ISSN 2347-1921

Bootstrap Estimation of Long-Range Dependence in Arfima Processes

Argjir Butka, Dhori Beta
"Fan S. Noli" University, Shetitorja "Rilindasit" Korce, Albania
argjirbutka@yahoo.com
"Fan S. Noli" University, Shetitorja "Rilindasit" Korce, Albania
dhoribeta@yahoo.com

ABSTRACT

Long-memory or long-range dependence (LRD) denotes the property of a time series to exhibit a significant dependence between very distant observations. The presence of long-memory is established in diverse fields, for example, in hydrology, finance, network traffic, psychology, etc. From a practical point of view an important issue is the estimation of LRD parameters, especially the estimation of the Hurst parameter proposed in [19]. The Hurst parameter, usually referred as the *H* parameter, indicates the intensity of LRD.

Various methods for estimating the *H* parameter in a time series data are available some of which are described in [3]. They are validated by appealing to an asymptotic analysis supposing that the sample size of the time series converges to infinity. Taqqu et al. [37] investigated empirical performance of nine methods for estimating the LRD in simulated auto regressive fractionally integrated moving average (ARFIMA) processes for a fairly large sample size of 10000 terms.

The bootstrap is a general statistical procedure for statistical inference,originally introduced by Efron [11] that generally yields better estimations in finite sample sizes, than the classical methods. In this paper, we consider a bootstrap technique, which uses blocks composed of cycles, to estimate the *H* parameter in ARFIMA processes. The bias and the root mean square error (RMSE) of five estimators and their corresponded bootstrap estimators are obtained by a Monte Carlo study.

Keyword: Time series; long-range dependence; Hurst parameter; block bootstrap; cycling blocks.

Mathematics Subject Classifications: 62G05, 62G09, 62M10.

ISSN 2347-1921

1. INTRODUCTION

Time series with long-memory or long-range dependence (LRD) appear in many contexts, for example, in hydrology [28], econometrics [2], network traffic [24], finance [26], etc. LRD is the dependence structure across long time periods. It denotes the property of a time series to exhibit the persistent behavior. However, there is not a unique theoretical definition of the LRD processes (see[10]). Most of the definitions appearing in the literature are based on the second order properties of a stochastic process. Such properties include asymptotic behavior of the autocovariance function or the spectral density function. In the time domain, if a stationary time series exhibits LRD, its autocovariances decay of hyperbolic order sothey are not summable. In frequency domain, if a stationary time series exhibits LRD, its spectral density function is unbounded at frequency zero.

The importance of LRD processes as stochastic models lies in the fact that they provide an elegant explanation and interpretation of an empirical law that is commonly referred to as Hurst's law or the Hurst effect. This feature of many natural phenomena has been called the "Joseph effect" by Mandelbrot and Wallis [29], referring to the Old Testament prophet who foretold of the seven years of plenty followed by the seven years of famine that Egypt was to experience. The intensity of the long memory can be measured by a parameter called the Hurst parameter or the *H* parameter, which is closely related to the order of the decay of the autocovariance function. Various methods for estimating the *H* parameter in a time series data are available. However, they are validated by appealing to an asymptotic analysis where one supposes that the sample size of the time series converges to infinity. The finite sample properties of these estimators can be quite different from their asymptotic properties. Also, they are biased in the case of the presence of the short memory component.

The Autoregressive Fractionally Integrated Moving Average, ARFIMA model has widely been used to represent a time series with long memory. The intensity of the memory can be measured by the fractional difference parameter *d*, related to the *H* parameter by the equation $H = d + 0.5$.Taqquet al. [37] investigated empirical performance of nine methods for estimating the LRD in simulated ARFIMA(0,*d*,0) processes for a relatively large sample size of 10000 terms. In this paper, we investigate the performance of five estimators and their corresponding bootstrap estimator in simulated ARFIMA(1,*d*,0) processes.

Bootstrap method, originally introduced by Efron [11] is a resample technique that provides better estimations in smallsample sizes than the classical methods. The Efron's bootstrap uses single observations as the resampling units and requires independence among the data. Referring to this bootstrap, Singh [36] noted that "*the bootstrapshould not be expected to provide a consistent approximation even in the case of weak, dependent processes"*([36], Remark 2.1, page 1192). When the data on the hand are not independent and identically distributed (i.i.d.), such as a time series, the resample technique must be carried out in such a way that the dependence structure of the original time series to be preserved in the bootstrap time series. The most common are the block bootstrap methods. The block bootstrap for time series consists of randomly resampling blocks of consecutive values of the given data and aligning these blocks into a bootstrap sample.

In this paper, we consider a bootstrap technique, which uses blocks composed of cycles.A cycle is defined as a pair of alternating high and low data that is created when the terms of the time series cross the sample mean. Then we randomly resample blocks composed of a fixed number of consecutive cycles and concatenate them to form the bootstrap series.

The remainder of the paper is as follows. In the next section we describe the notion of LRD and give details of the five estimators, used in a Monte Carlo study. In section 3 we describe the block bootstrap method with blocks composed of cycles. Section 4 reports the Monte Carlo experiment and presents the results.

