Available online at www.sciencedirect.com

SCIENCE DIRECT

Mathematics and Computers in Simulation 70 (2005) 172–179

MATHEMATICS AND COMPUTERS IN SIMULATION

www.elsevier.com/locate/matcom

The rescaled variance statistic and the determination of the Hurst exponent

Daniel O. Cajueiro [a], Benjamin M. Tabak [b,*]

[a] *Universidade Católica de Brasília, Mestrado em Economia de Empresas, SGAN 916, Módulo B, Asa Norte, DF 70790-160, Brazil*

[b] *Banco Central do Brasil, SBS Quadra 3, Bloco B, 9 Andar, DF 70074-900, Brazil*

Received 14 January 2005; received in revised form 29 June 2005; accepted 29 June 2005
Available online 6 September 2005

Abstract

A major issue in statistical physics literature is the study of the long range dependence phenomenon usually presented in natural, social and financial processes. In particular, a big part of this literature relies on the determination of a parameter known as the Hurst exponent. Although many methods have been proposed to deal with this task, none of them are suitable for any time series and sometimes when applied to the same time series present conflicting results. In this context, this paper presents a new method based on the rescaled variance statistic which can be used efficiently to this end.

Keywords: Hurst exponent; Long range dependence; R/S; V/S

1. Introduction

A major issue in the empirical literature is the study of the long range dependence phenomenon usually presented in natural [2,15], industrial [1], economic [6,19], internet trafic [21] and financial processes [3,10].

In this context, several methods have been introduced to take this phenomena into account.[1] The literature can be actually divided in two different strands: (1) a parametric approach to determine the

* Corresponding author. Tel.: +55 61 4143092; fax: +55 61 4143045.
E-mail address: benjamin.tabak@bcb.gov.br (B.M. Tabak).

[1] A survey of these methods may be found in [20].

0378-4754/$30.00
doi:10.1016/j.matcom.2005.06.005

parameter known as the Hurst exponent or a parameter related to it (see, for example [5,8,9]) and (2) a non-parametric approach that develops statistics to test, through a hypothesis, the presence of long range dependence (see, for example [7,11,12]).

In the former approach, the most used methods seem to be the R/S method due to [9] and the fractional ARIMA due to [8]. On the other hand, in the second approach, the R/S [12] is the most common choice.

However, in spite of the existence of several methods, the task of analyzing the long range dependence phenomena is not straightforward. Moreover, the methods above cited sometimes present incompatible estimates. For instance, the problem with the fractional ARIMA specification [8] is that it is highly dependent of the short run dynamics (ARMA component), which has to be specified in the model a priori. However, it is known that the R/S [9] is seriously biased. Besides, the R/S statistics due to [12] presents a preference for the null hypothesis of absence of long range dependence (for details, see [22]).

In this context, this paper presents a new method for the evaluation of the Hurst exponent H. This method is based on the variance rescaled statistic (V/S statistic) [7] which has been recently introduced in the econometric literature and whose purpose has been to test whether there is, or not, long range dependence based on the asymptotic behavior of its distribution.[2] As one may see along this paper, due to its valuable properties, this method seems to be very interesting to estimate the Hurst exponent H and is an improvement to the very popular R/S method used extensively in the statistical physics literature.

This paper is organized as follows. In Section 2, the R/S analysis [9] is revised and V/S analysis is introduced. In Section 3, a numerical algorithm which may be used to evaluate the Hurst exponent is defined. In Section 4, Monte Carlo simulations are performed to show the efficiency of the method and some results are presented. Finally, Section 6 presents some conclusions of this work.

2. R/S and V/S analysis

The V/S analysis is the focus of this paper. However, before introducing the V/S analysis, let us firstly revise the classical R/S analysis.

The R/S statistic [9] is the range of partial sums of deviations of times series from its mean, rescaled by its S.D. (Standard Deviation).[3] So, consider a sample $\{x_1, x_2, \ldots, x_\tau\}$ of a stationary time series and let $\bar{x}_\tau$ denote the sample mean $\frac{1}{\tau}\sum_{t=1}^{\tau} x_\tau$,where τ is the time span considered. Then the R/S statistic is given by:

$$\left(\frac{\mathrm{R}}{\mathrm{S}}\right)_\tau \equiv \frac{1}{s_\tau}\left[\max_{1\le t\le\tau}\sum_{k=1}^{t}(x_k - \bar{x}_\tau) - \min_{1\le t\le\tau}\sum_{k=1}^{t}(x_k - \bar{x}_\tau)\right] \tag{1}$$

[2] In the non-parametrical approach considered in [7], one is not interested in the determination of the Hurst exponent. The aim is to develop a statistical test. Unfortunately, the literature has shown, in general, these tests have preference for the null of absence of long range dependence [22].

