Available online at www.sciencedirect.com

SCIENCE DIRECT®

Physica A 363 (2006) 393–403

PHYSICA A

www.elsevier.com/locate/physa

# A Markov model of financial returns

M. Serva[a,b], U.L. Fulco[b,c], I.M. Gléria[b], M.L. Lyra[b], F. Petroni[a], G.M. Viswanathan[b,*]

[a]Dipartimento di Matematica and I.N.F.M. Università dell'Aquila, I-67010 L'Aquila, Italy
[b]Departamento de Física, Universidade Federal de Alagoas, Maceió– AL, 57072-970, Brazil
[c]Departamento de Fsica, Universidade Federal do Piau, Terezina– PI, 64049-550, Brazil

Received 2 February 2005; received in revised form 28 June 2005
Available online 17 October 2005

**Abstract**

We address the general problem of how to quantify the kinematics of time series with stationary first moments but having non stationary multifractal long-range correlated second moments. We show that a Markov process is sufficient to model important aspects of the multifractality observed in financial time series and propose a kinematic model of price fluctuations. We test the proposed model by analyzing index closing prices of the New York Stock Exchange and the DEM/USD tick-by-tick exchange rates obtained from Reuters EFX. We show that the model captures the characteristic features observed in actual financial time series, including volatility clustering, time scaling and fat tails in the probability density functions, power-law behavior of volatility correlations and, most importantly, the observed nonuniversal multifractal singularity spectrum. Motivated by our finding of strong agreement between the model and the data, we argue that at least two independent stochastic Gaussian variables are required to adequately model price fluctuations.


*Keywords:* Times series analysis; Volatility correlations; Multifractals

## 1. Introduction

The quantitative description of economic [1] phenomena has recently become a new challenge for a large community of physicists [1–9], due to the many similarities with collective phenomena in physics and the increasing availability of data records [10–23]. Prices have no stationary moments; however, price increments have approximately stationary mean. Let $S_t$ represent the market price at a given instant $t$, e.g., of a stock, a stock market index or the exchange rate between two currencies. We define the log-returns $r_t$ at time $t$ as

$$r_t = \log_{10} \frac{S_{t+1}}{S_t}. \tag{1}$$

Such returns have nonstationary multifractal long-range power-law correlated local variances [1,4,7]. The study of long-range memory is an important issue in finance. During the last several years, research in physics,

*Corresponding author.
*E-mail address:* gandhi@df.ufal.br (G.M. Viswanathan).

0378-4371/$ - see front matter 
doi:10.1016/j.physa.2005.08.070

mathematical finance and econometrics has attempted to give an explanation to this long-range memory of the markets [21–23]. The issue is appealing to physicists since it involves non linear dynamics, noise modeling, collective phenomena and fractals. Here we investigate the kinematics of financial time series (see Fig. 1) and propose a model that, despite being Markovian, captures important features of actual financial time series, including the observed fat tails, their nonuniversal volatility multifractal singularity spectra, as well as fat tails and time scaling in the probability density functions (PDFs).

An important contribution in modeling time series with nonstationary second moments came from the study of auto-regressive conditional heteroscedasticity (ARCH) models, introduced by Engle in 1982 [24]. ARCH models refer to Markovian stochastic processes characterized by zero mean and nonconstant variances dependent on the past. Such models can simultaneously have global stationarity as well as local nonstationary behavior. The simplest ARCH model of returns has local variance defined by

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2,$$

with returns defined as

$$r_t = \sigma_t \omega_t.$$

The random Gaussian variable $\omega_t$ has unitary variance and zero mean, while $\alpha_0$ and $\alpha_1$ represent tunable parameters. Hence, the returns $r_t$ follow a Gaussian distribution with zero mean and a volatility given by $\sigma_t^2$. In more general ARCH processes, the local variance can depend not only on the previous value $r_{t-1}$, but on any finite number $n$ of previous values $r_{t-1} \dots r_{t-n}$. GARCH [25] models result from a further generalization in which the local variance can depend not only on previous values of the returns but also on the previous values of the locally measured variances (see Eq. (8)). All such stochastic kinematic models neglect the causes underlying the price variations and only consider the equations of motion governing the fluctuating returns. Specifically, ARCH–GARCH models neglect the interacting buying and selling processes. Such models have nevertheless succeeded in capturing some features of actual financial data, e.g., volatility clustering. The three most important (and difficult to model) features not well accounted in ARCH–GARCH models are (i) the time scaling of the PDF of the returns, (ii) the known long-range power-law correlations in the volatility, and (iii) the multifractality of the returns, i.e., the volatility power-law correlation exponents depend on the magnitude of the events considered. Whereas "monofractals" have a single fractal dimension, on the other

