



Comparison of Hurst exponent estimation methods

Amjad H. Hamza⁽¹⁾

University of Baghdad, Baghdad, Iraq

amjad.hibtallah@uobaghdad.edu.iq

Munaf Y. Hmood⁽²⁾

University of Baghdad, Baghdad, Iraq

munaf.yousif@coadec.uobaghdad.edu.iq

Received:15/2/2021

Accepted: 18/4/2021

Published: JUNE/ 2021



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

Abstract:

Through recent years many researchers have developed methods to estimate the self-similarity and long memory parameter that is best known as the Hurst parameter. In this paper, we set a comparison between nine different methods. Most of them use the deviations slope to find an estimate for the Hurst parameter like Rescaled range (R/S), Aggregate Variance (AV), and Absolute moments (AM), and some depend on filtration technique like Discrete Variations (DV), Variance versus level using wavelets (VVL) and Second-order discrete derivative using wavelets (SODDW) were the comparison set by a simulation study to find the most efficient method through MASE. The results of simulation experiments were shown that the performance of the methods is relatively close, except for the SODDW method was the most efficient in MASE.

Paper type Categorize your paper under one of these classifications: *General review*.

Key Words: Fractional Brownian motion, Hurst exponent, Short memory, Long memory, Self-similarity, Discrete Wavelet, Long-range dependence, Short-range dependence.

⁽¹⁾ Assistant Instructor, Master in statistic

⁽²⁾ Professor, Ph.D. in Statistics

1- Introduction:

Fractional Brownian motion (FBM) provides an appropriate modeling framework for non-stationary self-similar stochastic processes with stationary increments. It has been widely used to model random phenomena related to different research fields.

The fractional Brownian motion was defined for the first time within Hilbert's space by Kolmogorov in (1940) where he called it Wiener Helix and was studied more broadly by Yaglom in (1958), but the designation belongs to the researcher's Mandelbrot and Van Ness wherein (1968) they explained the random integration of this process in its standard form when the Hurst parameter is 0.5, where the fractional Brownian is a generalization of this case, and unlike the standard form of this random process, the increase is not necessarily independent, as this process is a continuous-time, with zero mean and variance equal to (Nourdin, (2012)):

$$E[B_H(t)B_H(s)] = \frac{1}{2} (|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \quad (1)$$

We seek in this paper to set a comparison between the most known estimation methods in the field. The paper is structured as follows: Section 2, the fractional Brownian motion will be presented, in Section 3, we give an overview of Hurst parameter and present the estimation methods, and in Section 4, the simulation study will be conducted finally Section 5 the conclusions.

2- Fractional Brownian Motion:

Fractional Brownian motion (FBM) is a centered stochastic Gaussian process. Mandelbrot & Van Ness called it Brownian according to the biologist Robert Brown for his work in 1827 on pollen grains of *Clarkia pulchella* plant where Brown observed fine particles suspended in the water moving in a form Constant fluctuating movements, which he did not explain at the time., and to define it they use the different fractional integral of white noise:

$$B_H(t) = B_H(0) + \frac{1}{\Gamma(H+\frac{1}{2})} \left\{ \int_{-\infty}^0 [(t-s)^{H-1/2} - (-s)^{H-1/2}] dB_s + \int_0^t (t-s)^{H-1/2} dB_s \right\} \quad (2)$$

The most crucial difference between Brownian motion and fractional Brownian motion is in the increments, so in the classic process, the increments are independent but in fractional Brownian motion (FBM) are not, when $H > 1/2$ there is positive autocorrelation on other hands if $H < 1/2$ the autocorrelation is negative.

Where H is a Hurst parameter which was defined by both Mandelbrot and Van Ness and its value ranges between [0, 1] and shows the extent of the roughness of the movement, as its value increase smoother the signal becomes so the type of the random process is related to the value of this parameter.

The existence of the fractional Brownian motion comes from the presence of the centered Gaussian process. If we assume that there is a fundamental parameter $H > 0$, then there is a continuous central Gaussian process $B_H = (B_t^H)_{t \geq 0}$ With a known positive variance if and only if $H \leq 1$ (Nourdin, (2012)).

The main properties of fractional Brownian motion are:

1- Self-similarity is visually seen as the same pattern repeating both seen up close and seen from afar. In other words, there are minor versions of the larger pattern repeated inside larger patterns. This property of a random process is achieved if:

$$B_H(at) \sim |a|^H B_H(t) \quad (3)$$

This property is because the covariance function is homogeneous of order $2H$, making the process a fractional character.

2- Stationary increments which satisfied when:

$$B_H(t) - B_H(s) \sim B_H(t - s) \quad (4)$$

3- Long-range dependence or the long memory is a phenomenon that appears in the analysis of time-related data and expresses the slow exponential decay of the autocorrelation coefficient, and it appears clearly in FBM when $H > 1/2$.

$$\sum_{n=1}^{\infty} E [B_H(1)(B_H(n+1) - B_H(n))] = \infty \quad (5)$$

4- The regularity where the sample-paths are non-derivative at almost every point, i.e., all trajectories are locally Holder continuous for any function have d -dimensional Euclidian space is Holder continuous when there is a positive constant C and $\alpha > 0$ such that $|f(x) - f(y)| \leq C|x - y|^\alpha$ That is, for each trajectory path, for each $T > 0$, and each $\varepsilon > 0$, there is a constant value of C so that:

$$|B_H(t) - B_H(s)| \leq C|t - s|^{H-\varepsilon} \quad (6)$$

5- Semi-Martingale in the theory of probability, Martingale is a series of random variables (any random process) so that at any time, the conditional expectation of the present value excluding all previous values is equal to the first value.

$$E[X_{n+1}|X_1, \dots, X_n] = X_n \quad (7)$$

This property occurs in many random processes such as classic Brownian motion, i.e., when $H = 1/2$. It is so apparent, so it should be that this property is not realized in fractional Brownian motion because the covariance of this type not equal to zero, and many researchers have demonstrated this property when $H > 1/2$ and when $H < 1/2$.

