
WAVELETS AND ESTIMATION OF LONG MEMORY IN LOG VOLATILITY AND TIME SERIES PERTURBED BY NOISE

Milan Bašta*

1. Introduction

Volatility is a key variable on financial markets. Among others, it serves as an important input for the assessment of investment risk and for pricing of derivative securities. Several empirical studies (see, e.g., Labato & Savin, 1998; Granger & Ding, 1996; and Breidt et al., 1998) suggest that volatility of financial markets has a long memory. Traditional methods of estimating long memory characteristics include the rescaled range analysis (Hurst, 1951) and the methodology of Geweke & Porter-Hudak (1983). Recently, wavelets have also been used for the estimation of the long memory parameter. For example, Percival & Walden (2002, Chapter 9) propose and summarize several approaches based on the wavelet analysis.

However, volatility is not directly observable and has to be estimated from data. This leads to the task of estimation of the long memory parameter in time series perturbed by noise. Traditional methods of the estimation of the long memory parameter are not suitable in this case as they lead to biased estimates of the long memory parameter; see, e.g., Deo & Hurvich (2001) and Artech (2004). Several methods have been suggested instead; see, e.g., Sun & Phillips (2003) or Artech (2006). To remove the bias, we propose a wavelet estimator of the long memory parameter that combines (in its construction) the frequency-domain approach of Sun & Phillips (2003) and the approach of Percival & Walden (2002). Monte Carlo simulations are used to illustrate the properties of the proposed estimator.

* University of Economics, Prague, Faculty of Informatics and Statistics (milan.basta@vse.cz).

This paper was supported by means of institutional support to long-term conceptual advancement of science and research at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic.

The paper is organized as follows: an introduction to the maximal overlap discrete wavelet transform and the wavelet variance is given in Section 2. The logarithm of volatility and its estimators are defined and the long memory property of volatility is discussed in Section 3. The notion of long memory is introduced and the spectrum of log volatility is discussed in Section 4. Section 5 presents the methodology of estimation of the long memory parameter with the least squares method and wavelets of Percival & Walden (2002) and proposes a new estimator suitable for perturbed long memory processes. Monte Carlo simulations are used to assess the properties of this estimator.

2. MODWT and wavelet variance

In these paragraphs, we will give a brief introduction to the maximal overlap discrete wavelet transform (MODWT) and the wavelet variance, which will be used for the estimation of the long-memory parameter in the perturbed time series of the logarithm of volatility.

Our summary of the wavelet theory is based on Percival & Walden (2002) and Bašta (2010), which encompass further details and proofs of the statements given below. Further books may be recommended: e.g., Vidakovic (1999) gives a more technical and abstract introduction to wavelets. Gençay & Selçuk & Whitcher (2002) present a less technical introduction with applications in finance and economics.

2.1 Spectrum of the stochastic process and its properties

Let $\{X_t; t = \dots, -1, 0, 1, \dots\}$ be a discrete-time second-order stationary (hereafter, stationary) stochastic process. Let $\{s_{X,\tau}; \tau = \dots, -1, 0, 1, \dots\}$ be the autocovariance sequence of this process. The spectrum of this process may be defined as

$$S_X(f) \equiv \sum_{\tau=-\infty}^{\infty} s_{X,\tau} \exp(-i2\pi f \tau), \quad \text{for } f \in R, \quad (1)$$

where i is the imaginary unit and f is a real-valued variable called frequency. Since spectrum is an even function of frequency and a periodic function of frequency with the period equal to unity, it is sufficient to study and explore it only in the range of frequencies $[0, 1/2]$. We may say informally that the spectrum $S_X(\cdot)$ informs us about the „frequency content“ of the stochastic process $\{X_t\}$ in the sense that

$$s_{X,0} = 2 \int_0^{1/2} S_X(f) df. \quad (2)$$

The contribution of the frequency interval $(f, f + df)$, where $0 \leq f \leq f + df \leq 1/2$, to the variance $s_{X,0}$ of the process $\{X_t\}$ is given as

$$S_X(f) df. \quad (3)$$

2.2 Linear filtering

Let the finite sequence $\{a_t; t = 0, \dots, M-1\}$ of real-valued (non-random) variables be called the *linear filter* of length M . For ease of notation, we will use the notation $\{a_t\}$ whenever the values that t can assume are clear from the text. Let the linear filtering of the discrete-time stationary stochastic process $\{X_t; t = \dots, -1, 0, 1, \dots\}$ with the linear filter $\{a_t\}$ be a discrete-time stationary stochastic process $\{Y_t; t = \dots, -1, 0, 1, \dots\}$ defined as

$$Y_t \equiv \sum_{m=0}^{M-1} a_m X_{t-m}, \quad \text{for } t = \dots, -1, 0, 1, \dots \quad (4)$$

Let $S_X(\cdot)$ and $S_Y(\cdot)$ be the spectrum of the process $\{X_t\}$ and of the process $\{Y_t\}$ defined in the sense of Eq. 1. The effect of the linear filtering with the linear filter $\{a_t\}$ of the stochastic process $\{X_t\}$ is given by the characteristics of the linear filter $\{a_t\}$. An important characteristics is the *frequency response* $F_a(\cdot)$ of this filter defined as the Fourier transform of $\{a_t\}$, i.e.,

$$F_a(f) \equiv \sum_{t=0}^{M-1} a_t \exp(-i2\pi ft), \quad \text{for } f \in R. \quad (5)$$

The function $|F_a(\cdot)|^2$, i.e., the square of the modulus of $F_a(\cdot)$, is called the squared gain function of the filter $\{a_t\}$. The squared gain function is an even and periodic function of frequency, with the period equal to unity. Therefore, further on it is sufficient to explore it only in the frequency range $[0, 1/2]$. The squared gain function informs us about frequencies that are passed and about frequencies that are attenuated by the effect of the linear filtering. More specifically, it can be shown that

$$S_Y(f) = |F_a(f)|^2 S_X(f), \quad \text{for } 0 \leq f \leq 1/2. \quad (6)$$

2.3 Wavelet and scaling filters

To introduce the maximal overlap discrete wavelet transform, hereafter abbreviated to MODWT, we first need to define a special set of linear filters. This special set may be created with the knowledge of just *two* elementary filters: the filter $\{h_{1,t}; t = 0, \dots, L_1 - 1\}$ and the filter $\{g_{1,t}; t = 0, \dots, L_1 - 1\}$, both of lengths L_1 . The linear filters $\{h_{1,t}; t = 0, \dots, L_1 - 1\}$ and $\{g_{1,t}; t = 0, \dots, L_1 - 1\}$ are interconnected via the so-called *quadrature mirror relationship*, i.e.,

