



ELSEVIER

Physica A 312 (2002) 285–299

PHYSICA A

www.elsevier.com/locate/physa

Estimating long-range dependence: finite sample properties and confidence intervals

Rafał Weron *

*Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology,
50-370 Wrocław, Poland*

Received 31 January 2002

Abstract

A major issue in financial economics is the behavior of asset returns over long horizons. Various estimators of long-range dependence have been proposed. Even though some have known asymptotic properties, it is important to test their accuracy by using simulated series of different lengths. We test R/S analysis, Detrended Fluctuation Analysis and periodogram regression methods on samples drawn from Gaussian white noise. The DFA statistics turns out to be the unanimous winner. Unfortunately, no asymptotic distribution theory has been derived for this statistics so far. We were able, however, to construct empirical (i.e. approximate) confidence intervals for all three methods. The obtained values differ largely from heuristic values proposed by some authors for the R/S statistics and are very close to asymptotic values for the periodogram regression method. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Long-range dependence; Hurst exponent; R/S analysis; Detrended Fluctuation Analysis; Periodogram regression; Confidence interval

1. Introduction

Time series with long-range dependence are widespread in nature and for many years have been extensively studied in hydrology and geophysics [1–4]. More recently, “long memory” or “ $1/f$ noise” (as it is often called) has been observed in DNA sequences [5,6], cardiac dynamics [7,8], internet traffic [9], meteorology [10,11], geology [12] and even ethology [13].

In economics and finance, long-range dependence also has a long history (for a review see Refs. [14,15]) and still is a hot topic of active research [16–24]. Historical

* Corresponding author. Fax: +48-71-320-26-54.

E-mail address: rweron@im.pwr.wroc.pl (R. Weron).

records of economic and financial data typically exhibit nonperiodic cyclical patterns that are indicative of the presence of significant long memory. However, the statistical investigations that have been performed to test long-range dependence have often become a source of major controversies, especially in the case of stock returns. The reason for this are the implications that the presence of long memory has on many of the paradigms used in modern financial economics [19,25].

Various estimators of long-range dependence have been proposed. Even though some have known asymptotic properties, it is important to test their accuracy by using simulated series of different lengths. Such a study was presented in Refs. [26,27] using “ideal” models that display long-range dependence, i.e. fractional Brownian noise (fBn) and Fractional ARIMA(0,d,0). However, the authors tested estimators on rather long time series (10000 elements), whereas in practice we often have to perform analysis of much shorter data sets. For example, Lux [16] used daily time series comprising 1949 DAX returns, Lobato and Savin [18] originally performed tests on 8178 S&P500 daily returns, but later due to the non-stationarity of the process (which influenced the estimates) divided the sample into smaller subsamples, Grau-Carles [20] analyzed various stock index data including 4125 Nikkei and 1555 FTSE daily returns, Weron and Przybyłowicz [22] estimated the Hurst exponent using only 670 daily returns of the spot electricity price in California. Moreover, Taqqu et al. [26] based their statistical conclusions on only 50 simulated trajectories of fBn or FARIMA. This may be enough for the estimation of the mean or median, but certainly not enough for the estimation of very high or low quantiles, which can be used to construct empirical confidence intervals [28].

In this paper, we analyze rescaled range analysis, Detrended Fluctuation Analysis and periodogram regression methods on samples drawn from Gaussian white noise. In Section 2 we precisely describe all three methods and later, in Section 3, present a comparison based on Monte Carlo simulations. The DFA statistics turns out to be the unanimous winner. Unfortunately, no asymptotic distribution theory has been derived for this statistics so far. We were able, however, to construct empirical (i.e. approximate) confidence intervals for all three methods. These results are presented in Section 4. The obtained values differ largely from heuristic values proposed by some authors for the R/S statistics and are very close to asymptotic values for the periodogram regression method. In Section 5 we apply the results of Section 4 to a number of financial data illustrating their usefulness.

2. Methods for estimating H

2.1. R/S analysis

We began our investigation with one of the oldest and best-known methods, the so-called R/S analysis. This method, proposed by Mandelbrot and Wallis [29] and based on previous hydrological analysis of Hurst [1], allows the calculation of the self-similarity parameter H , which measures the intensity of long-range dependence in a time series.

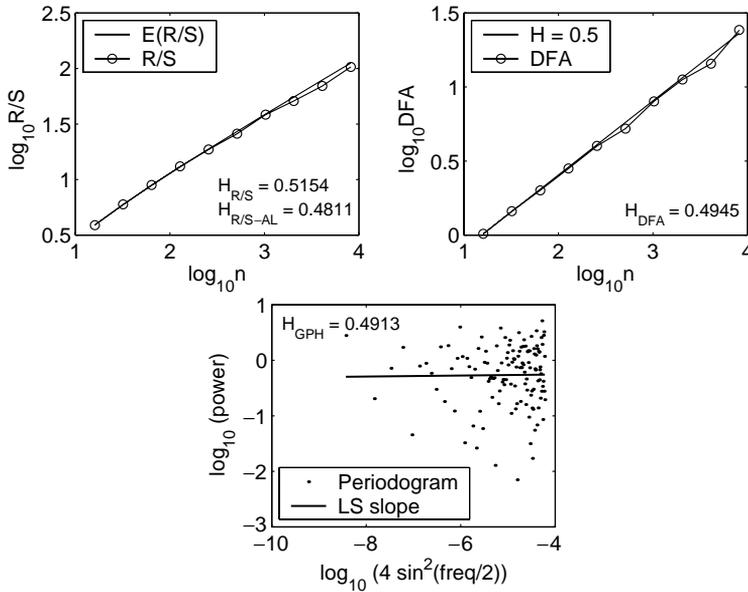


Fig. 1. Estimation of the Hurst exponent for a white noise ($H = 0.5$) sequence of $2^{14} = 16384$ observations via the Hurst R/S analysis (top left panel), the Detrended Fluctuation Analysis (top right panel) and the periodogram Geweke–Porter–Hudak method (bottom panel).