3- Hurst parameter

This parameter is used to measure long memory in time series, indicating the autocorrelation in time series and decreasing rate as more excellent the lag between values become. This parameter was developed in hydrology to measure the Nile River flood rates and drought conditions, which were observed over a long period. Harold Edwin Hurst (1880–1978), whose study first revealed the behavior of this parameter, has been used extensively in fractional signals and is sometimes called a dependence index where it measures the direction or the relative slope for the series, which is known as a cluster in direction when the value of this parameter is $H > 1/2$, it indicates a positive correlation in the long-range series, which means that a high value in the series may be followed by another high value and tends to remain high in the future. The value of parameter $H < 1/2$ will have a succession of high and low values in adjacent

pairs, i.e., the high value will probably be followed by a low value, which may be followed by a low value after that, and this trend alternately may last for some time, but if the value of parameter $H = 1/2$ It refers to a series that is not entirely correlated may be the correlations in short periods is positive or negative in terms of absolute values of relationships degrade doubly quickly to zero as opposed to the energy law in the previous two cases (Mandelbrot & Van Ness, (1968)). Many researchers try to estimate this parameter, and we may consider the following methods that are the most common:

1- Rescaled range (R/S)

In 1951 Hurst use this method to estimate the value of the parameter that took his name, which is based on the self-similarity property to divide the time series $X_i, i = 1, \dots, n$ into several series and then calculate the average for each series as a first step.

$$M = \frac{1}{n} \sum_{i=1}^n X_i \quad (8)$$

After that, a modified series is constructed with the mean adjusted series $Y_i = X_i - Mean$ and then find the cumulative deviation of the series $Z_t = \sum_{i=1}^t Y_i, t = 1, \dots, n$ using the cumulative deviation we can calculate the range

$$R(n) = \max(Z_1, Z_2, \dots, Z_n) - \min(Z_1, Z_2, \dots, Z_n) \quad (9)$$

Then we calculate the standard deviation

$$S(n) = \sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2} \quad (10)$$

Finally, we calculate the rescaled range $R(n)/S(n)$ and find the mean; Now, it is possible to calculate the Hurst parameter according to (Nourdin, 2012)).

$$\hat{H} = \frac{\log R/S}{\log n} \quad (11)$$

2- Aggregate Variance (AV)

This method is based on the self-similarity property of the aggregate process $X^{(m)}$ to begin dividing the series into m of subseries and then calculate the mean for each series.

$$X_k^{(m)} = \frac{1}{m} (X_{km} + \dots + X_{(k+1)m-1}) \quad (12)$$

Where m can be calculated as follow

- $L = \frac{n}{5}$, size of each subseries
- $Q = 10^L$, generate equally spaced points
- $m =$

unique(Q), of length L where m take only unique equally spaced points from Q
For $k = 0, 1, \dots, L$. And because the self-similarity $X^{(m)}$ have the identical finite-dimensional distributions, so the next step is calculating the average of all mean values

$$\overline{X^{(m)}} = \frac{1}{M} \sum_{i=0}^{M-1} X_i^{(m)} \quad (13)$$

Where M is an integer and represents N / m for all subseries, and now we calculate the variance of the means:

$$Var(\overline{X_k^{(m)}}) = \frac{1}{M} \sum_{i=0}^{M-1} (X_i^{(m)} - \overline{X^{(m)}})^2 \quad (14)$$

This estimator considers as a biased estimator in the presence of autocorrelation. However, this bias disappears when M is large, i.e., when N large and m small.

The estimate of H can be found by plotting $Var(\widehat{X}_k^{(m)})$ with $\log m$ On the log-log scale. The slope is estimated by matching the fitting of the straight line to the points, and the H estimator is found by the slope estimator where $slope = 2\widehat{H} - 2$ (Abry et al., (2000)).

3- Absolute moments (AM)

This method was proposed by Taqqu in 1995 that the absolute moment's method is a generalization of the aggregate variance method where it uses the same principle so that $X^{(m)}$ has the same distribution of the finite dimension when $m^{H-1}X$ for m is significant and so:

$$AM_m = \frac{1}{M} \sum_{i=0}^{M-1} |X_i^{(m)} - \overline{X^{(m)}}|^n \quad (15)$$

For $n > 0$ represent the polynomial degree, the following constraint will hold.

$$\begin{aligned} E(AM_m) &\sim E \left| X_i^{(m)} - \overline{X^{(m)}} \right|^n (1 - C_n M^{n(H-1)}) \\ &= m^{n(H-1)} \left| X_i^{(m)} - E(X_i^{(m)}) \right|^n (1 - C_n M^{n(H-1)}) \end{aligned} \quad (16)$$

When m is large and $M \rightarrow \infty$ where C_n is constant, $E(AM_m)$ is proportional to $m^{n(H-1)}$ and the estimate of H can be calculated from the $slope = \widehat{H} - 1$ and estimated from a regression on the log-log scale. Also, for any value when $n > 0$, an estimate can be obtained and is usually used when $n = 1$, and the aggregate variance method is used if $n = 2$ (Abry et al. (2000)).

4- Discrete Variations (DV)

This method was proposed by Coeurjolly (2001) and used a filtering technique. Consider a stochastic process $X = X(1), \dots, X(n)$ at time $i = 1, \dots, n$ and define (a_q) as a filter of length $l + 1$ and the $p \geq 1$ order is the sequence that satisfies.

$$\sum_{q=0}^l a_q q^r = 0 \text{ for } r = 0, \dots, p-1 \quad (17)$$

Where it is not satisfying when $r = p$ since $\sum_{q=0}^l a_q q^p \neq 0$, for instance, we shall consider the following filters (Coeurjolly, 2001):

- $a = (-1, 1)$
- (18)
- $a = (1, -2, 1)$
- $a = (-0.09150635, -0.15849365, 0.59150635, 0. -034150635)$
-

Then define the process V^a as the vector V filtered with a and given for $i = l + 1, \dots, n$ as the sample filtration $Y = \{Y_i = \sum_{k=0}^i X_k: i = 0, \dots, N-1\}$ then

$$V_k^a = \sum_{q=0}^l a_q Y_{k-q}, \text{ for } k = l, \dots, N-1 \quad (19)$$

As an example of a filter with the order $p = 1$ is $a = (-1, 1)$ where V_k^a reduces to X_k , and this method is based on k-th final moments of discrete variations; now we define the mean:

$$S(k, a) = \frac{1}{N-l} \sum_{i=l}^{N-1} |V_i^a|^k \quad (20)$$

For $k > 0$ and assuming the sample distribution is Gaussian then the standard formula for kth moment for a Gaussian variable is:

$$E[S(k, a)] = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi}} N^{-kH} \{2Var(N^H V_1^a)\}^{k/2} \quad (21)$$

Now H can be estimated by solving the equation above for H where $E[S(k, a)]$ is replaced by its estimated value, and these calculations are possible because the $Var(N^H V_1^a)$ It does not depend on H , and this method gives a good estimate of the parameter, but the bias is high when $H < 1/2$ (Coeurjolly, 2001).