2. LONG-RANGE DEPENDENCE

2.1. Definitions and models

Long-memory or long-range dependence (LRD) refers to the property of a time series to exhibit a significant dependence between very distant observations. Let $\{X_t\}_{t\in Z}$ be a stationary time series with real values. It is intuitively expected that the autocorrelation function vanishes when the distance between the data becomes large. So we can suppose that the data are asymptotically independent and the autocorrelations are absolutely summable. It is the case in most of stationary time series included the vast class of Autoregressive Integrated Moving Average, ARMA models, which are in general classified asshort-memory processes [4, 5]. In contrast, LRD is generally defined by the fact that the autocorrelations are absolutely non-summable [31]. However, there are several possible definitions of the property of LRD (see [10] for a collection of definitions of the LRD notion). Most of the definitions appearing in the literature are based on the second order properties of a stochastic process (for more details see [3]).

For a stationary process $\{X_t\}_{t\in Z}$, the autocovariance function is defined by equation $\gamma(k) = Cov(X_t, X_{t+k}), k = 1,2,...$, and does not depend on the moment t. Then the autocorrelation function is defined by the equation $\rho(k) = \frac{\gamma(k)}{\gamma(0)}, k = 0,1,...$, and the spectral density function is defined by equation $f(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma(k)e^{ik\omega}$, where ω is the angular frequency. Let us mention two common definitions of LRD (given in[3], page 42).

Definition in time domain:

Suppose that there exists a real number $\alpha \in (0,1)$ and a constant $c_\rho > 0$ not depended on *k* such that $\lim_{k\to\infty} \frac{\rho(k)}{c_\rho k^{-\alpha}} = 1$. Then $\{X_t\}_{t\in Z}$ is called a stationary process with long memory or long-range dependence or strong dependence, or a stationary process with slowly decaying or long-range correlations.

Definition in spectral domain:

Suppose that there exists a real number $\beta \in (0,1)$ and a constant $c_f > 0$ not depended on ω such that $\lim_{\omega\to 0} \frac{f(\omega)}{c_f |\omega|^{-\beta}} = 1$. Then $\{X_t\}_{t\in Z}$ is called a stationary process with long memory or long-range dependence or strong dependence.

A number of models have been proposed to describe the long-memory feature of time series. The fractional Gaussian noise (FGN) model is the first model with LRD introduced by Mandelbrot and Van Ness [28]. Then Hosking [18] and Granger and Joyeux [14] proposed the fractionally integrated autoregressive and moving average model, denoted by ARFIMA(*p*,*d*,*q*). It is defined by $\Phi(B)(1-B)^d X_t = \theta(B)\varepsilon_t$ where $\{X_t\}_{t\in Z}$ is the time series; *B* is the backshift operator, that is, $BX_t = X_{t-1}$; $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$ and $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + ... + \theta_p B^p$ represent the ordinary autoregressive and moving average components; ε_t is a white noise process with zero mean and variance σ_ε^2. When $-0.5 < d < 0.5$, the ARFIMA(*p*,*d*,*q*) is stationary, and if $0 < d < 0.5$ the process presents long-memory behavior. The larger *d*, the longer the memory the stationary process has. When $d = 0$ ARFIMA(*p*,*0*,*q*) model is reduced to an short-memory ARMA(*p*,*q*) model, popularized by Box and Jenkins [4]. Instead of using d, we may use $H = d + 0.5$ which is known as the Hurst parameter to measure the long-memory in $\{X_t\}_{t\in Z}$.

2.2. Estimation of the Hurst parameter

(a)*The rescaled adjusted range (R/S) estimator.*

The rescaled adjusted range method, denoted by R/S was proposed by Hurst [19] and discussed in details in [3, 27, 29, 30]. For a time series $\{X_t, t = 1,2,...,N\}$, let $Y_t = \sum_{j=1}^{t} X_j$ be the aggregated time series. Then we define the adjusted range $R(t,k)$ by $R(t,k) = \max_{0\le i\le k}\left[(Y_{t+i} - Y_t) - \frac{i}{k}(Y_{t+k} - Y_t)\right] - \min_{0\le i\le k}\left[(Y_{t+i} - Y_t) - \frac{i}{k}(Y_{t+k} - Y_t)\right]$. This range can be interpreted as the ideal capacity of a water reservoir with in flows at time i denoted by X_i for the time span between *t* and $t+k$ (see [3]).

For the adjusted range $R(t,k)$ be independent of a scale, it is standardized by standard deviation $S(t,k) = \sqrt{\frac{1}{k}\sum_{i=t+1}^{t+k}\left(X_i - \bar{X}_{t,k}\right)^2}$, where $\bar{X}_{t,k} = \frac{1}{k}\sum_{i=t+1}^{t+k} X_i$. Then the rescaled adjusted range is given by $R/S = \frac{R(t,k)}{S(t,k)}$ for all possible values of *k* and *t*.

To estimate the LRD parameter *H*, we plot the logarithm R/S against logarithm of *k* for all possible values of *t* and *k*. Hurst observed that many empirical records may be well represented by the relation $\log(R/S) \approx c + H\log(k)$ as *k* becomes large, with typically values of the Hurst parameter in the interval $]0.5,1[$ and *c* a finite positive constant that does not depend on *k.* Then *H* can be estimated as the slope ordinary least squares (OLS) estimator.

Let $\{X_t, t = 1,2,...,N\}$ be an observed time series of length *N*. Then, the algorithm consists asbellow for estimation of the *H* parameter.

1. Divide the time series into *b* blocks of length *k* such that $bk = N$ (supposing that $\frac{N}{k}$ is an integer).

2. For each block compute $(R/S)_{ki} = \frac{R(t_i,k)}{S(t_i,k)}$ for $t_i = 1, k+1, 2k+1, ..., (b-1)k+1$. Then, the average value of R/S for length *k* is computed, that is $(R/S)_k = \frac{1}{b}\sum_{i=1}^{b}(R/S)_{ki}$

3. Repeat the steps above by increasing the value of *k* until $k = \frac{N}{2}$ (such that $\frac{N}{k}$ is an integer).

