Physica A 389 (2010) 2770–2781

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Use and misuse of some Hurst parameter estimators applied to stationary and non-stationary financial time series

Francesco Serinaldi *

Dipartimento GEMINI, Università della Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy

ARTICLE INFO

Article history:
Received 15 December 2009
Received in revised form 9 February 2010
Available online 26 February 2010

This study is dedicated to Prof. Fabrizio Savi who passed away untimely in December 2009 after a two year battle with cancer. Friends, colleagues and students will never forget his kindness, the availability to share his profound knowledge of hydraulic engineering disciplines, and his love for scientific research.

Keywords:
Hurst parameter
Long range dependence
Fractional Gaussian noise
Fractional Brownian motion
Energy market prices
Stock market prices
Anti-persistence

ABSTRACT

The detection of long range dependence (LRD) is an important task in time series analysis. LRD is often summarized by the well-known Hurst parameter (or exponent) $H \in [0, 1]$, which can be estimated by a number of methods. Some of these techniques are designed to be applied to signals behaving as a stationary fractional Gaussian noise (fGn), whereas others imply that the analyzed time series behave as a non-stationary fractional Brownian motion (fBm). Moreover, some estimators do not yield the Hurst parameter but indexes related to H and ranging outside the unit interval. Therefore, the fGn or fBm nature of the studied time series has to be preliminarily analyzed before applying any estimation method, and the relationships between H and the indexes resulting from the analyses have to be taken into account to obtain coherent results. Since fGn-like series represent the increments of fBm-like processes and both the signals are characterized by the same H value by definition, estimators designed for fGn-like series can be applied to fBm-like sequences after preventive differentiation, and conversely estimators designed for fBm-like processes can be applied to fGn-like series after preventive integration. The signal characterization is particularly important when H is estimated on financial time series because the returns represent the first difference of price time series, which are often assumed to behave like self-affine sequences. The analysis of simulated fGn and fBm time series shows that all the considered methods yield comparable H values when properly applied. The reanalysis of several market price time series already studied in the literature points out that a correct application of the estimators (supported by a preventive signal classification) yields homogeneous H values allowing for a useful cross-validation of results reported in different works. Moreover, some conclusions reported in the literature about the anti-persistence of some financial series are shown to be incorrect because of the inappropriate application of the estimation methods.

1. Introduction

The study of long range dependence (LRD), i.e. the dependence between observations far away in time, has received increasing attention in many disciplines. Given its important consequences in describing and interpreting physical phenomena, the utility of LRD has been analyzed in economics and finance (e.g., Refs. [1,2]), hydrology and water resources (e.g., Refs. [3,4]), geophysics (e.g., Refs. [5–7]), DNA sequences (e.g., Refs. [8]), internet traffic (e.g., Ref. [9]) and many other fields. The strength of LRD is traditionally measured by the Hurst parameter H introduced by Hurst [10] during his studies on Nile discharges and problems related to water storage. The H parameter ranges in $[0, 1]$. For an uncorrelated white noise signal $H = 0.5$, whereas values greater than (less than) 0.5 are associated to persistent (anti-persistent) processes [11].

* Tel.: +39 0761 357 356.
E-mail address: f.serinaldi@unitus.it.

0378-4371/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2010.02.044

Theoretical processes exhibiting LRD are the self-affine processes [12]. The simplest example of these processes is the Brownian motion or random walk, which is defined as the cumulative sum of a Gaussian white noise signal. The Gaussian noise and Brownian motion can be generalized by introducing the concept of fractional differentiation, leading to the so-called fractional Gaussian noise (fGn) and fractional Brownian motion (fBm) [12,13,2]. Since in the original definition given by Mandelbrot and Van Ness [12], a unique H value characterizes both fBm and its increments, which represent the corresponding fGn (see also Ref. [11]), it is necessary to complement H with a piece of information about the type of signal we are dealing with (fGn or fBm) to avoid ambiguity [14]. Distinguishing between fBm- and fGn-like series is particularly important for correctly estimating the H parameter. Indeed, as discussed in more detail in the next sections, a number of commonly applied graphical estimators available in the literature rely explicitly or implicitly on the assumption that the analyzed series are fBm- or fGn-like. As pointed out by Cannon et al. [14], invalid results and conclusions about the estimation of H are due to the above-mentioned duality between fBm and fGn.