The analysis begins with dividing a time series (of returns) of length L into d subseries of length n . Next, for each subseries $m = 1, \dots, d$: (1) find the mean (E_m) and standard deviation (S_m); (2) normalize the data ($Z_{i,m}$) by subtracting the sample mean $X_{i,m} = Z_{i,m} - E_m$ for $i = 1, \dots, n$; (3) create a cumulative time series $Y_{i,m} = \sum_{j=1}^i X_{j,m}$ for $i = 1, \dots, n$; (4) find the range $R_m = \max\{Y_{1,m}, \dots, Y_{n,m}\} - \min\{Y_{1,m}, \dots, Y_{n,m}\}$; and (5) rescale the range R_m/S_m . Finally, calculate the mean value of the rescaled range for all subseries of length n

$$(R/S)_n = \frac{1}{d} \sum_{m=1}^d R_m/S_m .$$

It can be shown [30] that the R/S statistics asymptotically follows the relation:

$$(R/S)_n \sim cn^H .$$

Thus, the value of H can be obtained by running a simple linear regression over a sample of increasing time horizons

$$\log(R/S)_n = \log c + H \log n .$$

Equivalently, we can plot the $(R/S)_n$ statistics against n on a double-logarithmic paper, see Fig. 1. If the returns process is white noise then the plot is roughly a straight line with slope 0.5. If the process is persistent then the slope is > 0.5 ; if it is anti-persistent

then the slope is < 0.5 . The “significance” level is usually chosen to be one over the square root of sample length, i.e. the standard deviation of a Gaussian white noise [31].

However, it should be noted that for small n there is a significant deviation from the 0.5 slope. For this reason the theoretical (i.e. for white noise) values of the R/S statistics are usually approximated by

$$\mathbf{E}(R/S)_n = \begin{cases} \frac{n-\frac{1}{2}}{n} \frac{\Gamma((n-1)/2)}{\sqrt{\pi}\Gamma(n/2)} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & \text{for } n \leq 340, \\ \frac{n-\frac{1}{2}}{n} \frac{1}{\sqrt{n\pi/2}} \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} & \text{for } n > 340, \end{cases} \tag{1}$$

where Γ is the Euler gamma function. This formula is a slight modification of the formula given by Anis and Lloyd [32]; the $(n - \frac{1}{2})/n$ term was added by Peters [31] to improve the performance for very small n .

Formula (1) was used as a benchmark in all empirical studies of the R/S statistics presented in this paper, i.e. the Hurst exponent H was calculated as 0.5 plus the slope of $(R/S)_n - \mathbf{E}(R/S)_n$. The resulting statistics was denoted by R/S -AL.

A major drawback of the R/S analysis is the fact that no asymptotic distribution theory has been derived for the Hurst parameter H . The only known results are for the rescaled (but not by standard deviation) range R_m itself [25].

2.2. Detrended Fluctuation Analysis

The second method we used to measure long-range dependence was the Detrended Fluctuation Analysis (DFA) proposed by Peng et al. [6]. The method can be summarized as follows. Divide a time series (of returns) of length L into d subseries of length n . Next for each subseries $m = 1, \dots, d$: (1) create a cumulative time series $Y_{i,m} = \sum_{j=1}^i X_{j,m}$ for $i = 1, \dots, n$; (2) fit a least-squares line $\tilde{Y}_m(x) = a_m x + b_m$ to $\{Y_{1,m}, \dots, Y_{n,m}\}$; and (3) calculate the root mean square fluctuation (i.e. standard deviation) of the integrated and detrended time series

$$F(m) = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{i,m} - a_m i - b_m)^2}.$$

Finally, calculate the mean value of the root mean square fluctuation for all subseries of length n

$$\bar{F}(n) = \frac{1}{d} \sum_{m=1}^d F(m).$$

Like in the case of R/S analysis, a linear relationship on a double-logarithmic paper of $\bar{F}(n)$ against the interval size n indicates the presence of a power-law scaling of the form cn^H [6,26]. If the returns process is white noise then the slope is roughly 0.5. If

the process is persistent then the slope is > 0.5 ; if it is anti-persistent then the slope is < 0.5 .

Unfortunately, no asymptotic distribution theory has been derived for the DFA statistics so far. Hence, no explicit hypothesis testing can be performed and the significance relies on subjective assessment.

2.3. Periodogram regression

The third method is a semi-parametric procedure to obtain an estimate of the fractional differencing parameter d . This technique, proposed by Geweke and Porter-Hudak [33] (GPH), is based on observations of the slope of the spectral density function of a fractionally integrated series around the angular frequency $\omega=0$. Since they showed that the spectral density function of a general fractionally integrated model (e.g. FARIMA) with differencing parameter d is identical to that of a fractional Gaussian noise with Hurst exponent $H = d + 0.5$, the GPH method can be used to estimate H .

The estimation procedure begins with calculating the periodogram, which is a sample analogue of the spectral density. For a vector of observations $\{x_1, \dots, x_L\}$ the periodogram is defined as

$$I_L(\omega_k) = \frac{1}{L} \left| \sum_{t=1}^L x_t e^{-2\pi i(t-1)\omega_k} \right|^2,$$

where $\omega_k = k/L$, $k = 1, \dots, [L/2]$ and $[x]$ denotes the largest integer $\leq x$. Observe that I_L is the squared absolute value of the Fourier transform and if the observations vector is of appropriate length (even or a power of 2) then we can use fast algorithms to calculate the Fourier transform.