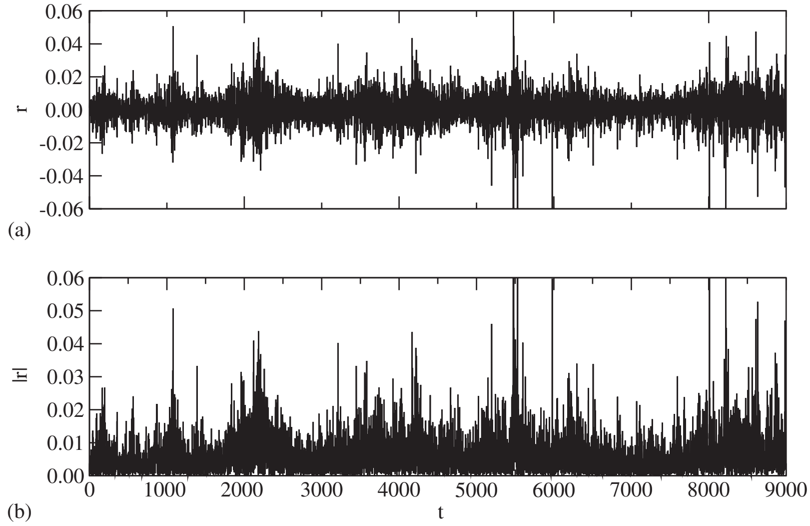


Fig. 1. (a) Daily log-returns (see Eq. (1)) and (b) absolute log-returns for the New York Stock Exchange (NYSE) data set. Such time series have stationary means but nonstationary second moments (hence, the "volatility"). Moreover, the absolute returns have long-range power-law correlations. The values of the scaling exponents depend on the moment studied, i.e., the correlations have multifractal rather than monofractal characteristics.

hand a multifractal typically has a fractal dimension that depends on the moment studied. We discuss multifractality more thoroughly in Section 2.4.

As an alternative model, Lévy [26] and truncated Lévy distributions [27] have been proposed to fit the observed fat tailed distributions. The recently proposed [28] BNS stochastic volatility (SV) continuous-time model attempts to describe returns in terms of a differential equation with the volatility $\sigma_t^2$ following a non Gaussian Ornstein–Uhlenbeck process driven by a Lévy stochastic process with no correlations (i.e., independent Lévy variables). Moreover, to better characterize financial time series, a number of other models have been proposed, including TARCH [21] and FIGARCH [29] processes, multifractal random walks [23], and other SV and similar models [30–33]. Lévy-based models can approximately (but not correctly [6,34]) reproduce the time scaling of the PDF. However, such models cannot account for known multifractal long-range correlations in volatility. Some studies have suggested that such correlations may generate volatility clustering and lead to the persistence of fat tails over long time lags [6,34].

Motivated by the need to adequately address these concerns, we investigate here a new model that does not suffer from the drawbacks mentioned [35]. We construct the model by taking into account the following recent findings: (i) the PDF of the local variance $\sigma_t$ is similar to a log-normal [4], (ii) long-range correlations found in the absolute value of the returns; the multifractal behavior of the returns [7]. Based on these findings, and liberally extending a model proposed in Ref. [36], we propose the following map for the evolution of the volatility:

$$\sigma_{t+1} = \mathrm{e}^{[a+b\omega_{t+1}]}(\overline{\sigma_t})^d, \tag{2}$$

with returns defined as

$$r_t = \sigma_t \tilde{\omega}_t, \tag{3}$$

where $\tilde{\omega}_t$ and $\omega_t$ are independently distributed unitary Gaussian variables with zero mean that are also independent from each other. The term $\overline{\sigma_t}$ represents the average with respect to the last $T$ times, from $t-T$ to $t$. Since the average involves a number $T$ of lags (typically in the range $2 \leqslant T \leqslant 10$), therefore Eqs. (2), (3) form a Markov process. We find, remarkably, that the correlation time for the model far exceeds $T$, by more than an order of magnitude. We note that while we have used an unweighted average for calculating $\overline{\sigma_t}$, in principle we could have used any kind of moving average, e.g., an exponentially weighted moving average.

