

COMPARATIVE ANALYSIS FOR ESTIMATING OF THE HURST EXPONENT FOR STATIONARY AND NONSTATIONARY TIME SERIES

Ludmila Kirichenko, Tamara Radivilova, Zhanna Deineko

Abstract: Estimating of the Hurst exponent for experimental data plays a very important role in the research of processes which show properties of self-similarity. There are many methods for estimating the Hurst exponent using time series. The aim of this research is to carry out the comparative analysis of the statistical properties of the Hurst exponent estimators obtained by different methods using model stationary and nonstationary fractal time series. In this paper the most commonly used methods for estimating the Hurst exponents are examined. There are: *R/S*-analysis, variance-time analysis, detrended fluctuation analysis (DFA) and wavelet-based estimation. The fractal Brownian motion that is constructed using biorthogonal wavelets have been chosen as a model random process which exhibit fractal properties.

In this paper, the results of a numerical experiment are represented where the fractal Brown motion was modelled for the specified values of the exponent H . The values of the Hurst exponent for the model realizations were varied within the whole interval of possible values $0 < H < 1$. The lengths of the realizations were defined as 500, 1000, 2000 and 4000 values. For the nonstationary case model time series are presented by the sum of fractional noise and the trend component, which are a polynomial in varying degrees, irrational, transcendental and periodic functions. The estimates of H were calculated for each generated time series using the methods mentioned above. Samples of the exponent H estimates were obtained for each value of H and their statistical characteristics were researched.

The results of the analysis have shown that the estimates of the Hurst exponent, which were obtained for the stationary realisations using the considered methods, are biased normal random variables. For each method the bias depends on the true value of the degrees self-similarity of the process and length of time series. Those estimates which are obtained by the DFA method and the wavelet transformation have the minimal bias. Standard deviations of the estimates depending on the estimation method and decrease, while the length of the series increases. Those estimates which are obtained by using the wavelet analysis have the minimal standard deviation.

In the case of a nonstationary time series, represented by a trend and additive fractal noise, more accurate evaluation is obtained using the DFA method. This method allows estimating the Hurst exponent for experimental data with trend components of virtually any kind. The greatest difficulty in estimating, presents a series with a periodic trend component. It is desirable in addition to investigate the spectrum of the wavelet energy, which is demonstrated in the structure of the time series. In the presence of a slight trend, the wavelet-estimation is quite effective.

Keywords: Hurst exponent, estimate of the Hurst exponent, self-similar stochastic process, nonstationary time series, methods for estimating the Hurst exponent.

ACM Classification Keywords: G.3 Probability and statistics - Time series analysis, Stochastic processes, G.1 Numerical analysis, G.1.2 Approximation - Wavelets and fractals.

Introduction

Nowadays problems of nonlinear physics, radio electronics, control theory and image processing require the development and employment of new mathematical models, methods and algorithms for data analysis. At present it has been generally accepted, that many stochastic processes in nature and in engineering exhibit a long-range dependence and fractal structure. The most suitable mathematical method for research of the dynamics and structure of such processes is fractal analysis.

The Stochastic process $X(t)$ is statistically self-similar if the process $a^{-H}X(at)$ shows the same second-order statistical properties as $X(t)$. Long-range dependence means slow (hyperbolic) decay in the time of the autocorrelation function of a process. The parameter H ($0 < H < 1$) is called the Hurst exponent and is a measure of self-similarity or a measure of duration of long-range dependence of a stochastic process.

Let us consider the most well-known examples of the self-similar processes. One of the first real stochastic processes, that have been found with self-similar properties is informational data traffic in telecommunication networks. For self-similar traffic methods for calculating the characteristics of a computer network ([channel capacity](#), buffer size, etc.), that is based on classical models, don't meet the necessary requirements and don't allow to estimate adequately, the [load](#) of the network. There are a large number of publications, which are devoted to the fractal properties analysis and their influence on the functioning and quality of service in the telecommunication network. Another example of fractal stochastic structures is the modern financial market. The modern fractality hypothesis of a financial time series supposes that the market is a self-regulating macro-economic system with feedback which uses information about past events to affect decisions in the present, and contains long-term correlations and trends. The market remains stable, as long as it retains its fractal structure. Analyzing the dynamics of occurrence of time section with various fractal structures, we can diagnose and predict unstable states (crises) of a market. It has become generally accepted in the recent years that a lot of bioelectric signals have fractal structure, so in researches of cerebral and cardio processes are increasingly important roles played by fractal analysis. Distinct changes in fractal characteristics of cardiograms and encephalograms manifest in various diseases, in changes of mental and physical load on the body. Fractal analysis of bioelectrical signals can be the basis for statistical researches, which will allow to formulate a methodology that will be significant for clinical practice.

It is obvious that Hursts' exponent estimation for experimental data plays an important role in the study of processes which exhibit properties of self-similarity. There are many methods for the Hurst exponent evaluation for a time series. Sufficient review of these methods is represented in [Willinger, 1996; Clegg, 2005]. However, most methods of the Hurst exponent estimation is applied only to stationary time series, while a lot of natural, technical and information processes are nonstationary. The main type of nonstationarity, which occurs in practice, is the existence of trend and cyclical components.

Nevertheless, at the present time there is no proper summary research where the results of the Hurst exponent estimation \hat{H} using stationary and nonstationary fractal time series with different methods would be generalized and the comparative analysis of statistical properties of estimations obtained for a small amount of sample data would be. The given research is an attempt to carry out such analysis for the most commonly used methods of estimation of self-similarity.

The aim of this research is to carry out the comparative analysis of the statistical properties of the Hurst exponent estimates obtained by different methods, using a short length model fractal time series (the number of values being less than 4000). In this paper the most commonly used methods for estimating the Hurst exponent are researched. There are: *R/S*-analysis (rescaled range method) (see, for instance [Feder, 1988; Peters, 1996; Stollings, 2003; Sheluhin, 2007]), variation in time of variance of an aggregate time series (variance-time analysis), see [Stollings, 2003; Sheluhin, 2007], detrended fluctuation analysis (see [Kantelhardt, 2001; Chen, 2002; Gu, 2006; Kantelhardt, 2008]) and estimation using the wavelet analysis (see [Mallat, 1998; Abry, 1998; Abry, 2003]) The fractal Brownian motion has been chosen as a model random process which exhibit fractal properties.

Methods of estimating the Hurst exponent

Rescaled range method. This empirical method suggested by G. Hurst is still one of the most popular methods of research of fractal series of different nature. According to this method for the time series $x(t)$ of the length τ the rate $\frac{R(\tau)}{S(\tau)}$ is defined, where $R(\tau)$ is the range of the cumulative deviate series $x^{cum}(t, \tau)$, $S(\tau)$ is standard deviation of the initial series:

$$R/S = \frac{\max(x^{cum}(t, \tau)) - \min(x^{cum}(t, \tau))}{\sqrt{\frac{1}{\tau-1} \sum_{t=1}^{\tau} (x(t) - \bar{x})^2}}, \quad t = \overline{1, \tau}, \quad (1)$$

where $\bar{x}(\tau) = \frac{1}{\tau} \sum_{t=1}^{\tau} x(t)$, $x^{cum}(t, \tau) = \sum_{i=1}^t x(i) - \bar{x}(\tau)$.