5- Higuchi method (HM)

This method was proposed by Higuchi in 1988, and the method is similar to the final moment with $n = 1$; instead of using non-crossing blocks, it uses a sliding window which makes the calculation more complicated than the other methods (Higuchi, (1988)). Beginning with binning the series into k subseries (can be calculated as in AV method) and calculate the normalized length for each subseries

$$L_m(k) = \frac{N-1}{k^2 M} \sum_{i=1}^M |X_{m+ik} - X_{m+(i-1)k}| \quad (22)$$

Where $M = \frac{N-m}{k}$ and N represent the original series length; after this step, we have to calculate the normalized length for the whole series

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k) \quad (23)$$

Where m is the number of observations inside every k , and by the self-similarity property

$$E[L(k)] \sim C_H k^{-D} \quad (24)$$

The estimate of H is found by plotting $L(k)$ in the log-log scale versus m and by adding 2 to the slope of the corresponding straight line.

6- Periodogram method (PM)

The parameter is estimated in this method by fitting the straight line in the spectral field based on observations that follow the behavior of the spectrum function $C_f |\lambda|^{1-2H}$ for $|\lambda| \rightarrow 0$ where (Abry et al., (2000)):

$$I(\lambda) = \sum_{j=-(N-1)}^{N-1} \hat{\gamma}(j) e^{ij\lambda} \quad (25)$$

Also, we can rewrite the previous formula of $I(\lambda)$ to calculate the periodogram as follows:

$$I(\lambda) = \frac{1}{N} \left| \sum_{k=0}^{N-1} (X_k - \bar{X}) e^{ij\lambda} \right|^2 \quad (26)$$

Where the Periodogram is symmetric around zero just as a spectrum function, and its function is unbiased to the spectrum function.

$$\lim_{N \rightarrow \infty} E|I(\lambda)| = f(\lambda) \quad (27)$$

$I(\lambda)$ is calculated for $k = 1, \dots, N$ where $\lambda_k = \frac{\pi k}{N}$ and the periodogram values are plotted on the log-log scale and the parameter is estimated by drawing a straight line corresponding to the data, which has a slope angle of $1 - 2H$.

7- Variance of the regression residuals (VRR)

In 1994, Peng et al proposed this method, which is based on dividing the series into m -sized segments within each segment k . The partial sum are regressed on the line $\hat{\alpha}^{(k)} + \hat{\beta}^{(k)} X_i$ where the residuals of the regression line is calculated under different aggregate level, as follows (Peng et al, (1994)):

$$\hat{\epsilon}_i^{(k)} = \sum_{j=km}^{km+i-1} Y_j - \hat{\alpha}^{(k)} - \hat{\beta}^{(k)} X_i \quad (28)$$

Where: $\hat{\epsilon}_i^{(k)}$: residuals, Y_j : dependent variable, $\hat{\alpha}^{(k)}$: intercept, $\hat{\beta}^{(k)}$: slope, X_i : explanatory variable

Thus, the variance of the residuals is calculated for each segment as the sample mean variation on all segments is proportional to m^{2H} . Where $\hat{H} = \frac{1}{2} * slope$ can be found by plotting the variance of the residuals versus m on a log-log scale. This method consider as one of the scaled windowed variance methods.

8- Variance versus level using wavelets (VVL)

The slow decomposition of the correlation structure in the long memory and the absence of a normal description of processes of type $1/f$ until the appearance of the model proposed by Mandelbrot and Van Ness, which they called the fractional Brownian motion. Also, this signal was characterized by self-similarity and here Flandrin looks at the possibility of studying such signals from the wavelet view, which may give a new perspective on this topic. Flandrin in (1992) propose variance versus level using wavelets (Haar's wavelet) and here we will discuss in this topic details of how to calculate the estimator in this method.

As we mentioned in the previews section, FBM is the normal extension of classic Brownian motion where it is Gaussian with zero mean and non-stationary and it depends on one parameter on its behavior, which is the Hurst, and this non-stationary can be noted through the covariance structure, which take the following form in equation (1). The variance can be concluded as

$$var(B_H(t)) = |t|^{2H} \quad (29)$$

Because this process is non-stationary in covariance, it does not have a spectrum, but an average spectrum, that can be written as follows (Flandrin, 1992)

$$P_{B_H}(\omega) = \frac{1}{|\omega|^{2H+1}}, (\omega \text{ represent the observations}) \quad (30)$$

Also, this process has stationary-increment despite its non-stationary, as these increases are self-similar, and the parameter H indicates the extent of the roughness of this signal and when analyzing the fractional Brownian motion. Now suppose there is a FBM $B_H(t)$ transformed using DWT and the wavelet coefficients for the signal was

$$d_{j,k} = 2^{-j/2} \int_{-\infty}^{+\infty} B_H(t) \psi(2^{-j}t - k) dt \quad (31)$$

As the wavelet function justify the admissibility condition (vanishing moments)

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (32)$$

Wavelet coefficients represent the details of the differences between approximate coefficients, so instead of calculating coefficients through inner product, they can be calculated repeatedly from separate consecutive filters starting with a preliminary event $a_{0,k}$ at a specific accuracy level.