Since the proposed model represents a variation of a multiplicative process, it is not surprising that it should lead to multifractal features. The tunable constants $a$, $b$ and $d$ allow for diverse phenomena and nonuniversal market characteristics, and relate to the nearly log-normal form of $\sigma_t$. The model requires $d<1$ to guarantee globally stationary returns. While $a$ relates to the mean of the daily volatility, $b$ relates to the variance of the volatility. Notice that this model, with two mutually independent Gaussian variables, contrasts with GARCH models, which only have one Gaussian variable. Below, we discuss why this feature might better correspond to actual markets. As a final remark we emphasize the fact that time is discrete in this model and takes integer values, as in actual markets.

There are two important differences between our model and the families of ARCH–GARCH models. The first is the presence of two independent Gaussian variables, one for returns and one for volatility. This choice, besides giving reasonable agreement with experimental data, has a deeper underlying motivation. In ARCH–GARCH models there is an explicit dependence between the present-day volatility and the previous day's absolute return while in our model the temporal evolution of volatility and returns is separated. In fact, it is commonly accepted that today's sentiment about volatility does not depend on the previous day's variation of price but rather on other quantities, such as previous day intraday volatility, implicit volatility in derivative products and public (or less public) information. Therefore, the present day volatility is not influenced by the previous day's absolute return (which can be small after a fearful market day with enormous variation of prices), and so the evolution of the first should be kept separate, exactly as we have done in the model proposed (see also Ref. [30]). Indeed, such a separation can become necessary whenever we encounter nonstationary second moments. Audio time series also have non-stationary variance [37] and also effectively require two independent variables, one to model the audio frequency signal and another to vary independently the sub-audio frequency loudness or intensity modulation [38] (i.e., the local variance).

The proposed model bears some resemblance to a latent volatility model proposed in Ref. [36] in two aspects: (i) the volatility term contains a nonlinear power-law feedback on the previous day's volatility, and (ii) both models use two independent Gaussian variables. Since our focus here lies in the returns PDF and multifractality, we have extended the model to allow the present volatility to depend nonlinearly on a nonuniformly weighted moving average of the previous days' volatility.

Although volatility correlations exist, return correlations must decay rapidly for a market to be efficient. The efficient market hypothesis (EMH) holds that markets tend to reflect the available information. Specifically, to the extent that price changes appear non-random and hence forecastable, to that extent a profit-seeking arbitrageurs can make appropriate purchases and sales of assets to exploit this information [39–41]. An expected consequence is that correlations in the returns decay very rapidly in time.

## 2. Method and results

### 2.1. Data

We base our study on two significantly different data sets: the New York Stock Exchange (NYSE) daily composite index price closes from January 1966 to September 2001 (a total of some 9000 data points) and preprocessed [42] DEM/USD tick-by-tick exchange rates taken from Reuters EFX (provided by Olsen & Associates) during a period of 1 year from January to December 1998, corresponding to $1.620843 \times 10^6$ quote entries in the EFX system. We thus use data obtained from two entirely different markets: the DEM/USD data set contains quotes for approximately every 20 s while the NYSE closes daily. Note the difference in the data set sizes. By using two substantially different data sets, we can obtain a more rigorous testing of the proposed kinematic model.

### 2.2. Long-range power-law volatility correlations

We test the model against the NYSE time series by generating 9000 artificial returns using Eqs. (2), (3). We have empirically estimated the choice of constants $a = -0.1$, $b = 0.1$, $d = 0.98$ and $T = 10$ days. To measure correlations efficiently in non-stationary time series, we apply a variation of the well documented detrended fluctuation analysis (DFA) [43,44]. Consider the cumulative returns $\phi_t(L)$, defined as the mean value of $L$ successive returns:

$$\phi_t(L) \equiv \frac{1}{L}\sum_{i=1}^{L} r_{t+i}.$$

One can define $N/L$ nonoverlapping variables of this type, and compute the associated variance $\sigma^2(\phi(L))$, where $N$ denotes the number of data points in the data set. According to the central limit theorem, if $r_t$ are uncorrelated variables (i.e., white noise), then the cumulative return fluctuations will scale as a power-law,

$$\sigma^2(\phi(L)) \sim L^{-\alpha}, \tag{4}$$

with exponent $\alpha = 1$ for large $L$, while $\alpha = 0$ for $1/f$-noise and $\alpha = -1$ for Brown noise. Indeed, $\alpha = 2 - 2H$, where $H$ is the Hurst exponent. We find that the exponent $\alpha$ both for the NYSE index and the model proposed here falls near 1 (see Fig. 2). In fact, daily returns have no auto-correlations for lags larger than a single day, in agreement with the EMH.

