance $\sigma_t^2 \equiv Var(X_t|\mathcal{F}_{t-1})$ defined by

$$\sigma_t^2 = \omega + \alpha(\mathcal{B})X_t^2 + \beta(\mathcal{B})\sigma_t^2, \tag{2.18}$$

where $\omega > 0$ is a real constant, $\alpha(\mathcal{B}) = \sum_{j=1}^{p} \alpha_j \mathcal{B}^j$ and $\beta(\mathcal{B}) = \sum_{k=1}^{q} \beta_k \mathcal{B}^k$. For a FIGARCH process (see Baillie et al. [2] and Bollerslev and Mikkelsen [7]) the σ_t, in expression (2.17), is defined as

$$\begin{aligned} \sigma_t^2 &= \omega\,(1-\beta(\mathcal{B}))^{-1} + \{1 - (1-\beta(\mathcal{B}))^{-1}[1-\alpha(\mathcal{B})-\beta(\mathcal{B})](1-\mathcal{B})^d\}X_t^2 \\ &= \omega\,(1-\beta(\mathcal{B}))^{-1} + \{1 - (1-\beta(\mathcal{B}))^{-1}\phi(\mathcal{B})(1-\mathcal{B})^d\}X_t^2 \\ &= \omega\,(1-\beta(\mathcal{B}))^{-1} + \lambda(\mathcal{B})X_t^2, \end{aligned} \tag{2.19}$$

where

$$\lambda(\mathcal{B}) = \sum_{k=0}^{\infty} \lambda_k \mathcal{B}^k = 1 - (1 - \beta(\mathcal{B}))^{-1} \phi(\mathcal{B})(1 - \mathcal{B})^d, \tag{2.20}$$

$\phi(\mathcal{B}) = 1 - \alpha(\mathcal{B}) - \beta(\mathcal{B}))$ and the binomial series expansion in $\mathcal{B}$, denoted by

$$(1 - \mathcal{B})^d \equiv 1 - \delta_d(\mathcal{B}) = 1 - \sum_{k=1}^{\infty} \delta_{d,k} \, \mathcal{B}^k, \tag{2.21}$$

is given by (2.2), with $d \in [0, 1]$.

The coefficients $\delta_{d,k} = d\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(1-d)}$, in expression (2.21), are such that

$$\delta_{d,k} = \delta_{d,k-1} \left(\frac{k - 1 - d}{k} \right),$$

for all $k \geq 1$, where $\delta_{d,0} \equiv 1$.

The following proposition totally characterizes any FIGARCH(p, d, q) process and also gives a recurrent formula for the coefficients $\{\lambda_k\}_{k \geq 0}$ given in expression (2.20).

Proposition 2.8. *Let $\{X_t\}_{t \in \mathbb{Z}}$ be any FIGARCH(p, d, q) process, for $d \in [0, 1]$, defined by expressions (2.17) and (2.19). Then, the coefficients $\{\lambda_k\}_{k \geq 0}$, in expression (2.20), are given by*

$$\begin{aligned}
\lambda_0 &= 0 \\
\lambda_k &= \sum_{j=1}^{p} \beta_j \lambda_{k-j} + \alpha_k + \delta_{d,k} - \sum_{m=1}^{\max\{p,q\}} \gamma_m \delta_{d,k-m}, \quad \textit{if} \quad 1 \leq k \leq p \\
\lambda_k &= \sum_{l=1}^{q} \beta_l \lambda_{k-l} + \delta_{d,k} - \sum_{m=1}^{\max\{p,q\}} \gamma_m \delta_{d,k-m}, \quad \textit{if} \quad k > p,
\end{aligned}$$

where

$$\gamma_m = \begin{cases} \alpha_m, & \textit{if} \quad p > q, \\ \alpha_m + \beta_m, & \textit{if} \quad p = q, \\ \beta_m, & \textit{if} \quad p < q, \end{cases}$$

with α_j, $1 \leq j \leq p$, and β_l, $1 \leq l \leq q$, are given in expression (2.18) and $\delta_{d,k}$, for $k \geq 0$, given in expression (2.21).

Proof. For a proof see Lopes and Mendes [26]. □

2.4. FISV(p, d, q) Processes

In this section we shall consider another natural extension of the ARFIMA framework towards volatility models. We consider the Fractionally Integrated Stochastic Volatility (FISV) model, introduced by Breidt et al. [9]. Lopes and Mendes [26] also consider this model when analyzing the *long-range dependence in volatility*.

Let $\{Y_t\}_{t\in\mathbb{Z}}$ be the stochastic process such that

$$Y_t = \sigma_\epsilon \, g(X_t)\epsilon_t, \tag{2.22}$$

where X_t is a long memory in mean process, $g(\cdot)$ is a continuous function and $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is a white noise process with zero mean and unit variance. Since $Var(Y_t|X_t) = \sigma_\epsilon^2 \, g(X_t)^2$, for certain functions $g(\cdot)$ the process defined by (2.22) may be described as a long memory stochastic volatility process (see Robinson and Zaffaroni [43]). This large class of volatility models include the long memory nonlinear moving average models of Robinson and Zaffaroni [43] and the FISV process introduced by Breidt et al. [9].

In a FISV(p, d, q) process $\{Y_t\}_{t\in\mathbb{Z}}$, the function $g(\cdot)$ in (2.22) is given by

$$g(X_t) = \exp\left(\frac{X_t}{2}\right), \tag{2.23}$$

where $\{X_t\}_{t\in\mathbb{Z}}$ is an ARFIMA(p, d, q) process given by (2.1), and ϵ_t and ε_t are independent and identically distributed standard normal, and mutually independent. One observes that $Var(Y_t|X_t) = \sigma_\epsilon^2 \exp(X_t)$. In particular, squaring both sides of equation (2.22), with the function $g(\cdot)$ given by expression (2.23), and taking logarithms,

$$\ln(Y_t^2) = \mu_\xi + X_t + \xi_t, \tag{2.24}$$

where $\mu_\xi = \ln(\sigma_\epsilon^2) + \mathbb{E}[\ln(\epsilon_t^2)]$, and $\xi_t = \ln(\epsilon_t^2) - \mathbb{E}[\ln(\epsilon_t^2)]$. Hence, $\ln(Y_t^2)$ is the sum of a Gaussian ARFIMA process and independent non-Gaussian noise with zero mean. Consequently, the autocovariance function of the process $\ln(Y_t^2)$, when $d \in (-0.5, 0.5)$, is such that

$$\gamma_{\ln(Y_t^2)}(k) \sim k^{2d-1}, \quad \text{when } k \to \infty, \tag{2.25}$$

while its spectral density function has the property that

$$f_{\ln(Y_t^2)}(w) \sim w^{-2d}, \quad \text{when } w \to 0. \tag{2.26}$$

For $d \in (0.0, 0.5)$, the spectral density function in expression (2.26) is unbounded, when $w \to 0$. This point is of great importance for the application of the traditional regression estimation procedures, based on the periodogram function, given in Section 4. Lopes and Mendes [26] also present the performance of the estimators of all parameters in FISV models when the white noise process $\{\epsilon_t\}_{t\in\mathbb{Z}}$ has standard normal distribution or t-Student distribution with 4 degrees of freedom.