$$g_{1,t} \equiv (-1)^{t+1} h_{1,L_1-1-t}, \quad \text{for } t = 0, \dots, L_1 - 1 \quad (7)$$

and fulfil certain specific conditions, e.g.,

$$\sum_{t=0}^{L_1-1} h_{1,t} = 0, \quad \sum_{t=0}^{L_1-1} h_{1,t}^2 = \frac{1}{2}, \quad \sum_{t=0}^{L_1-1} h_{1,t} h_{1,t+2n} = 0, \quad \text{for any nonzero integer } n, \quad (8)$$

$$\sum_{t=0}^{L_1-1} g_{1,t} = 1, \quad \sum_{t=0}^{L_1-1} g_{1,t}^2 = \frac{1}{2}, \quad \sum_{t=0}^{L_1-1} g_{1,t} g_{1,t+2n} = 0, \text{ for any nonzero integer } n. \quad (9)$$

An ideal filter with the nominal pass band $[d, u]$, where d and u are real parameters satisfying $0 \leq d \leq u \leq 1/2$, is such a filter whose squared gain function is given as

$$\left| F_i^{d,u}(f) \right|^2 \equiv \begin{cases} 1, & \text{for } d \leq f \leq u \\ 0, & \text{for } 0 \leq f < d \text{ and for } u < f \leq 1/2 \end{cases} \quad (10)$$

It can be shown that $\{h_{1,t}\}$ is approximately an ideal filter with the nominal pass band $[1/4, 1/2]$. On the other hand, $\{g_{1,t}\}$ is approximately an ideal low-pass filter with the nominal pass band $[0, 1/4]$.

For a given pair of filters $\{h_{1,t}\}$ and $\{g_{1,t}\}$, two sets of filters may be created via the so-called *pyramid algorithm*. This algorithm will not be specified here because it is technical and will be of no use for the purpose of this text (details on the algorithm can be found, e.g., in Percival & Walden, 2002; Gençay & Selçuk & Whitcher, 2002; Vidakovic, 1999; and Bašta, 2010). Only the properties of the filters from both sets will be discussed in this text.

First, let us call the filters from the first set the *wavelet* filters and let us denote them as $\{h_{1,t}; t = 0, \dots, L_1 - 1\}$, $\{h_{2,t}; t = 0, \dots, L_2 - 1\}$, $\{h_{3,t}; t = 0, \dots, L_3 - 1\}$, etc., where the filter $\{h_{1,t}\}$ is the filter already defined above, and where L_1, L_2, L_3 are the lengths of these filters. The expression “etc.” in the previous sentence is to mean that the set of the wavelet filters is infinite. The wavelet filters are indexed by positive integers and generally the j th wavelet filter for $j = 1, 2, \dots$ is denoted as $\{h_{j,t}; t = 0, \dots, L_j - 1\}$ and has the length $L_j = (2^j - 1)(L_1 - 1) + 1$.

Similarly, let us call the filters from the second set the *scaling* filters and let us denote them as $\{g_{1,t}; t = 0, \dots, L_1 - 1\}$, $\{g_{2,t}; t = 0, \dots, L_2 - 1\}$, $\{g_{3,t}; t = 0, \dots, L_3 - 1\}$, etc., where the filter $\{g_{1,t}\}$ is the filter already defined in the previous paragraphs and where L_1, L_2, L_3 are the lengths of these filters. Again, the set of the scaling filters is infinite and generally the j th filter of this set is denoted as $\{g_{j,t}; t = 0, \dots, L_j - 1\}$ and has the same length as filter $\{h_{j,t}\}$, i.e., L_j .

It is important to stress that the filter $\{h_{j,t}\}$ (for $j = 1, 2, \dots$) is approximately an ideal filter with the nominal pass band $[2^{-(j+1)}, 2^{-j}]$ and $\{g_{j,t}\}$ (for $j = 1, 2, \dots$) is approximately an ideal low-pass filter with the nominal pass band $[0, 2^{-(j+1)}]$.

There are different families of sets of wavelet and scaling filters, for example the Haar family, the D(4) family, the LA(8) family, etc. (for details see, e.g., Percival & Walden, 2002).

2.4 MODWT wavelet coefficients and wavelet variance

Let $\{X_t; t = \dots, -1, 0, 1, \dots\}$ be a discrete-time stationary stochastic process with the variance $\sigma_X^2 \equiv s_{X,0}$ and the spectrum $S_X(\cdot)$. The linear filtering of $\{X_t\}$ with the wavelet filter $\{h_{j,t}; t = 0, \dots, L_j - 1\}$ (for $j = 1, 2, \dots$) results in a discrete-time stationary stochastic process defined as

$$\bar{W}_{X,j,t} \equiv \sum_{m=0}^{L_j-1} h_{j,m} X_{t-m}, \quad \text{for } t = \dots, -1, 0, 1, \dots, \quad (11)$$

called the process of the j th level *MODWT wavelet coefficients*. j th level MODWT wavelet coefficients possess certain intriguing properties: they are stationary and it can be shown that their unconditional mean is zero and their unconditional variance – called the j th level wavelet variance – is defined as

$$\nu_{X,j}^2 \equiv E\left(\bar{W}_{X,j,t}^2\right). \quad (12)$$

It can be shown that

$$\sigma_X^2 = \sum_{j=1}^{\infty} \nu_{X,j}^2. \quad (13)$$

Moreover, let $S_{W_j}(\cdot)$ be the spectrum of j th level wavelet coefficients. It holds that (see Eq. 6)

$$S_{W_j}(f) = |F_{h_j}(f)|^2 S_X(f), \quad (14)$$

where $|F_{h_j}(f)|^2$ is the squared gain function of the filter $\{h_{j,t}\}$. Therefore (see Eq. 2),

$$\nu_{X,j}^2 = 2 \int_0^{1/2} S_{W_j}(f) df = 2 \int_0^{1/2} |F_{h_j}(f)|^2 S_X(f) df \approx 2 \int_{1/2^{j+1}}^{1/2^j} S_X(f) df. \quad (15)$$

The approximation \approx in Eq. 15 is meant in the sense that $\{h_{j,t}\}$ is approximately an ideal filter with the nominal pass band $[2^{-(j+1)}, 2^{-j}]$ (see Section 2.3) and thus, $|F_{h_j}(f)|^2$ may be approximately replaced with $|F_i^{d,u}(f)|^2$ with $d = 2^{-(j+1)}$ and $u = 2^{-j}$. Eq. 13 and Eq. 15 tell us that $\nu_{X,j}^2$ is associated with the contribution of the frequency band $[2^{-(j+1)}, 2^{-j}]$ to the variance of the stochastic process $\{X_t\}$.