At each scale 2^j , the wavelet coefficients constitute a discrete sequence of random coefficients, and so the entire coefficients will be an orthogonal system, and there is no previous reason for the coefficients to be uncorrelated and the covariance can be written as

$$E(d_{j,k} d_{m,n}) = \frac{1}{2} (- \int_{-\infty}^{+\infty} A_\psi(2^{j-m}, \tau - (2^{j-m}k - n)) |\tau|^{2H} d\tau) (2^m)^{2H+1} \quad (33)$$

Where

$$A_\psi(\alpha, \tau) = \sqrt{\alpha} \int_{-\infty}^{+\infty} \psi(t) \psi(\alpha t - \tau) dt \quad (34)$$

Which is called the wide-band ambiguity function of $\psi(t)$ or the reproducing kernel, in other words, the wavelet transformation of the wavelets themselves. This leads to two results

1- In a time field that is self-similar and stationary, each $E(d_{j,k}d_{j,n})$ is the unique function of $k - n$

$$E(d_{j,k}d_{j,n}) = \frac{1}{2} \left(- \int_{-\infty}^{+\infty} \gamma_{\psi}(\tau - (k - n)) |\tau|^{2H} d\tau \right), \text{ where } \gamma_{\psi} = A_{\psi}(1, \tau) \quad (35)$$

2- In the scale field which is stationary because each k and n is equal to $2^{j-m}k$ then $E(d_{j,k}d_{m,n})$ is the unique function of $j - m$

$$E(d_{j,k}d_{m,2^{j-m}k}) = \frac{1}{2} \left(- \int_{-\infty}^{+\infty} A_{\psi}(2^{j-m}, \tau) |\tau|^{2H} d\tau \right) (2^{j-m})^{-(H+\frac{1}{2})} \quad (36)$$

It seems that the self-similarity and stationary in the scale field are natural, but the stationary in time for wavelet coefficients of the fractional Brownian motion occurred because the mother wavelet necessarily has a zero mean according to the admissibility condition, and so the FBM was centered around the zero frequency and could not appear in a band-pass analysis method despite hence, the non-stationary of the low frequency of the FBM if the wavelet analysis includes the scale function we will obtain

$$\text{var}(a_{j,k}) = \frac{1}{2} \left(- \int_{-\infty}^{+\infty} (\gamma_{\psi}(\tau) - 2\phi(\tau - k)) |\tau|^{2H} d\tau \right) (2^j)^{2H+1} \quad (37)$$

Which is self-similar but time-dependent, that is, the detail coefficients are stationary, but approximate coefficients are not. Thus, the variance of the detail coefficients is

$$\text{var}(d_{j,k}) = \frac{1}{2} V_{\psi}(H) (2^j)^{2H+1} \quad (38)$$

Whereas, the constant $V_{\psi}(H)$ depends on the selected wavelet and the input signal

$$V_{\psi}(H) = - \int_{-\infty}^{+\infty} \gamma_{\psi}(\tau) |\tau|^{2H} d\tau \quad (39)$$

From this equation and from the power-law distribution of wavelet coefficients

$$\log_2(\text{var}(d_{j,k})) = (2H + 1)j + \text{constant} \quad (40)$$

Thus, the Hurst parameter of the FBM can be easily found by the slope of the covariance drawn as a function of scales in the log-log plot (least square method).

Now, if we assume that the signal is a FBM, when the parameter is less than half, we will find the correlations quickly decay, but if they are greater than half, we will find that they are slowly decay. In general, the decay of correlations depends mostly on the number of vanishing moments in the wavelet used. The higher the number of vanishing moments, we will notice that we obtain coefficients free of correlations.

Algorithm

- 1- Transform the FBM signal using Haar wavelet.
- 2- Find the detail coefficients for the transformed signal.
- 3- Calculate the variance for the detail coefficients.
- 4- Perform regression to calculate the slope for the detail coefficients variance.
- 5- Use the following equation to find the Hurst exponent.

$$\hat{H} = \frac{(\text{slope} - 1)}{2} \quad (41)$$

9- Second order discrete derivative using wavelets (SODDW)

In recent years, many mathematical and physical processes and systems have emerged to cover natural phenomena. In the biology, DNA chains, change in heartbeat and nerve connections to the spine appeared and in physics turbulence modeling and hydrology and physics of solid-state materials as well as human activities such as Internet traffic in communications and financial. These phenomena have shown the behavior of the invariance in scale, which means that there is a relationship between different scales. Usually two main characteristics appear in this behavior, self-similarity and long-range dependence, and here researchers have found a way to drive a method capable of estimating the scaling parameter, which is the Hurst parameter, and in this direction, Abry et al in (2000) have found a way to estimate this parameter using discrete wavelets (Symlet 5) and here we will discuss in this topic details of how to calculate the estimator in this method.

Beginning with the definition of wavelet function

$$\psi_{j,k}(t) = \frac{1}{2^{j/2}} \psi(2^{-j}t - k) = 2^{-\frac{j}{2}} \psi(2^{-j}(t - 2^j k)), j, k \in \mathbb{Z} \quad (42)$$

Where j, k is dilations and translations of ψ , so the function $\psi(2^{-1}t)$ for example is the dilation of ψ by two units and the function $\psi(t - k)$ is the translation of ψ to the right by k units so if ψ have support in the interval $[0,1]$ then the function $\psi(2^{-1}t - 3) = \psi(2^{-1}(t - 6))$ support will be in the interval $[6,8]$ where 2^j and j is the scale and octave respectively. Note that the positive values of j is for expansion and negative values for contraction, where $2^{j/2}$ is the normalization that ensure for every $j, k \in \mathbb{Z}$

$$\int \psi_{j,k}^2(t) dt = \int \psi^2(t) dt \quad (43)$$

That allow the norm of $L^2(\mathbb{R})$ to preserve the signal energy, where the detail coefficient for the wavelet is

$$d_{j,k} = \int_{\mathbb{R}} X(t) \psi_{j,k}(t) dt \quad (44)$$

Where it's encoding the difference information between adjacent scales around 2^j in the time $2^j k$.

The discrete wavelet transformation of Daubchies is a multiresolution analysis, which is a special class of wavelets because of its mathematical properties, which also gave the possibility to the emergence of a pyramidal algorithm, which is one of the important characteristics of this transformation in addition to the number of zero moments, regularity, time support and frequency, which can be easily controlled. Below are two important features in this transformation that will be useful for the proposed method.