[3] In this paper, we choose s_τ as the classical S.D.—this is actually a significant difference from [7] and also [12]. We believe that since it is possible to find expressions for variance in the denominator as well as in the numerator of the V/S statistic, it is not appropriate to correct one and not to correct the other by means of a square root of a consistent estimator of the partial sum's variance introduced by [12,17]. Moreover, if one believes that there is short range dependence presented in the data, it seems that the shuffling procedure presented in [4] is the best way to deal with this problem (see also [22]).

where s_τ is the usual S.D. estimator.

$$s_\tau \equiv \left[\frac{1}{\tau}\sum_{t=1}^{\tau}(x_t - \bar{x}_\tau)^2\right]^{1/2} \tag{2}$$

Hurst [9] found that the rescaled range, R/S, for many records in time is very well described by the following empirical relation:

$$\left(\frac{\mathrm{R}}{\mathrm{S}}\right)_\tau = \left(\frac{\tau}{2}\right)^H \tag{3}$$

By means of the R/S analysis, the Hurst exponent may be evaluated by plotting the data $(\mathrm{R/S})_\tau$ versus τ in a log–log plot and measuring the slope of the straight line.

The idea behind the V/S analysis is quite similar. The only difference is that the range given by $\left[\max_{1\le t\le\tau}\sum_{k=1}^{t}(x_k - \bar{x}_\tau) - \min_{1\le t\le\tau}\sum_{k=1}^{t}(x_k - \bar{x}_\tau)\right]$ in the R/S analysis is replaced by the sample variance of $\sum_{k=1}^{t}(x_k - \bar{x}_\tau)$.

The V/S statistic[4] is given by:

$$\left(\frac{\mathrm{V}}{\mathrm{S}}\right)_\tau \equiv \frac{1}{\tau s_\tau^2}\left[\sum_{t=1}^{\tau}\left(\sum_{k=1}^{t}(x_k - \bar{x}_\tau)\right)^2 - \frac{1}{\tau}\left(\sum_{t=1}^{\tau}\sum_{k=1}^{t}(x_k - \bar{x}_\tau)\right)^2\right] \tag{4}$$

It is easy to notice that[5]:

$$\left(\frac{\mathrm{V}}{\mathrm{S}}\right)_\tau \sim \tau^{2H} \tag{5}$$

By means of the V/S analysis, the Hurst exponent may be also evaluated by plotting the data $(\mathrm{V/S})_\tau$ versus τ in a log–log plot, measuring the slope of the straight line and dividing it by two.

Although, the definitions of the R/S statistic and the V/S statistic seem to be very close, the V/S statistic is a very nice improvement in relation to the R/S statistic. It is possible to notice that: (1) while the R/S statistics is highly influenced by outliers, this does not likely happen to the V/S statistics; (2) while it is known that the R/S statistic presents a biased evaluation of the Hurst exponent, since the V/S statistic is a finer measure of the oscillation around the mean, this does not happen with the V/S statistic.

3. A numerical algorithm

The numerical algorithm that may be used to determine the Hurst exponent H here is the usual one. Let N be the total number of points of the given time series. So, for each scale τ, one should calculate:

$$\overline{\left(\frac{\mathrm{V}}{\mathrm{S}}\right)}_\tau = \frac{1}{N-\tau}\sum_{i=0}^{N-\tau}\left(\frac{\mathrm{V}}{\mathrm{S}}\right)_{\tau i} \tag{6}$$

[4] In order to agree with the R/S statistic (1), the definition of the V/S statistic presented in [7] was multiplied by τ.