2.5 Estimation of wavelet variance

Let us be given a portion X_0, X_1, \dots, X_{N-1} of the stationary process $\{X_t; t = \dots, -1, 0, 1, \dots\}$ and let us define $\{W_{X,j,t}; t = 0, \dots, N-1\}$ as

$$W_{X,j,t} \equiv \sum_{m=0}^{L_j-1} h_{j,m} X_{(t-m) \bmod N}, \quad \text{for } t = 0, \dots, N-1, \quad (16)$$

where mod denotes the operation modulo. The modulo operation affects the construction of $W_{X,j,t}$ for $0 \leq t < L_j - 1$, whereas the construction of $W_{X,j,t}$ for $L_j - 1 \leq t \leq N - 1$ would remain unaffected if the modulo operation was left out in the definition of Eq. 16. Therefore, we say that $W_{X,j,t}$ for $0 \leq t < L_j - 1$ are affected by the circularity assumptions, whereas $W_{X,j,t}$ for $L_j - 1 \leq t \leq N - 1$ are not.

An unbiased estimator of the wavelet variance $v_{X_j}^2$ of Eq. 12 for $N - L_j \geq 0$ is defined as

$$n_{X,j}^2 \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} W_{X,j,t}^2, \quad (17)$$

where

$$M_j \equiv N - L_j + 1, \quad (18)$$

i.e., only $W_{X,j,t}$ not affected by the circularity assumption are involved in the construction of the estimator.

3. Log volatility

In this section, we introduce the notion of volatility and its two estimators. For a given trading day t (i.e., from the open of the day t to the close of the day t), let us assume that the *logarithm* of the stock price is a Brownian motion. Further, let $Close_t$ be the closing stock price for the trading day t and $Open_t$ be the opening stock price for the trading day t . Let the return for the trading day t be defined as

$$R_t \equiv \log\left(\frac{Close_t}{Open_t}\right) = \log(Close_t) - \log(Open_t). \quad (19)$$

Furthermore, let V_t be the variance of R_t , hereafter called volatility, i.e.,

$$V_t \equiv E(R_t^2) \quad (20)$$

and let log volatility H_t be defined as the logarithm of V_t , i.e.,

$$H_t \equiv \log(V_t). \quad (21)$$

Further on, we will assume that log volatility is constant for a given trading day but can differ from one trading day to another.

3.1 Log-squared-return estimator

With respect to the definitions and assumptions of the previous section, we may write $R_t = (V_t)^{0.5} U_t$, where U_t is a standard normal random variable independent of V_t . It follows that $\log(R_t^2) = H_t + \log(U_t^2)$ and (see, e.g., Breidt et al., 1998) $\gamma_1 \equiv E(\log(U_t^2)) = -1.27$, $\text{var}(\log(U_t^2)) = \pi^2/2$. Thus, $A_t \equiv \log(R_t^2) - \gamma_1$ is an unbiased estimator of H_t in the sense that $E(A_t|H_t) = H_t$. The variance of this estimator is $\text{var}(A_t|H_t) = \pi^2/2$. We will call A_t the *log-squared-return* estimator.

3.2 Log-Garman-Klass estimator

If the information on the opening, closing, high and low prices for the trading day t is available, it is possible to obtain another unbiased estimator with a smaller variance than that of the log-squared-return estimator. Let us define (see Garman & Klass, 1980)

$$P_{0,t} \equiv \log[0.511(M_t - B_t)^2 - 0.019[c_t(M_t + B_t) - 2M_t B_t] - 0.383C_t^2], \quad (22)$$

where $M_t \equiv \log(Max_t) - \log(Open_t)$, $B_t \equiv \log(Min_t) - \log(Open_t)$, $C_t \equiv \log(Close_t) - \log(Open_t)$, where Max_t , Min_t , $Close_t$, $Open_t$ are the high, low, closing and opening stock prices for the day t . It can be shown that

$$P_{0,t} = H_t + Z_{0,t}, \quad (23)$$

where $Z_{0,t}$ has approximately a Gaussian distribution with (see for example Molnár, 2011) $\gamma \equiv E(Z_{0,t}) = -0.13$, $\text{var}(Z_{0,t}) = 0.26$ and is independent of H_t . Let us define $P_t \equiv P_{0,t} - \gamma$, $Z_t \equiv Z_{0,t} - \gamma$. Then, $\sigma_z^2 \equiv \text{var}(Z_t) = 0.26$ and $P_t = H_t + Z_t$ is an unbiased estimator of H_t in the sense that $E(P_t|H_t) = H_t$. The variance of this estimator is $\text{var}(P_t|H_t) = 0.26$. We will call P_t the *log-Garman-Klass* estimator.

3.3 Stochastic processes

So far we have analyzed the estimation of the log volatility for a single trading day t based on the information for this day. Now, let us generalize our approach to a sequence of trading days indexed by $\dots, -1, 0, 1, \dots$. Each of these trading days is characterized by H_t , $\log(U_t^2)$ and Z_t . Again, we stress that log volatility is assumed to be constant during a given trading day but can differ from one trading day to another. Specifically, let $\{H_t; t = \dots, -1, 0, 1, \dots\}$, $\{\log(U_t^2); t = \dots, -1, 0, 1, \dots\}$ and $\{Z_t; t = \dots, -1, 0, 1, \dots\}$ be discrete-time stochastic processes. From the assumptions of the previous sections it can be shown that the process $\{H_t\}$ is independent of $\{\log(U_t^2)\}$ and also of $\{Z_t\}$. Moreover, $\{\log(U_t^2)\}$ and $\{Z_t\}$ are white noise processes. In practice realizations of $\{H_t\}$ and $\{\log(U_t^2)\}$ are not observable.

For each trading day $t = \dots, -1, 0, 1, \dots$ either the log-squared-return estimator A_t or the log-Garman-Klass P_t estimator may be constructed to estimate log volatility H_t of the day. We may write

$$A_t = H_t + \log(U_t^2) - \gamma_1, \quad t = \dots, -1, 0, 1, \dots \quad (24)$$

$$P_t = H_t + Z_t, \quad t = \dots, -1, 0, 1, \dots \quad (25)$$

Thus, $\{A_t; t = \dots, -1, 0, 1, \dots\}$ and $\{P_t; t = \dots, -1, 0, 1, \dots\}$ are discrete-time stochastic processes.