For a self-similar process and big values of τ this ratio has the following characteristics:

$$M \left[\frac{R}{S} \right] \sim (c \cdot \tau)^H, \quad (2)$$

where c is a constant.

The log-log diagram dependence of $\frac{R(\tau)}{S(\tau)}$ on τ represents a line approximated by the least square method.

Then the estimate of the exponent H is a tangent of the angle of slope of the line which represents the dependence $\log \frac{R(\tau)}{S(\tau)}$ on $\log(\tau)$ (see Fig. 1, where theoretical values of $H = 0.8$).

Variance-time analysis is most often used to processes researches in telecommunication networks. When someone mentions the aggregation of a time series $x(t)$ of the length τ on the time scale with the parameter m , they mean the transition to the process $x^{(m)}$, where $x_k^{(m)} = \frac{1}{m} \sum_{t=km-m+1}^{km} x(t)$, $k = \overline{1, \tau/m}$. For self-similar process the variance of the aggregated time series $x^{(m)}$ for big values of m follows the formula:

$$\text{Var}(x^{(m)}) \sim \frac{\text{Var}(x)}{m^\beta}. \quad (3)$$

In this case, the parameter of self-similarity $H = 1 - \frac{\beta}{2}$ can be obtained if we generate an aggregated process on different levels of aggregation m and calculate the variance for each level. The dependence diagram of $\log(\text{Var}(x^{(m)}))$ on $\log(m)$ will represent a line with a slope equal to $-\beta$. Fig. 2 shows log-log diagram of $\text{Var}(x^{(m)})$ on m (dependence (3)), where theoretical values of $H = 0.8$.

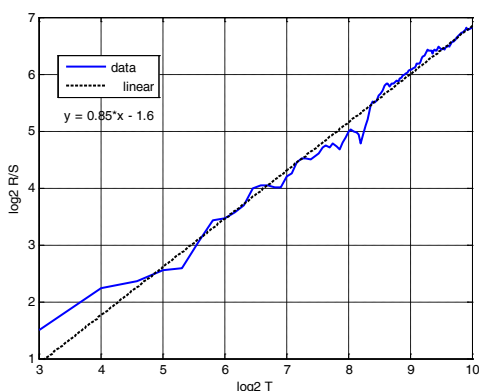


Figure 1. log-log diagram of $\frac{R(\tau)}{S(\tau)}$ on τ

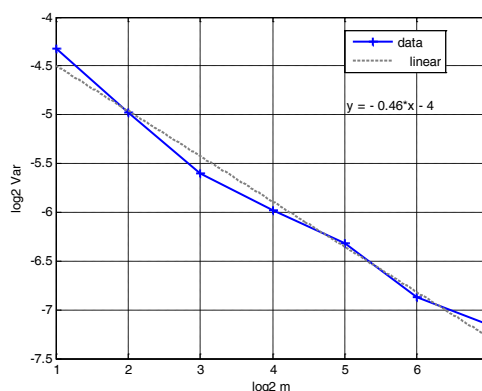


Figure 2. log-log diagram of $\text{Var}(x^{(m)})$ on m

Detrended fluctuation analysis (DFA). DFA originally suggested by [Peng, 1994], is the main method of determining self-similarity for nonstationary time series nowadays. This method is based on the ideology of one-dimensional random walks. Assume that the autocorrelation function (ACF) of the process shows the power dependence for the values of time argument $\tau > \tau_1$.

When τ is large, the ACF receives small values (near zero), that leads to the increase of statistical errors. The main idea of the fluctuation analysis consists in transformation of decay of the ACF which will be less sensitive to statistical errors. This analysis is applied for the purpose of detection effects of the long-term correlations in researched process. There are many variations of fluctuation analysis methods. From the very beginning the DFA method has used in biological and medical researches, but recently is being applied more often in analysis of financial time series [Kantelhardt, 2001, Kantelhardt, 2008].

According to the DFA(m) method, for the initial time series $x(t)$ the cumulative time series $y(t) = \sum_{i=1}^t x(i)$ is constructed which is then divided into N segments of length τ , and for each segment $y(t)$ the following fluctuation function is calculated:

$$F^2(\tau) = \frac{1}{\tau} \sum_{t=1}^{\tau} (y(t) - Y_m(t))^2, \tag{4}$$

where $Y_m(t)$ is a local m -polynomial trend within the given segment. The averaged on the whole of the time series $y(t)$ function $F(\tau)$ depends on the length of the segment: $F(\tau) \propto \tau^H$.

In some interval the diagram of dependence of $\log F(\tau)$ on $\log \tau$ represents a line approximated by the least square method. Estimate of the exponent \hat{H} is a tangent of the angle of slope of the line which represents the dependence of $\log F(\tau)$ on $\log(\tau)$ (see Fig. 3, where theoretical values of $H = 0.7$).

Wavelet-based estimation. Of recent, the effective tool for a time series analysis is the multiresolution wavelet-analysis, which's main idea consists in the expansion of a time series on an orthogonal base, formed by shifts and the multiresolution copies of the wavelet function. Base functions $\psi(t)$ are named wavelets, if they satisfy a number of conditions, in particular they should be defined in place of complex-valued functions with restricted energy, which oscillate around an abscissa axis, converging rapidly to zero and having a vanishing moment of the first order. Discrete wavelet-transform (DWT) is a continuous and discrete form of wavelet-transformation consisting of a two decomposition researched series : approximating and detailing, with their successive separation for the purpose of increasing the decomposition level.

Discrete wavelets are used, as a rule, together with scaling-functions connected to them. Scaling-functions with wavelets have the general definitional domain and a determined relation between values. At a given mother wavelet ψ and corresponding scaling-function ϕ with approximate coefficients $a(j,k)$ and detailing coefficients $d(j,k)$ are defined as follows:

$$a(j,k) = \int_{-\infty}^{\infty} X(t)\phi_{j,k}(t)dt, \quad d(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt, ,$$

where

$$\begin{aligned} \phi_{j,k} &= 2^{-j/2} \phi(2^{-j}t - k); \\ \psi_{j,k} &= 2^{-j/2} \psi(2^{-j}t - k). \end{aligned}$$

According to the DWT the time series is represented as the sum of detailing and approximating components:

$$\begin{aligned} X(t) &= \text{approx}_J(t) + \sum_{j=1}^{j=J} \text{detail}_j(t) = \\ &= \sum_k a(j,k)\phi_{j,k}(t) + \sum_{j=1}^J \sum_k d(j,k)\psi_{j,k}(t) \end{aligned} \tag{5}$$

For estimation of the Hurst exponent in applied research, the method described in [Abry, 1998] is the commonly used. The mentioned method is based on the statement that the averaged squared values of the wavelet

coefficients $E_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_x(j,k)|^2$ obey the scaling law:

$$E_j \sim 2^{(2H-1)j}, \tag{6}$$

where H is the Hurst exponent. The following equation represents the practicable method of the estimation of the Hurst exponent:

$$\log_2 E_j = \log_2 \left(\frac{1}{n_j} \sum_{k=1}^{n_j} |d(j, k)|^2 \right) \sim (2H - 1)j + \text{const} . \tag{7}$$

From this formula it can be concluded that if there is the long-range dependence of the time series $x(t)$ then the Hurst exponent H can be obtained by estimating the slope of the graph of the function $\log_2(E_j)$ from j . Fig. 4 shows dependence $\log_2(E_j)$ on j (dependence 7), where theoretical values of $H = 0.7$.