1- Wavelet function can be defined through the scale function.

2- Wavelet and scale function justify the two-scale equations.

$$\phi\left(\frac{t}{2}\right) = \sqrt{2} \sum_k h_k \phi(t - k)$$

(45)

$$\psi\left(\frac{t}{2}\right) = \sqrt{2} \sum_k g_k \psi(t - k)$$

From these two equations it's clear that the approximate and detail coefficients $a_{j,k}$ and $d_{j,k}$ respectively at the scale j can be calculated through the approximate coefficient $a_{j-1,k}$. Where these coefficients are of great importance in this method to estimate the Hurst parameter through the following properties:

1- The detail coefficients $d_{j,k}$ is the same for $X(t)$ and $X(t) + P(t)$ where P is a polynomial of $N - 1$ and ψ have N vanishing moments.

$$\int_{\mathbb{R}} P(t) \psi_{j,k}(t) dt = 2^{\frac{j}{2}} \int_{\mathbb{R}} P(2^j(u+k)) \psi(u) du = 0 \quad (46)$$

We notice from this equation that $P(2^j(u+k))$ is a polynomial of $N - 1$, and that is mean that DWT does not affect by polynomials.

2- If we have a stochastic process $X(t)$ that have stationary increment where the finite probabilities for $X(t+n) - X(t)$ does not depend on t then $d_{j,k}$ is a stationary sequence that can be proved easily by assuming $j = 0$ for the ease of the calculation. We see that the sequence does not depend on k , which indicates it is stationary.

3- When we have a self-similar stochastic process with the parameter H

$$d_{j,k} = \int_{\mathbb{R}} X(2^j u) 2^{-j/2} \psi(u-k) 2^j du = 2^{j(H+\frac{1}{2})} d_{0,k} \quad (47)$$

4- When we have a self-similar and stationary increment stochastic process with the parameter H with zero mean and known variance, then

$$E(d_{j,k}) = 0 \text{ and } E(d_{j,k})^2 = C 2^{j(2H+1)} \quad (48)$$

Where $C = E(d_{0,0})^2$, and by taking the logarithm of the two sides, we get a linear function of j with a slope of $2H + 1$. As for the covariance structure of the wavelet coefficients for the parameter H , we find a correlation, but this correlation tends continuously and repeatedly to zero at the more considerable lags if N is large enough.

5- Let $Y(t)$ be a stationary stochastic process and v_Y and p_Y variance and spectrum function respectively then the wavelet function $\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k)$ justify Fourier transform and the covariance for wavelet coefficients of the process Y is

$$\begin{aligned} E(d_{j,k} d_{j',k'}) &= \int_{\mathbb{R}} \int_{\mathbb{R}} v_Y(u) \psi_{j,k}(\omega) \psi_{j',k'}(u+\omega) d\omega du \\ (49) \quad &= \int_{\mathbb{R}} p_Y(\omega) 2^{\frac{j}{2}} 2^{\frac{j'}{2}} f(2^j \omega) f^*(2^{j'} \omega) e^{-i2\pi(k2^j - k'2^{j'})\omega} d\omega \end{aligned}$$

Where f^* the complex conjugate for Fourier transform, therefor the variance of the details is

$$E(d_{j,k}^2) = \int_{\mathbb{R}} p_Y(\omega) 2^j |f(\omega)|^2 d\omega \quad (50)$$

Now if the process $Y(t)$ has a long memory, and by adopting the previous equation and the definition of the spectrum function, it will be seen that the variance of wavelet coefficients redistributes the convergence using the definition of long memory

$$E(d_{j,k}^2) \sim c_g 2^{j\gamma} \int_{\mathbb{R}} |\omega|^{-\gamma} |f(\omega)|^2 d\omega \quad (51)$$

This convergence in long memory is the base of estimating the parameter γ where $\gamma = 2H - 1$.

Thus, from this relationship, we note that the number of vanishing moments should not be less than $N > H - 1/2$. So if we use a wavelet and $N = 2$ or 3 will allow a faster decrease, that is, a very large N is undesirable since the wavelets will become less localized.

From the points above, we note that wavelet coefficients for self-similar and long memory signals have the same basic properties and can be summarized as follows:

- Stationary in fixed scales.
- Short memory when the following equations are achieved

$$E d_{j,k_1} d_{j,k_2} \leq C |k_1 - k_2|^{2(H-N)}$$

(52)

$$E(d_{j,k} d_{j',k'}) \approx |k - k'|^{\gamma-2N-1}$$

- Power-law in the wavelet defines the phenomenon of scale change, according to the two equations

$$E(d_{j,k}^2) = C 2^{j(2H+1)} \quad (53)$$

$$E(d_{j,k}^2) \sim c_g 2^{j\gamma} \int_{\mathbb{R}} |\omega|^{-\gamma} |f(\omega)|^2 d\omega$$

These properties appear in the study of phenomena that possess the properties of self-similarity and long memory using wavelets because they contain at least one vanishing moment and that the wavelets are built through dilating the signal. These properties can be used when studying any type 1 / f signal and fractional processes, and stochastic multiplicative cascades.