4. Plot logarithm of $(R/S)_k$ against logarithm of *k*. This plot is sometimes called the *pox plot* for the R/S statistic.

5. The estimated parameter $\hat{H}$ is the estimated slope of the line in the *pox plot*.

The R/S method requires cutting off both the low and high end of the plot to make reliable estimates.

(b) *The detrended fluctuation analysis(DFA) estimator*

Detrended fluctuation analysis (DFA) estimator is introduced by Peng et.al. [32] as an improvement of classical fluctuation analysis method and is supposed to deal with power-law correlations in non-stationary time series. The DFA method has proven useful in revealing the extent of LRD in time series. The construct of this method consists as bellow.

1. First, we integrate the time series $\{X_t, t=1,2,...,N\}$ obtaining the cumulative sum time series $\{Y_t, t=1,2,...,N\}$ (called the *profile function*) where $Y_t = \sum_{i=1}^{t}\left(X_i - \bar{X}\right)$ and $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$.

2. Divide the cumulative sum time series into *k* windows (non-overlapping intervals) of length *m* observations. Then in any box, Y_t is fitted by an *r* order polynomial, usually with $r=1$ or $r=2$ obtaining the fitted values $\hat{Y}_t^{(m)}$ in any box.

3. Detrend the $\{Y_t, t=1,2,...,N\}$ locally in any box obtaining $Y_t^{DET} = Y_t - \hat{Y}_t^{(m)}$.

4. Calculate the variance about the detrended series for each box, and calculate the average of these variances over all boxes of size *m*, $F^2(m) = \frac{1}{N}\sum_{t=1}^{N}(Y_t^{DET})^2$.

5. Repeat steps 1-4 for increasing values of *m*.

6. Regress $\log F^2(m)$ on $\log m$ and estimate the slope parameter. If $\hat{\beta}$ is the OLS slope estimate, then $\hat{H} = \frac{\hat{\beta}}{2}$ is the estimate of *H*.

(c) *The discrete wavelet transforms (DWT) estimator*

Wavelet methods are used in diverse applications in several fields (see [33]). One scientific area where wavelet methods have been finding many applications is that of the analysis of discrete time series. The discrete wavelet transform (DWT) method for estimating long memory was introduced by Jensen [20] and consists as below (see [1] and [38]) for the estimation of long-memory parameter *H*.

1. Assume that $N = 2^J$ for some positive integer *J*.

The discrete wavelet filter *h* of support $L_j = 2^{J-j+1}$ (length of the filter) is defined so that satisfies the following conditions.

$\sum_{i=1}^{L_j} h_i = 0$, $\sum_{i=1}^{L_j} h_i^2 = 1$ and $\sum_{i=1}^{L_j} h_i h_{i+2m} = 0$, with *m* an integer other than zero.

Similarly, the scaling filter of support L_j satisfies the following three conditions.

$\sum_{i=1}^{L_j} g_i = 2^{j/2}$, $\sum_{i=1}^{L_j} g_i^2 = 1$ and $\sum_{i=1}^{L_j} g_i g_{i+2m} = 0$

2. Using these filters the original time series $\{X_t, t=1,2,...,N\}$ is transformed to two new series $W_{j,t}$ and $V_{j,t}$ called the wavelet coefficient and the scaling coefficient respectively, by recursive equations $V_{1,t} = \{X_t\}$, $V_{j,t} = \sum_{i=1}^{L_j} g_i V_{(j-1),\{(2t-i)\bmod(N/2^{j-1})\}+1}$ and $W_{j,t} = \sum_{i=1}^{L_j} h_i V_{(j-1),\{(2t-i)\bmod(N/2^{j-1})\}+1}$, for $j = 2,3,...,J-1$, and $t = 1,2,...,L_j$.

3. Using the wavelet coefficient $W_{j,t}$ the adjusted wavelet variance is defined by $\hat{\sigma}_j^2 = \frac{\sum_{t=1}^{L_j} W_{j,t}^2}{2^{j-1}}$. If we denote $\tau_j = 2^j - 1$, then the relation $\hat{\sigma}_j^2 \approx \tau_j^{2H-1}$ holds.

4. By taking the logarithm of both sides and regressing $\log \hat{\sigma}_j^2$ on $\log \tau_j$ we can calculate the OLS slope estimator $\hat{\beta}$ obtaining an estimate of *H* by equation $\hat{H} = \frac{\hat{\beta}+1}{2}$.

It is supposed that the data are generated by an ARFIMA (*p*,*d*,*q*) model for the two following estimators to be valued.