Modelling of fractal Brownian motion

One of the well known and simple models of stochastic dynamics which exhibits fractal properties is fractal Brownian motion (fBm). It is widely used in physics, chemistry, biology, economics and theory of network traffic.

Gaussian process $X(t)$ is called fractal Brownian motion with the parameter H , $0 < H < 1$, if the increments of the random process $\Delta X(\tau) = X(t + \tau) - X(t)$ are distributed in the following way

$$P(\Delta X < x) = \frac{1}{\sqrt{2\pi\sigma_0\tau^H}} \cdot \int_{-\infty}^x \text{Exp} \left[-\frac{z^2}{2\sigma_0^2\tau^{2H}} \right] dz , \tag{8}$$

where σ_0 is diffusion coefficient.

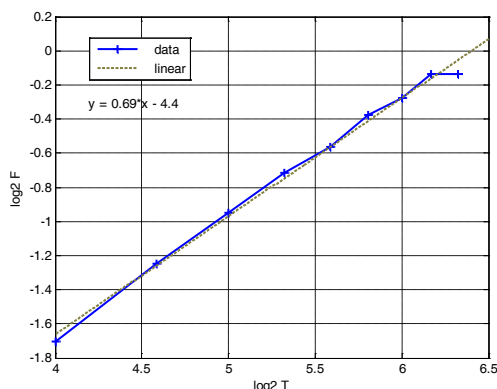


Figure 3. log-log diagram of $F(\tau)$ on τ

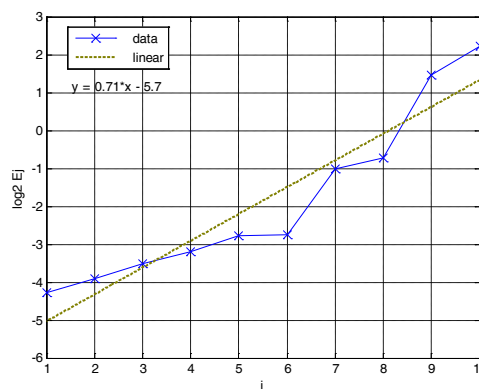


Figure 4. Dependence $\log_2(E_j)$ on j

fBm with the parameter $H = 0.5$ coincides with the classic Brownian motion. Increments of fBm are called fractal Gaussian noise and its dispersion can be described by the formula $D[X(t + \tau) - X(t)] = \sigma_0^2 \tau^{2H}$.

There are many methods of construction of fBm for the case of discrete time, which have been considered in [Mandelbrot, 1983; Feder, 1988; Voss, 1988; Cronover, 2000]. These models have some weak sides. One of them is underestimating/overestimating of the degree of self-similarity of a process for small and big theoretical values of the Hurst exponent and the short length of a model realisation [Jeongy, 1998; Cronover, 2000; Sheluhin, 2007].

One of the methods which can help to resolve the mentioned problems is the construction of fBm using biorthogonal wavelets [Sellan, 1995; Abry, 1996; Sellan, 1995; Meyer, 1999; Bardet, 2003]. In this case the fBm

realization is constructed using discrete wavelet transform where the detail wavelet coefficients on each level are independent normal random values and approximation wavelet coefficients are obtained using fractal autoregression and moving average process FARIMA:

$$B_H(t) = \sum_{k=-\infty}^{\infty} \Phi_H(t-k) S_k^{(H)} + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{-jH} \Psi_H(2^j t - k) \varepsilon_{j,k} - b_0, \tag{9}$$

where Ψ_H is biorthogonal base wavelet function, Φ_H is corresponding Ψ_H scaling function, $S_k^{(H)}$ is stationary Gaussian process FARIMA with the fractal differentiation parameter $d = H - 0.5$, $\varepsilon_{j,k}$ – independent standard Gaussian random values, b_0 – constant where $B_H(0) = 0$.

The program implementation of the given algorithm is accessible in the mathematical packet MatLab, from version 7 onwards. Fig. 5 shows a time series fBm with different Hurst exponents ($H = 0.3, 0.5, 0.8$). In fig. 6 we can see corresponding realizations of fGn.

Investigation Results: Stationary time series.

In this paper the results of a numerical experiment are represented where a fractal Brownian motion with the specified value of exponent \hat{I} have been simulated. The values of \hat{I} for the model realisation were varied within the interval $0 < H < 1$. The length of realizations was accepted equal to 500, 1000, 2000 and 4000. For each received realization, estimates \hat{I} have been obtained using the methods described above: R/S -analysis (\hat{Hrs}), variance-time analysis ($\hat{H}d$), DFA method ($\hat{H}fa$) and discrete wavelet transform ($\hat{H}w$). For each value of H samples of its estimates have been computed, and their statistical characteristics have been investigated.

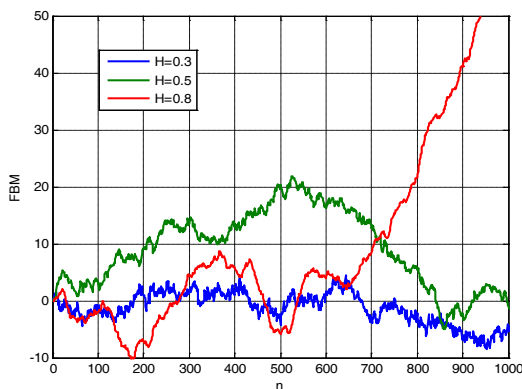


Figure 5. Realizations of fBm

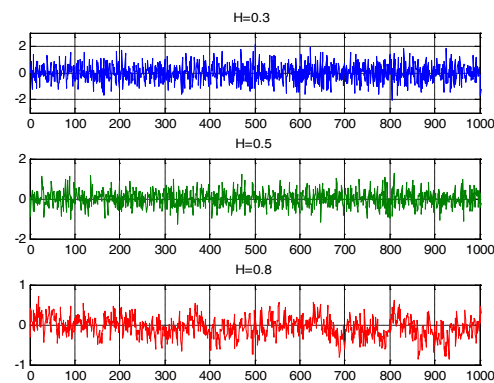


Figure 6. Realizations of fGn

Investigation of bias of estimates

Fig. 7 shows dependence of average values of estimates of the Hurst exponent on its theoretical value. The model realizations contained 1000 values. The Solid line refers to the theoretical values of H . Obviously, the average values of the estimates are biased, where the bias depends on the theoretical values of the Hurst exponent.

Obviously, the estimates of the Hurst exponent are biased in a region of persistence as well as in an antipersistence one. Since most of the fractal processes have a long-range dependence, we will be considering results only for the interval $0,5 < H < 1$. From Fig. 7 we can see that the estimates obtained by the methods of R/S -analysis and variance-time analysis are the most biased.

Let us consider the results of estimation of the exponent H by the method of R/S -analysis. The method of the rescaled range proposed by Hurst is, perhaps, the most popular one and is used in all fields of scientific research. Its main merit is its robustness. Actually, this method works even on non-stationary data. But also, as it was noticed by Hurst, the estimates of H below $H \approx 0,75$ obtained by the R/S -method are overestimated, and the estimate of H over $H \approx 0,75$ are understate.