3. Durbin-Levinson Algorithm

The partial autocorrelation function of a process $\{X_t\}_{t\in\mathbb{Z}}$ with zero mean and autocovariance function $\gamma_X(\cdot)$, such that $\gamma_X(k) \to 0$, as $k \to 0$, is defined below. For more details, see Brockwell and Davis [11]. This definition also holds for any ARFIMA or SARFIMA processes and item (v) of Theorem 2.4, in Section 2, presents the partial autocorrelation function of SARFIMA$(0, D, 0)$ processes.

Definition 3.1. Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stochastic process with zero mean and autocovariance function $\gamma_X(\cdot)$ such that $\gamma_X(k) \to 0$, as $k \to 0$. The *partial autocorrelation function*, denoted by $\phi_X(k, j)$, $k \in \mathbb{Z}_{\geqslant}$ and $j = 1, \cdots, k$, are the coefficients in the equation

$$\mathcal{P}_{\overline{sp}(X_1,X_2,\cdots,X_k)}(X_{k+1}) = \sum_{j=1}^{k} \phi_X(k, j) X_{k+1-j},$$

where $\mathcal{P}_{\overline{sp}(X_1,X_2,\cdots,X_k)}(X_{k+1})$ is the orthogonal projection of X_{k+1} in the closed span $\overline{sp}(X_1,\, X_2, \cdots, X_k)$ generated by the previous observations. Then, from the equations

$$\langle X_{k+1} - \mathcal{P}_{\overline{sp}(X_1,X_2,\cdots,X_k)}(X_{k+1}), X_j\rangle = 0, \;\; j = 1, \cdots, k,$$

where $\langle\cdot,\cdot\rangle$ defines the internal product on the Hilbert space $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ given by $\langle X, Y\rangle = \mathbb{E}(XY)$, we obtain

$$\begin{bmatrix} 1 & \rho_X(1) & \rho_X(2) & \cdots & \rho_X(k-1) \\ \rho_X(1) & 1 & \rho_X(1) & \cdots & \rho_X(k-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_X(k-1) & \rho_X(k-2) & \rho_X(k-3) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_X(k,1) \\ \phi_X(k,2) \\ \vdots \\ \phi_X(k,k) \end{bmatrix} = \begin{bmatrix} \rho_X(1) \\ \rho_X(2) \\ \vdots \\ \rho_X(k) \end{bmatrix}, \tag{3.1}$$

with $\rho_X(\cdot)$ the autocorrelation function of the process $\{X_t\}_{t\in\mathbb{Z}}$. The coefficients $\phi_X(k, j)$, $k \in \mathbb{Z}_{\geqslant}$, $j = 1, \cdots, k$, are uniquely determined by (3.1).

The definition of partial autocorrelation function plays an important role in the Durbin-Levinson algorithm (see expressions (3.2) and (3.3) below) and its expression for seasonal fractionally integrated processes is given in item (v) of Theorem 2.4.

Brietzke et al. [10] give a closed formula for the Durbin-Levinson Algorithm for the partial autocorrelation function, defined by the expression (2.12), for seasonal fractionally integrated processes. This is a crucial algorithm and a summary of its description is given as follows.

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a SARFIMA$(0, D, 0)_s$ process, given in expression (2.6), with mean μ equal to zero. We want to show that its partial autocorrelation function $\phi_X(\cdot,\cdot)$, given in item (v) of Theorem 2.4, satisfies the following systems

$$\phi_X(sl, sl) = \frac{\rho_X(sl) - \sum_{j=1}^{sl-1} \phi_X(sl-1, j)\rho_X(sl-j)}{1 - \sum_{j=1}^{sl-1} \phi_X(sl-1, j)\rho_X(j)} \tag{3.2}$$

and

$$\phi_X(k+1, sl) = \phi_X(k, sl) - \phi_X(k+1, k+1)\phi_X(k, k+1-sl), \tag{3.3}$$

for any $l \in \mathbb{Z}_{\geqslant}$ such that $sl < k+1$, where $k+1$ may or may not be a multiple of s, where $\rho_X(\cdot)$ is given in item (iv) of Theorem 2.4.

Recurrence relations (3.2) and (3.3) are known as the Durbin-Levinson algorithm and they explain how to go from lag k to lag $(k+1)$ for the partial autocorrelation function $\phi_X(\cdot, \cdot)$. Brietzke et al. [10] prove the recurrence relation (3.2)-(3.3) for any $D \in (-0.5, 0.5)$, with $D \neq 0$.

Lemma 3.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a process given by* (2.6)*. For any $k, l \in \mathbb{Z}_{\geqslant}$, the partial autocorrelation function of $\{X_t\}_{t\in\mathbb{Z}}$, denoted by $\phi_X(\cdot, \cdot)$, satisfies the system given in* (3.3)*, whenever $l < k+1$.*

Proof. For a proof see Brietzke et al. [10]. □

Lemma 3.3. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a process given by* (2.4)*, where $D \in (-0.5, 0.5)$ with $D \neq 0$. Then, the quocient in expression* (3.2) *is given by*

$$\frac{\rho_X(l) - \sum_{j=1}^{l-1} \phi_X(l-1, j)\rho_X(l-j)}{1 - \sum_{j=1}^{l-1} \phi_X(l-1, j)\rho_X(j)} = \frac{\displaystyle\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-D)\Gamma(l-j-D)\Gamma(l-j+D)}{\Gamma(l-j-D+1)}}{\displaystyle\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-D)\Gamma(l-j-D)\Gamma(j+D)}{\Gamma(j-D+1)}}. \tag{3.4}$$

Proof. For a proof see Brietzke et al. [10]. □

We still need to show that (3.4) is equal to $\phi_X(l, l)$. This follows from Theorem 3.7 below. One can show that the numerator of the left-hand side of expression (3.4) times $(l-D)$ (or its denominator times D) is equal to $\phi_X(l, l)$, that is,

$$(l-D)\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-D)\Gamma(l-j-D)\Gamma(l-j+D)}{\Gamma(l-j-D+1)}$$
$$= D\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-D)\Gamma(l-j-D)\Gamma(j+D)}{\Gamma(j-D+1)}. \tag{3.5}$$

Moreover, one can also show that

$$
\begin{aligned}
&(l-D)\sum_{j=0}^{l-1}\binom{l-1}{j}\frac{\Gamma(j-D)\Gamma(l-j-D)\Gamma(l-j+D)}{\Gamma(l-j-D+1)} \\
&= D\Gamma(-D)\Gamma(D-l+1)(l-1)!\,2^{l-1}\cdot\prod_{i=0}^{l-2}\Big(D-\frac{i+1}{2}\Big). \qquad (3.6)
\end{aligned}
$$

The equalities (3.5) and (3.6) follow, respectively, from Corollaries 3.10 and 3.8, below. The system (3.2) follows immediately from expressions (3.5) and (2.13). The Durbin-Levinson algorithm is a consequence of Theorem 2.4 and equality (3.6) above. We shall first define the *hypergeometric function*.