The next and final step is to run a simple linear regression

$$\log\{I_L(\omega_k)\} = a - \hat{d} \log\{4 \sin^2(\omega_k/2)\} + \varepsilon_k, \tag{2}$$

at low Fourier frequencies ω_k , $k = 1, \dots, K \leq [L/2]$. The least squares estimate of the slope yields the differencing parameter d through the relation $d = \hat{d}$, hence $H = \hat{d} + 0.5$. A major issue on the application of this method is the choice of K . Geweke and Porter-Hudak [33], as well as a number of other authors, recommend choosing K such that $K = [L^{0.5}]$, however, other values (e.g. $K = [L^{0.45}]$, $[L^{0.2}] \leq K \leq [L^{0.5}]$) have also been suggested.

Periodogram regression is the only of the presented methods, which has known asymptotic properties. Inference is based on the asymptotic distribution of the estimate \hat{d} of d , which is normally distributed with mean d and variance

$$\frac{\pi^2}{6 \sum_{k=1}^K (x_t - \bar{x})^2}, \tag{3}$$

where $x_t = \log\{4 \sin^2(\omega_k/2)\}$ is the regressor in Eq. (2).

3. Comparison of estimators

In order to test the presented estimation methods we performed Monte Carlo simulations. We generated samples of Gaussian white noise sequences (independent and $N(0, 1)$ distributed) of length $L=2^N$, where $N=8, 9, \dots, 16$, i.e. $L=256, 512, \dots, 65\,536$. For each L 10 000 trajectories were produced. Next, we applied all three estimation procedures—Anis–Lloyd corrected R/S statistics, Detrended Fluctuation Analysis and periodogram regression—to the data series and compared the results.

In Fig. 2 we plotted the mean values (over all 10 000 samples) of the estimated Hurst exponents. For illustrative purposes we also included the results of the classical R/S statistics (R/S), i.e. without the Anis–Lloyd correction. As can be seen in Fig. 2, on average, this method overestimates the true Hurst exponent to a great extent. On the other hand, all other methods have a slight negative bias.

We analyzed the R/S and DFA methods for two cutoffs of the subinterval length: $n > 10$ and $n > 50$. Despite the corrections to the original rescaled range statistics, R/S -AL still possesses a large variance if very small n are included in the calculations. Thus, we decided to use only subintervals of length $n > 50$ (to be more precise: $n = 64, 128, 256, \dots$, since the length of our samples was a power of 2). On the other hand, the DFA statistics behaves nicely even for small subintervals. So, we calculated H_{DFA} for $n > 10$ (more precisely: $n = 16, 32, 64, \dots$) and—to be consistent with the results for the rescaled range analysis—separately for $n > 50$.

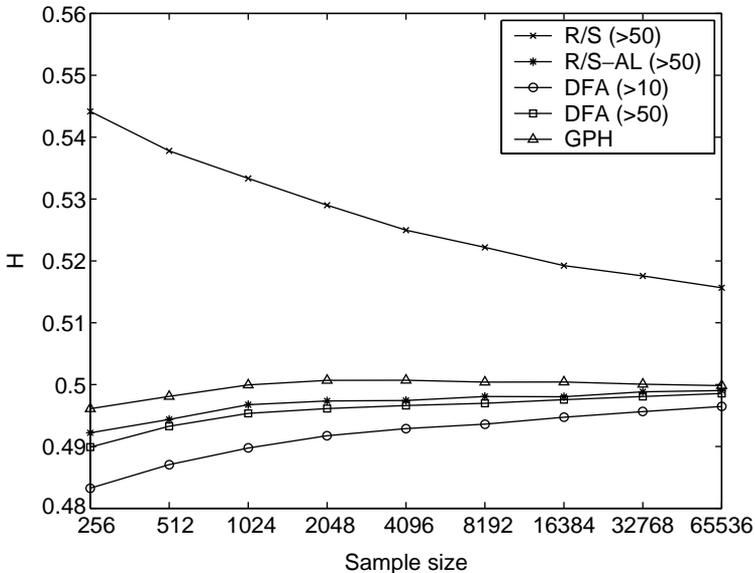


Fig. 2. Mean values (over 10 000 samples) of the estimated Hurst exponents for Gaussian white noise sequences of length $L = 256, 512, \dots, 65\,536$. The methods used are the classical R/S analysis (R/S) for subintervals of length $n > 50$, the Anis–Lloyd corrected R/S statistics (R/S -AL) for $n > 50$, the Detrended Fluctuation Analysis for $n > 10$ and $n > 50$, and the periodogram Geweke–Porter–Hudak method for cutoff $K = \lfloor L^{0.5} \rfloor$.