Fig. 8 represents a dependence of average values of estimates \hat{H}_N on theoretical values of H for model series of different length. Obviously, the average values of estimates can be approximates quite well with lines $\hat{H}_N = k_N H + b_N$, where coefficients k_N and b_N depend on the realization N where the estimation is done. This lines cross the line of the theoretical values of H at around $H \approx 0,75$; and are overestimated below this values and are underestimated above this value. The results of the performed research confirm the results obtained by analysing the estimation of other models [Feder, 1988; Jeongy, 1998; Кириченко, 2005; Sheluhin, 2007]. With the increase of the realisation length N the angle of slope k_N of the approximated line increases slowly and approaches the theoretical value $\pi/4$.

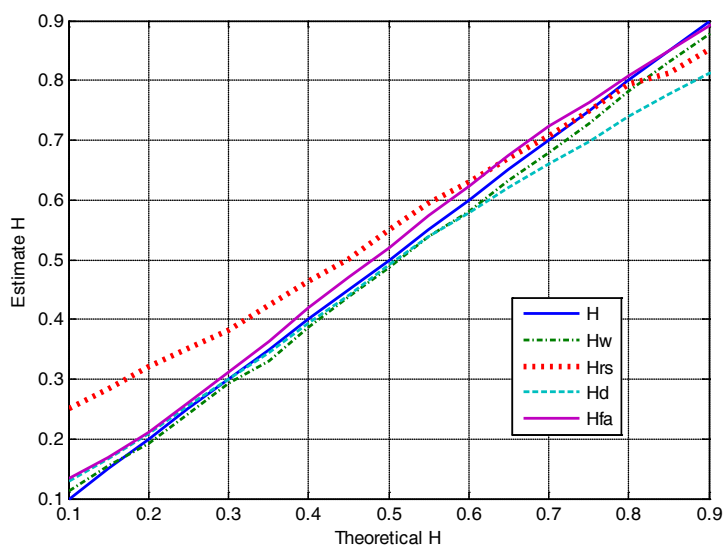


Figure 7. Dependence of the average values of estimates obtained by various methods on the theoretical H

Due to its simplicity and easy understanding of its results the method of variance-time analysis is the most commonly used for assessment of self-similarity of the information network traffic. Nevertheless, for processes with a long-range dependence, this method gives undervalued estimates. [Jeongy, 1998; Кириченко, 2005; Sheluhin, 2007]. This can be unacceptable, for instance, in the case of assessment of the network load during the transmission of the self-similar traffic. [Stollings, 2003].

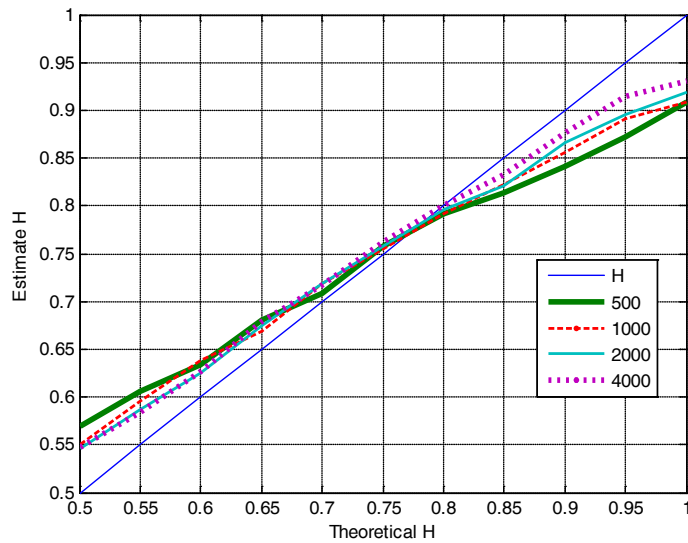


Figure 8. Dependence of average values of estimates \hat{H}_{rs} on the theoretical value of H

Fig. 9 shows the dependence of the average values of estimates $\hat{H}d$ obtained by the method of variance-time analysis. These dependence can also be approximated by the lines $\hat{H}_N = k_N H + b_N$ where coefficients k_N and b_N depend on the length of realization N . In this case the approximating lines cross the line of theoretical values at around $H \approx 0,5$ (see Fig. 7) and actually the estimates of the exponent H are undervalued within the whole interval of persistence. The bias increases with the growth of the Hurst exponent, particularly for $H > 0,9$. It can be noted that the bias of the estimates $\hat{H}d$ are greater than the appropriate bias of \hat{H}_{rs} . With the increase of the realization length N , the bias slowly decreases.

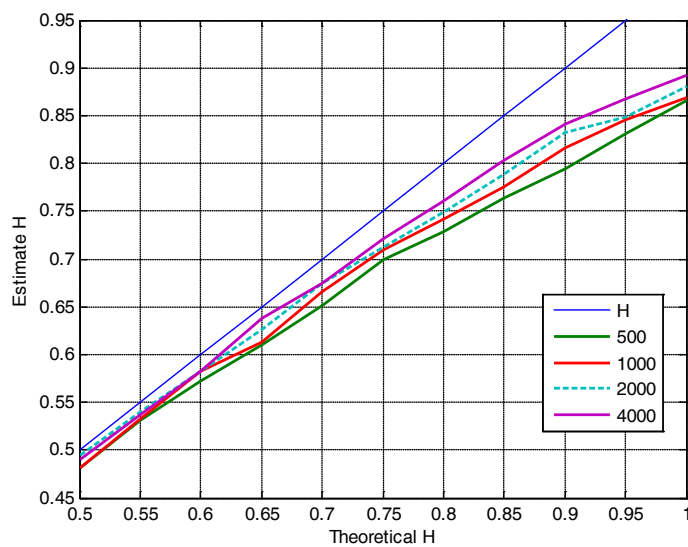


Figure 9. Dependence of average values of estimates $\hat{H}d$ on the theoretical value of H

The DFA method is based on the ideology of one-dimensional random walk and is widely used in the analysis of bioelectric signals. The estimates \hat{H}_{fa} obtained by the DFA method can be characterised by a small bias within the interval $0,5 < H < 0,9$ (see Fig. 10) even for realizations of short length. The sign of this bias reverses and increases for $H > 0,9$. It should be brought into focus that most of the natural and information fractal processes have a degree of self-similarity less than 0,9.