(d) *The Geweke and Porter-Hudak(GPH) estimator*

One possibility of estimating the memory parameter is a log-periodogram regression introduced by Geweke and Porter-Hudak [13], known as a GPH estimator in literature. The GPH estimator is the slope coefficient estimated in the least-squares regression of $\log I(\omega_j) = a - d\log[4\sin^2(\omega_j/2)] + e_j$ for $j = 1,2,...,g$, where $I(\omega) = \frac{1}{2\pi}\left[\hat{\gamma}_0 + 2\sum_{k=1}^{N-1}\hat{\gamma}_k \cos(k\omega)\right]$ is the periodogram of the observed time series; $\hat{\gamma}_k = \frac{1}{N}\sum_{i=1}^{N-k}(X_{i+k} - \bar{X})(X_i - \bar{X})$, for $k = 1,2,...,N-1$ are the sample autocovariances; and $\omega_j = \frac{2\pi j}{N}$, $j = 1,2,...,g < \frac{N}{2}$ denote the harmonic frequencies.Then, the GPH estimator is given by $\hat{d}_{GPH} = \frac{\sum_{j=1}^{g}(v_j - \bar{v})\log I(\omega_j)}{\sum_{j=1}^{g}(v_j - \bar{v})^2}$, where $v_j = -\log[4\sin^2(\omega_j/2)]$ and $\bar{v} = \frac{1}{g}\sum_{j=1}^{g} v_j$. Then we have the estimator $\hat{H} = 0.5 + \hat{d}_{GPH}$ The parameter $g = g(N)$ is a function of N and has to take an appropriate value since it determines the bias and variance of the GPH estimator.

(e) *The smoothed periodogram(SP) estimator.*

Hassler [16], Chen et al. [9] and Reisen [35] all independently suggested an estimate of d based on the regression procedure for the logarithm of the smoothed periodogram. The SP estimator, usually denoted by $\hat{d}_{SP}$, is obtained by GPH estimator by replacing the periodogram by a smoothed periodogram function. The SP estimate of the spectral density function is $I_{SP} = \frac{1}{2\pi}\left[\lambda_0\hat{\gamma}_0 + 2\sum_{k=1}^{m}\lambda_k\hat{\gamma}_k \cos(k\omega)\right]$, where $\{\lambda_k, k = 1,2,...,m\}$ is a set of weights called the lag window. Reisen[35] used the Parzen lag window given by

$$\lambda_k = \begin{cases} 1 - 6\left(\frac{k}{m}\right)^2 + 6\left(\frac{k}{m}\right)^3, 0 \le k \le \frac{m}{2} \\ 2\left(1 - \frac{k}{m}\right)^3, \quad \frac{m}{2} \le k \le m \end{cases}$$

Here, m is a function of N and is usually referred to as the truncation point. Then the $\hat{d}_{SP}$ is given by $\hat{d}_{SP} = \frac{\sum_{j=1}^{g}(v_j - \bar{v})\log I_{SP}(\omega_j)}{\sum_{j=1}^{g}(v_j - \bar{v})^2}$. Then we have the estimator $\hat{H} = 0.5 + \hat{d}_{SP}$. In this case, we have to choose two tuning parameters, the truncation parameter of the regress, $g = g(N)$ and the truncation point of the lag window, $m = m(N)$.

3. THE BLOCK BOOTSTRAP METHOD WITH CYCLING BLOCKS

Efron's bootstrap classical application in the context of dependent data, such as a time series, fails to work. In this case, the bootstrap technique must be carried out in such a way that the dependence structure of the original time series to be preserved in the bootstrap time series. Block bootstrap methods for dependent data were introduced by Hall [15], Carlstein [7] and Kunsch [21] among others. The block bootstrap for time series consists of randomly resampling blocks of consecutive values of the given data and aligning these blocks into a bootstrap sample. Different methods differ in the way as blocks are constructed. These methods are highly adaptive, or non parametric, in the spirit of bootstrap methods. The block bootstrap relies on a compromise between preserving dependence within blocks, and destroying it between blocks. Kunsch [21] and Liu and Singh [25] have independently formulated a resampling procedure, called the moving-block bootstrap (MBB), for general weakly-dependent observations. For a theoretical comparison of some methods see [23].

Bootstrap with blocks was used for the estimation of parameters of $ARFIMA(p,d,q)$ model (see for example [12]). Lahiri [22] showed that in general, the MBB procedure fails to provide a valid approximation to the distribution of normalized sample mean under LRD. One of the reasons behind this is that joining independent bootstrap blocks to form the bootstrapped statistic destroys the strong dependence of the underlying observations. Carlstein et al. [8] sampled blocks dependently, attempting to follow each block with one that might realistically follow it in the underlying process, to better match the dependence structure of the data. Hesterberg [17] used the matched-block bootstrap for LRD processes, investigating block matching rules, based on linear combinations of observations in the block.In this paper, we consider blocks that are composed of one or more consecutive cycles.

Let $X_t, t = 1,2,...,N$ be a time series. A cycle is defined as a pair of alternating high and low runs of the data that are created when the terms of the time series cross the sample mean. Let $C_1, C_2, ..., C_k$ denote the created cycles and let $n_i, i = 1,2,...,k$ be the number of terms of the cycle C_i. It is clear enough that $n_1 + n_2 + ... + n_k = N$. We suppose that $X_1 > \bar{X}$. Then we can write $C_1 = \{X_1, X_2, ..., X_{n_1}\}$, $C_2 = \{X_{n_1+1}, X_{n_1+2}, ..., X_{n_1+n_2}\}$,…, $C_i = \{X_{n_1+n_2+...+n_{i-1}+1}, ..., X_{n_1+n_2+...+n_{i-1}+n_i}\}$,…, $C_k = \{X_{n_1+n_2+...+n_{k-1}+1}, ..., X_N\}$, where $X_{n_1+...+n_k} < \bar{X}$ and $X_{n_1+...+n_k+1} > \bar{X}$, for $i = 1,2,...,k$. Then we define a block composed of a predetermined number of consecutive cycles.