Table 1

Estimation results for $H = 0.5$ using 10 000 independent realizations of different length. The best results in each category are in bold

Sample size	Method			
	R/S-AL ($n > 50$)	DFA ($n > 10$)	DFA ($n > 50$)	GPH
<i>Standard deviation</i>				
256	0.1739	0.0972	0.1222	0.2205
512	0.1043	0.0724	0.0716	0.1705
1024	0.0739	0.0559	0.0497	0.1401
2048	0.0535	0.0436	0.0355	0.1085
4096	0.0422	0.0359	0.0278	0.0898
8192	0.0332	0.0293	0.0217	0.0727
16 384	0.0277	0.0254	0.0181	0.0622
32 768	0.0237	0.0223	0.0154	0.0511
65 536	0.0193	0.0184	0.0125	0.0424
<i>Mean absolute error</i>				
256	0.1396	0.0800	0.1003	0.1733
512	0.0838	0.0589	0.0582	0.1357
1024	0.0594	0.0455	0.0401	0.1108
2048	0.0428	0.0355	0.0287	0.0861
4096	0.0337	0.0291	0.0224	0.0711
8192	0.0266	0.0240	0.0176	0.0580
16 384	0.0224	0.0209	0.0147	0.0494
32 768	0.0189	0.0181	0.0124	0.0402
65 536	0.0155	0.0150	0.0101	0.0335

As we have mentioned in the previous section, results of the Geweke–Porter–Hudak method depend on the choice of the cutoff value K , which determines how many of the low Fourier frequencies ω_k are taken into account. In our simulations we decided to use the standard value, i.e. $K = \lfloor L^{0.5} \rfloor$, because lower powers of L introduce larger estimation errors, while larger powers force us to move away from the region for which the theoretical results hold.

The large number of simulated trajectories allowed us to compare the methods. For a given method we obtained 10 000 estimated values of H , called $\{H_i, i=1, 2, \dots, 10\,000\}$. Apart from calculating their mean (see Fig. 2), we computed their standard deviation and mean absolute error, i.e.

$$MAE = \frac{1}{10\,000} \sum_{i=1}^{10\,000} |H_i - 0.5|,$$

which provides some information on the bias. The results are presented in Table 1. To have a complete picture, in Fig. 3 we also plotted the 2.5% and 97.5% sample quantiles for all methods. A $p\%$ sample quantile, denoted by \hat{x}_p , is such a value that $p\%$ of the sample observations are $< \hat{x}_p$. Equivalently we can say that the empirical distribution F_e evaluated at \hat{x}_p equals $p\%$, i.e. $F_e(\hat{x}_p) = p\%$.

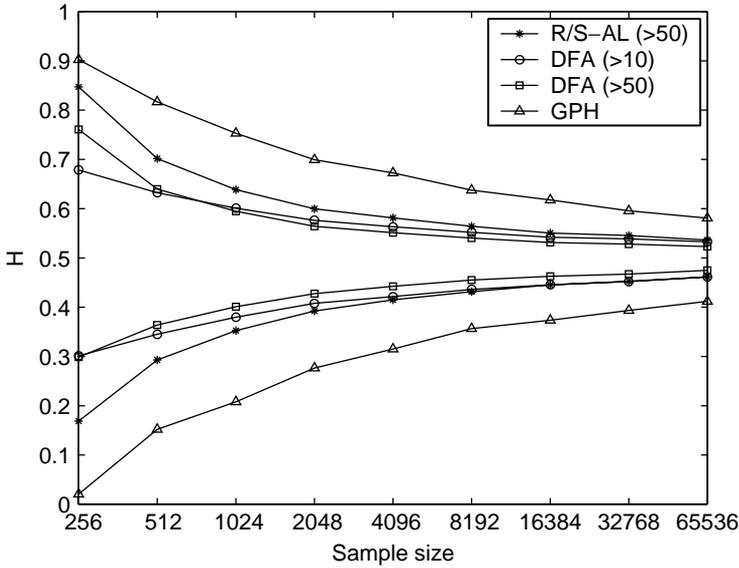


Fig. 3. 2.5% and 97.5% sample quantiles (over 10 000 samples) of the estimated Hurst exponents for Gaussian white noise sequences of length $L = 256, 512, \dots, 65\,536$. The methods used are the Anis–Lloyd corrected R/S statistics (R/S -AL) for $n > 50$, the Detrended Fluctuation Analysis for $n > 10$ and $n > 50$, and the periodogram Geweke–Porter–Hudak method for cutoff $K = \lfloor L^{0.5} \rfloor$.

In all three tests the clear winner for $L > 500$ is the DFA statistics with $n > 50$ as it gives the estimated values closest to the initial Hurst exponent ($H = 0.5$). The reason it performs worse than the DFA statistics with $n > 10$ for the smallest tested samples is probably the fact that $L = 256$ has only three divisors > 50 : $n = 64, 128, 256$ and that regression based on only three points can yield large errors.

Unfortunately, no asymptotic distribution theory has been derived for the DFA statistics so far. However, using Monte Carlo simulations we were able to construct empirical (i.e. approximate) confidence intervals for all three analyzed methods.

4. Construction of confidence intervals

The estimation of the Hurst exponent H alone is not enough. We also need a measure of the significance of the results. Traditionally, the statistical approach is to test the null hypothesis of no or weak dependence vs. the alternative of strong dependence or long memory at some given significance level. However, to construct a test the asymptotic distribution of the test statistics must be known. Of the three analyzed methods only the spectral one has well know asymptotic properties. In this section we will thus construct empirical confidence intervals.