The variations of the details at all scales 2^j , when they are stationary, they represent a random process X , which is a kind of wavelet spectrum. The large values of j are due to the low frequencies and their small values to the high frequencies. The properties we mentioned earlier pave the way for the logarithm of this spectrum, since a straight line in this range if represented by scale, its slope will be the parameter to be estimated, which depends on $s_j = \log_2(E(d_{j,\cdot}^2))$, this method is called an exact logscale diagram. Whereas, $E(y_j) = s_j$ and y_j are the random variables under study, and to obtain the estimator first, we assume the existence of an unbiased estimate of the variance, as follows:

$$v_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_{j,k}|^2 \quad (54)$$

The logarithm of this variable will be an estimate of s_j , but it will be biased due to nonlinearity, and therefore the logarithm will be $E[\log(\cdot)] \neq \log(E[\cdot])$ to address this problem, a small amount of bias will be added to correct it, which is $g(j)$ Thus, we will define y_j as follows:

$$y_j = \log_2(v_j) - g(j) \quad (55)$$

It is worth mentioning that Abry follows Istas & Lang (1997) careful steps in estimation but uses wavelet filters instead for more accuracy, which stabilizes the estimator in the sense of unique characteristics wavelet has. Istas formula count on finding the slope using a quadratic variation, and Abry count on finding an unbiased estimate for the detail coefficients variance that can be used to calculate an estimate for the Hurst exponent as follow

$$\hat{H} = \frac{1}{2} \log_2 (E(d_{j,k}^2)) \quad (56)$$

Algorithm

- 1- Find the high decomposition filter of sym5 wavelet type.
- 2- Find the half-band high decomposition filter of sym5 wavelet type.
- 3- Transform the FBM signal independently using (1) and (2) filters.
- 4- Cut the first length of the new signals using the length of the filter.

5- Find the variance for both transformed signals using (55).

6- Calculate the Hurst exponent using:

$$\hat{H} = \frac{1}{2} \log_2 (E(d_2^2)/E(d_1^2)) \quad (57)$$

$$v_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_{j,k}|^2 \quad (54)$$

The logarithm of this variable will be an estimate of s_j , but it will be biased due to nonlinearity, and therefore the logarithm will be $E[\log(\cdot)] \neq \log(E[\cdot])$ to address this problem, a small amount of bias will be added to correct it, which is $g(j)$ Thus, we will define y_j as follows:

$$y_j = \log_2(v_j) - g(j) \quad (55)$$

It is worth mentioning that Abry follows Istas & Lang (1997) careful steps in estimation but uses wavelet filters instead for more accuracy, which stabilizes the estimator in the sense of unique characteristics wavelet has. Istas formula count on finding the slope using a quadratic variation, and Abry count on finding an unbiased estimate for the detail coefficients variance that can be used to calculate an estimate for the Hurst exponent as follow

$$\hat{H} = \frac{1}{2} \log_2 (E(d_{j,k}^2)) \quad (56)$$

Algorithm

7- Find the high decomposition filter of sym5 wavelet type.

8- Find the half-band high decomposition filter of sym5 wavelet type.

9- Transform the FBM signal independently using (1) and (2) filters.

10- Cut the first length of the new signals using the length of the filter.

11- Find the variance for both transformed signals using (55).

12- Calculate the Hurst exponent using:

$$\hat{H} = \frac{1}{2} \log_2 (E(d_2^2)/E(d_1^2)) \quad (57)$$

4- A simulation study

In this section, we will conduct simulation experiments of fractional Brownian motion, where the wavelet synthesis method proposed by Sellan, Mayer, and Abry (1996) will be used for generating the mentioned random process, and then we will the fore mentioned methods in section three and compare them through MASE. The length of the series to be generated will be as $n = 100, 200$; we will repeat the process with $rep = 500$ for the sake of increasing the accuracy in the estimation process. As for the wavelet used for the generation, $s = 2^{10}, 2^{12}$. Taking into account that the estimation process for the Hurst parameter will be for the levels $H = [0.1, 0.5, 0.9]$. The mean average square error and the bias of the estimator will be calculated as follows

$$\begin{aligned} MASE(\hat{H}) &= Var(\hat{H}) + Bias^2(\hat{H}) \\ &= \left(\frac{1}{4} \sum_j \sigma_j^2 \theta_j^2\right) + (E(\hat{H}) - H)^2 \end{aligned} \quad (58)$$

Where $\theta_j = \frac{\sum(S*j - S_j)/\sigma_j^2}{S S_{jj} - S_j^2}$, and $S = \sum \frac{1}{\sigma_j^2}$, $S_j = \sum \frac{j}{\sigma_j^2}$, $S_{jj} = \sum \frac{j^2}{\sigma_j^2}$

$$Var(y_j) = \sigma_j^2 = \frac{\zeta(2, \frac{N_j}{2})}{\ln^2(2)}, E(y_j) = 2(H + 1) + constant \quad (59)$$

Where $\zeta\left(2, \frac{N_j}{2}\right)$ is a generalized Riemann Zeta function (Power and Turvey, (2010)). which is a function of a complex variable and can be defined as

$$\zeta\left(2, \frac{N_j}{2}\right) = 2^{-\frac{N_j}{2}} \quad (60)$$

Also, a new random series had been generated each time before the estimation in order to get the best view of the performance of the method used and at different levels of the random process.