4. Long memory and volatility

In this section, we introduce long memory processes and one example of these processes – the fractionally differenced process – and argue that log volatility processes are long memory ones. We discuss their spectra.

4.1 Long memory and fractionally differenced processes

First, let us assume a discrete-time stationary stochastic process $\{X_t; t = \dots, -1, 0, 1, \dots\}$ and its spectrum $S_X(\cdot)$. This process (by definition; see Beran, 1994) is called a stationary

process with long memory if real-valued constants $K > 0$ and $-1 < \alpha < 0$ exist such that

$$\lim_{f \rightarrow 0^+} \frac{S_X(f)}{K \times f^\alpha} = 1. \quad (26)$$

One example of such a process is the fractionally differenced process defined as (see Granger & Joyeux, 1980; Hosking, 1981)

$$(1 - B)^\delta X_t = E_t, \quad \text{for } t = \dots, -1, 0, 1, \dots \quad (27)$$

where B is the lag operator defined as $BX_t \equiv X_{t-1}$, $\{E_t; t = \dots, -1, 0, 1, \dots\}$ is a Gaussian white noise process with the variance σ_E^2 , and δ is a real-valued parameter satisfying $0 < \delta < 0.5$. Hereafter, the parameter δ will be called the long memory parameter. It can be shown that the spectrum $S_X(\cdot)$ of the fractionally differenced process is given as

$$S_X(f) = \frac{\sigma_E^2}{[4\sin^2(\pi f)]^\delta}, \quad \text{for } 0 \leq f \leq \frac{1}{2}. \quad (28)$$

4.2 Spectrum of log volatility

Let $S_H(\cdot)$, $S_P(\cdot)$ and $S_Z(\cdot)$ be the spectra of $\{H_t\}$, $\{P_t\}$ and $\{Z_t\}$. The empirical analysis of the time series of squared returns (e.g., Lobato & Savin, 1998), absolute returns (e.g., Granger & Ding, 1996) or log squared returns (e.g., Breidt et al., 1998) suggests that the processes of volatility have a long memory. Therefore, let us assume that $\{H_t; t = \dots, -1, 0, 1, \dots\}$ is a stationary fractionally differenced process with the long memory parameter δ , i.e., (see Eq. 27)

$$(1 - B)^\delta H_t = E_t. \quad (29)$$

The spectrum $S_H(\cdot)$ of $\{H_t\}$ is thus given as (see Eq. 28; we assume the frequency range $[0, 1/2]$)

$$S_H(f) = \frac{\sigma_E^2}{[4\sin^2(\pi f)]^\delta}. \quad (30)$$

The logarithm of Eq. 30 results in

$$\log S_H(f) = -\delta \log(4\sin^2(\pi f)) + \log(\sigma_E^2). \quad (31)$$

For „ f close to zero“ we can use the approximation

$$\log S_H(f) \approx -2\delta \log(2\pi f) + \log(\sigma_E^2). \quad (32)$$

4.3 The perturbed case

Now, let us discuss the spectrum of $\{P_t; t = \dots, -1, 0, 1, \dots\}$. Since $\{H_t; t = \dots, -1, 0, 1, \dots\}$ is independent of $\{Z_t; t = \dots, -1, 0, 1, \dots\}$ we can write (assuming the frequency range $[0, 1/2]$)

$$S_P(f) = S_H(f) + S_Z(f). \quad (33)$$

Thus, we can write

$$S_P(f) = \frac{\sigma_E^2}{[4\sin^2(\pi f)]^\delta} + \sigma_Z^2. \quad (34)$$

Following the idea of Sun & Phillips (2003) we can write

$$S_P(f) = \frac{1}{[4\sin^2(\pi f)]^\delta} \left[\sigma_E^2 \left(1 + \frac{\sigma_Z^2}{\sigma_E^2} \times [4\sin^2(\pi f)]^\delta \right) \right]. \quad (35)$$

Taking the logarithm of both sides results in

$$\log S_P(f) = -\log[4\sin^2(\pi f)]^\delta + \log(\sigma_E^2) + \log \left(1 + \frac{\sigma_Z^2}{\sigma_E^2} \times [4\sin^2(\pi f)]^\delta \right). \quad (36)$$

For

$$\frac{\sigma_Z^2}{\sigma_E^2} \times [4\sin^2(\pi f)]^\delta \quad (37)$$

„close to zero“ and f „close to zero“ we can write

$$\log S_P(f) \approx -2\delta \log(2\pi f) + \log(\sigma_E^2) + \frac{\sigma_Z^2}{\sigma_E^2} \times (2\pi f)^{2\delta}. \quad (38)$$

5. Long memory parameter estimation

In this section, we build up on the methodology by Percival & Walden (2002), who present the least squares estimation of the long memory parameter with the use of wavelets in unperturbed long memory time series. We enhance this methodology so that it enables us to estimate the long memory parameter in the case of long memory processes perturbed with independent white noise (which is the case of, e.g., the log-Garman-Klass estimates of volatility of financial series).

5.1 Percival & Walden estimator of the long memory parameter

Percival & Walden (2002) present a general methodology that allows the estimation of the long memory parameter δ in unperturbed fractionally differenced processes. Let us assume that log volatility $\{H_t; t = \dots, -1, 0, 1, \dots\}$ is a stationary fractionally differenced

process with the long memory parameter δ . Let $S_H(\cdot)$ be the spectrum of $\{H_t\}$, $v_{H,j}^2$ the j th level wavelet variance of $\{H_t\}$ and $n_{H,j}$ the estimate of the j th level wavelet variance.