FBm, fGn and LRD play a prominent role in financial time series analysis. Brownian motion is particularly important in finance and economics because it naturally arises from the basic hypotheses of the early definition of the "efficient market hypothesis", which was assumed to imply that successive price changes (or more usually, successive one-period returns) are independent [15]. Even though a Brownian (or geometric Brownian) behavior can be empirically observed in price time series (e.g., Ref. [16]), some market prices exhibit auto-correlated returns and anti-persistence (also referred to as mean reversion). Hence, the LRD assessment has received increasing attention due to its relevant implications in time series modeling and forecasting (e.g., Refs. [17–23], among many others).

This paper explores the importance of preventive signal classification (before applying H estimators) with a focus on price time series analysis, extending the works by Cannon et al. [14], Malamud and Turcotte [24] and Li [6] in three issues: (i) the relationships between H and the indexes returned by the estimators are studied for additional methods; (ii) results of the appropriate application of the methods are shown for a number of financial time series as well as for simulated series, (iii) an indirect method for signal classification based on the properties of H estimators is suggested. Moreover, the reanalysis of real-world market data shows that all methods, when properly applied, yield homogeneous and coherent H values, allowing for a cross-validation of results and a more reliable assessment of possible anti-persistence.

The remainder of the paper is structured as follows. In the next section, some properties of self-affine sequences are recalled and the literature coping with the problem of signal classification for a proper H estimation is reviewed. Next, the relationships between H and the indexes returned by seven graphical H estimators are discussed, and an indirect method for signal classification is suggested in Section 3. Section 4 reports the reanalysis of several price time series already studied in the literature, pointing out some misleading results due to an incorrect application of the methods. Final remarks and conclusions close the work.

2. Signal classification and *H* estimation

As previously mentioned, among the available H estimators, some widely applied graphical methods require an fGn-like input signal, whereas others require the corresponding fBm-like series. An fGn sequence is characterized by the same H value of the corresponding fBm sequence. However, these processes are rather different because fGn represents the first difference of fBm. To point out the nature of fGn and fBm, it is worth recalling some properties of self-affine processes. For a discrete self-affine time series x_n, $n = 1, \dots, N$, such as fGn and fBm, the power spectrum density (PSD) S_m is defined to have a power law dependence on frequency:

$$S_m(f_m) \propto f_m^{-\beta}, \tag{1}$$

where $m = 1, \dots, N/2$, $f_m = m/N$, and the value of β is a measure of the strength of persistence in a time series [24]. For a white noise and a Brownian motion the theoretical values of β are 0 and 2, respectively. In general, the cumulative summation (integration) of a self-affine series shifts the PSD exponent β by $+2$. Conversely, taking the first difference shifts the PSD exponent β by -2. Since fGn has $\beta \in (-1, 1)$ and is the incremental process of fBm, then fBm exhibits a PSD exponent $\beta \in (1, 3)$. Following Malamud and Turcotte [24], a self-affine series can be classified as:

- stationary and anti-persistent if $\beta < 0$,
- stationary and uncorrelated if $\beta = 0$ (e.g., white noise),
- stationary and weak persistent if $\beta \in (0, 1)$,
- non-stationary and strong persistent if $\beta > 1$.

Therefore, even though fGn and the corresponding fBm exhibit the same H value, the first one is stationary whereas the latter is non-stationary. The parameter H can be estimated on the fGn or on the corresponding fBm. However, as discussed later, methods designed for stationary fGn can fail to provide correct values if they are applied to non-stationary fBm and vice versa. To the best of our knowledge, the problem of detecting the type of signal before applying an H estimator was firstly tackled by Cannon et al. [14]. These authors provide a theoretical discussion and show the effect of the incorrect application of the dispersional analysis (also known as aggregated variance method) [25] and three versions of the scaled windowed variance (SWV) method to misspecified signals. Cannon et al. [14] also analyze the reliability of those methods by Monte Carlo simulation of fGn and fBm sequences. Malamud and Turcotte [24] compare a variety of techniques to measure the strength of persistence of simulated self-affine time series and conclude that the semivariogram method is effective