The cycle lengths and the number of cycles are random variables so long as they are determined automatically by the data. The number of cycles of a block is a tuning parameter and is analogous with the block length in MBB. We consider the circular moving-block bootstrap approach, which amounts to wrapping the data around in a circle before the blocks are created [21, 34]. We treat a cycle as an inseparable observation. Since the cycles, consequently the blocks, are created automatically by the series' crossings of the sample mean, we expect that the transition across joint points of the blocks in the bootstrap pseudo series to be more realistic.

Figures1 and 2 illustrate the bootstrap method with blocks composed of cycles.We simulated a time series of size $N = 40$ generated by an AR(1) process with coefficient $\phi = 0.5$. Figure 1(b) illustrates how the cycles are created (the first term of the simulated series coincided to be below the sample mean in the case showed in figure). Then we applied the block bootstrap method as we described above. For the clarity of the plot,we used a single cycle per block. So, the block is the same as the cycle in this example. In figure 2(a) a bootstrap sample is plotted while, the resampled cycles are plotted in figure 2(b). Observing figures 1(b) and 2(b) we see that the block joint points in the bootstrap sample have a similar structure with the joint points of the blocks in the original sample. In both original and bootstrap samples, the joint points have approximate levels, with respect to the sample mean.

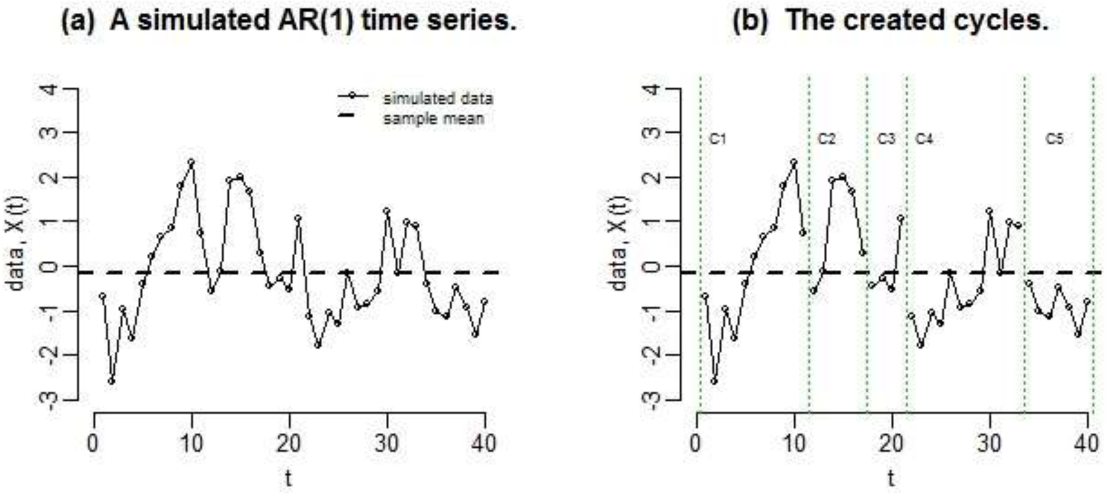

Figure 1: Illustration of cycle definition for a simulated AR(1) model

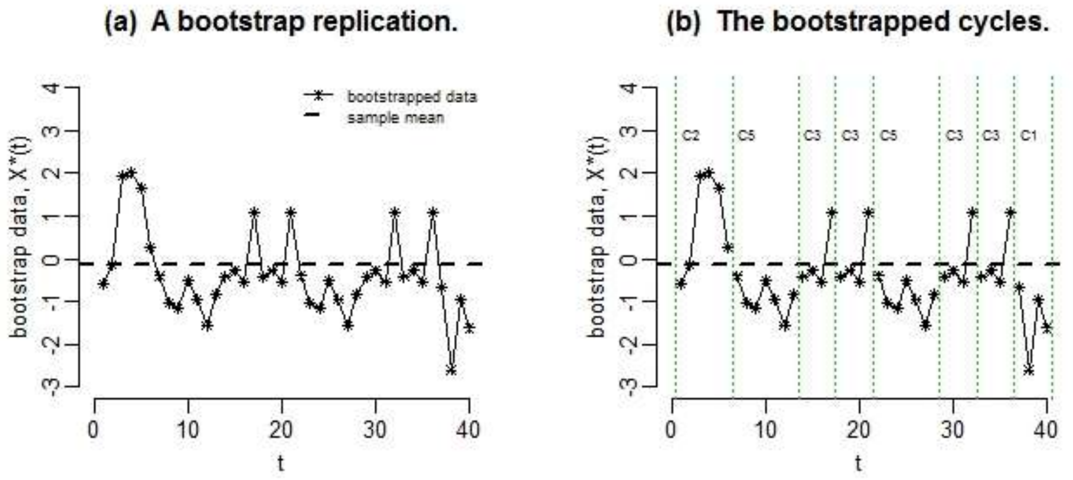

Figure 2: Illustration of block bootstrap method with blocks composed of cycles

4. SIMULATION STUDY

4.1. Experiment design

We conducted a relatively extensive Monte Carlo study and investigated the performance of all five estimators described in section 2.2 for estimating *H* in simulated ARFIMA(*p,d,q*) models. We ran 500 Monte Carlo replications of simulated ARFIMA(1*,d,*0) processes with different lengths, different values of *d* (consequently of *H*) and different values of the autoregressive coefficient in order to obtain the bias and the root mean square error (RMSE) for each estimator. The sample lengths were 128, 256, 512, 1024 and 2048 data points.The *d* values were between 0 and 0.45 in steps of 0.15 corresponding to short memory processes for $d = 0$, and to LRD processes for other values of *d*. The autoregressive coefficient values were $\phi = 0, 0.15, 0.3, 0.5, 0.8, 0.95$ corresponding to the absence of the short memory part for $\phi = 0$, and to the presence the short memory part for other values of ϕ. Specifically, we simulated ARFIMA(1,*d*,0) models of the form $(1-\phi B)^d X_t = \varepsilon_t$ for all combinations of values of ϕ and *d*. In each simulated model we discarded 300 first generated terms in order to reduce the effect of initial values. The white noise process ε_t was generated from the standard normal distribution.