Definition 3.4. If a_i, b_i and x are complex numbers, with $b_i \notin \mathbb{Z}_{\leqslant}$, we define the *hypergeometric function* by

$$
{}_3F_2\big(a_1,a_2,a_3;b_1,b_2;x\big)=\sum_{n=0}^{\infty}\frac{\big(a_1\big)_n\big(a_2\big)_n\big(a_3\big)_n}{\big(b_1\big)_n\big(b_2\big)_n}\,\frac{x^n}{n!}\,,
$$

where $(a)_n$ stands for the *Pochhammer symbol*

$$
(a)_n=\frac{\Gamma(a+n)}{\Gamma(a)}=\begin{cases} a(a+1)\cdots(a+n-1), & \text{if } n\ge 1\\ 1, & \text{if } n=0.\end{cases}
$$

This series is absolutely convergent for all $x \in \mathbb{C}$ such that $|x| < 1$, and also for $|x| = 1$, provided $\Re\big(b_1+b_2\big) > \Re\big(a_1+a_2+a_3\big)$, where $\Re(z)$ means the real part of $z \in \mathbb{C}$. Furthermore, it is said to be *balanced* if $b_1+b_2 = 1+a_1+a_2+a_3$. Note that in case some a_i is a nonpositive integer the above sum is finite and it suffices to let n range from 0 to $-a_i$.

The following identity for a *terminating balanced hypergeometric sum* is very useful. For the identity's proof we refer the reader to Andrews et al. [1], Thm. 2.2.6, page 69.

Theorem 3.5 (Identity of Pfaff–Saalschütz). *Let $k \in \mathbb{Z}_{\geqslant}$, and a, b, and c be complex numbers such that $c, 1+a+b-c-k \notin \mathbb{Z}_{\leqslant}$. Then,*

$$
{}_3F_2\big(-k,a,b;c,1+a+b-c-k;1\big)=\frac{\big(c-a\big)_k\big(c-b\big)_k}{\big(c\big)_k\big(c-a-b\big)_k}\,. \qquad (3.7)
$$

Remark 3.6. If $\big(c_n\big)_{n\ge 0}$ is a sequence of complex numbers satisfying

$$
\frac{c_{n+1}}{c_n}=\frac{(a_1+n)(a_2+n)(a_3+n)x}{(n+1)(b_1+n)(b_2+n)} \text{ for all } n,
$$

straightforward computations show that

$$\sum_{n=0}^{\infty} c_n = c_0 \cdot {}_3F_2\big(a_1, a_2, a_3; b_1, b_2; x\big)\,. \tag{3.8}$$

The identity (3.7) and the above remark are fundamental for the proof of the following theorem.

Theorem 3.7. *Let x and z be complex numbers, with $x \notin \mathbb{Z}$ and $z \notin \mathbb{Z}_{\geqslant}$. Then,*

$$\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-x)\,\Gamma(l-j+x)}{z-j} = \frac{\Gamma(-x)\,\Gamma(1+x)\,\Gamma(1-z)}{z\,\Gamma(l-z)} \cdot (l-1)! \prod_{i=1}^{l-1} \big(x-z+i\big)\,. \tag{3.9}$$

For $z \in \{l, l+1, \ldots\}$ the right-hand side of expression (3.9) has a removable singularity and by analytic continuation the result is still true.

Proof. For a proof see Brietzke et al. [10]. □

Corollary 3.8. *If $l \in \mathbb{N} - \{1\}$ and D is a noninteger complex number, then*

$$\begin{aligned}(l-D) &\sum_{j=0}^{l-1} \binom{l-1}{j} \frac{\Gamma(j-D)\,\Gamma(l+D-j)}{l-D-j} \\ &= D\,\Gamma(-D)\,\Gamma(D-l+1)\,(l-1)!\,2^{l-1} \cdot \prod_{i=0}^{l-2} \left(D - \frac{i+1}{2}\right).\end{aligned} \tag{3.10}$$

Corollary 3.9. *If $l \in \mathbb{N} - \{1\}$ and D is a noninteger complex number, then*

$$\begin{aligned}D &\sum_{k=0}^{l-1} \binom{l-1}{k} \frac{\Gamma(l-1-k+D)\,\Gamma(k-D+1)}{l-1-k-D} \\ &= D\,\Gamma(-D)\,\Gamma(D-l+1)\,(l-1)!\,2^{l-1} \cdot \prod_{i=0}^{l-2} \left(D - \frac{i+1}{2}\right).\end{aligned} \tag{3.11}$$

Corollary 3.10. *If $l \in \mathbb{N}$ and D is a noninteger complex number, then*

$$(l-D)\sum_{j=0}^{l-1}\binom{l-1}{j}\frac{\Gamma(j-D)\,\Gamma(l+D-j)}{l-D-j}$$
$$= D\sum_{k=0}^{l-1}\binom{l-1}{k}\frac{\Gamma(l-1-k+D)\,\Gamma(k-D+1)}{l-1-k-D}. \qquad (3.12)$$

4. Classical and Robust Estimation Procedures

In the literature of the stochastic long memory processes, there exist several estimation procedures for the fractional parameter d. We now summarize some of these estimation procedures both for long memory models in mean and in volatility: here we present one parametric and four semi-parametric methods. For these methods we also consider their robust versions. For a non-parametric method based on wavelet theory applied to the fractional parameter estimation, in ARFIMA processes, we refer the reader to Lopes and Pinheiro [25]. We also refer the reader to Olbermann et al. [35] for another work where a non-parametric method based on wavelet theory is used to estimate the hyperbolic rate decay parameter for the autocorrelation function of Manneville-Pomeau processes.

The methodology in this section will be presented based on ARFIMA(p, d, q) processes which are the simplest among all the others analyzed in Section 2.

We recall that when $\{Y_t\}_{t\in\mathbb{Z}}$ follows a FISV process with $d \in (-0.5, 0.5)$, $\ln(Y_t^2)$ is the sum of a zero mean Gaussian ARFIMA process and an independent non-Gaussian innovation process. Also, the FIGARCH(p, d, q) process, with $d \in [0, 1]$, has been defined in expression (8) of Baillie et al. [2] as an ARFIMA process on the squared data with a more complicated error structure. Thus, the regression based methods described below also apply to the other processes considered in Section 2.3 and 2.4.

In this section we summarize six methods for the estimation of the fractional differencing parameter:

• The semi-parametric regression method based on the periodogram function proposed by Geweke and Porter-Hudak [16]. This estimator is denoted hereafter by GPH;

• The semi-parametric regression method based on the smoothed periodogram when one considers the Bartlett lag window. This estimator is denoted hereafter by BA;

• The semi-parametric regression method based on GPH with trimming l and bandwidth $g(n)$ proposed by Robinson [44]. This estimator is denoted here by R;

• The semi-parametric method based on the partial sum process proposed by Mandelbrot and Taqqu [33], based on Hurst [21] estimator. This estimator is largely known as the R/S statistics;

• The cosine-bell tapered data method, denoted in the sequel by $GPHT$, considers the cosine-bell function as a transformation of the data and follows similarly to the GPH method. It was proposed by Hurvich and Ray [22];

• The parametric approximated maximum likelihood method, proposed by Fox and Taqqu [15], based on the approximation given by Whittle [50], is denoted hereafter by W.