For „ f close to zero“ we may write (see Eq. 32; only the frequency range $[0, 1/2]$ is assumed)¹

$$\log S_H(f) = -2\delta \log(2\pi f) + \log(\sigma_E^2). \quad (39)$$

Let us further assume that the spectrum $S_H(\cdot)$ is approximately constant and equal to C_j in the frequency range $2^{-(j+1)} < f < 2^{-j}$ for $j = 1, 2, \dots$. From Eq. 15 it follows that

$$C_j \approx 2^j v_{H,j}^2. \quad (40)$$

Again, for j such that „ $2^{-(j+1/2)}$ is close to zero“ we can write²

$$\log(v_{H,j}^2) = -2\delta \log(2\pi 2^{-(j+1/2)}) - \log(2^j) + \log(\sigma_E^2), \quad (41)$$

where $2^{-(j+1/2)}$ is the centre of the frequency interval $[2^{-(j+1)}, 2^{-j}]$ on the logarithmic scale. Furthermore, we get

$$\log(v_{H,j}^2) = (1 - 2\delta) \log(2\pi 2^{-(j+1/2)}) + \log(\sigma_E^2) - \log(\sqrt{2}\pi). \quad (42)$$

Now, let us assume (at least for a while) that realizations of $\{H_t\}$ be observable, which is not the case in reality. Then, let us calculate the estimates $n_{H,j}^2$ of the wavelet variances $v_{H,j}^2$ in the sense of Eq. 17 for all integers j satisfying $\alpha \leq j \leq \beta$, where α and β are positive integers chosen in such a way so that:

- $2^{-(\alpha+1/2)}$ is „close to zero“. This condition ensures that $2^{-(j+1/2)}$ is „close to zero“ for every integer j satisfying $j \geq \alpha$.
- $N - L_\beta \geq 0$, where L_β is the length of the wavelet filter $\{h_{\beta,t}\}$. This ensures that the variance of the j th level wavelet coefficients can be estimated for all integers j satisfying $1 \leq j \leq \beta$, since $M_j \geq 1$ for $1 \leq j \leq \beta$ (see also Eq. 17).

For further on, let us assume that the integer j satisfies $\alpha \leq j \leq \beta$. Further, we can use the approximation (see, e.g., Percival & Walden, p. 313)³

$$n_{H,j}^2 = \frac{d v_{H,j}^2 \chi^2[\eta_j]}{\eta_j}, \quad (43)$$

where $\chi^2[\eta_j]$ is a chi-square random variable with η_j degrees of freedom, where

$$\eta_j \equiv \max\{M_j / 2^j, 1\}, \quad (44)$$

1 For ease of notation we use the equality sign (=) instead of the approximation sign (\approx) having in mind that the equality is only approximate.

2 Again, we use the equality sign (=) instead of the approximation sign (\approx) having in mind that the equality is only approximate.

3 d stands for equality in distribution

where M_j is defined in Eq. 18. Thus

$$\log(n_{H,j}^2) \stackrel{d}{=} \log(v_{H,j}^2) - \log(\eta_j) + \log(\chi^2[\eta_j]). \quad (45)$$

Therefore,

$$\log(n_{H,j}^2) + \log(\eta_j) \stackrel{d}{=} (1 - 2\delta) \log(2\pi 2^{-(j+1/2)}) + \log(\sigma_E^2 / (\sqrt{2}\pi)) + \log(\chi^2[\eta_j]). \quad (46)$$

We can write (see Bartlett & Kendall, 1946; Percival & Walden, 2002)

$$E(\log \chi^2[\eta_j]) = \psi(\eta_j / 2) + \log(2), \quad (47)$$

$$\text{var}(\log \chi^2[\eta_j]) = \psi'(\eta_j / 2), \quad (48)$$

where $\psi(\cdot)$ and $\psi'(\cdot)$ is the digamma and trigamma function. Thus, for

$$Q_j^* \equiv \log(n_{H,j}^2) - \psi(\eta_j / 2) + \log(\eta_j / 2), \quad (49)$$

we can write

$$Q_j^* = (1 - 2\delta) \log(2\pi 2^{-(j+1/2)}) + \log(\sigma_E^2) - \log(\sqrt{2}\pi) + e_j^*, \quad (50)$$

where

$$e_j^* \stackrel{d}{=} \log(\chi^2[\eta_j]) - \psi(\eta_j / 2) - \log(2) \quad (51)$$

is a random variable with a zero mean and variance equal to $\psi'(\eta_j/2)$. Moreover, e_j^* is approximately normally distributed (for $\eta_j > 10$) and e_j^* and e_i^* (for integer j and integer i satisfying $\alpha \leq j \leq \beta$, $\alpha \leq i \leq \beta$ and $i \neq j$) are approximately uncorrelated (see Percival & Walden, 2002). Afterwards, an estimate of δ – let us call it the PW estimate – can be obtained from the estimate of the slope of the linear regression model of Eq. 50. Moreover, the estimate of $\log(\sigma_E^2)$ can be obtained from the estimate of the intercept in the linear regression model of Eq. 50. The parameters of the regression model can be estimated by the weighted least squares with the weight at the level j given as $1/\psi'(\eta_j/2)$.

However, one major difficulty occurs. During the construction of the PW estimator we have assumed that the realizations of $\{H_i\}$ be observable so that $n_{H,j}^2$ could be calculated in the sense of Eq. 17. However, this is not the case as realizations of $\{H_i\}$ are not observable since $\{H_i\}$ is buried in noise. We will tackle this problem in the next section.

5.2 An estimator for perturbed time series – Approach A

First let us give some notations. Let $S_p(\cdot)$ and $v_{p,j}^2$ (for $j = 1, 2, \dots$) be the spectrum of $\{P_i\}$ and the j th level wavelet variance of $\{P_i\}$. Furthermore, let $n_{p,j}^2$ be the estimate of $v_{p,j}^2$ calculated in the sense of Eq. 17 from a given realization of a portion of $\{P_i\}$.

The Approach A assumes that $n_{p,j}^2$ equals $n_{H,j}^2$. Afterwards, the estimate of δ may be easily obtained based on the model of Eq. 50. Thus, said in a word, in Approach A we estimate the long memory parameter with the approach of Percival & Walden as if the time series was *not* perturbed by noise.

5.3 An estimator for perturbed time series – Approach B

In contrast to the Approach A, the Approach B takes into account a first-order correction for the presence of noise. More specifically, the Approach B assumes that

$$\frac{\sigma_Z^2}{\sigma_E^2} \times [4\sin^2(\pi f)]^\delta \quad (52)$$

is „close to zero“ and f is „close to zero“. Then (see Eq. 38)

$$\log S_P(f) = -2\delta \log(2\pi f) + \log(\sigma_E^2) + \frac{\sigma_Z^2}{\sigma_E^2} \times (2\pi f)^{2\delta}. \quad (53)$$

Afterwards, we follow ideas and steps similar to those used in the construction of the PW estimator above (see Section 5.1). At this moment, the positive integers α and β have to be chosen in such a way that

- $2^{-(\alpha+1/2)}$ is „close to zero“.
- $\sigma_Z^2 / \sigma_E^2 \times [4\sin^2(\pi 2^{-(\alpha+1/2)})]^\delta$ is „close to zero“. Since the ratio σ_Z^2 / σ_E^2 and δ are not known in advance, the upper bound of $\sigma_Z^2 / \sigma_E^2 \times [4\sin^2(\pi 2^{-(\alpha+1/2)})]^\delta$ has to be estimated.
- $N - L_\beta \geq 0$, where L_β is the length of the wavelet filter $\{h_{\beta,t}\}$.