For each simulated case we calculated the value of *d* (and therefore of *H*), estimated by five estimation methods under consideration in this paper and by their corresponded bootstrap estimators. We applied the circular cycling block bootstrapas we described in section 3, using 800 bootstrap replications for each Monte Carlo simulation.Then we calculated the average bias and the RMSE for each estimator based on 500 Monte Carlo replications. The number of cycles per block was chosen such that the average block length was at order $O(N^{0.5})$. We used $g(n) = N^{0.8}$ terms in the periodogram regression for GPH estimator (see [6]). For SP estimator we used $g(n) = N^{0.5}$ regression terms and the value $m = N^{0.9}$ in Parzen lag window (see [35]). For the other three estimators, we used the parameters' values as they are default in the package "*fArma*" of **R**-project. We used sample lengths expressible at a power of 2 for the DWT method to use all the series terms (the DWT method is forced to truncate the time series to the nearest power of 2)

4.2. Results

As the main result, we note that RMSE of bootstrap estimators is almost always smaller than RMSE of the corresponded classical estimators. This is true even in the cases when the bootstrap estimator has clearly larger bias in absolute value than the corresponded classical estimator. Also, we note that there is no estimator that out performs the others for all cases under consideration.

ISSN 2347-1921

Table 1. Bias of estimator (bias) and of bootstrap estimator (B.bias). The ARFIMA(1,0,0) case. (H=0.5)

Model		R/S		DFA		DWT		GPH		SP	
AR.coef	**N**	**bias**	**B.bias**	**bias**	**B.bias**	**bias**	**B.bias**	**bias**	**B.bias**	**bias**	**B.bias**
0.15	128	.1357	.1275	.0310	.0156	.0128	**-.0099**	.0845	.0691	-.0460	-.0681
	256	.1134	.1108	.0238	.0111	**.0043**	-.0072	.0692	.0598	-.0388	-.0572
	512	.1112	.1053	.0186	.0111	.0143	**.0044**	.0601	.0552	-.0227	-.0339
	1024	.0978	.0911	.0150	.0112	.0085	**-.0017**	.0447	.0415	-.0223	-.0289
	2048	.0903	.0929	.0172	.0155	-.0030	**.0001**	.0364	.0332	-.0209	-.0207
0.3	128	.1861	.1732	.1093	.0884	.1267	.0965	.2043	.1788	**-.0200**	-.0541
	256	.1684	.1488	.0771	.0640	.0799	.0647	.1630	.1496	**-.0257**	-.0488
	512	.1394	.1343	.0666	.0567	.0647	.0513	.1360	.1279	**-.0215**	-.0370
	1024	.1211	.1208	.0561	.0496	.0481	.0418	.1114	.1072	**-.0215**	-.0272
	2048	.1282	.1223	.0582	.0545	.0390	.0328	.0881	.0863	**-.0178**	-.0201
0.5	128	.2612	.2362	.2356	.2064	.2892	.2610	.3694	.3400	.0483	**-.0006**
	256	.2140	.2108	.1805	.1647	.2229	.2060	.3143	.3014	**.0080**	-.0161
	512	.1917	.1884	.1443	.1357	.1730	.1603	.2769	.2678	**-.0070**	-.0244
	1024	.1763	.1695	.1196	.1138	.1371	.1268	.2404	.2330	**-.0063**	-.0189
	2048	.1716	.1697	.1260	.1220	.1017	.0990	.2011	.1976	**-.0119**	-.0208
0.8	128	.3885	.3668	.5375	.5018	.6660	.6483	.7124	.6781	.3059	**.2388**
	256	.3442	.3415	.4587	.4360	.5832	.5688	.6651	.6374	.2280	**.1797**
	512	.3217	.3112	.3910	.3746	.4892	.4778	.6082	.5962	.1228	**.0943**
	1024	.2925	.2864	.3357	.3269	.4041	.3972	.5648	.5575	.0651	**.0494**
	2048	.2948	.2883	.3385	.3339	.3341	.3255	.5172	.5122	.0307	**.0187**
0.95	128	.4565	**.4338**	.7805	.7351	.8550	.8529	.9350	.8904	.7356	.6272
	256	.4485	**.4335**	.7570	.7216	.8476	.8475	.9090	.8834	.6637	.5908
	512	.4353	**.4263**	.7163	.6906	.8321	.8162	.8796	.8644	.5552	.5007
	1024	.4191	.4126	.6587	.6443	.7681	.7600	.8542	.8443	.4456	**.4112**
	2048	.4216	.4160	.6661	.6581	.6903	.6836	.8272	.8213	.3268	**.3049**

Table 1 and 2 report the results of the case of short-memory AR(1) models. The least bias in absolute value of each case and the smallest RMSE of each case are printed in bold. As can be seen, all estimators except SP exhibit bias towards over estimating *H* (that equals zero in this case) for any series length. The bias is very serious for short series lengths and large values of the AR coefficient. The magnitude of bias appears to increase with increasing AR coefficient and to decrease with increasing the sample length. Exception are the bias of DFA estimator and the bias of bootstrap DFA estimator that slightly increase when the sample length increases from 1024 to 2048 data points. The same trends are obteined for RMSE of all estimators. DWT estimator or bootstrap DWT estimator has the smallest bias in absolute value for $\phi = 0.15$ but DFA has the smallest RMSE in this case. The bootstrap SP estimator exhibits the smallest RMSE for $\phi = 0.3, 0.5$ and 0.8. For $\phi = 0.95$, the bootstrap R/S estimator performs in general better than the others regarding to both bias and RMSE.