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a ARFIMA(p,d,q) process with $d \in (-0.5, 0.5)$, given by (2.1). Its spectral density function is given by

$$f_X(w) = f_U(w)\left[2\,\sin\left(\frac{w}{2}\right)\right]^{-2d}, \text{ for } 0 < w \leq \pi, \tag{4.1}$$

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

Consider the set of harmonic frequencies $w_j = \frac{2\pi j}{n}$, $j = 0, 1, \cdots, \lceil n/2 \rceil$, where n is the sample size and $\lceil x \rceil$ means the integer part of x. By taking the logarithm of the spectral density function $f_X(\cdot)$ given by (4.1), and adding $\ln(f_U(0))$, and $\ln(I(w_j))$ to both sides of this expression we obtain

$$\ln(I(w_j)) = \ln(f_U(0)) - d\ln\left[2\,\sin\left(\frac{w_j}{2}\right)\right]^2 + \ln\left\{\frac{f_U(w_j)}{f_U(0)}\right\} + \ln\left\{\frac{I(w_j)}{f_X(w_j)}\right\}, \tag{4.2}$$

where $I(\cdot)$ is the periodogram function given by

$$I(w) = \frac{1}{2\pi}\left(\hat{\gamma}_X(0) + 2\sum_{l=1}^{n-1}\hat{\gamma}_X(l)\cos(l\,w)\right), \tag{4.3}$$

with $\hat{\gamma}_X(h) = \frac{1}{n}\sum_{k=1}^{n-h}(x_k - \bar{x})(x_{k+h} - \bar{x})$, for $h \in \{0, 1, \cdots, n-1\}$, is the sample autocovariance function and $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$ is the sample mean of the process $\{X_t\}_{t\in\mathbb{Z}}$ in (2.1).

When considering only the frequencies close to zero, the term $\ln\{\frac{f_U(w_j)}{f_U(0)}\}$ may be discarded. Then, we may rewrite (4.2) in the context of a simple linear regression model

$$y_j = a - d\,x_j + e_j,\ j = 1, \cdots, m, \tag{4.4}$$

where $m = g(n) = n^{\alpha}$, for $0 < \alpha < 1$, $(a, -d)$ are the regression coefficients, $a = \ln(f_U(0))$, $y_j = \ln(I(w_j))$, $x_j = \ln\{2\sin(w_j/2)\}^2$ and the errors $e_j = \ln\{\frac{I(w_j)}{f_X(w_j)}\}$ are uncorrelated random variables centered at zero with constant variance.

A semi-parametric regression estimator may be obtained by minimizing some loss function of the residuals $r_j = y_j - \hat{a} + \hat{d}\,x_j$. We will consider three different loss functions. They give rise to the classical *Ordinary Least Squares* method (OLS), and two high breakdown point robust methods, the *Least Trimmed Squares* method (LTS), and the MM*-estimation* method.

The OLS estimators are the values $(\hat{a}, -\hat{d})$ which minimize the loss function

$$L_1(m) = \sum_{j=1}^{m} (r_j)^2, \tag{4.5}$$

where $r_j = y_j - \hat{a} + \hat{d}\,x_j$ is the residual related to the regression (4.4).

Whenever the errors e_i follow a normal distribution, the OLS estimates have the minimum variance among all unbiased estimates. In fact, it is well known (see Huber [20]) that regression outliers, leverage points, and gross errors are responsible for considerable bias and inefficiency (even in the Gaussian environment) in the OLS estimates.

Robust alternatives to OLS may be obtained by minimizing a robust version of the dispersion of the residuals. The *Least Trimmed Squares* (LTS) estimates of Rousseeuw [45] minimize the loss function

$$L_2(m) = \sum_{j=1}^{m^*} \left(r^2\right)_{j:m} , \tag{4.6}$$

where $\left(r^2\right)_{j:m}$ are the squared and then ordered residuals, that is, $\left(r^2\right)_{1:m} \leq \cdots \leq \left(r^2\right)_{m^*:m}$, and m^* is the number of points used in the optimization procedure. The constant m^* is responsible both for the breakdown point value and the efficiency. When m^* is approximately $m/2$ the breakdown point is approximately 50%. The LTS estimates have been previously used by Taqqu et al. [47] for the estimation of the long-range parameter in ARFIMA models and by Lopes and Mendes [26] for the estimation of long-range parameter both in mean and in volatility models.

The MM*-estimates* (see Yohai [51]) may present simultaneously high breakdown point and high efficiency. They are defined as the solution $(\hat{a}, -\hat{d})$ which minimizes the loss function

$$L_3(m) = \sum_{j=1}^{m} \rho_2 \left(\frac{r_j}{\kappa}\right)^2 , \tag{4.7}$$

subject to the constraint

$$\frac{1}{m} \sum_{j=1}^{m} \rho_1\left(\frac{r_j}{\kappa}\right) \leq b , \tag{4.8}$$

where ρ_2 and ρ_1 are symmetric, bounded, nondecreasing functions on $[0, \infty)$ with $\rho_j(0) = 0$ and $\lim_{u\to\infty} \rho_j(u) = 1$, for $j = 1, 2$, κ is a scale parameter, and b is a tuning constant. The breakdown point of the MM-estimator only depends on ρ_1 and it is given by $\min(b, 1 - b)$.

4.1. Classical and Robust GPH Estimators

The first estimation method based on the periodogram function was introduced in the pioneer work of Geweke and Porter-Hudak [16].

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a ARFIMA(p,d,q) process with $d \in (-0.5, 0.5)$, given by the expression (2.1). From the linear regression given by (4.4) the classical *GPH-LS* estimator of d is then given by

$$GPH - LS = -\frac{\sum_{j=1}^{g(n)}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{g(n)}(x_j - \bar{x})^2}, \tag{4.9}$$

where the trimming value $g(n)$ is usually $g(n) = n^{\alpha}$, for $0 < \alpha < 1$, y_j is based on (4.3) and x_j is as previously defined in expression (4.4). Lopes et al. [28] considered α in the interval $[0.55, 0.65]$, and Porter-Hudak [40] considered $\alpha \in \{0.62, 0.75\}$ for the case of seasonal fractionally integrated time series data. Lopes and Mendes [26] consider $\alpha \in \{0.50, 0.52, \cdots, 0.84, 0.86\}$. The version not tunned by α, that is, based on the $\lceil \frac{n}{2} \rceil$ data points, equivalent to set $\alpha = 0.8997$, was also considered in Lopes and Mendes [26].

To obtain the robust versions of the GPH estimator we just apply the LTS and the MM methodologies to the regression model (4.4) with $m = n^{\alpha}$, based on (4.3). This gives rise to the *GPH-LTS* and the *GPH-MM* estimators.