The procedure is quite simple, although burdensome, and consists of the following: (1) for a set of sample lengths (in our case: $L = 256, 512, \dots, 65\,536$) generate a large

Table 2
Empirical confidence intervals for the Anis–Lloyd corrected R/S statistics and sample length $L = 2^N$

Level	Confidence intervals for $n > 50$	
	Lower bound	Upper bound
90%	$0.5 - \exp(-7.35 \log(\log N) + 4.06)$	$\exp(-7.07 \log(\log N) + 3.75) + 0.5$
95%	$0.5 - \exp(-7.33 \log(\log N) + 4.21)$	$\exp(-7.20 \log(\log N) + 4.04) + 0.5$
99%	$0.5 - \exp(-7.19 \log(\log N) + 4.34)$	$\exp(-7.51 \log(\log N) + 4.58) + 0.5$

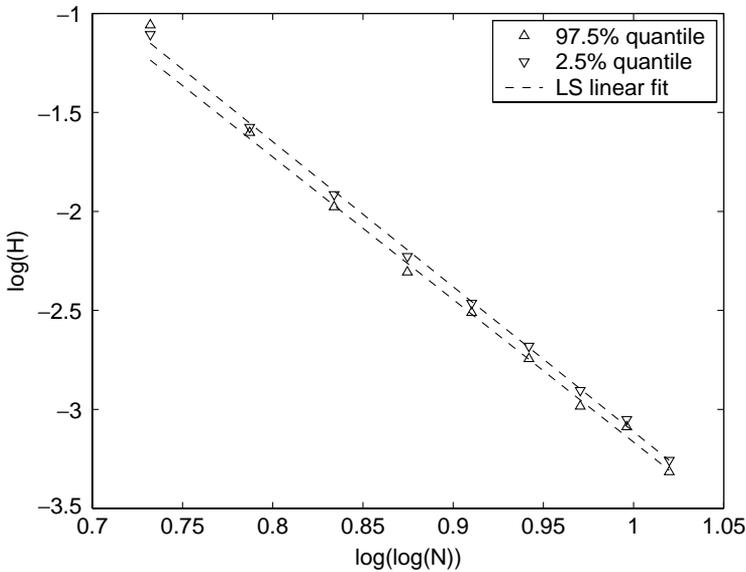


Fig. 4. Fitting R/S -AL quantiles ($n > 50$): a plot of $\log(\hat{x}_{97.5} - 0.5)$ and $\log(0.5 - \hat{x}_{2.5})$ vs. $\log(\log N)$, where $N = \log_2 L$. For the smallest analyzed samples ($L = 256$) the 97.5% quantile introduced large estimation errors and was not used in the analysis.

number (here: 10 000) of realizations of an independent or a weakly dependent time series (here: Gaussian white noise); (2) compute the lower (0.5%, 2.5%, 5%) and upper (95%, 97.5%, 99.5%) sample quantiles for all sample lengths; and (3) plot the sample quantiles vs. sample size and fit them with some functions. These functions can be later used to construct confidence intervals. The 5% and 95% quantiles designate the 90% (two-sided) confidence interval, 2.5% and 97.5% quantiles—the 95% confidence interval and 0.5% and 99.5% quantiles—the 99% confidence interval.

Results of the above procedure for the Anis–Lloyd corrected R/S statistics are presented in Table 2 and Figs. 4 and 5. To find the 95% confidence interval we plotted 2.5% and 97.5% sample quantiles vs. sample size. The only satisfactory results were obtained for $\log(\hat{x}_{97.5} - 0.5)$ and $\log(0.5 - \hat{x}_{2.5})$ vs. $\log(\log N)$, where $N = \log_2 L$, see

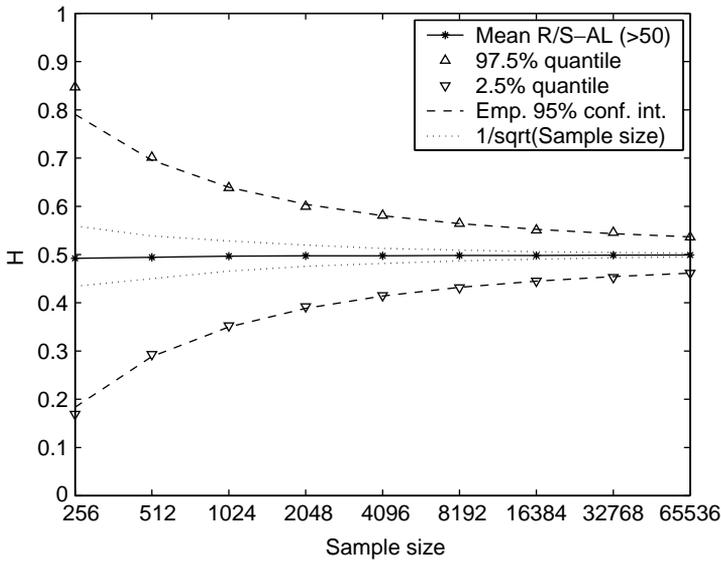


Fig. 5. Mean value of the Anis-Lloyd corrected R/S statistics, 95% empirical confidence intervals and $1/\sqrt{L}$ intervals for samples of size $L = 256, 512, \dots, 65536$.

Fig. 4. For sample size $L = 256$ the 97.5% quantile introduced large estimation errors, probably due to the fact that $L = 256$ has only three divisors > 50 : $n = 64, 128, 256$ and that regression based on only three points can yield large errors. Thus, we decided to use 97.5% quantiles (in fact 95% and 99.5% quantiles as well) coming only from samples of at least 512 observations. The obtained fit was very good, the R^2 statistics for the lower quantile was 0.9987 and for the upper 0.9972. The complete formulas, also for 90% and 99% confidence intervals are given in Table 2. In Fig. 5 we plotted the mean value of H and the 95% confidence interval vs. sample size. For comparison we also added the heuristic “significance level” (one over the square root of sample length) given in Peters [31]. It is clearly seen that this “significance level” rejects the null hypothesis too often compared to the 95% confidence interval (in fact too often even compared to the 90% confidence interval).