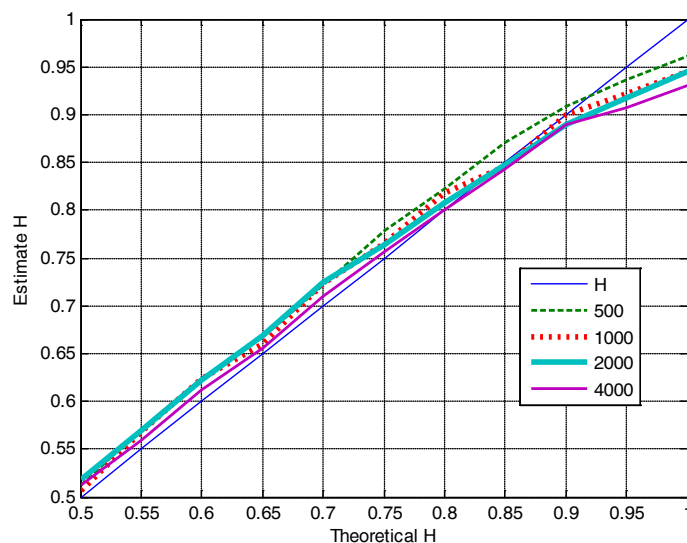


Figure 10. Dependence of the average values of the estimates \hat{H}_{fa} on the theoretical value of H

The methods of estimation of the Hurst exponent by the use of wavelet analysis are the most recent and still have not been commonly used. Nevertheless, their merits are obvious. In the paper [Abry, 1998] it has been shown that the estimates are asymptotically unbiased if base wavelets are chosen in a proper way. Fig 11 represents the dependence of the average values of \hat{H}_w obtained by the method of discrete wavelet expansion with the base wavelet function of Daubechies D4. It is obvious that with the increase of the time series length N the bias decreases and is actually equal to 0 for $N \approx 4000$.

Research of the standard deviations of the estimates

In this paper, the dependence of standard deviations of estimates of the Hurst exponent on the values of H and length of the model fractal series has been investigated for each method. In Table 1 the values of the standard deviations of the estimates of the Hurst exponent which have been received for the series of length for 1000 values are represented.

Table 1. Standard deviations of the estimates of the Hurst exponent

Estimation method	Range $S_{\hat{H}}$	Dependence on H
R/S -analysis	$0.03 \leq S_{\hat{H}} \leq 0.08$	Increases along with H
Variation of dispersion	$S_{\hat{H}} \approx 0.06$	No obvious trend
DFA	$S_{\hat{H}} \approx 0.07$	No obvious trend
Wavelet analysis	$S_{\hat{H}} \approx 0.045$	No obvious trend

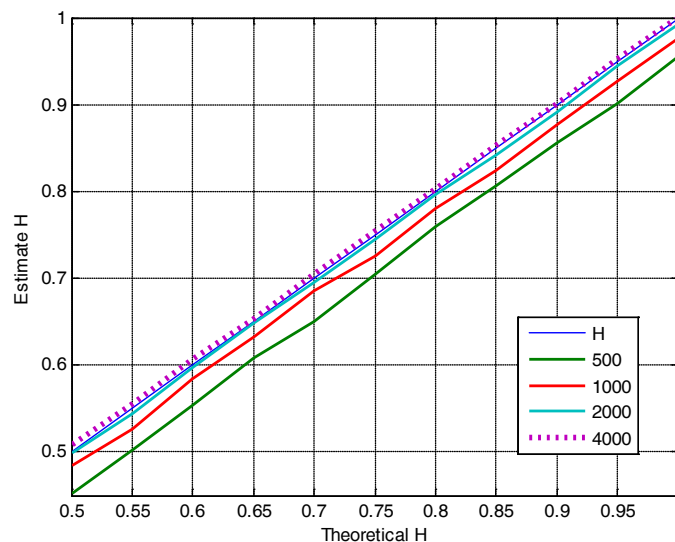


Figure 11. Dependence of the average values of the estimates \hat{H}_w on the theoretical values of H

Table 2 demonstrates how standard deviations obtained during the estimation of \hat{H} (rounded value) decrease with the increase of the length of time series. In this specific case the Hurst model exponent $H = 0.8$.

The problem of distribution law of the estimates of \hat{I} was considered in a number of works where it was shown, numerically and analytically, that the estimates are normal for a specific method or specific values of the Hurst exponent (see, for instance [Feder, 1988, Peters, 1996; Abry, 1998]). In this paper the distribution of the estimates \hat{H} have been researched for each method and different values of the parameter. For all considered methods, the hypothesis of normal distribution of sample values of the estimates with parameters $N(\bar{H}, S_{\hat{H}})$ have been suggested. For nearly all sample data, the hypothesis has been accepted with the confidence level $\alpha = 0.05$ by a narrow criterion.

Table 2. Standard deviations of the estimates depending on length of time series

$S_{\bar{H}}$	500	1000	2000	4000
S_{Hrs}	0.08	0.06	0.05	0.04
S_{Hd}	0.07	0.06	0.05	0.045
S_{Hfa}	0.085	0.07	0.055	0.045
S_{Hw}	0.065	0.045	0.03	0.02

Thus, the estimates of the Hurst exponent which are obtained by the methods considered above are biased normal random variables. For each method, the bias depends on a true value of degree of self-similarity and the length of a time series. Standard deviations of the estimates depend on the estimation method and decrease with the growth of the series length.

Research results: nonstationary time series

We investigated different model time series, presented by the sum of fractional Brownian noise the specified value of the Hurst exponent and the trend component, which is a polynomial in varying degrees, irrational, transcendental and periodic functions (see Fig. 12). The total signal can be written as

$$Y(t) = k * T(t) + fgn(t),$$

where $T(t)$ is a trend, $fgn(t)$ is a fractal Gaussian noise, and k a is a factor that regulates the ratio of trend to the noise.

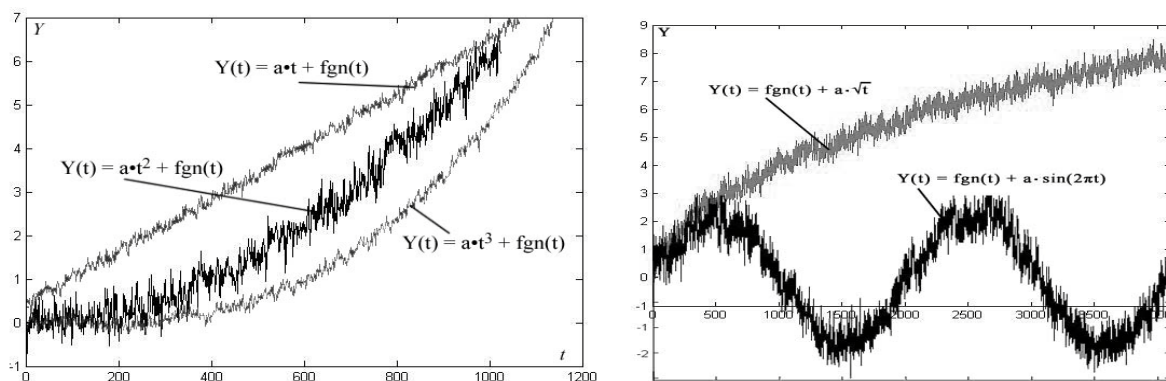


Figure 12. Model nonstationary time series

Detrended fluctuation analysis. The DFA(m) method is traditionally used in analyzing the fractal structure and estimating the degree of self-similarity of a time series with trends (for example, the implementation of encephalograms) or cumulative series with nonstationary increments (for example, financial series).

In this paper, in each case in the construction of the fluctuation functions, a local polynomial trend of increasing orders was considered [Kirichenko, 2010]. The numerical study of a fractal series with a polynomial trend

component of the order ρ showed that an adequate valuation of the Hurst parameter is achieved by using a local polynomial trend order $m > \rho$. Fig. 13 presents the fluctuation function $F(\tau)$ constructed of the different order polynomials for the fractal series (the theoretical value of the Hurst exponent equals 0.7) with linear (a) and cubic (b) trends. In the first case, the adequate valuation H is achieved for the local trend of the order $\rho = 2$, i.e. it is sufficient DFA(2), and in the second case adequate valuation is started with order $\rho = 4$, i.e. we need to use DFA(4).