4.2. Classical and Robust BA Estimators

The periodogram function is not a consistent estimator for the spectral density function (see Brockwell and Davis [11]). Lopes and Lopes [29] analyzes the convergence in distribution sense for the periodogram function based on a time series of a stationary process. This process is obtained from the iterations of a continuous transformation invariant for an ergodic probability. In this later work, the authors only assume a certain rate of convergence to zero for the autocovariance function of the stochastic process, that is, it is assumed that there exist $C > 0$ and $\xi > 2$ such that $|\gamma_X(k)| \leq C|k|^{-\xi}$, for all $k \in \mathbb{Z}$, where $\gamma_X(\cdot)$ is the autocovariance function of the process. This result can be applied to a time series obtained from the iteration of a certain class of deterministic transformations (or its natural extension; see, for instance, Lopes and Lopes [30]) whose initial point is distributed according to an ergodic probability.

Returning to the general setting, by considering the Bartlett lag window, a consistent estimator for the spectral density function may be obtained. This smoothed version of the periodogram function is defined by

$$I_{\text{smooth}}(w) = \frac{1}{2\pi} \sum_{j=-\nu}^{\nu} \kappa\left(\frac{j}{\nu}\right) \hat{\gamma}_X(j) \cos(j\, w), \tag{4.10}$$

where $\kappa(\cdot)$ is the Bartlett lag window given by

$$\kappa(x) = \begin{cases} 1 - |x|, & \text{if } |x| \le 1 \\ 0, & \text{otherwise,} \end{cases} \tag{4.11}$$

with ν being the truncation point of the weighted function.

The classical and robust versions are obtained by applying the OLS, the LTS and the MM methodologies to the regression model (4.4) based on (4.10) and (4.11), producing the BA-LS, the BA-LTS, and the BA-MM estimators.

4.3. Classical and Robust R Estimators

The regression estimator R, proposed by Robinson [44] is obtained by applying the Ordinary Least Squares method in (4.4) based on (4.3), but considering only the frequencies $j \in \{l, l+1, \cdots, g(n)\}$, where $l > 1$ is a trimming value that tends to infinity more slowly than $g(n)$.

It is interesting to compare the R and the LTS concepts. The R concept trims the extreme x_j values associated with the frequencies close to zero, which we know are the important ones. On the other hand, the LTS concept trims the extreme ordered residuals which may or may be not associated to small frequencies, but certainly are associated to leverage points. In other words, the LTS procedure identifies which data points associated with small frequencies are outliers and, if they exist, excludes them from the calculations. The R-LTS and R-MM versions are obtained by applying the robust methodologies, as previously.

4.4. Classical and Robust R/S Estimators

Various methods for estimating the self-similarity parameter H or the intensity of long-range dependence in a time series are available, some of which are described in detail in Beran [4]. In a pioneer work by Mandelbrot and van Ness [31] the authors describe the self-similarity parameter H through the fractional Brownian motion processes. The R/S statistics, the so-called *rescaled adjusted range*, was firstly considered by Mandelbrot and Taqqu [33], based on Hurst [21] estimator. The self-similarity parameter H is related to the fractional parameter d by the equation $d = H + \frac{1}{2}$. Lo [24] proposes the use of the R/S statistics with a different normalization that makes the estimator more robust to some form of short-range dependence. It is based on the range of the partial sum process $S_k = \sum_{j=1}^{k}(X_j - \bar{X}_n)$ and it is defined by

$$R/S(q) = \frac{\max_{1\le k\le n} S_k - \min_{1\le k\le n} S_k}{\hat{\sigma}(q)}, \tag{4.12}$$

where $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$, $\hat{\sigma}^2(q) = \hat{\gamma}_S(0) + 2\sum_{j=1}^{q}\omega_j(q)\hat{\gamma}_S(j)$, the sample autocovariances $\gamma_S(h) = \frac{1}{n}\sum_{l=1}^{n-h}(S_l - \bar{S}_n)(S_{l+h} - \bar{S}_n)$, for $0 \le h < n$, account for the possible short-range dependence up to the qth order and the weights $\omega_j(q) = 1 - \frac{1}{q+1}$ correspond to the Bartlett window.

4.5. Classical and Robust $GPHT$ Estimators

The $GPHT$ method (see Hurvich and Ray [22] and Velasco [49]) uses a modified periodogram function given by

$$I(w_j) = \frac{1}{\sum_{t=0}^{n-1} g(t)^2} \left| \sum_{t=0}^{n-1} g(t) X_t e^{-i\, w_j\, t} \right|^2, \tag{4.13}$$

where the tapered data is obtained from the cosine-bell function $g(\cdot)$ defined by

$$g(t) = \frac{1}{2}\left[1 - \cos\left(\frac{2\pi(t+0.5)}{n}\right)\right]. \tag{4.14}$$

We obtain the classical $GPHT$-LS and the robust versions $GPHT$-LTS and $GPHT$-MM by applying the classical and the robust methodologies on model (4.4) based on (4.13) and (4.14), and setting $m = n^{\alpha}$.

4.6. Classical W Estimator

The W estimator was proposed by Whittle [50]. He considered the function

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

where $\boldsymbol{\eta}$ denotes the vector of unknown parameters, and $f_X(\cdot\, ; \boldsymbol{\eta})$ is the spectral density function of $\{X_t\}_{t\in\mathbb{Z}}$, given by (4.1) and $I(\cdot)$ is the periodogram function given by (4.3).

The W estimator is the value of $\boldsymbol{\eta}$ which minimizes the function $Q(\cdot)$. Here $\boldsymbol{\eta}$ is the vector $(\phi_1, \cdots, \phi_p, d, \sigma_\varepsilon, \theta_1, \cdots, \theta_q)$. The estimation procedure is carried out by finding the value $\hat{\boldsymbol{\eta}}$ which minimizes

$$\mathcal{B}_n(\boldsymbol{\eta}) = \sum_{j=1}^{\lceil \frac{n-1}{2} \rceil} \frac{I(w_j)}{f_X(w_j; \boldsymbol{\eta})}. \tag{4.15}$$

More details of this estimator can be found in Fox and Taqqu [15]. Differently from the previous four estimators, the W estimator is in the parametric class.

We point out that the estimation procedures considered for the ARFIMA processes in this section can be easily extended to the stochastic volatility models given in Sections 2.3 and 2.4.