Results for the DFA statistics are presented in Table 3 and Figs. 6 and 7. Again in order to find the 95% confidence interval we plotted 2.5% and 97.5% sample quantiles vs. sample size. The only satisfactory results were obtained for $\log(\hat{x}_{97.5} - 0.5)$ and $\log(0.5 - \hat{x}_{2.5})$ against $\log N$, where $N = \log_2 L$. Like in the case of the R/S -AL statistics, for sample size $L = 256$ the 97.5% quantile introduced large estimation errors when the subinterval size was restricted to $n > 50$. Thus, we decided to use 97.5% quantiles (95% and 99.5% quantiles as well) coming only from samples of at least 512 observations when the higher ($n > 50$) cutoff was applied. The obtained fit was very good. In the $n > 10$ case the R^2 statistics for the lower quantile was 0.9985 and for the upper 0.9972. In the $n > 50$ case the R^2 statistics for the lower quantile was 0.9973 and for

Table 3
Empirical confidence intervals for the DFA statistics and sample length $L = 2^N$

Level	Confidence intervals for $n > 10$	
	Lower bound	Upper bound
90%	$0.5 - \exp(-2.33 \log N + 3.09)$	$\exp(-2.44 \log N + 3.13) + 0.5$
95%	$0.5 - \exp(-2.33 \log N + 3.25)$	$\exp(-2.46 \log N + 3.38) + 0.5$
99%	$0.5 - \exp(-2.20 \log N + 3.18)$	$\exp(-2.45 \log N + 3.62) + 0.5$
	<i>Confidence intervals for $n > 50$</i>	
90%	$0.5 - \exp(-2.99 \log N + 4.45)$	$\exp(-3.09 \log N + 4.57) + 0.5$
95%	$0.5 - \exp(-2.93 \log N + 4.45)$	$\exp(-3.10 \log N + 4.77) + 0.5$
99%	$0.5 - \exp(-2.67 \log N + 4.06)$	$\exp(-3.19 \log N + 5.28) + 0.5$

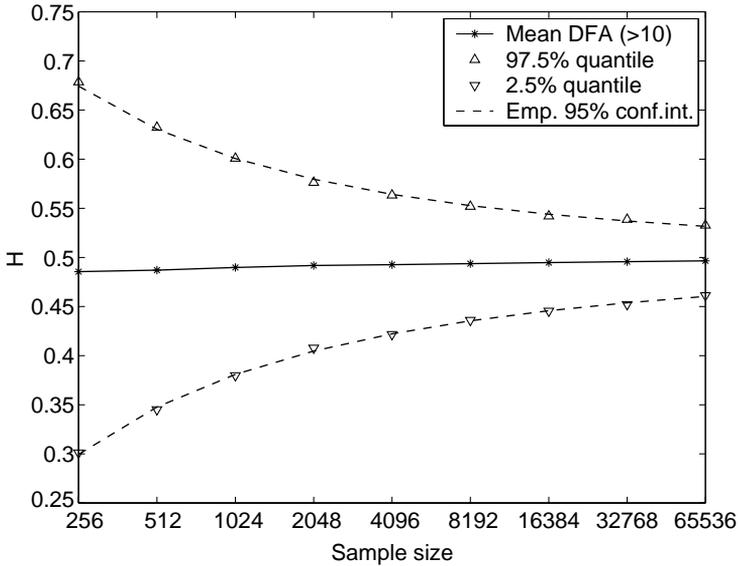


Fig. 6. Mean value of the DFA statistics ($n > 10$) and 95% confidence intervals for samples of size $L = 256, 512, \dots, 65536$.

the upper 0.9918. The complete formulas, also for 90% and 99% confidence intervals are given in Table 3. In Figs. 6 and 7 we plotted the mean value of H and the 95% confidence interval vs. sample size.

Finally, results for the GPH statistics are presented in Table 4 and Fig. 8. The only satisfactory results for the 95% confidence interval were obtained by plotting $\log(\hat{x}_{97.5} - 0.5)$ and $\log(0.5 - \hat{x}_{2.5})$ against $N^{2/3}$, where $N = \log_2 L$. The power $\frac{2}{3}$ was chosen arbitrarily, however, comparably good results were possible in the range (0.6, 0.7). The obtained fit was very good, the R^2 statistics for the lower quantile was 0.9953 and for

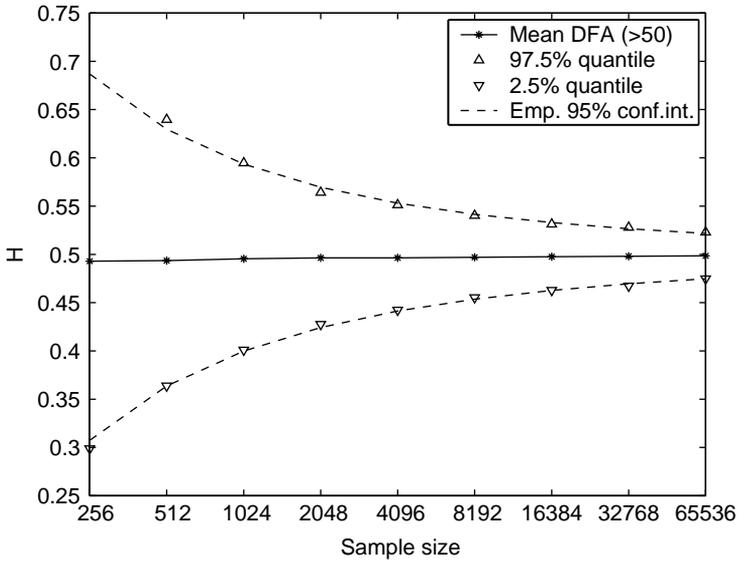


Fig. 7. Mean value of the DFA statistics ($n > 50$) and 95% confidence intervals for samples of size $L = 256, 512, \dots, 65536$.