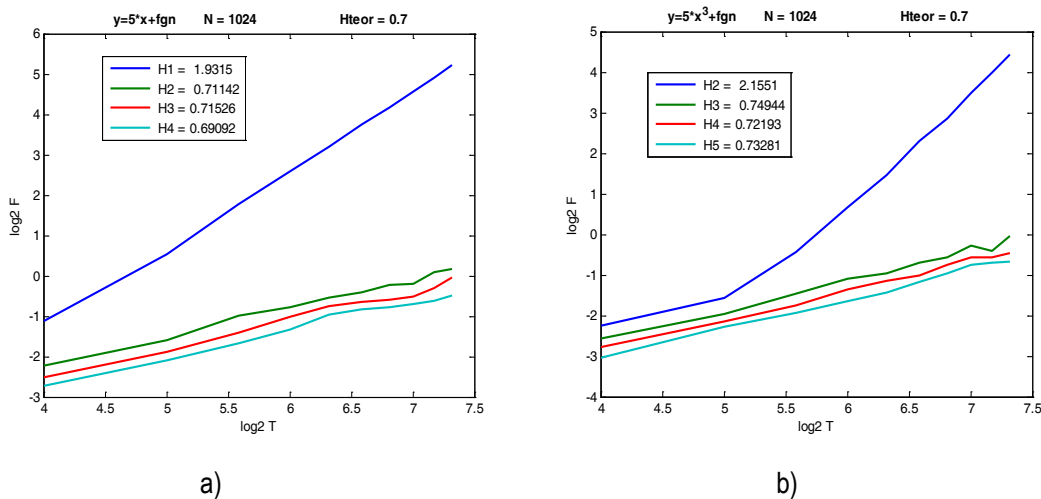


Figure 13. Fluctuation functions for time series with linear (a) and cubic (b) trend

The estimation of Hurst's exponent has been analyzed for time realizations with different trend-to-fractal noise ratio. Analysis of results showed that with local polynomial trend $m > \rho$, the value of the trend-to-noise ratio does not matter. An investigation conducted for the fractal noise with rational polynomials and transcendental trends showed an adequate estimation using the local polynomial trend order $m > 2$.

The greatest difficulty in estimating the degree of self-similarity presents a series with a periodic trend component. Fig. 14 presents a fluctuation function for a series including sinusoidal components with the number of periods equal to two and four. A numerical analysis showed that the greater the numbers of periods, the greater the degree of local polynomial should be used.

Summarizing the numerical results we can conclude that the DFA method is a convenient and reliable method for determining the degree of self-similar stochastic processes with trending components of different types.

Wavelet-based estimation. It is shown, that the influential parameter in the selection of the mother wavelet for estimating the Hurst parameter of a time series with a polynomial trend order m is the existence of its vanishing moments of order $\rho > m$ [Flandrin, 2009]. However, numerical studies have shown that the above method of wavelet estimation gives an adequate valuation in the case where the trend-to-fractal noise ratio is not great [Кириченко, 2010]. The estimator of the Hurst parameter H is calculated correctly or not, depending on value of

the ratio of trend and noise $\text{Ratio} = \frac{S_{trend}}{S_{noise}}$, where S_{trend} is the standard deviation of the trend component,

S_{noise} is the standard deviation fractal noise.

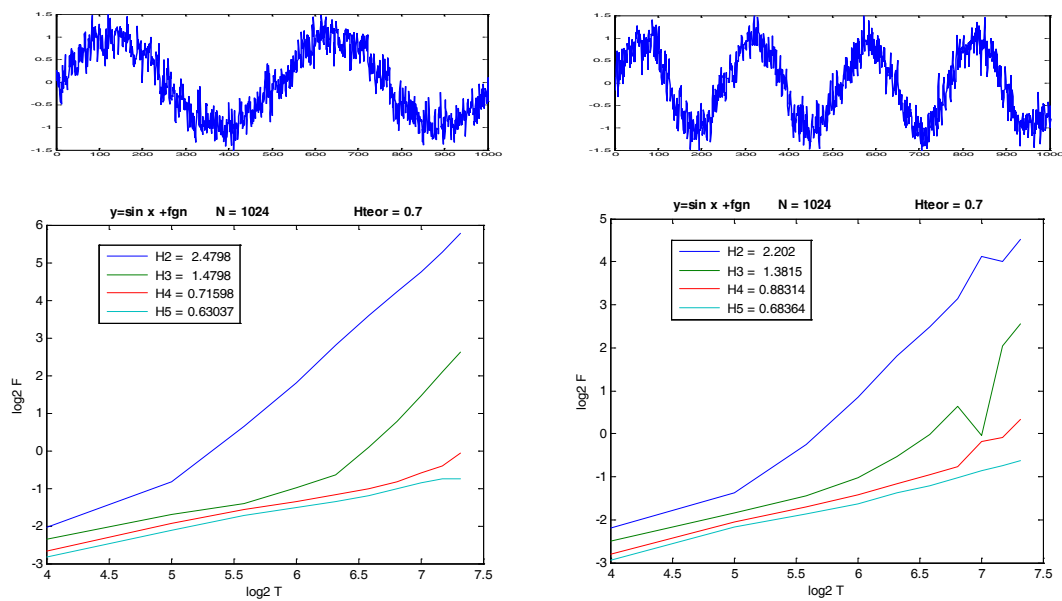


Figure 14. Model time series with sinusoidal trends and corresponding fluctuation functions for DFA(2)- DFA(5)

The numerical investigation have shown that the value Ratio^* , from which the effect of trend becomes significant, does not depend on the length of the model realization, but only on the standard deviation of the fractal process and factor k in the expression $k \cdot T(t)$. The table 3 shows the experimental results for various model signals and the value of Ratio^* .

Table 3. Model trends and the value of Ratio^*

Trend	Value of Ratio^*	Trend	Value of Ratio^*
$T(t) = k \cdot t$	9	$T(t) = k \cdot t^{\frac{3}{4}}$	0,45
$T(t) = k \cdot t^2$	7	$T(t) = k \cdot \ln(t)$	0,30
$T(t) = k \cdot t^3$	5	$T(t) = k \cdot \sin(2\pi t)$, 1 period	0,18
$T(t) = k \cdot \sqrt{t}$	2	$T(t) = k \cdot \sin(2\pi t)$, 2 periods	0,08

To understand why the trend component leads to a deterioration of estimation, let us consider in more detail the spectrum of the wavelet energy of the signal E_j , which obeys the scaling law $E_j \sim 2^{(2H-1)j}$. The wavelet energy

is equal to the amount of energy at a given level of wavelet decomposition $E_j = \frac{1}{N_j} \sum_{k=1}^{N_j} d(j, k)^2$. The Fig. 15

shows a spectrum of the wavelet energy fractal noise (a), noise with a polynomial trend (b) and noise with a periodic component (c). The lower figures show the graphs of wavelet energy in a logarithmic scale, in which the slope gives an estimate of the Hurst exponent.