Then, for integer j satisfying $\alpha \leq j \leq \beta$ we can write

$$Q_j = (1 - 2\delta) \log(2\pi 2^{-(j+1/2)}) + \log(\sigma_E^2 / \sqrt{2}\pi) + \frac{\sigma_Z^2}{\sigma_E^2} \times [2\pi 2^{-(j+1/2)}]^{2\delta} + e_j, \quad (54)$$

where

$$Q_j \equiv \log(n_{P,j}^2) - \psi(\eta_j / 2) + \log(\eta_j / 2), \quad (55)$$

and

$$e_j = \log(\chi^2[\eta_j]) - \psi(\eta_j / 2) - \log(2), \quad (56)$$

where $\chi^2[\eta_j]$ is a chi-square random variable with η_j degrees of freedom where η_j is defined in Eq. 44. The error term e_j in Eq. 54 has a zero mean and variance equal to $\psi'(\eta_j/2)$. Moreover, e_j and e_i (for integer j and integer i satisfying $\alpha \leq j \leq \beta$, $\alpha \leq i \leq \beta$ and $i \neq j$) are approximately uncorrelated. The parameter δ can be estimated within the non-linear regression model of Eq. 54 via the weighted least squares estimator with the weight at the j th level given as $1/\psi'(\eta_j/2)$.

5.3.1 An iterative approach

We propose a way of estimating δ within the model of Eq. 54 which avoids using non-linear least squares. Since several approximations have already been used in the construction of the model of Eq. 54, this model is unlikely to be an exact description of reality. Therefore, it might be sufficient – instead of using non-linear regression – to apply a procedure consisting of the following steps:

Let us obtain the starting estimate d_0 of the long memory parameter δ as well as the starting estimate $s_{E,0}$ of the parameter σ_E . More specifically, d_0 is related to the weighted least squares estimate of the slope in the linear regression model of Eq. 50, whereas $s_{E,0}$ can be obtained as a function of the weighted least squares estimate of the intercept of the model of Eq. 50.

Consequently, the corrected estimates d_1 and $s_{E,1}$ of the long memory parameter δ and of the parameter σ_E are obtained from the following linear regression model

$$Q_j - \frac{\sigma_Z^2}{s_{E,0}^2} \times [2\pi 2^{-(j+1/2)}]^{2d_0} = (1-2\delta) \log[2\pi 2^{-(j+1/2)}] + \log(\sigma_E^2 / \sqrt{2\pi}) + e_j, \quad (57)$$

estimated by ordinary weighted least squares with the weight at the level j given as $1/\psi'(\eta_j/2)$.

Consequently, the final estimate d of δ is obtained from

$$Q_j - \frac{\sigma_Z^2}{s_{E,1}^2} \times [2\pi 2^{-(j+1/2)}]^{2d_1} = (1-2\delta) \log[2\pi 2^{-(j+1/2)}] + \log(\sigma_E^2 / \sqrt{2\pi}) + e_j, \quad (58)$$

estimated by ordinary weighted least squares with the weight at the level j given as $1/\psi'(\eta_j/2)$.

5.4 Monte Carlo simulations

We illustrate the properties of the Approach A (i.e., the approach of Percival & Walden applied to the time series as if the time series were not perturbed by noise) and the Approach B (which takes into account a first-order correction for the presence of noise) via a Monte Carlo simulation.

Several realizations of fractionally differenced processes of Eq. 29 (with different values of δ and σ_E^2) perturbed with a Gaussian white noise with variance $\sigma_Z^2 = 0.26$ (see Section 3.2) were generated. The following settings of the simulation were chosen:

- Length N : 512.
- Long memory parameter δ : 0.1, 0.2, 0.3, 0.4, 0.45, 0.49.
- Ratio σ_E^2/σ_Z^2 (where σ_E^2 is the variance of process $\{E_t\}$ of Eq. 29 and $\sigma_Z^2 = 0.26$; see Section 3.2): 0.1, 0.5, 1, 1.5, 2, 2.5, 3, 5, 10.
- Wavelet family: Haar. On the one hand, Haar wavelet filters are the shortest ones. Therefore, only a few wavelet coefficients are affected by the circularity assumption. On the other hand, Haar wavelet filters might be quite poor approximations to ideal band-pass filters for the respective frequency bands. Our choice of the Haar wavelet family is due to its simplicity.

- For each possible combination of $\{\delta, \sigma_E^2/\sigma_Z^2\}$ 500 realizations of length N of the process of Eq. 29 were generated perturbed with a Gaussian white noise with a variance $\sigma_Z^2 = 0.26$.
- $\alpha = 2, \beta = 9$.⁴ This choice of α and β corresponds to the case where it is assumed that the time series can be modelled well as a realization of a perturbed long memory process (with parameters δ, σ_E^2 and σ_Z^2) in the low frequency region as well as in the high frequency region.
- $\log(n_{Pj}^2)$ of this perturbed series was calculated for $\alpha \leq j \leq \beta$. The long memory parameter was estimated with Approach A and B.

In Fig. 1, we present the bias of the estimation of δ for both the approaches (Approach A and Approach B). Clearly, the Approach A suffers from a downward bias in the estimation of δ . This bias is especially large for small values of σ_E^2/σ_Z^2 . For example, for $\delta \geq 0.4$ and $\sigma_E^2/\sigma_Z^2 \leq 1.5$, the absolute bias is larger than 0.1. For a given δ the bias of the Approach A tends approximately to zero with increasing σ_E^2/σ_Z^2 . On the other hand, the Approach B seems to be superior to the Approach A in terms of bias. Even for $0.1 \leq \sigma_E^2/\sigma_Z^2 \leq 1.5$, the absolute bias is bounded by about 0.06 for $\delta = 0.1, 0.5, 1, 1.5, 2, 2.5, 3, 5, 10$. For $\sigma_E^2/\sigma_Z^2 \geq 2$ the absolute bias is not larger than 0.02 for $\delta = 0.1, 0.5, 1, 1.5, 2, 2.5, 3, 5, 10$.