5. Forecasting in Long Memory Processes

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a SARFIMA$(0, D, 0)_s$ process with $D \in (-0.5, 0.5)$, given by the expression (2.6). Suppose one wants to forecast the value X_{t+h} for h-step-ahead. The *minimum mean squared error forecasting value* is given by

$$\widehat{X}_t(h) \equiv \mathbb{E}\,(X_{t+h}\,|\mathcal{F}_t), \tag{5.1}$$

where $\mathcal{F}_t$ is the σ-field of events generated by $\{X_\ell; \ell \leqslant t\}$. This minimizes the *mean squared error of forecasting* $\mathbb{E}(X_{t+h} - \widehat{X}_t(h))$. In this case, the *forecasting error* is given by

$$e_t(h) = X_{t+h} - \widehat{X}_t(h). \tag{5.2}$$

To calculate the forecasting values one uses the following facts:

(a). $\mathbb{E}(X_{t+h}|\mathcal{F}_t) = \begin{cases} X_{t+h}, & \text{if} \quad h \leqslant 0, \\ \widehat{X}_t(h), & \text{if} \quad h > 0; \end{cases}$

(b). $\mathbb{E}(\varepsilon_{t+h}|\mathcal{F}_t) = \begin{cases} \varepsilon_{t+h}, & \text{if} \quad h \leqslant 0, \\ 0, & \text{if} \quad h > 0. \end{cases}$

Therefore, to calculate the forecasting values, one

(a). substitutes the past expectations ($h \leqslant 0$) for known values, X_{t+h} and ε_{t+h};
(b). substitutes the future expectations ($h > 0$) for forecasting values $\widehat{X}_t(h)$ and 0.

The following theorem presents some results for forecasting a future value of a SARFIMA$(0, D, 0)_s$ process, given by the expression (2.6).

Theorem 5.1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a SARFIMA$(0, D, 0)_s$ process, with zero mean and seasonality $s \in \mathbb{N}$, given in expression (2.6). Consider $D > -0.5$. Then, for all $h \in \mathbb{N}$:*

(i). *The minimum mean squared error forecasting value is given by*

$$\widehat{X}_n(h) = -\sum_{k\geq 0} \pi_k \, \widehat{X}_n(h - sk), \tag{5.3}$$

where π_k is given in expression (2.7).

(ii). *The forecasting error is given by $e_n(h) = \sum_{k=0}^{\lceil \frac{h}{s} \rceil - 1} \psi_k \, \varepsilon_{n+h-sk}$, where ψ_k is given by expression (2.8).*

(iii). *The theoretical and sample variances of the forecast error are given, respectively, by*

$$Var(e_n(h)) = \sigma_\varepsilon^2 \sum_{k=0}^{\lceil \frac{h}{s} \rceil - 1} \psi_k^2, \quad \text{and} \quad \widehat{V}ar(e_n(h)) = \widehat{\sigma}_\varepsilon^2 \sum_{k=0}^{\lceil \frac{h}{s} \rceil - 1} \widehat{\psi}_k^2,$$

where $\widehat{\psi}_k$ is given by expression (2.8) when D is replaced by one of its estimated values, through some of the estimation procedures proposed in Section 4.

(iv). *The bias and the percentage bias to estimate the theoretical variance of the forecasting error are given by*

$$bias(h) = \widehat{V}ar(e_n(h)) - Var(e_n(h))$$

and

$$perbias(h) = \frac{|\widehat{V}ar(e_n(h)) - Var(e_n(h))|}{Var(e_n(h))} \times 100\ \%.$$

(v). *The mean squared error of forecasting is given by* $msef_n = \dfrac{1}{h}\sum_{k=1}^{h}(e_n(k))^2$.

(vi). *Moreover, if the process* $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ *is such that* $\varepsilon_t \sim \mathcal{N}(0,\sigma_\varepsilon^2)$, *for any* $t \in \mathbb{Z}$, *then an* $100(1-\gamma)\%$ *confidence interval for* X_{n+h} *is given by*

$$\widehat{X}_n(h) - z_{\frac{\gamma}{2}}\,\widehat{\sigma}_\varepsilon \left[\sum_{k=0}^{\lceil \frac{h}{s}\rceil - 1} \widehat{\psi}_k^2\right]^{\frac{1}{2}} \leqslant X_{n+h} \leqslant \widehat{X}_n(h) + z_{\frac{\gamma}{2}}\,\widehat{\sigma}_\varepsilon \left[\sum_{k=0}^{\lceil \frac{h}{s}\rceil - 1} \widehat{\psi}_k^2\right]^{\frac{1}{2}},$$

where $z_{\frac{\gamma}{2}}$ *is the value such that* $\mathbb{P}(Z \geqslant z_{\frac{\gamma}{2}}) = \frac{\gamma}{2}$, *with* $Z \sim \mathcal{N}(0,1)$, *and* $\widehat{\psi}_k$ *is given by the above item* (iii).

Proof. For a proof see Bisognin and Lopes [5]. □

We point out that a similar result to Theorem 5.1 can be stated for ARFIMA processes.

6. An Application

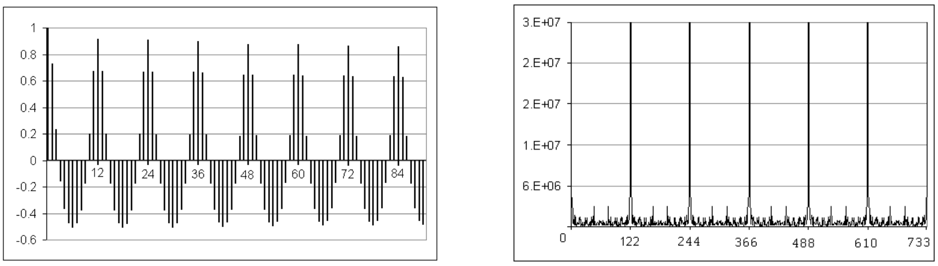
In this section we analyze an observed time series data, and also a simulated seasonal fractionally integrated ARMA time series. Our goal is to give an application of the SARFIMA methodology, analyzing these two time series in order to detect whether seasonal long memory is present in these data.

In Section 6.1 we analyze an observed time series as an application to the SARFIMA$(p,d,q)\times(P,D,Q)_s$ process. As it is not easy to find observed examples modeled by the pure SARFIMA$(0,D,0)_s$ process, we simulate a time series and analyzed it in Section 6.2.

6.1. Nile River Monthly Flows Data

We consider the time series reporting the Nile River monthly flows at Aswan, kindly provided by A. Montanari (for the graphic of the data we refer the reader to Montanari et al. [34]). This time series consists of 1,466 observations, from August of 1872 to September of 1994, and it is approximately a Gaussian time series.

Figures 6.1 (a) and (b) present, respectively, the sample autocorrelation function and the periodogram function for the Nile River flows at Aswan. From these figures one can see long memory features for this time series, since its sample autocorrelation has a slowly hyperbolic decay, and its periodogram function exhibits periodic pattern caused by an annual cycle. Figure 6.1 (b) shows the peaks on the Fourier frequencies w_j, where $j = \left[\frac{n}{s}\right] i = \left[\frac{1{,}466}{12}\right] i = 122i$, for $i = 0, 1, \cdots, 6$. These features are also reported in Montanari et al. [34].



(a) (b)

Figure 6.1: Nile River Monthly Flows Data at Aswan: (a) sample autocorrelation function; (b) periodogram function.

The model that best fits the original data is a SARFIMA$(p,d,q)\times(P,D,Q)_s$ with $p = q = P = Q = 1$, $d = 0$, $D = W$ and $s = 12$ (for a complete analysis, we refer the reader to Bisognin and Lopes [5]). The long memory parameter is estimated by the approximated maximum likelihood method proposed by Fox and Taqqu [15] (see Section 4), with $W = 0.1980$.