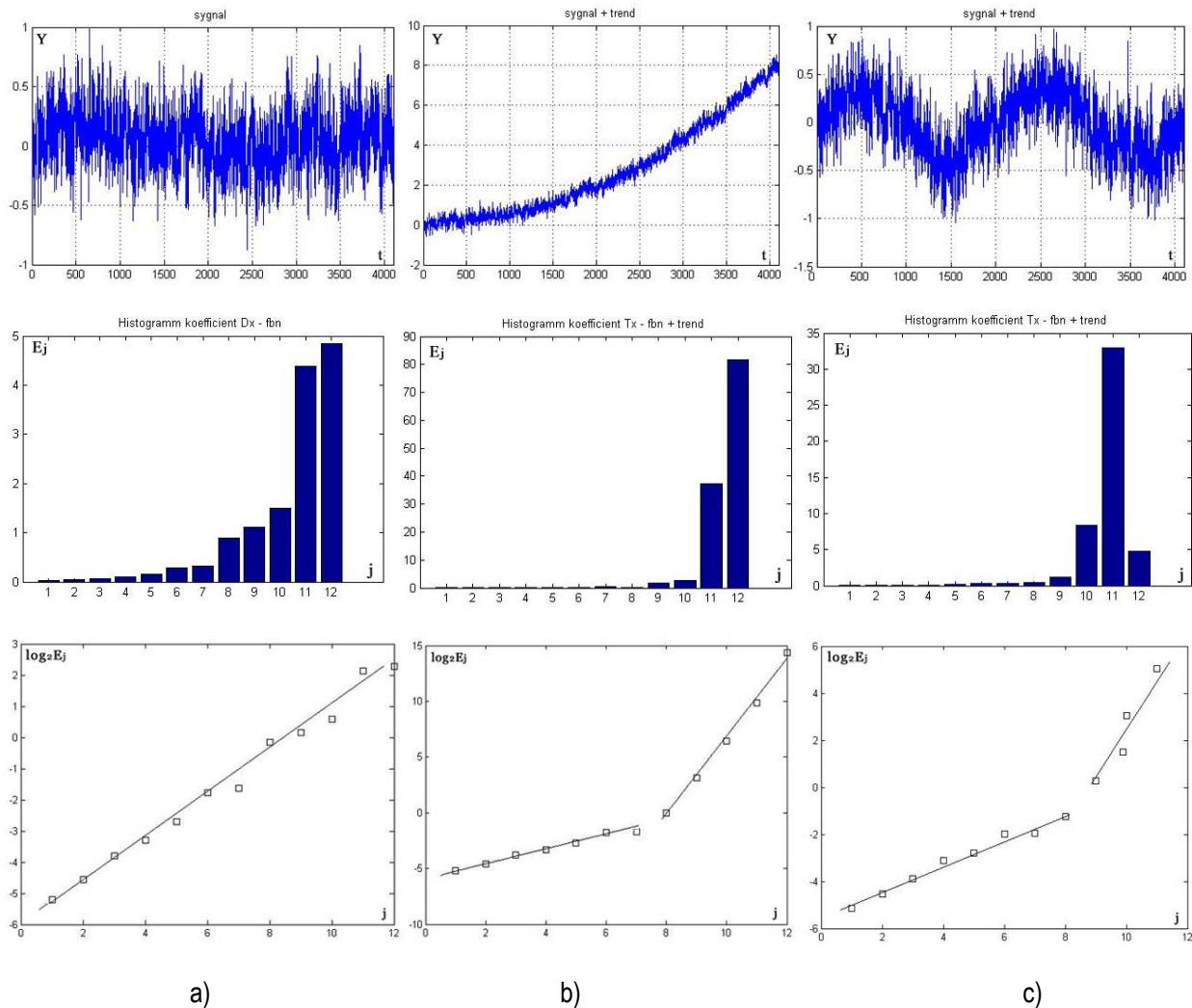


Figure 15. Model time series and corresponding spectrum of wavelet energy and dependence $\log_2(E_j)$ on j

Obviously, the high value of energy at levels determined by the trend components overstates the estimator of the Hurst parameter. In cases where these levels are low frequency components (which is typical), they can be discarded, and carry out evaluation of high-frequency levels.

Examples of real time series

In the Fig.16 below are three time realizations of processes that possess self-similar properties: site traffic protocol TCP, finance index S&P 500 for 2005 and the encephalogram of a laboratory animal in the awaking state.

The table 4 below shows the values of estimator of the Hurst parameter, obtained by the methods discussed in this paper. It should be noted that the financial series is cumulative, so in this case it was considered a series of increments.

Conclusion

Thus, the estimates of the Hurst exponent which are obtained for the stationary realisations by the methods considered above are biased normal random variables. For each method the bias depends on a true value of the degree of self-similarity and the length of a time series. Standard deviations of the estimates depend on the estimation method and decrease with the growth of the series length.

Table 4. Estimates of the Hurst exponent

Method	R/S -analysis	Variance-analysis	DFA(1)	Wavelet- estimation
Traffic (N=2000)	$\hat{H} = 0.69; S = 0.05$	$\hat{H} = 0.81; S = 0.05$	$\hat{H} = 0.75; S = 0.055$	$\hat{H} = 0.78; S = 0.03$
Index S&P 500 (N=250)	$\hat{H} = 0.65; S = 0.09$	$\hat{H} = 0.51; S = 0.08$	$\hat{H} = 0.57; S = 0.09$	$\hat{H} = 0.6; S = 0.07$
Encephalogram (N=1000)	$\hat{H} = 0.72; S = 0.06$	$\hat{H} = 0.63; S = 0.06$	$\hat{H} = 0.68; S = 0.07$	$\hat{H} = 0.65; S = 0.04$

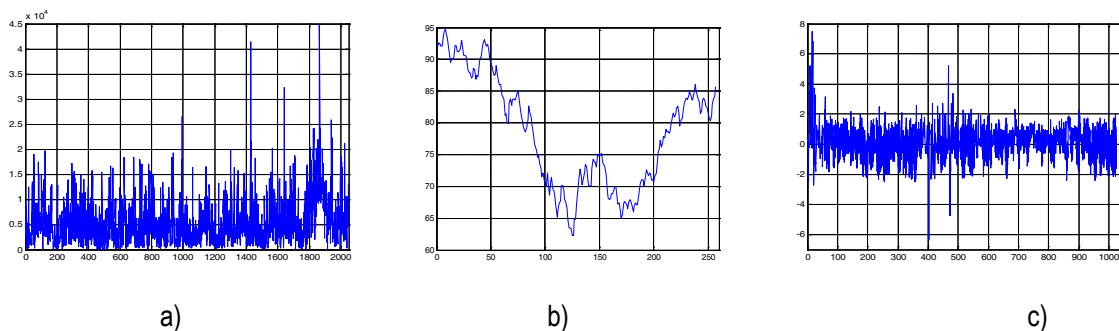


Figure 16. Time series of traffic protocol TCP, finance index and encephalogram

Summarizing the results of the numerical research we can make a conclusion that the estimates with the least bias and standard deviation can be given by the method which uses wavelet analysis. Also, other methods have some merits which can be significant in relation to some aims and ways of the research. For instance, R/S -analysis allows to estimate the degree of self-similarity of a time series for which the wavelet estimation nearly cannot be applied, The DFA method gives the best results for short series. Thus, in most cases for the estimation of the Hurst exponent it makes sense to use various methods and comparison of the results provides extra information.