We may thus conclude that estimating the long memory parameter following the Approach A (i.e., the Percival & Walden's approach as if the time series were not perturbed by noise) may lead to a large bias in the estimation of this parameter. This bias is substantially reduced if the Approach B is applied (i.e., a first-order correction for the presence of noise is applied). The reduction of the bias is not complete in the Approach B due to several approximations in the model building.

In Fig. 2, we present the root mean squared error (RMSE) of the estimation of δ for the Approaches A and B. Both the approaches generally lead to a higher RMSE for smaller values of σ_E^2/σ_Z^2 . We can see that the Approach B is superior to the Approach A in terms of RMSE for $\delta \geq 0.2$ for any considered value of σ_E^2/σ_Z^2 . This result (together with the results of the bias study above) suggests that the application of the Approach B (i.e., the first-order correction for the presence of noise) is desirable if the long memory parameter is close to 0.5. In the case of the long memory parameter being close to 0, the Approach A (i.e., no correction for the presence of noise) might be well acceptable.

In this paper, we have not explored the characteristics of both the approaches for other settings (i.e., for other choices of N, δ , the ratio σ_E^2/σ_Z^2 , the wavelet family, α and β). In spite of this fact, it can be largely expected that the Approach B (first-order correction for the presence of noise) might be superior to the Approach A (no correction for presence of noise) under many circumstances.

4 For small values of σ_E^2/σ_Z^2 , this choice of α and β does not satisfy the condition of $\sigma_Z^2 / \sigma_E^2 \times [4\sin^2(\pi 2^{-(\alpha+1/2)})]^\delta$ being close to zero. Despite of this, it might be instructive to see what the properties of the proposed estimators are in this case (i.e., for $\alpha = 2$ and $\beta = 9$ and for small values of σ_E^2/σ_Z^2).

6. Conclusions

Wavelets are an intriguing tool of data analysis. They may be utilized for the estimation of the long memory parameter in long memory processes: Percival & Walden (2002) summarize a wavelet methodology of the least squares estimation of the long memory parameter for fractionally differenced processes. Building up on their work, we have suggested an approach based on wavelets that may be used in the estimation of the long memory parameter in time series perturbed by noise such as log-Garman-Klass estimates of log volatility of financial markets. The proposed approach makes a first-order correction for the presence of noise. Our Monte Carlo simulations suggest that this correction leads to superior estimation when compared to a case where no correction for the perturbation is taken into account. In future research, it might be possible to exploit the fact that wavelets are capable of depicting time-varying properties of time series and our approach might be extended to the estimation of the time-varying long memory parameter in time series perturbed by noise.

Figure 1

The bias (on the y-axis) of the estimation of the long memory parameter δ in fractionally differenced processes (see the text for further specifications). Points connected with the solid line correspond to the Approach B, whereas points connected with the dashed line correspond to the Approach A. The ratio σ_E^2/σ_Z^2 is plotted on the x-axis. Different subplots correspond to different values of the long memory parameter δ . The top row corresponds to (from left to right) $\delta = 0.1$, $\delta = 0.2$ and $\delta = 0.3$. The bottom row corresponds to (from left to right) $\delta = 0.4$, $d = 0.45$ and $\delta = 0.49$.

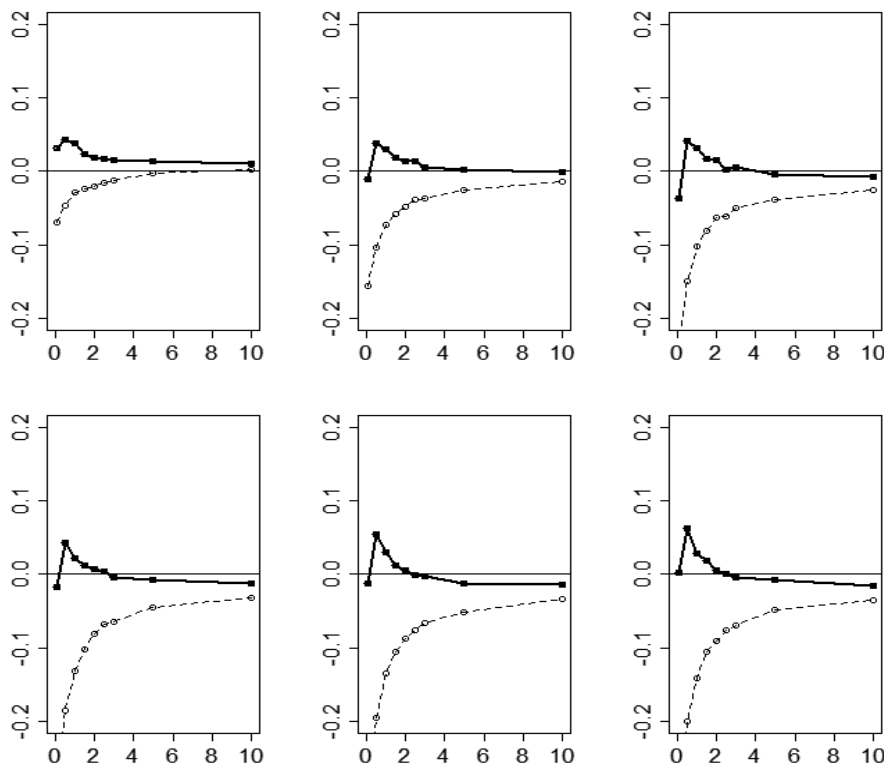
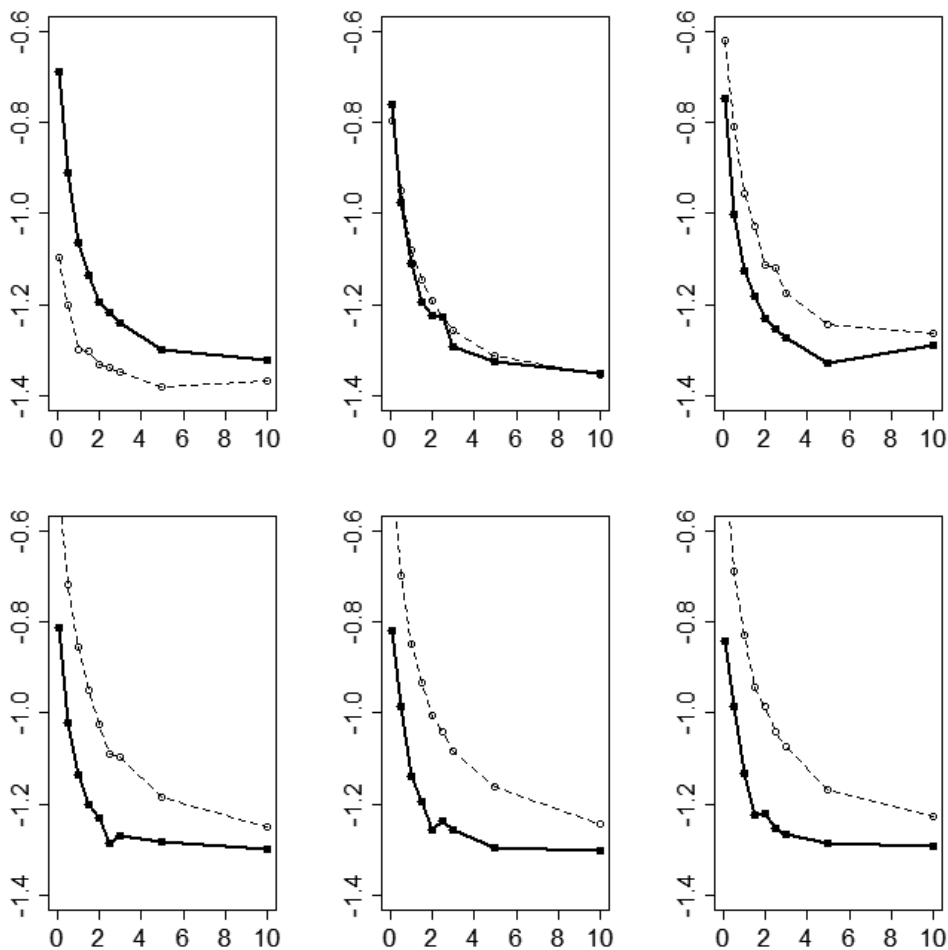