Table 4
Empirical confidence intervals for the GPH statistics and sample length $L = 2^N$

Level	Confidence intervals for $K = \lfloor L^{0.5} \rfloor$	
	Lower bound	Upper bound
90%	$0.5 - \exp(-0.71 \cdot N^{2/3} + 1.87)$	$\exp(-0.68 \cdot N^{2/3} + 1.62) + 0.5$
95%	$0.5 - \exp(-0.71 \cdot N^{2/3} + 2.04)$	$\exp(-0.68 \cdot N^{2/3} + 1.78) + 0.5$
99%	$0.5 - \exp(-0.73 \cdot N^{2/3} + 2.45)$	$\exp(-0.65 \cdot N^{2/3} + 1.92) + 0.5$

the upper 0.9987. The complete formulas, also for 90% and 99% confidence intervals are given in Table 4. In Fig. 8 we plotted the mean value of H and the 95% confidence interval vs. sample size. For comparison we also added the theoretical 95% confidence intervals, see formula (3). The empirical and theoretical intervals are quite close to each other, especially for large samples. For small samples the GPH statistics has a slight negative bias and the empirical values are shifted downward. Recall, however, that the theoretical results are derived from the limiting behavior and obviously cannot take into account finite sample properties.

The small difference between the empirical and theoretical confidence intervals justifies our approach and permits us to use in practical applications the empirical confidence intervals given in Tables 2–4.

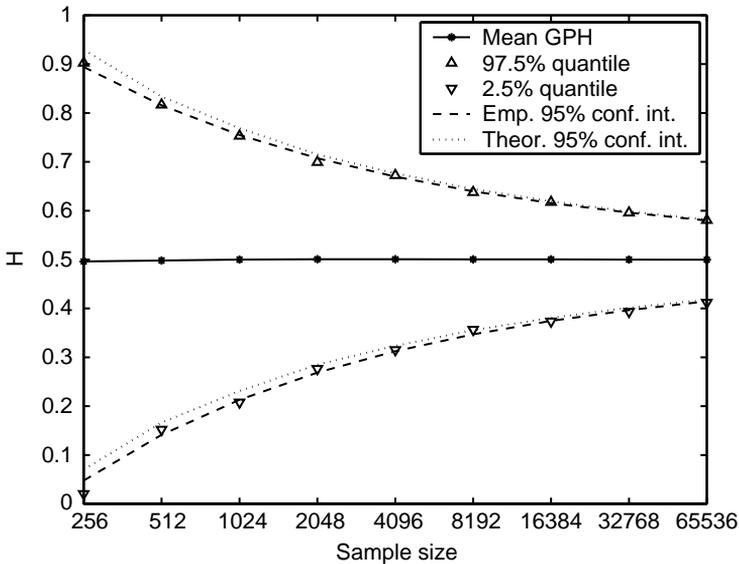


Fig. 8. Mean value of the GPH statistics and 95% confidence intervals for samples of size $L = 256, 512, \dots, 65536$. For comparison the theoretical quantiles are also plotted. The small difference between the empirical and theoretical confidence intervals justifies our approach.

5. Applications

Having calculated the empirical confidence intervals we are now ready to test for long-range dependence in financial time series. For the analysis we selected four data sets (two stock indices and two power market benchmarks):

- 2526 daily returns of the Dow Jones industrial average (DJIA) index for the period 2 January 1990–30 December 1999;
- 1560 daily returns of the WIG20 Warsaw stock exchange index (based on 20 blue chip stocks from the Polish capital market) for the period 2 January 1995–30 March 2001;
- 728 daily returns of electricity traded in the California Power Exchange (CalPX) spot market for the period 1 April 1998–29 March 2000;
- 690 daily returns of firm on-peak power (OTC spot market) in the Entergy region (Louisiana, Arkansas, Mississippi and East Texas) for the period 2 January 1998–25 September 2000.

The results of the Anis–Lloyd corrected R/S analysis (for $n > 50$), the Detrended Fluctuation Analysis (for $n > 50$ only) and the periodogram Geweke–Porter–Hudak method (for $K = \lfloor L^{0.5} \rfloor$) for these time series are summarized in Table 5. The significance of the results is based on values obtained from Tables 2–4. For example, the 90%

Table 5

Estimates of the Hurst exponent H for financial data. *, ** and *** denote significance at the (two-sided) 90%, 95% and 99% level, respectively

Date	Method		
	R/S-AL	DFA	GPH
<i>Stock indices</i>			
DJIA returns	0.4585	0.4195**	0.3560
DJIA absolute value of returns	0.7838***	0.9080***	0.8357***
WIG20 returns	0.5030	0.4981	0.4604
WIG20 absolute value of returns	0.9103***	0.9494***	0.8262***
<i>Electricity price benchmarks</i>			
CalPX returns	0.3473*	0.2633***	0.0667***
Entergy returns	0.2995**	0.3651**	0.0218***

Table 6

Confidence intervals for 2526 DJIA returns

Level	Method			
	R/S-AL	DFA	GPH	GPH (theoretical)
90%	(0.4138, 0.5810)	(0.4392, 0.5538)	(0.3184, 0.6645)	(0.3305, 0.6695)
95%	(0.3980, 0.5965)	(0.4297, 0.5641)	(0.2847, 0.6931)	(0.2981, 0.7019)
99%	(0.3686, 0.6258)	(0.4106, 0.5858)	(0.2067, 0.7583)	(0.2346, 0.7654)

confidence intervals of the R/S-AL statistics for DJIA returns ($N = \log_2 2526 = 11.3026$) were calculated as follows:

$$\text{lower bound : } 0.5 - \exp(-7.35 \log(\log N) + 4.06) = 0.4138,$$

$$\text{upper bound : } \exp(-7.07 \log(\log N) + 3.75) + 0.5 = 0.5810.$$

Similarly, we obtained confidence intervals for the 95% and 99% two-sided levels and other methods. All obtained values are presented in Table 6. For comparison we included the theoretical confidence intervals for the periodogram Geweke–Porter–Hudak method. The differences between the empirical (obtained from Table 4) and theoretical confidence intervals are small and the significance of the results is the same for both sets of values.