In the case of a nonstationary time series represented, a trend and additive fractal noise, more accurate evaluation is obtained using the DFA method. This method allows estimation of the Hurst exponent for experimental data with trend components of virtually any kind. The greatest difficulty in estimating presents a series with a periodic trend component. It is desirable in addition to investigate the spectrum of the wavelet energy, which is demonstrates the structure of the time series. It should be noted that in the presence of a slight trend, the wavelet-estimation is quite effective.

Bibliography

- [Abry, 1996] P. Abry. The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer. Remarks and fast implementation. P. Abry, F. Sellan. *Applications and Computing Harmonic Analise* (V.3(4)), 1996.
- [Abry, 1998] P. Abry. Wavelet analysis of long-range dependent traffic. P. Abry, D. Veitch. *IEEE/ACM Transactions Information Theory* (№ 1(44)), 1998.
- [Abry, 2003] P. Abry. Self-similarity and long-range dependence through the wavelet lens. P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch. *Theory and applications of long-range dependence*, Birkhäuser, 2003.
- [Bardet, 2003] J.-M. Bardet. Generators of long-range dependence processes: a survey, *Theory and applications of long-range dependence*. Bardet J.-M., G. Lang, G. Oppenheim, A. Philippe, S. Stoev, M.S. Taqqu. *Theory and applications of long-range dependence*, Birkhäuser, 2003.
- [Chen, 2002] Z. Chen. Effect of non-stationarities on detrended fluctuation analysis. Z. Chen, P.Ch. Ivanov, K. Hu, H.E. Stanley. *Phys. Rev. E* 65, 041107, 2002.
- [Clegg, 2005] R.G. Clegg. A practical guide to measuring the Hurst parameter. R. G. Clegg. *Computing science technical report* (№ CS-TR-916), 2005.
- [Cronover, 2000] R. M. Cronover. *Introduction to fractals and chaos*. R. M. Cronover. John and Bartlett Publishers, N.Y., 2000.
- [Feder, 1988] J. Feder. *Fractals*. J. Feder. Plenum, New York, 1988.
- [Flandrin, 2009] Patrick Flandrin. *Scale Invariance and Wavelets*. Patrick Flandrin, Paulo Gonzalves and Patrice Abry in *Scaling, Fractals and Wavelets*. Ed. by P. Abry, P. Gonçaves, J. Lévy Véhel. John Wiley & Sons, London, 2009.
- [Gu, 2006] G.-F. Gu. Detrended fluctuation analysis for fractals and multifractals in higher dimensions. G.-F. Gu, W.-X. Zhou. *Phys. Rev. (E)* 74, 061104, 2006.
- [Jeongy, 1998] H.-D. J. Jeongy. *A Comparative Study of Generators of Synthetic Self-Similar Teletraffic*. H.-D. J. Jeongy, McNickle D., Pawlikowski K. Department of Computer Science and Management, University of Canterbury, 1998.
- [Kantelhardt, 2001] J.W. Kantelhardt. Detecting long-range correlations with detrended fluctuation analysis. J.W. Kantelhardt, E. Koscielny-Bunde, H.H.A. Rego, S. Havlin, A. Bunde. *Phys (A)* 295, 441, 2001.
- [Kantelhardt, 2008] J. W. Kantelhardt. *Fractal and Multifractal Time Series*. J. W. Kantelhardt. <http://arxiv.org/abs/0804.0747>, 2008.
- [Kirichenko, 2010] L. Kirichenko. Application of DFA method in fractal analysis of time series of different nature L. Kirichenko, T. Radivilova, Zh. Deineko. *Abstracts of International Conference Modern Stochastics: Theory and Application II*, Kyiv, 2010.
- [Mallat, 1998] S. Mallat. *A wavelet tour of signal processing*. S. Mallat. Academic Press, San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto, 1998.
- [Mandelbrot, 1983] B. B. Mandelbrot. *The Fractal Geometry of Nature*. B. B. Mandelbrot. W.H. Freeman, San Francisco, 3rd Edition, 1983.
- [Meyer, 1999] Y. Meyer. Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. Y. Meyer, F. Sellan, M.S. Taqqu. *The Journal of Fourier Analysis and Applications* 5(5), Birkhäuser, Boston 1999.
- [Peng, 1994] C.-K. Peng. Mosaic organization of DNA nucleotides. C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger. *Phys. Rev. (E)* 49, 1685, 1994.
- [Peters, 1996] Edgar E. Peters. *Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility*. Edgar E. Peters. Wiley, 2 edition, 1996.

- [Sellan, 1995] F. Sellan. Synthèse de mouvements browniens fractionnaires par l'aide de la transformation par ondelettes. F. Sellan. Comptes Rendus de l'Académie des Sciences de Paris, Série I, 1995.
- [Sheluhin, 2007] Oleg I. Sheluhin. Self-similar processes in telecommunications. Oleg I. Sheluhin, Sergey M. Smolskiy, Andrey V. Osin. John Wiley & Sons Ltd, Chichester, 2007.
- [Stollings, 2003] W. Stollings. High-speed networks and Internets. Performance and quality of service. W. Stollings. New Jersey, 2002.
- [Voss, 1988] R. F Voss. Fractals in nature: From characterization to simulation in The Science of Fractal Images. R. F Voss. Springer-Verlag, New York, 1988.
- [Willinger, 1996] W. Willinger. Bibliographical guide to self-similar traffic and performance modeling for modern high-speed network in «Stochastic networks: theory and applications». W. Willinger, M. S. Taqqu, A. A. Erramilli. Clarendon Press (Oxford University Press), Oxford, 1996.
- [Кириченко, 2005] Л. О Кириченко. Сравнительный анализ методов оценки параметра Херста самоподобных процессов Л. О. Кириченко, Т. А. Радивилова, М. И. Синельникова. Системи обробки інформації Вип. 8(48), Харків, 2005.
- [Кириченко, 2010] Л. О Кириченко. Оценивание параметра Хёрста для временных рядов с трендом методом вейвлет-преобразования. Л. О Кириченко, Ж.В. Дейнеко. Системи управління навігації та зв'язку. Вип 4 (16), Київ, 2010.

Authors' Information



Ludmila Kirichenko – Ph. D., Associate professor, Kharkiv National University of Radio Electronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine; e-mail: ludmila.kirichenko@gmail.com.

Major Fields of Scientific Research: Time series analysis, Stochastic self-similar processes, Wavelets, Fractals, Chaotic systems



Tamara Radivilova – Ph. D., Associate professor, Kharkiv National University of Radio Electronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine; e-mail: tomachka_7@yahoo.com.

Major Fields of Scientific Research: Wavelets and fractals, Computer systems and networks



Zhanna Deineko - lecturer, Kharkiv National University of Radio Electronics; 14 Lenin Ave., 61166 Kharkiv, Ukraine; e-mail: shanna@kture.kharkov.ua.

Major Fields of Scientific Research: Wavelets and fractals