Table (1) of Hurst est. bias and MASE comparison

method	H	$s = 2^{10}$				$s = 2^{12}$			
		n=100		n=200		n=100		n=200	
		Bias	MASE	Bias	MASE	Bias	MASE	Bias	MASE
R/S	0.1	0.0621	0.1325	0.0922	0.2121	0.0876	0.2230	0.0988	0.2111
	0.5	0.2141	0.2235	-0.1780	0.5123	0.2274	0.1530	-0.1532	0.3421
	0.9	0.3101	0.1215	0.2280	0.1190	0.2905	0.2230	0.2309	0.1099
AV	0.1	-0.0145	0.2564	0.0243	0.1971	0.0412	0.2231	0.0359	0.1432
	0.5	0.2004	0.1190	-0.1999	0.3098	0.2220	0.4003	-0.1088	0.2177
	0.9	-0.1131	0.2168	0.2510	0.1299	-0.1409	0.2531	-0.1092	0.2001
AM	0.1	0.0510	0.1327	-0.0910	0.2911	0.0544	0.1422	-0.0732	0.2555
	0.5	0.1475	0.2711	0.1672	0.3141	-0.2310	0.1821	0.1141	0.1950
	0.9	-0.1289	0.3120	0.2100	0.2309	0.2109	0.2120	0.1661	0.2275
DV	0.1	-0.1405	0.0828	0.1003	0.0984	-0.1155	0.0900	-0.1170	0.1208
	0.5	-0.1912	0.2768	-0.3002	0.2731	-0.1971	0.0221	-0.1655	0.1417
	0.9	0.0954	0.0606	0.1138	0.0478	0.1073	0.0190	-0.1847	0.1148
HM	0.1	-0.1535	0.6028	-0.3673	0.2984	-0.1925	0.0990	-0.1120	0.1198
	0.5	0.1002	0.1748	0.3222	0.1831	-0.1271	0.1121	-0.1565	0.1397
	0.9	-0.1094	0.5906	0.1118	0.1578	0.1813	0.3099	-0.1147	0.2128
PM	0.1	0.0445	0.3638	-0.4773	0.1004	-0.0755	0.0344	-0.1432	0.1291
	0.5	-0.1599	0.3218	-0.3072	0.2611	-0.1521	0.1117	-0.1292	0.3590
	0.9	0.2954	0.1746	0.1218	0.4878	0.2593	0.1799	-0.1007	0.1948
VRR	0.1	-0.0545	0.1148	-0.0973	0.4404	0.0285	0.2560	-0.1200	0.1098
	0.5	-0.1874	0.3268	-0.3112	0.6540	-0.1721	0.3421	-0.2165	0.1491
	0.9	0.4154	0.1716	0.1237	0.2570	-0.1193	0.3490	-0.1077	0.3628
VVL	0.1	-0.1045	0.0628	-0.5673	0.1984	0.1055	0.0300	-0.1470	0.0298
	0.5	-0.1512	0.0768	-0.3802	0.2731	-0.1271	0.0421	-0.1865	0.0497
	0.9	-0.0454	0.0706	0.1038	0.0578	-0.1093	0.0399	-0.1147	0.0128
SODDW	0.1	0.0192	0.0323	0.0878	0.0351	0.0069	0.0209	0.0462	0.0143
	0.5	0.0029	0.0251	0.0044	0.0211	-0.1002	0.0275	-0.0280	0.0110
	0.9	0.0228	0.0192	-0.0014	0.0248	-0.0916	0.0448	-0.0148	0.0233

Through Table (1), we notice that all the methods give the same results relatively where they deepened on the deviations slope in estimation. In addition, the Discrete Variations (DV) and variance versus level (VVL) also give the same results where they depend on filtration technique except for Second-order discrete derivative using wavelet (SODDW) that use both ways by filtering the data into two sets band and half band, and by finding the variances slope, it was capable of finding the lowest MASE among all estimation methods.

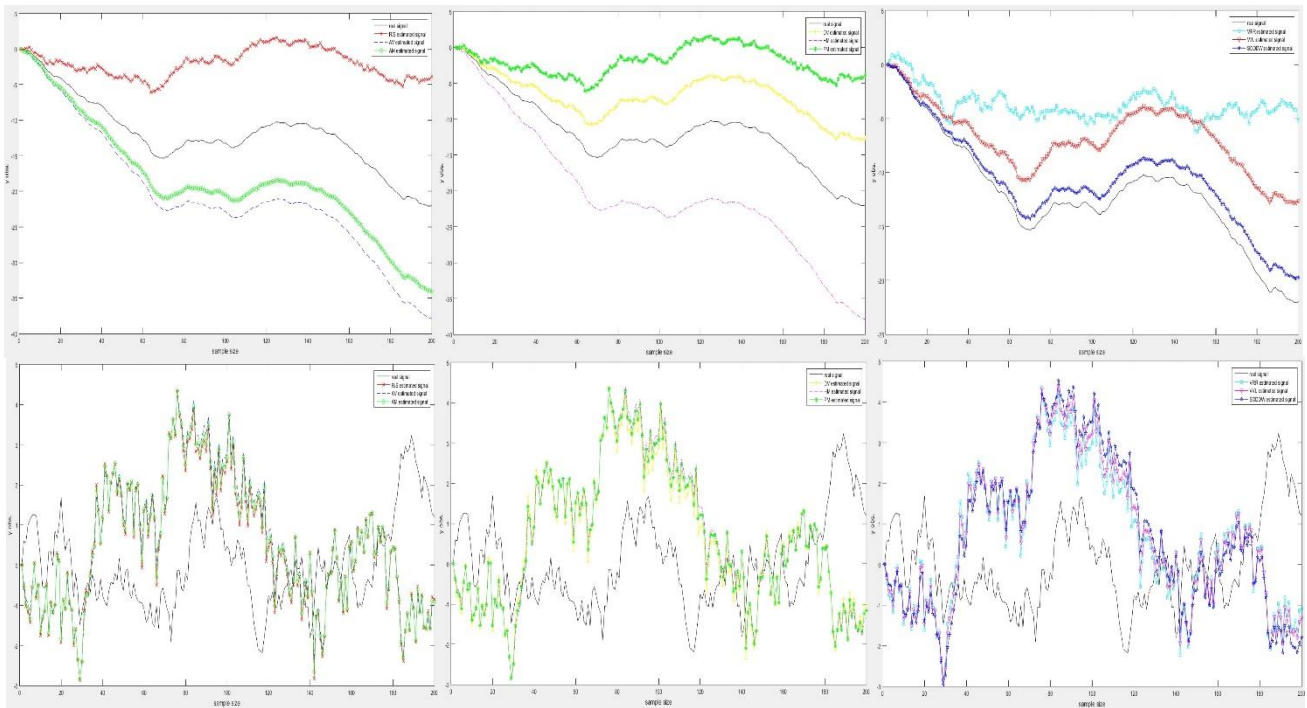


Figure 1: the left figure in the first row is $\hat{y}(t)$ for self-similarity signal at sample size $n=200$ and FBM wavelet synthesis scale $s = 2^{10}$ with Hurst exponent .9 as black line compared to R/S as red and AV as blue and AM as green lines respectively, the following figure in the middle is $\hat{y}(t)$ for DV as yellow and HM as magnet and PM as green lines, respectively, the last figure on the first row is $\hat{y}(t)$ for VRR as cynic and VVL as red and SODDW as blue lines, respectively, the left figure in the second row is $\hat{y}(t)$ for self-similarity signal at sample size $n=200$ and FBM wavelet synthesis scale $s = 2^{12}$ with Hurst exponent .5 as black line compared to R/S as red and AV as blue and AM as green lines respectively, the following figure in the middle is $\hat{y}(t)$ for DV as yellow and HM as magnet and PM as green lines, respectively, the last figure on the first row is $\hat{y}(t)$ for VRR as cynic and VVL as red and SODDW as blue lines, respectively.