Correlations do appear, however, if we consider the absolute returns rather than the returns themselves. In order to perform the appropriate multifractal scaling analysis, we introduce the generalized cumulative absolute returns, defined as the mean value of $L$ successive absolute return powers $|r_t|^q, \ldots, |r_{t+L-1}|^q$,

$$\phi_t(L, q) \equiv \frac{1}{L}\sum_{i=1}^{L} |r_{t+i}|^q,$$

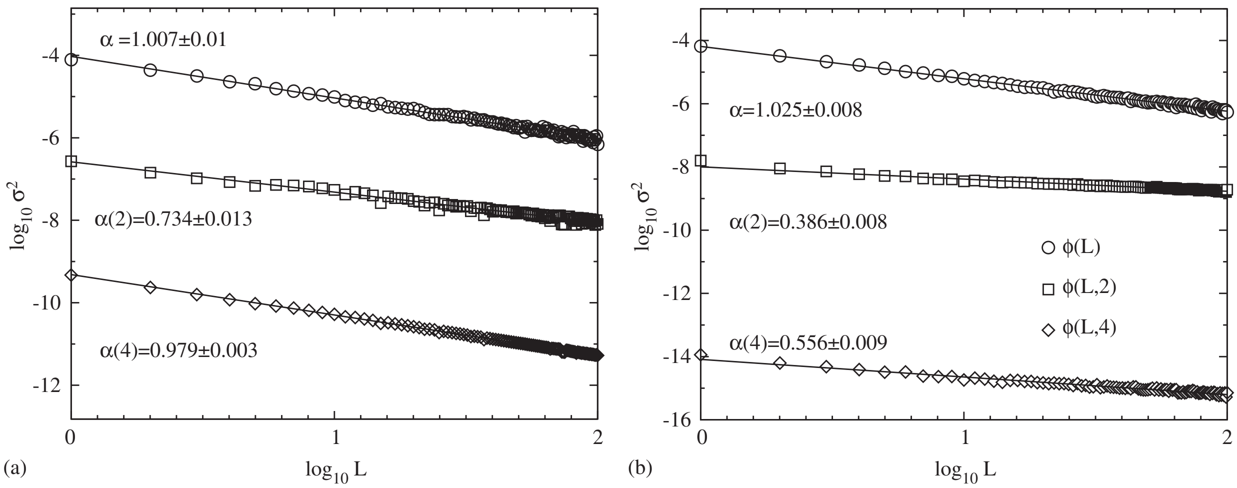


Fig. 2. Double log plot of the variance $\sigma^2(\phi(L,q)) \sim L^{-\alpha(q)}$ defined in Eq. (5) of the generalized cumulative absolute returns for (a) the NYSE index and (b) the proposed kinematic model. Data for the absolute moments with $q = 2$ (square) and $q = 4$ (diamond) are compared with the variance $\sigma^2(\phi(L)) \sim L^{-\alpha}$ (see Eq. (4)) of the cumulative returns (circles). Both data and model show multifractal behavior, since there is no unique scaling exponent. The ability of the proposed kinematic model to generate multifractal behavior distinguishes it from the well-known families of models.

where $q$ is a real exponent. Note that we use nonoverlapping quantities. If the absolute returns have long-range power-law correlations, then

$$\sigma^2(\phi(L,q)) \sim L^{-\alpha(q)}. \tag{5}$$

However, if the $|r_t|^q$ are short-range correlated or power-law-anticorrelated with an exponent $\alpha(q) > 1$, then we would not detect anomalous scaling in the analysis of variance, because it is not possible to detect $\alpha > 1$ using the method discussed here. This situation rarely or never arises for absolute returns.

We find that both the actual data and the model (Fig. 2) show apparent multifractal behavior with non-unique scaling exponents, i.e., $\alpha(q) \neq$ constant. Note that the function $\alpha(q)$ is not universal [4,7] but depends on the particular asset considered. Therefore, it is not surprising that the values $\alpha(q)$ we compute do not coincide for the actual data and the model. Indeed, we expect the model and the data not to have identical $\alpha(q)$ due to the nonuniversality inherent in finance. The important point to note is the (nonuniversal) multifractality.