Table 6.1: Estimated Values of D for: (a) Nile River Monthly Flows Data; (b) Simulated Time Series Data.

SARFIMA$(0,D,0)_s$ with $s = 12$ and $\alpha = 0.55$						
Estimator	GPH	BA	R	R/S	$GPHT$	W
(a) Nile River Monthly Flows Data	0.2399	0.3126	0.2381	0.1638	0.4196	0.1980
(b) Simulated Time Series Data	0.4219	0.4216	0.4398	0.3893	0.4185	0.3834

Table 6.1 (a) gives the estimation results for this time series with seasonality $s=12$, since Figures 6.1 (a) and (b) exhibits this periodic pattern. All the semi-parametric estimation procedures select the number of regressors $m = g(n)$, in the expression (4.4), from the first seasonal frequency, no matter what value one uses for s.

Table 6.2 gives the estimators and its standard deviation (denoted here by Std. Dev.) values for the parameters in the SARFIMA$(p,d,q)\times(P,D,Q)_s$ model, that best fitted the Nile River monthly flows data at Aswan.

The residual analysis was also performed for the fitted model and it indicates that the errors are approximately Gaussian white noise.

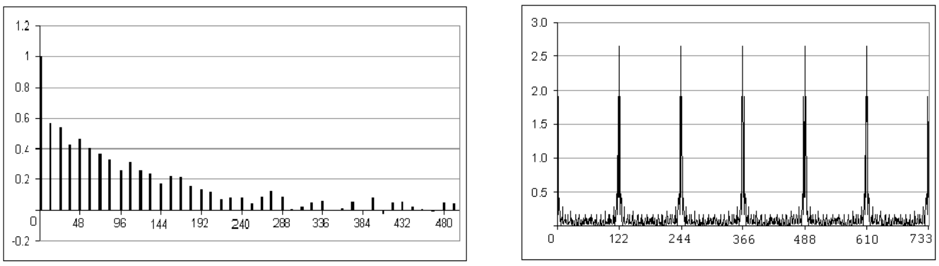
Table 6.2: Fitted Model for the Nile River Flows Data.

SARFIMA$(p,d,q)\times(P,D,Q)_s$ with $p=q=P=Q=1$, $d=0$, $D=W$ and $s=12$					
	ϕ_1	Φ_1	D	θ_1	Θ_1
Estimator	0.6147	0.9944	0.1980	-0.2238	0.9207
Std. Dev.	0.0291	0.0295	0.0011	0.0357	0.0145

6.2. Simulated Time Series

Here we consider a complete estimation, and also the forecasting analysis, for a simulated seasonal fractionally integrated time series as in expression (2.6), when $n = 1,466$, $D = 0.4$, and $s = 12$.

Figures 6.2 (a), and (b) show the sample autocorrelation, and the periodogram functions of this simulated time series: there exist long memory characteristics in this time series. By analyzing the periodogram function we also observe a periodic pattern with seasonality $s = 12$.



(a) (b)

Figure 6.2: Simulated Time Series Data: (a) sample autocorrelation function; (b) periodogram function.

Table 6.1 (b) gives the estimators of the parameter D for a SARFIMA$(0, D, 0)_s$ with $s = 12$, that best fits the simulated time series.

The best estimator for the simulated time series is $W = 0.3834 \simeq 0.4$. In the semi-parametric estimator class the total number of regressors $m = g(n)$, in the expression (4.4), was selected from the first seasonal frequency.

Figure 6.3 shows the confidence interval at 95% confidence level for the 5-step ahead forecasting values based on all estimation procedures considered in Section 4 for the simulated time series data.

We refer the reader to Lopes and Nunes [27] and Pinheiro and Lopes [39] for a long-range dependence studies in DNA sequences. For a self-similar analysis in the Ethernet traffic we refer the work by Leland et al. [23]. We also mention Beran [4], Doukhan et al. [12] and Palma [38] for a series of examples and applications on *long-range dependence*.

Figure 6.3: Confidence interval at 95% confidence level for the 5-step ahead forecasting in the simulated time series data.

7. Conclusions

In this paper we addressed the issue of modeling the *long-range dependence* through a series of different stochastic processes.

We considered models with *long memory in mean* (ARFIMA and SARFIMA processes) and *in volatility* (FIGARCH and FISV processes), with innovations following either a Gaussian or a non-Gaussian distribution.

In studying SARFIMA$(0, D, 0)_s$ processes we emphasize Theorems 2.1-2.3 and 5.1. Theorem 2.2 presents the ergodicity property while Theorem 2.3 presents the conditional expectation, and conditional variance for these processes. Theorem 2.3 is very important for generating any SARFIMA$(0, D, 0)_s$ process or its complete version SARFIMA$(p, d, q) \times (P, D, Q)_s$ process. Theorem 5.1 gives some properties for forecasting the value X_{n+h}, when $h \geqslant 1$, in SARFIMA$(0, D, 0)_s$ processes.

Based on the Pfaff–Saalschütz's Identity and some properties of the hypergeometric functions, we derived a compact and closed formula for the Durbin-Levinson algorithm in order to obtain the partial autocorrelation functions of order k for SARFIMA$(0, D, 0)_s$ processes.

In Section 4 we presented one parametric and four semiparametric methods for estimating the fractional differencing parameter. The classical *Ordinary Least Squares* (*OLS*) method and two robust methodologies (that is, the *Least Trimmed Squares* (*LTS*) and the *MM-estimation*) were presented for each estimation method in the semiparametric class.

We recall that in a FISV process the logarithm transformation of its squared value is the sum of a zero mean Gaussian ARFIMA process and an independent non-Gaussian innovation process. Also, the FIGARCH(p, d, q) process is an ARFIMA process on the squared data with a more complicated error structure. In view of this, all the estimation methods proposed in Section 4, or in any of the author's papers mentioned in the references, can also be applied to FIGARCH and FISV processes besides the ARFIMA process.

As an illustration of the SARFIMA methodology we presented an application in hydrology.

References

[1] G. E. Andrews, R. Askey and R. Roy, *Special Functions, Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 1999.

[2] R. T. Baillie, T. Bollerslev and H. O. Mikkelsen, *Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity.* J. of Econometrics **74** (1996), 3–30.

[3] N. K. Bary, *A Treatise on Trigonometric series.* Pergamon Press, 1964.

[4] J. Beran, *Statistics for Long-Memory Process.* Chapman and Hall, 1994.

[5] C. Bisognin and S. R. C. Lopes, *Estimating and Forecasting the Long Memory Parameter in the Presence of Periodicity.* J. of Forecasting **26** (2007), 405–427.

[6] C. Bisognin, *Estimação e Previsão em Processos SARFIMA*$(p, d, q) \times (P, D, Q)_s$ *na Presença de Outliers.* Ph.D. Thesis in the Mathematics Graduate Program. Federal University of Rio Grande do Sul, Porto Alegre, 2007. URL Address: www.mat.ufrgs.br/∼slopes.