Table 2. RMSE of estimator (rmse) and of bootstrap estimator (B.rmse). The ARFIMA(1,0,0) case. (H=0.5)

Model		R/S		DFA		DWT		GPH		SP	
AR.coef	N	rmse	B.rmse	rmse	B.rmse	rmse	B.rmse	rmse	B.rmse	rmse	B.rmse
0.15	128	.2059	.1367	.1025	**.0721**	.2672	.1131	.1435	.1044	.1967	.1153
	256	.1655	.1162	.0748	**.0529**	.1619	.0656	.1034	.0806	.1652	.0927
	512	.1439	.1082	.0548	**.0374**	.1095	.0424	.0843	.0671	.1338	.0539
	1024	.1285	.0927	.0400	**.0265**	.0714	.0300	.0600	.0500	.1058	.0436
	2048	.1179	.0938	.0316	.0245	.0616	**.0200**	.0490	.0400	.0837	.0300
0.3	128	.2406	.1797	.1493	**.1149**	.3022	.1483	.2345	.1957	.2071	.1241
	256	.2074	.1530	.1034	**.0825**	.1780	.0959	.1792	.1597	.1634	.0954
	512	.1652	.1364	.0837	.0663	.1257	.0671	.1473	.1334	.1323	**.0608**
	1024	.1473	.1225	.0693	.0574	.0964	.0529	.1183	.1114	.1105	**.0480**
	2048	.1503	.1229	.0656	.0583	.0714	.0387	.0938	.0889	.0860	**.0316**
0.5	128	.3040	.2417	.2604	.2245	.4039	.2893	.3857	.3506	.2045	**.1245**
	256	.2456	.2145	.1954	.1752	.2737	.2202	.3254	.3077	.1625	**.0975**
	512	.2102	.1903	.1539	.1414	.2007	.1682	.2832	.2711	.1261	**.0678**
	1024	.1921	.1706	.1261	.1175	.1568	.1315	.2439	.2352	.1000	**.0520**
	2048	.1881	.1703	.1292	.1237	.1170	.1020	.2035	.1990	.0860	**.0387**
0.8	128	.4140	.3705	.5520	.5115	.7215	.6625	.7217	.6842	.3707	**.2960**
	256	.3624	.3437	.4673	.4422	.6082	.5751	.6698	.6403	.2827	**.2254**
	512	.3335	.3124	.3956	.3778	.5007	.4808	.6108	.5977	.1806	**.1349**
	1024	.3022	.2872	.3384	.3286	.4116	.3991	.5663	.5585	.1261	**.0906**
	2048	.3035	.2888	.3401	.3350	.3394	.3270	.5181	.5127	.0900	**.0566**
0.95	128	.4754	**.4379**	.7935	.7437	.9042	.8641	.9411	.8947	.7671	.6590
	256	.4581	**.4350**	.7641	.7267	.8656	.8529	.9127	.8859	.6826	.6089
	512	.4418	**.4271**	.7199	.6933	.8391	.8188	.8816	.8656	.5725	.5158
	1024	.4240	**.4130**	.6610	.6461	.7726	.7616	.8553	.8450	.4585	.4225
	2048	.4268	.4162	.6670	.6589	.6931	.6845	.8279	.8216	.3372	**.3134**

Based on all results including the results presented in Table 1 and 2 we can summarize as following

R/S and bootstrap R/S estimators are less influenced by the large values of AR coefficient than the other estimators.

In the case of ARFIMA(0,*d*,0) models, GPH estimator outperforms the others regarding to bias, whereas the bootstrap GPH estimator outperforms the others in respect to RMSE except the case of $d = 0.3$. In the case of $d = 0.3$ the bootstrap R/S estimator has the smallest RMSE.

The bias and RMSE of bootstrap estimators are smaller than the bias and RMSE of the corresponded classical estimators in most of the cases of ARFIMA(1,*d*,0) models with moderate and large values of *d* and of AR coefficient.

Bootstrap R/S estimator has the best performance almost in all cases of LRD processes with $\phi = 0.95$. In the extreme case of ARFIMA(0.95,0.45,0) model the R/S and bootstrap R/S estimators obviously outperform the other estimators.

In conclusion, the simulation results show that applying the block bootstrap method with blocks composed of cycles yields better estimation of the LRD parameter than applying a single classical method. At least for the methods and the models conducted in this paper, the bootstrap estimators appeared smaller RMSE than the corresponded classical estimators. However, there is no estimator that simultaneously outperforms the others in all cases under consideration.

REFERENCES

[1] Abry, P. and Veitch, D. 1998. Wavelet Analysis of Long-Range-Dependent Traffic. IEEE Transactions on Information Theory 44(1), 2-15.

[2] Baillie, R. T. 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.