Conclusion:

To detect the presence of long or short memory in a time series, the estimation of the Hurst parameter is necessary. Here we took the most using estimation methods to analyze fractional Brownian motion, a second-order stochastic process.

The current estimation methods depend totally on the slope of the deviations to calculate the Hurst parameter, which makes them sensitive to high noise and has weak accuracy. On the other hand, Coeurjolly uses a filtration technique to remove the noise from the signal before calculating the estimate; this modification makes his method dominate the others. However, using compact support orthonormal filters make another jump for the estimation of these signals, which make the SODDW give another view on finding a better estimate; at the same time, it noticed that the scale level used for simulation has a significant effect on this method because of the high smoothing level.

References

1. Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D., (2000), "Wavelets for the analysis, estimation, and synthesis of scaling data", In: Park K, Willinger W (eds) Self-similar network traffic and performance evaluation. Wiley, New York, pp 39–88.
2. Abry, P. and Sellan, F., (1996), "The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: remarks and fast implementation", Applied and computational harmonic analysis, Vol. 3, pp. 377-383.
3. Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W., (1995), "Long-range dependence in variable-bit-rate video traffic", IEEE Trans. on Communications, Vol. 43, pp. 1566-1579.
4. Coeurjolly, J. F., (2000), "Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study, Journal of Stat. Software, 5. Available from <http://www.jstatsoft.org>.
5. Coeurjolly, J. F., (2001), "Estimating the parameter of a fractional Brownian motion by discrete variations of its sample paths", Kluwer Academic Publishers, pp. 15-39.
6. Flandrin, P., (1992), "Wavelet analysis and synthesis of Fractional Brownian Motion", IEEE Transactions on Information Theory, Vol. 38, No. 2, pp. 910-917.
7. Higuchi, T., (1988), "Approach to an irregular time series on the basis of the fractal theory", Physica D, Vol. 31, pp. 277-283.
8. Hmood, M. Y. and Jumma, M., (2020), "Use of Le'vy's Model to Simulate Stock Dividends Actual Estimations of Some Iraqi Banks", May 2020 International Journal of Psychosocial Rehabilitation Vol. 24, No. 9, pp. 83-97.
9. Hmood, M. Y., (2016), "Using the Le'vy Model to estimate the stock returns for some Iraqi banks" April 2016, Journal of Economics and Administrative Sciences, Vol. 22, No. 94, pp. 460- 479.
10. Hmood, M. Y. and Amir, Y., (2020), "Estimation of return stock rate by using wavelet and kernel smoothers" Periodicals of Engineering and Natural Sciences (PEN), Vol. 8, No. 2, pp. 602-612.

11. Istas, J., Lang, G., (1997), “Quadratic variations and estimation of the local Holder index of a Gaussian process”, *Annales de l’I.H.P. Probabilités et statistiques*, Vol. 33, pp. 407-436
12. Mandelbrot, B. B. & Van Ness, J., (1968), “Fractional Brownian motion, fractional noises and applications”, *IBM Watson Research Center*, Vol. 10, No. 4, pp. 422-437.
13. Mielniczuk, J. and Wojdylo, P., (2007), “Estimation of Hurst exponent revisited”, *Comput. Stat. Data Anal.*, Vol. 51, pp. 4510-4525.
14. Nourdin, I., (2012), “selected aspects of fractional Brownian motion”, *B&SS – Bocconi & Springer Series*.
15. Peng, C.K., Buldyrev, S.V., Simons, M., Stanley, H.E. and Goldberger, A.L., (1994), “Mosaic organization of DNA nucleotides”, *Physical Review E*, Vol. 49, pp. 1685-1689.
16. Power, G. J. and Turvey, G. C., (2010), “Long-range dependence in the volatility of commodity futures prices: wavelet-base evidence”, *Phys. A.*, Vol. 389, pp. 79-90.





مقارنة اساليب تقدير أس هارست

أ.د. مناف يوسف حمود⁽²⁾
جامعة بغداد، بغداد، العراق

munaf.yousif@coadec.uobaghdad.edu.iq

م.م. أمجد هبة الله حمزة⁽¹⁾
جامعة بغداد، بغداد، العراق

amjad.hibtallah@uobaghdad.edu.iq

Received:15/2/2021

Accepted: 18/4/2021

Published: JUNE/ 2021

هذا العمل مرخص تحت اتفاقية المشاع الابداعي نسب المصنّف - غير تجاري - الترخيص العمومي الدولي 4.0

[Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)



المستخلص:

طور العديد من الباحثين خلال السنوات الأخيرة طرقاً لتقدير معامل التشابه الذاتي والذاكرة الطويلة التي اشتهرت باسم معلمة هارست. في هذا البحث ، قمنا بإجراء مقارنة بين تسع طرق مختلفة يستخدم معظمها منحدر الانحرافات للعثور على تقدير لمعامل Hurst مثل نطاق (Rescaled R / S) و Aggregate و (Variance AV) (Absolute moments AM) وبعضها يعتمد على تقنية الترشيح مثل (Discrete Variations DV) ، و (Variance versus level using wavelet VVL) و (Second order derivative using discrete wavelet SODDW) وقد تم إجراء المقارنة باستخدام دراسة المحاكاة للعثور على الطريقة الأكثر كفاءة من خلال MASE. أظهرت نتائج تجارب المحاكاة أن أداء الطرق قريب نسبياً باستثناء طريقة SODDW كانت الأكثر كفاءة في MASE. نوع البحث: تصنيف الورقة الخاصة بك تحت أحد هذه التصنيفات: مراجعة عامة. المصطلحات الرئيسية للبحث: الحركة البراونية الكسرية ، اس هارست ، الذاكرة القصيرة ، الذاكرة الطويلة ، التشابه الذاتي ، الموجات المتقطعة ، الاعتماد طويل المدى ، الاعتماد قصير المدى.

(1) مدرس مساعد، ماجستير في الاحصاء

(2) استاذ، دكتوراه في الاحصاء