[7] T. Bollerslev and H. O. Mikkelsen, *Modeling and pricing long memory in stock market volatility.* J. of Econometrics **73** (1996), 151–184.

[8] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity.* J. of Econometrics **31** (1986), 307–327.

[9] F. J. Breidt, N. Crato and P. de Lima, *The detection and estimation of long memory in stochastic volatility.* J. of Econometrics **83** (1998), 325–348.

[10] E. H. M. Brietzke, S. R. C. Lopes and C. Bisognin, *A Closed Formula for the Durbin-Levinson's Algorithm in Seasonal Fractionally Integrated Processes.* Mathematical and Computer Modelling **42** (2005), 1191–1206.

[11] P. J. Brockwell and R.A. Davis, *Time Series: Theory and Methods.* Springer-Verlag, 1991.

[12] P. Doukhan, G. Oppenheim and M. S. Taqqu (eds.) *Theory and Applications of Long-Range Dependence.* 1st Edition, Birkhäuser, 2003.

[13] R. Durret, *Probability: Theory and Examples.* Duxbury Press, 2004.

[14] R. F. Engle, *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation.* Econometrica **50** (1982), 987–1008.

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

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

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

[18] J. Hosking, *Fractional Differencing.* Biometrika **68** (1981), 165–167.

[19] J. Hosking, *Modelling Persistence in Hydrological Time Series using Fractional Differencing.* Water Resources Research **68** (1984), 1898–1908.

[20] P. J. Huber, *Robust Statistics.* John Wiley, 1981.

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

[22] C. M. Hurvich and B. K. Ray, *Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes.* J. of Time Series Analysis **16** (1995), 017–041.

[23] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, *On the self-similar nature of Ethernet traffic (Extended version).* IEEE/ACM Transactions on Networking **2** (1994), 01–15.

[24] A. W. Lo, *Long-term memory in stock market prices.* Econometrica **59** (1991), 1279–1313.

[25] S. R. C. Lopes and A. Pinheiro, *Wavelets for Estimating the Fractional Parameter in Non-stationary ARFIMA Process.* In revision, 2007.

[26] S. R. C. Lopes and B. V. M. Mendes, *Bandwidth Selection in Classical and Robust Estimation of Long Memory.* International Journal of Statistics and Systems **1** (2006), 177–200.

[27] S. R. C. Lopes and M. A. Nunes, *Long Memory Analysis in DNA Sequences.* Physica A: Statistical Mechanics and its Applications **361** (2006), 569–588.

[28] S. R. C. Lopes, B. P. Olbermann and V. A. Reisen, *A Comparison of Estimation Methods in Non-stationary Arfima Processes.* J. of Statistical Computation and Simulation **74** (2004), 339–347.

[29] A. O. Lopes and S. R. C. Lopes, *Convergence in Distribution of the Periodogram for Chaotic Processes.* Stochastics and Dynamics **2** (2002), 609–624.

[30] A. O. Lopes and S. R. C. Lopes, *Parametric Estimation and Spectral Analysis of Piecewise Linear Maps of the Interval.* Advances in Applied Probability **30** (1998), 757–776.

[31] B. B. Mandelbrot and J. van Ness, *Fractional Brownian Motion, Fractional Noises and Applications.* S.I.A.M. Review **10** (1968), 422–437.

[32] B. B. Mandelbrot and J. R. Wallis, *Computer Experiments with Fractional Gaussian Noises. Part 1, Averages and Variances.* Water Resources Research **5** (1969), 228–267.

[33] B. B. Mandelbrot and M. S. Taqqu, *Robust R/S analysis of long-run serial correlation.* In Proceedings of the 42nd Session of the International Statistical Institute, Manila. *Bulletin of the International Statistical Institute* **48** (1979), Book 2, 69–104.

[34] A. Montanari, R. Rosso and M. S. Taqqu, *A seasonal fractional ARIMA Model applied to the Nile River monthly flows at Aswan.* Water Resources Research **36** (2000), 1249–1259.

[35] B. P. Olbermann, S. R. C. Lopes and A. O. Lopes, *Parameter Estimation in Manneville-Pomeau Processes.* Submitted, 2007.

[36] B. P. Olbermann, S. R. C. Lopes and V. A. Reisen, *Invariance of the First Difference in ARFIMA Models.* Computational Statistics **21** (2006), 445–461.

[37] B. P. Olbermann, *Estimação em Classes de Processos Estocásticos com Decaimento Hiperbólico da Função de Autocorrelação.* Ph.D. Thesis in the Mathematics Graduate Program. Federal University of Rio Grande do Sul, Porto Alegre, 2002. URL Address: www.mat.ufrgs.br/∼slopes.

[38] W. Palma, *Long-memory Time Series: Theory and Methods.* 1st Edition, John Wiley, 2007.

[39] A. Pinheiro and S. R. C. Lopes, *The Use of Wavelets for Studying Long-Range Dependence in DNA Sequences.* Manuscript in preparation, 2007.

[40] S. Porter-Hudak, *An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates.* J. of American Statistical Association **85** (1990), 338–344.

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

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

[43] P. M. Robinson and P. Zaffaroni, *Nonlinear Time Series with Long Memory: a Model for Stochastic Volatility.* J. of Statistical Planning and Inference **68** (1998), 359–371.

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

[45] P. J. Rousseeuw, *Least Median of Squares Regression.* J. of the American Statistical Association **79** (1984), 871–880.

[46] M. R. Sena Jr, V. A. Reisen and S. R. C. Lopes, *Correlated Errors in the Parameters Estimation of ARFIMA Model: A Simulated Study.* Communications in Statistics: Simulation and Computation **35** (2006), 789–802.

[47] M. S. Taqqu, V. Teverovsky and W. Willinger, *Estimators for long range dependence: an empirical study.* Fractals **3** (1995), 785–798.

[48] C. Velasco, *Gaussian Semiparametric Estimation of Non-stationary Time Series.* J. of Time Series Analysis **20** (1999a), 87–127.

[49] C. Velasco, *Non-stationary log-periodogram regression.* J. of Econometrics **91** (1999b), 325–371.

[50] P. Whittle, *Estimation and information in stationary time series.* Arkiv för Mathematik **2** (1953), 423–434.

[51] V. J. Yohai, *High Breakdown point and high efficiency robust estimates for regression.* Annals of Statistics **15** (1987), 642–656.

Acknowledgements

S. R. C. Lopes was partially supported by CNPq-Brazil, by *Millennium Institute in Probability*, by Edital Universal *Modelos com Dependência de Longo Alcance: Análise Probabilística e Inferência* (CNPq-No. 476781/2004-3) and also by *Fundação de Amparo à Pesquisa no Estado do Rio Grande do Sul* (FAPERGS Foundation).

The author would like to thank an anonymous referee and both editors for their valuable comments and suggestions that improved the final version of the manuscript.

Sílvia R. C. Lopes
Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Av. Bento Gonçalves, 9500
91509-900 Porto Alegre - RS
Brazil
e-mail: `silvia.lopes@ufrgs.br`