Figure 2

The decadic logarithm (on the y-axis) of the root mean squared error of the estimation of the long memory parameter δ in fractionally differenced processes (see the text for further specifications). The points connected with the solid line correspond to the Approach B, whereas the points connected with the dashed line correspond to the Approach A. The ratio σ_E^2/σ_Z^2 is plotted on the x-axis. Different subplots correspond to different values of the long memory parameter δ . The top row corresponds to (from left to right) $\delta = 0.1$, $\delta = 0.2$ and $\delta = 0.3$. The bottom row corresponds to (from left to right) $\delta = 0.4$, $\delta = 0.45$ and $\delta = 0.49$.



References

- ARTECHE, J. 2004. Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics*. 2004, vol. 119, no. 1, s. 131–154.
- ARTECHE, J. 2006. Semiparametric estimation in perturbed long memory series. *Computational Statistics & Data Analysis*. 2006, vol. 51, no. 4, s. 2118–2141.
- BAŠTA, M. 2010. Waveletová transformace a její aplikace při analýze ekonomických a finančních časových řad /Doktorská dizertační práce/. Praha : VŠE, Fakulta informatiky a statistiky, 2010.
- BERAN, J. 1994. *Statistics for Long-Memory Processes*. 1994. Chapman-Hall, 1994. ISBN 9780412049019.
- BREIDT, F. J.; CRATO, N.; DE LIMA, P. 1998. The Detection and Estimation of Long Memory in Stochastic Volatility. *Journal of Econometrics*. 1998, vol. 83, no. 1–2, s. 325–348.
- DEO, R.; HURVICH, C. 2001. On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Economic Theory*. 2001, vol. 17, no. 4, s. 686–710.
- GARMAN, M. B.; KLASS, M. J. 1980. On the Estimation of Security Price Volatilities from Historical Data. *The Journal of Business*. 1980, vol. 53, no. 1, s. 67–78.
- GENÇAY, R.; SELÇUK, F.; WHITCHER, B. 2002. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, 2002. ISBN 9780122796708.
- GEWEKE, J.; PORTER-HUDAK, S. 1983. The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*. 1983, vol. 4, no. 4, s. 221–238.
- GRANGER, C. W. J.; DING, Z. 1996. Varieties of Long Memory Models. *Journal of Econometrics*. 1996, vol. 73, no. 1, s. 61–77.
- GRANGER, C. W. J.; JOYEUX, R. 1980. An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*. 1980, vol. 1, no. 1, s. 15–29.
- HOSKING, J. R. M. 1981. Fractional Differencing. *Biometrika*. 1981, vol. 68, no. 1, s. 165–176.
- HURST, H. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*. 1951, vol. 116, s. 770–808.
- LOBATO, I. N.; SAVIN, N. E. 1998. Real and Spurious Long-Memory Properties of Stock-Market Data. *Journal of Business & Economic Statistics*. 1998, vol. 16, no. 3, s. 261–268.
- MOLNÁR, P. 2011. Properties of range-based volatility estimators. *International Review of Financial Analysis*. Forthcoming.
- PERCIVAL, D. B.; WALDEN, A. T. 2002. *Wavelet Methods for Time Series Analysis*. Cambridge : Cambridge University Press, 2002 (reprint). ISBN 9780521640687.
- SUN, Y.; PHILLIPS, P. 2003. Nonlinear log-periodogram regression for perturbed fractional processes. *Journal of Econometrics*. 2003, vol. 115, no. 2, s. 355–389.
- VIDAKOVIC, B. 1999. *Statistical Modeling by Wavelets*. Wiley-Interscience, 1999. ISBN 9780471293651.

WAVELETS AND ESTIMATION OF LONG MEMORY IN LOG VOLATILITY AND TIME SERIES PERTURBED BY NOISE

Abstract: Percival and Walden (2002) present a wavelet methodology of the least squares estimation of the long memory parameter for fractionally differenced processes. We suggest that the general idea of using wavelets for estimating long memory could be used for the estimation of long memory in time series perturbed by noise. One prominent example thereof is the time series of log-Garman-Klass estimates of log volatility of financial markets. The estimator of Percival and Walden (2002) is biased if the long memory time series is perturbed by noise. We propose a new estimator of the long memory parameter which combines (in its construction) the frequency-domain approach of Sun & Phillips (2003) and the approach of Percival & Walden (2002). We illustrate the properties of the proposed estimator via Monte Carlo simulations. The results show that the estimator may be useful for the estimation of the long memory in volatility.

Keywords: time series, long memory, volatility, wavelets, finance

JEL Classification: C49, G10