[5] This happens since the V/S analysis is a measure of the square of dispersion of $\sum_{k=1}^{t}(x_k - \bar{x}_\tau)$.

where

$$\left(\frac{\mathrm{V}}{\mathrm{S}}\right)_{\tau i} = \frac{1}{\tau s_{\tau i}^2}\left[\sum_{t=1}^{\tau}\left(\sum_{k=1+i}^{t+i}(x_k - \bar{x}_{\tau i})\right)^2 - \frac{1}{\tau}\left(\sum_{t=1}^{\tau}\sum_{k=1+i}^{t+i}(x_k - \bar{x}_{\tau i})\right)^2\right] \tag{7}$$

$$S_{\tau i} = \left[\frac{1}{\tau}\sum_{t=1+i}^{\tau+i}(x_t - \bar{x}_{\tau i})^2\right]^{1/2} \tag{8}$$

and

$$\bar{x}_{\tau i} = \frac{1}{\tau}\sum_{t=1+i}^{\tau+i} x_t \tag{9}$$

This procedure is repeated for several scales τ. Then, the Hurst exponent is calculated by the slope of straight line $(1/2)\log\overline{(\mathrm{V/S})}_{\tau} \times \log\tau$. The rolling sample approach presented in Eqs. (6)–(9) is necessary due to the large variation of the variance in small scales τ. An algorithm similar to this one was presented in [16,18] to perform the so-called Detrended fluctuation analysis (DFA).

4. Results from the Monte Carlo simulation

To test the quality of the estimation of the Hurst exponent H evaluated by the V/S analysis, we have used two examples of fractal self-afim profiles: the discrete approximation of the fractional Brownian motion (FBM) provided by [13] and the Weierstrass–Mandelbrot cosine fractal function defined by [14]. For both sets, we have considered the exponents $H = 0.25$, 0.50 and 0.75. Also, for comparison purposes, we have also applied the R/S analysis to these same sets.

So, the fractional Brownian motion (FBM) used here is given by:

$$B_H(t) - B_H(t-1) = \frac{n^{-H}}{\Gamma\left(H+\frac{1}{2}\right)}\left\{\sum_{i=1}^{n}(i)^{H-1/2}\xi_{(1+n(M+t)-i)} + \sum_{i=1}^{n(M-1)}((n+i)^{H-1/2} - (i)^{H-1/2})\xi_{(1+n(M-1+t)-i)}\right\} \tag{10}$$

with parameters $M = 100$ and $n = 8$.

Since the FBM is generated by means of sequences of random numbers, for each exponent, we have performed a thousand of simulations for time series with size $N = 2000$. The histograms of the results for both R/S and V/S analysis are presented in Fig. 1. We have additionally evaluated the mean, the variance and mean square error for both V/S and R/S methods estimated Hurst exponents. These results are presented in Tables 1–3.

The estimated Hurst exponents of the Weierstrass–Mandelbrot cosine fractal function:

$$C(t) = \sum_{n=-\infty}^{\infty}\frac{1-\cos(b^n t)}{b^{nH}} \tag{11}$$

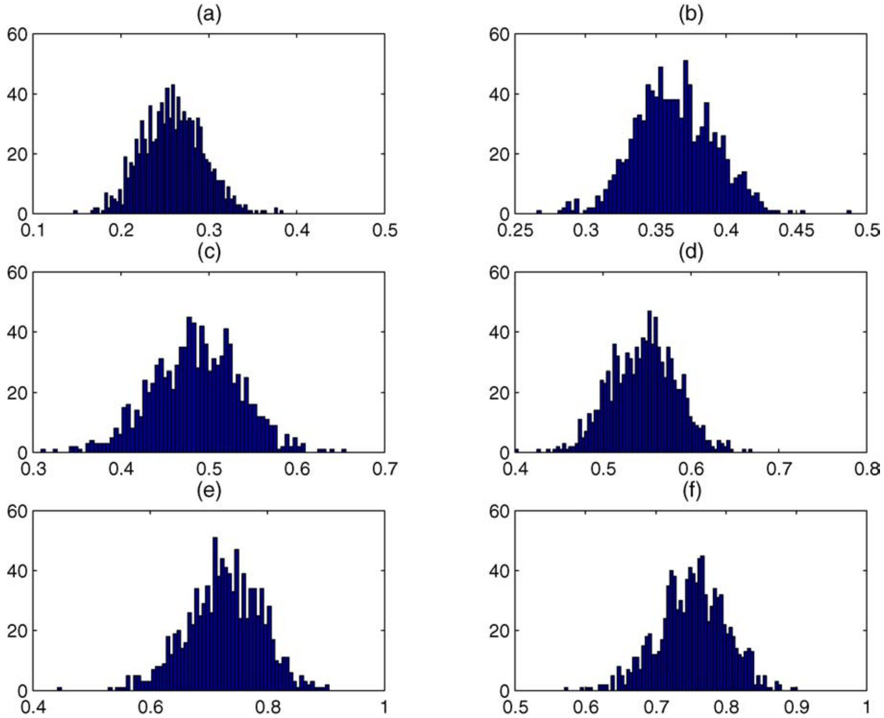

Fig. 1. Histograms of the results of the estimation of H of the simulated FBMs: (a) FBM with $H = 0.25$ using the V/S analysis; (b) FBM with $H = 0.25$ using the R/S analysis; (c) FBM with $H = 0.50$ using the V/S analysis; (d) FBM with $H = 0.50$ using the R/S analysis; (e) FBM with $H = 0.75$ using the V/S analysis; (f) FBM with $H = 0.75$ using the R/S analysis.

Table 1
The mean for both V/S and R/S methods estimated Hurst exponents

H	V/S	R/S
0.25	0.26	0.36
0.50	0.49	0.55
0.75	0.73	0.75

with parameter $b = 2.1$ and the infinity in the summation truncated in 50, by means of the V/S and R/S analysis are presented in Table 4.

5. Empirical implementation

In this section, we estimate the Hurst exponent using both R/S and V/S methodologies. We test for long range dependence in equity returns for Australia, Hong Kong, Singapore and Taiwan. Since these stock returns may possess short range dependence, we employ a shuffling procedure to purge this feature, and present both results for shuffled and unshuffled returns.[6]

[6] We use a block of size 20. Nonetheless, results are robust to blocks in that neighborhood.

Table 2
The variance for both V/S and R/S methods estimated Hurst exponents

H	V/S	R/S
0.25	0.0012	0.0008
0.50	0.0025	0.0014
0.75	0.0040	0.0023

Table 3
The mean square error for both V/S and R/S methods estimated Hurst exponents

H	V/S	R/S
0.25	0.0013	0.0136
0.50	0.0027	0.0035
0.75	0.0045	0.0023

Table 4
The estimated Hurst exponents of the Weierstrass–Mandelbrot cosine fractal function by means of both V/S and R/S methods

H	V/S	R/S
0.25	0.259 ± 0.004	0.361 ± 0.008
0.50	0.496 ± 0.006	0.563 ± 0.009
0.75	0.723 ± 0.006	0.765 ± 0.009

Table 5
Hurst exponents estimated using R/S and V/S statistics for equity returns with and without a shuffling procedure

	R/S		V/S	
	Without shuffling	With shuffling	Without shuffling	With shuffling
Panel A: Hurst exponents and S.E. (Standard Error)				
Taiwan	0.59, 0.024	0.59, 0.009	0.56, 0.009	0.561, 0.008
Hong kong	0.54, 0.021	0.57, 0.007	0.47, 0.008	0.482, 0.009
Singapore	0.58, 0.014	0.61, 0.008	0.56, 0.004	0.585, 0.005
Australia	0.53, 0.010	0.51, 0.010	0.45, 0.011	0.432, 0.010
Panel B: Wald statistics				
Taiwan	13.43, 0.000	105.77, 0.000	40.93, 0.000	62.76, 0.000
Hong kong	4.69, 0.030	95.87, 0.000	10.67, 0.001	3.85, 0.050
Singapore	32.02, 0.000	160.33, 0.000	213.67, 0.000	320.98, 0.000
Australia	10.60, 0.001	1.29, 0.256	22.80, 0.000	48.76, 0.000

The Wald tests for the null hypothesis $H = 0.5$ are provided in panel B.

These indices were studied in [10]. The authors employ the modified R/S and the GPH estimators for long range dependence to study these indices for the period from January 1975 through December 1994.[7] With the modified R/S they reject the null of no long range dependence at the 5% and 10% significance

[7] The exception is Australia for which the index begins in January 1981.

Table 6
Hurst exponents estimated using R/S and V/S statistics for volatility with and without a shuffling procedure

	R/S		V/S	
	Without shuffling	With shuffling	Without shuffling	With shuffling
Panel A: Hurst exponents and S.E.				
Taiwan	0.828, 0.037	0.818, 0.013	0.845, 0.006	0.854, 0.006
Hong kong	1.024, 0.034	0.925, 0.019	0.969, 0.015	0.979, 0.016
Singapore	0.913, 0.038	0.893, 0.022	0.913, 0.027	0.944, 0.026
Australia	0.738, 0.013	0.742, 0.014	0.775, 0.013	0.785, 0.014
Panel B: Wald statistics				
Taiwan	77.62, 0.000	600.31, 0.000	3,168.21, 0.000	3,325.99, 0.000
Hong kong	236.34, 0.000	499.41, 0.000	935.36, 0.000	859.16, 0.000
Singapore	117.54, 0.000	311.83, 0.000	226.52, 0.000	281.41, 0.000
Australia	347.50, 0.000	296.84, 0.000	452.81, 0.000	437.35, 0.000

The Wald tests for the null hypothesis $H = 0.5$ are provided in panel B.

level for Australia and Japan, respectively. However, with the GPH no evidence of long range dependence was found.