As expected [16,18–20] we found (almost) no evidence for long-range dependence in the stock indices returns and strong—i.e. significant at the two-sided 99% level for all three methods—dependence in the stock indices volatility (more precisely: in absolute value of stock indices returns). On the other hand, electricity price returns were found to exhibit a mean-reverting mechanism (the Hurst exponent was found significantly smaller than 0.5), which is consistent with our earlier findings [22].

Acknowledgements

Many thanks to Thomas Lux for stimulating discussions. Partial financial support for this work was provided by KBN Grant no. PBZ 16/P03/99.

References

- [1] H.E. Hurst, *Trans. Am. Soc. Civ. Eng.* 116 (1951) 770.
- [2] B.B. Mandelbrot, J.R. Wallis, *Water Resour. Res.* 4 (1968) 909.
- [3] B.B. Mandelbrot, J.R. Wallis, *Water Resour. Res.* 5 (1969) 321.
- [4] B.B. Mandelbrot, *Fractal Geometry of Nature*, W.H. Freeman, San Francisco, 1982.
- [5] C.-K. Peng, S.V. Buldyrev, A.L. Goldberger, S. Havlin, F. Sciortino, M. Simons, H.E. Stanley, *Nature* 356 (1992) 168.
- [6] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, *Phys. Rev. E* 49 (1994) 1685.
- [7] C.-K. Peng, J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, A.L. Goldberger, *Phys. Rev. Lett.* 70 (1993) 1343.
- [8] T.-H. Makikallio, J. Koistinen, L. Jordaens, M.P. Tulppo, N. Wood, B. Golosarsky, C.-K. Peng, A.L. Goldberger, H.V. Huikuri, *Am. J. Cardiol.* 83 (1999) 880.
- [9] W. Willinger, M.S. Taqqu, A. Erramilli, in: F.P. Kelly, S. Zachary, I. Ziedins (Eds.), *Stochastic Networks: Theory and Applications*, Clarendon Press, Oxford, 1996, pp. 339–366.
- [10] E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, H.-J. Schellnhuber, *Phys. Rev. Lett.* 81 (1998) 729.
- [11] A. Montanari, R. Rosso, M.S. Taqqu, *Water Resour. Res.* 36 (2000) 1249.
- [12] B.D. Malamud, D.L. Turcotte, *J. Stat. Plann. Inference* 80 (1999) 173.
- [13] C.L. Alados, M.A. Huffman, *Ethology* 106 (2000) 105.
- [14] R.T. Baillie, M.L. King (Eds.), *Fractional differencing and long memory processes (special issue)*, *J. Econometrics* 73, 1996.
- [15] B.B. Mandelbrot, *Fractals and Scaling in Finance*, Springer, Berlin, 1997.
- [16] T. Lux, *Appl. Econ. Lett.* 3 (1996) 701.
- [17] A. Weron, R. Weron, *Financial Engineering: Derivatives Pricing, Computer Simulations, Market Statistics*, WNT, Warsaw, 1998 (in Polish).
- [18] I.N. Lobato, N.E. Savin, *J. Bus. Econ. Stat.* 16 (1998) 261.
- [19] W. Willinger, M.S. Taqqu, V. Teverovsky, *Finance Stochastics* 3 (1999) 1.
- [20] P. Grau-Carles, *Physica A* 287 (2000) 396.
- [21] I.N. Lobato, C. Velasco, *J. Bus. Econ. Stat.* 18 (2000) 410.
- [22] R. Weron, B. Przybyłowicz, *Physica A* 283 (2000) 462.
- [23] J.B. Cromwell, W.C. Labys, E. Kouassi, *Empirical Econ.* 25 (2000) 563.
- [24] B.B. Mandelbrot, *Quant. Finance* 1 (2001) 113.
- [25] A.W. Lo, *Econometrica* 59 (1991) 1279.
- [26] M.S. Taqqu, V. Teverovsky, W. Willinger, *Fractals* 3 (1995) 785.
- [27] M.S. Taqqu, V. Teverovsky, in: R.J. Adler, R.E. Feldman, M.S. Taqqu (Eds.), *A Practical Guide to Heavy Tails*, Birkhäuser, Boston, 1998, pp. 177–217.
- [28] M.A. Stephens, *J. Am. Stat. Assoc.* 69 (1974) 730.
- [29] B.B. Mandelbrot, J.R. Wallis, *Water Resour. Res.* 5 (1969) 967.
- [30] B.B. Mandelbrot, *Zeit. Wahrsch. verwandte Gebiete* 31 (1975) 271.
- [31] E.E. Peters, *Fractal Market Analysis*, Wiley, New York, 1994.
- [32] A.A. Anis, E.H. Lloyd, *Biometrika* 63 (1976) 283.
- [33] J. Geweke, S. Porter-Hudak, *J. Time Ser. Anal.* 4 (1983) 221.