[3] Beran, J. 1994. Statistics for Long-Memory Processes. New York, Chapman Hall.

[4] Box, G.E.P., and Jenkins, G.M. 1976. TIME SERIES ANALYSIS, forecasting and control (revised edition). Oakland, California,HOLDEN-DAY.

[5] Brockwell, P.J., Davis, R.A. 2002. Introduction to Time Series and Forecasting (second edition). New York, Springer-Verlag.

[6] Brodsky,J. 1998. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. Journal of Time Series Analysis 19(1), 19-46.

[7] Carlstein, E. 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. The Annals Statistics 14(3), 1171-1179.

[8] Carlstein, E., Do, K-A., Hall, P., Hesterberg, T. and Kunsch, H. R. 1998. Matched-block bootstrap for dependent data. Bernoulli 4(3), 305-328

[9] Chen, G., Abraham, B. and Peiris, S. 1993. Lag window estimation of the degree of differencing in fractionally integrated time series model. Journal of Time Series Analysis 15, 473-487.

[10] Dominique, G. 2005. How can we define the concept of long memory? An econometric survey. Econometric Reviews 24(2), 113 – 149.

[11] Efron, B. 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics 7, 1-26.

[12] Franco, G.C. and Reisen, V.A. 2004. Bootstrap techniques in semiparametric estimation methods for ARFIMA models: a comparison study. Computational Statistics & Data Analysis 19, 243–259.

[13] Geweke, J., Porter-Hudak 1983. The estimation and application of long memory time series model. Journal of Time Series Analysis 4(4), 221-238.

[14] Granger, C. W. J. and Joyeux, R. 1980. An Introduction to Long-range Time Series Models and Fractional Differencing. Journal of Time Series Analysis 1, 15–30.

[15] Hall,P. 1985. Resampling a coverage pattern. Stochastic Processes and their Applications 20, 231-246.

[16] Hassler, U. 1993. Regression of spectral estimators with fractionally integrated time series. Journal of Time Series Analysis 14, 369-380.

[17] Hesterberg, T. 1997. Matched-Block Bootstrap for long Memory Processes, Technical Report No. 66, Research Department, MathSoft, Inc. 1700 Westlake Ave. N., Suite 500, Seattle, WA 98109.

[18] Hosking, J. 1981. Fractional differencing. Biometrika 68(1), 165-176.

[19] Hurst, H. E. 1951. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers 116, 770–808.

[20] Jensen, M. J. 1999. Using Wavelets to Obtain a Consistent Ordinary Least Square Estimator of the Long-memory Parameter. Journal of Forecasting 18, 17–32.

[21] Kunsch, H. 1989. The jackknife and bootstrap for general stationary observations. Annals of Statistics 17, 1217-1241.

[22] Lahiri, S. N. 1993. On the moving block bootstrap under long range dependence. Statistics & Probability Letters 18, 405-413.

[23] Lahiri, S. N. 1999. Theoretical comparisons of block bootstrap methods. The Annals of Statistics 27(1), 386-404.

[24] Leland, W. E., Taqqu, M. S., Willinger, W. and Wilson, D. V. 1994. On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM TRANSACTIONS ON NETWORKING, 2(1), 1-15.

[25] Liu, R. and Singh, K. 1992. Moving blocks jackknife and bootstrap capture weak dependence. Exploring the Limits of Bootstrap, Wiley, New York, 225–248.

[26] Lo, A. W. 1991. Long-Term Memory in Stock Market Prices. Econometrica 59(5), 1279-1313.

[27] Mandelbrot, B. B. and Taqqu, M. S. 1979. Robust R/S analysis of long-run serial correlation. In Proceedings of the 42nd Session of the International Statistical Institute, volume 48 of Bulletin of the International Statistical Institute, 69–104.

[28] Mandelbrot, B. B. and Van Ness, J. W. 1968. Fractional Brownian motions, fractional noises and applications. SIAM Review 10(4), 422-437.

[29] Mandelbrot, B. B. and Wallis, J. R. 1968. Noah, Joseph, and operational hydrology. Water Resources Research 4, 909-918.

[30] Mandelbrot, B. B. and Wallis, J. R. 1969. Global dependence in geophysical records. Water Resources Research 5, 321-340.

[31] McLeod, A.I. and Hipel,K.W. 1978. Preservation of the Rescaled Adjusted Range 1 A Reassessment of the Hurst Phenomenon, Water Resources Research 14(3), 491-508.

[32] Peng, C. K., Buldyrev, S. V., Simons, M., Stanley, H. E., and Goldberger, A. L. 1994. Mosaic organization of DNA nucleotides. Physical Review E 49, 1685–1689.

[33] Percival, D. B. and Walden, A. T. 2000. Wavelet methods for time series analysis, Cambridge University Press, New York, USA.

[34] Politis, D, Romano, J. A 1992. Circular block resampling procedure for stationary data. Exploring the Limits of Bootstrap, Wiley-New York, 263-270.

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

[36] Singh, K. 1981 On the Asymptotic Accuracy of Efron's Bootstrap, Annals of Statistics 9(6), 1187-1195.

[37] Taqqu, M.T., Teverovsky, V. and Willinger, W. (1995) Estimators for long-range dependence: an empirical study. Fractals 3(4), 785-798.

[38] Veitch, D. and Abry, P. 1999. A Wavelet-Based Joint Estimator of the Parameters of Long-Range Dependence. IEEE Transactions on Information Theory 45(3), 878–897.