Table 5 presents empirical results for these countries. Panel A presents the Hurst exponent with its associated S.E. (second column following the Hurst exponent). As we can see using the R/S statistics, we cannot reject the null of absence of long range dependence for Australia with unshuffled returns but this conclusion is reversed once we employ shuffled returns, controlling for short-term dependency. However, these results change when we study the behavior of the V/S statistic, which suggests that all of these indices possess long range dependence. Table 6 presents a similar picture for volatility. All indices possess strong long range dependence in volatility.

6. Conclusions

This paper presents a new method to estimate the Hurst exponent and test for long range dependence. By means of a Monte Carlo simulation it is shown that this methodology is superior to the usual R/S statistic, which is the most frequently used method and also the most popular. The simulation reveals that the performance of the V/S near the critical value 0.5 is much better than that of the R/S statistic. This is robust to two different ways of simulating fractional Brownian motions. The Hurst exponent estimated from the V/S does not have the shortcomings of the R/S methodology.[8] The implications are that one should use the V/S to estimate the Hurst exponent.

Empirical estimates of Pacific Basin stock returns suggests that these methods may lead to different qualitatively results. From the V/S, we infer that there is some evidence of long range dependence for Australia, Hong Kong, Singapore and Japan, in both equity returns and volatility.

[8] The R/S provides a biased H around the 0.5 value. The bias decreases as the value of the H increases.

Acknowledgements

The authors acknowledge useful comments from an anonymous referee and the editor Andre Preumont. Benjamin M. Tabak gratefully acknowledges financial support from CNPQ foundation. The views expressed here are solely the responsibility of the authors and do not reflect those of the Central Bank of Brazil or its members.

References

[1] R.F.S. Andrade, D.O. Cajueiro, C.S. Ferreira, Physica A 295 (2001) 323.
[2] M.C. Breslin, J.A. Belward, Math. Comput. Simul. 48 (1996) 437.
[3] D.O. Cajueiro, B.M. Tabak, Physica A 336 (2004) 521–537.
[4] A. Erramili, O. Narayan, W. Willinger, IEEE Trans. Neural Netw. 4 (1996) 209.
[5] J. Geweke, Porter-Hudak, J. Time Ser. Anal. 4 (1983) 221.
[6] L.A. Gil-Alana, Appl. Econ. 36 (2004) 1205.
[7] L. Giraitis, P. Kokoszka, R. Leipus, G. Teyssiere, J. Econom. 112 (2003) 265.
[8] J.R.M. Hosking, Frac. Differ. 68 (1981) 165.
[9] E. Hurst, Trans. Am. Soc. Civil Eng. 116 (1951) 770.
[10] C.S. Koong, A.K. Tsui, W.S. Chan, Math. Comput. Simul. 43 (1997) 445.
[11] D. Lee, P. Schimidt, J. Econom. 73 (1996) 285.
[12] A.W. Lo, Econometrica 59 (1991) 1279.
[13] B.B. Mandelbrot, J.R. Wallis, Water Resour. Res. 5 (1969) 228.
[14] B.B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman, New York, 1982.
[15] J.G.V. Miranda, R.F.S. Andrade, Physica A 295 (2001) 38.
[16] J.G. Moreira, J.K.L. Silva, S.O. Kamphorst, J. Phys. A: Math. Gen. 27 (1994) 8079.
[17] W.K. Newey, K.D. West, Econometrica 55 (1987) 703.
[18] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, Physica A 279 (2000) 443.
[19] F. Sowell, J. Monet. Econ. 29 (1992) 277.
[20] M.S. Taqqu, V. Teverovsky, W. Willinger, Fractals 3 (1999) 785.
[21] M.S. Taqqu, Scand. J. Stat. 29 (2002) 273.
[22] W. Willinger, M.S. Taqqu, V. Teverovsky, Financ. Stochastics 3 (1999) 1.