Physica A 447 (2016) 434–445

Contents lists available at ScienceDirect

Physica A

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

Anomalous volatility scaling in high frequency financial data

Noemi Nava [a,c,*], T. Di Matteo [b], Tomaso Aste [a,c]

[a] Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK
[b] Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK
[c] Systemic Risk Centre, London School of Economics and Political Sciences, London WC2A 2AE, UK

HIGHLIGHTS

- Empirical mode decomposition is used to study high frequency data.
- For fractional Brownian motion, we identify a variance scaling law.
- We measure volatility at different time horizons for different stock market indices.
- Some scaling deviations are identified for the less developed financial markets.

ARTICLE INFO

Article history:
Received 18 July 2015
Received in revised form 22 October 2015
Available online 28 December 2015

Keywords:
Empirical mode decomposition
Hurst exponent
Multi-scaling
Market efficiency

ABSTRACT

Volatility of intra-day stock market indices computed at various time horizons exhibits a scaling behaviour that differs from what would be expected from fractional Brownian motion (fBm). We investigate this anomalous scaling by using empirical mode decomposition (EMD), a method which separates time series into a set of cyclical components at different time-scales. By applying the EMD to fBm, we retrieve a scaling law that relates the variance of the components to a power law of the oscillating period. In contrast, when analysing 22 different stock market indices, we observe deviations from the fBm and Brownian motion scaling behaviour. We discuss and quantify these deviations, associating them to the characteristics of financial markets, with larger deviations corresponding to less developed markets.

1. Introduction

Over the last few years financial markets have witnessed the availability and widespread use of data sampled at high frequencies. The study of these data allows to identify the intra-day structure of financial markets [1,2]. Data at these frequencies have dynamic properties which are not generated by a single process but by several components that are superimposed on top of each other. These components are not immediately apparent, but once identified, they can be meaningfully categorized as noise, cycles at different time-scales and trends [1].

Since the early work of Mandelbrot [3,4], it was recognized that different time-scales contribute to the complexity of financial time series in a self-similar (fractal) manner. Empirical properties of financial data at various frequencies have been observed in a number of studies, see for example Refs. [5–9].

According to the random walk hypothesis [10], financial market dynamics can be described by a random walk, a self-similar process with scaling exponent (Hurst exponent) $H = 0.5$ [11]. Opposing this theory, Peters [12] introduced the

* Corresponding author at: Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK.
E-mail addresses: n.morales.11@ucl.ac.uk (N. Nava), tiziana.di_matteo@kcl.ac.uk (T. Di Matteo), t.aste@ucl.ac.uk (T. Aste).

http://dx.doi.org/10.1016/j.physa.2015.12.022
0378-4371/

fractal market hypothesis (FMH), representing financial market dynamics by fractional Brownian motion (fBm), a self-similar process with scaling exponent $0 < H < 1$. The focus of the FMH is on the interaction of agents with various investment horizons and differing interpretations of information. Based on this theory, heterogeneous market models have explained some stylized facts (such as volatility clustering, kurtosis, fat tails of returns, power law behaviours) observed in financial markets, see for example Refs. [13–16].

In self-similar uni-scaling process, such as fBm, all time-scales contribute proportionally and there is a specific relation that links statistical properties at different time-scales [17]. However, real financial time series have more complex scaling patterns, with some time-scales contributing disproportionally; these patterns characterize multi-scaling processes whose statistical properties vary at each time-scale [18–22].

The knowledge of scaling laws in financial data helps us to understand market dynamics [23,24], that can be interpreted to construct efficient and profitable trading strategies. In this paper, we use empirical mode decomposition (EMD), an algorithm introduced by Huang [25], to decompose intra-day financial time series into a trend and a finite set of simple oscillations. These oscillations, called intrinsic mode functions (IMFs), are associated with the time-scale of cycles latent in the time series. The EMD provides a tool for an exploratory analysis that takes into account both the fine and coarse structure of the data. This decomposition has been widely used in many fields, including the analysis of financial time series [26–30], river flow fluctuations [31], wind speed [32], heart rate variability [33], etc.

In this paper, we first apply EMD to fBm, uncovering a power law scaling between the period and variance of the IMFs with scaling exponent related to the Hurst exponent. We then apply EMD to 22 different stock market indices whose prices are sampled at 30 s intervals over a time span of 6 months. In this case, we encounter more complex scaling laws than in fBm. The deviations from the fBm behaviour are quantified and interpreted as an anomalous multi-scaling behaviour.

This paper is organized as follows. In Section 2, we introduce the EMD. In Section 3, we present the variance scaling properties of fBm. In Section 4, we present an application to high frequency financial data. We finally conclude in Section 5.

2. Empirical mode decomposition

The empirical mode decomposition is a fully data-driven decomposition that can be applied to non-stationary and non-linear data [25]. Different from the Fourier and the wavelet transform, the EMD does not require any a priori filter function [34]. The purpose of the method is to identify a finite set of oscillations with scale defined by the local maxima and minima of the data itself. Each oscillation is empirically derived from data and is referred as an intrinsic mode function. An IMF must satisfy two criteria:

1. The number of extrema and the number of zero crossings must either be equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The IMFs are obtained through a process that makes use of local extrema to separate oscillations starting with the highest frequency. Hence, given a time series $x(t)$, $t = 1, 2, \dots, T$, the process decomposes it into a finite number of intrinsic mode functions denoted as $IMF_k(t)$, $k = 1, \dots, n$ and a residue $r_n(t)$. If the decomposed data consist of uniform scales in the frequency space, the EMD acts as a dyadic filter and the total number of IMFs is close to $n = \log_2(T)$ [35]. The residue is the non-oscillating drift of the data. At the end of the decomposition process, the original time series can be reconstructed as:

$$x(t) = \sum_{k=1}^{n} IMF_k(t) + r_n(t). \quad (1)$$

The EMD comprises the following steps:

1. Initialize the residue to the original time series $r_0(t) = x(t)$ and set the IMF index $k = 1$.
2. Extract the kth IMF:
   (a) initialize $h_0(t) = r_{k-1}(t)$ and the iteration counter $i = 1$;
   (b) find the local maxima and the local minima of $h_{i-1}(t)$;
   (c) create the upper envelope $E_u(t)$ by interpolating between the local maxima (lower envelope $E_l(t)$ by interpolating the local minima, respectively);
   (d) calculate the mean of both envelopes as $m_{i-1}(t) = \frac{E_u(t)+E_l(t)}{2}$;
   (e) subtract the envelope mean from the input time series, obtaining $h_i(t) = h_{i-1}(t) - m_{i-1}(t)$;
   (f) verify if $h_i(t)$ satisfies the IMF's conditions:
   - if $h_i(t)$ does not satisfy the IMF's conditions, increase $i = i + 1$ and repeat the sifting process from step (b);
   - if $h_i(t)$ satisfies the IMF's conditions, set $IMF_k(t) = h_i$ and define $r_k(t) = r_{k-1}(t) - IMF_k(t)$.
3. When the residue $r_k(t)$ is either a constant, a monotonic slope or contains only one extrema stop the process, otherwise continue the decomposition from step 2 setting $k = k + 1$.