In order to obtain better agreement for the multifractal exponents for a specific choice of financial time series one can generalize the map as follows:

$$x_{t+1} = \mathrm{e}^{a+b\omega_{t+1}}(\overline{x_t^c})^d, \tag{6}$$

where the exponent $c$ in the volatility average would allow for a finer control of the effects caused by extremal events. The resulting returns could also be written as

$$r_t = (\alpha_v + x_t) \quad \tilde{\omega}_t \equiv \sigma_t\tilde{\omega}_t, \tag{7}$$

with $\alpha_v$ representing an intrinsic constant contribution to daily volatility. Using this generalized map we generate a data set of $1.6 \times 10^6$ returns to compare our model with high frequency data from the foreign exchange market. We use the following estimated parameters: $a = -0.29$, $b = 0.37$, $d = 0.98$, $c = 1$, $\alpha = 0.1 \times 10^{-4}$, $T = 10$. We also consider for completeness a data set of identical size generated using a GARCH(1,1) model defined as

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{8}$$

with returns defined as

$$r_t = \sigma_t \omega_t, \tag{9}$$

with parameters $\alpha_0 = 2.3 \times 10^{-5}$, $\alpha_1 = 0.085$ and $\beta_1 = 0.9$ [1]. We chose the parameters for the GARCH(1,1) model on the basis of the kurtosis and variance estimated from actual data.

### 2.3. Autocorrelation function

Unlike the NYSE index data, which have only around 9000 data points, the relatively large number of data points available in the DEM/USD data set permits a direct estimation of autocorrelations and a quantitative comparison of the multifractal power-law correlations with the models. We define the autocovariances $C(L,q)$ as

$$C(L,q) \equiv \langle |r_t|^q |r_{t+L}|^q \rangle - \langle |r_t|^q \rangle^2 \tag{10}$$

for time lag $L$. Results for different values of $q$ and for the two data sets are presented in Fig. 3. We find that while the power-law behavior of autocovariances is preserved when data are generated using our model, a GARCH(1,1) model does not show this behavior [1], as seen in Fig. 3(c). Moreover the power-law behavior of autocovariances shows multifractality since the exponent $\alpha(q)$ is different for each $q$. This behavior is qualitatively identical for the actual data and the proposed model. However, multifractal power-law behavior appears inconsistent with the GARCH(1,1) model.

### 2.4. Multifractal singularity spectra

Mandelbrot pioneered research into multifractal measures [45,46]. Later, a number of studies applied multifractal concepts to describe quantitatively the behavior of financial processes [47,48]. More recently, multifractal models incorporating elements of Mandelbrot's past research have been proposed [49–51] as alternatives to ARCH-type representations. Indeed, price changes in financial markets appear to share many features characterizing turbulent flows and multifractal models seem to provide a surprisingly good fit of the PDF of the data [50]. We follow a similar approach with our focus on mutifractality, fat tails and long range correlations in volatility. Most previous studies have investigated the multifractality of returns, however our interest here lies in the multifractality of the absolute returns.

Motivated thus to quantify further the multifractal properties of the time series, we now study their multifractal spectra. The multifractal spectrum of a time series contains information about $n$-point correlations [52] and thus provides more information compared to two-point correlation functions. We apply the multifractal detrended fluctuation analysis (MF-DFA) method to study the multifractal spectrum [53]. Like the wavelet transform modulus maximus (WTMM) method [54], the MF-DFA method avoids the difficulties of managing diverging negative moments. We briefly summarize the MF-DFA method as follows:

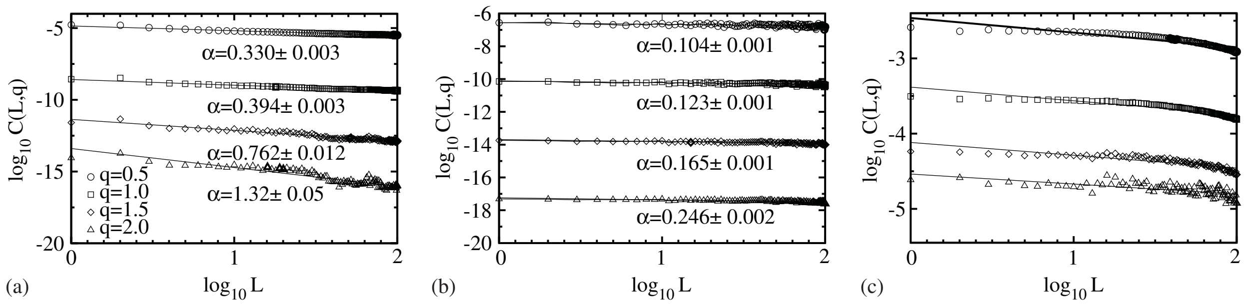


Fig. 3. Double log plot of the covariances $C(L,q)$ defined in Eq. (10) for (a) DEM/USD rates, (b) the kinematic model of the exchange rate, and (c) for the GARCH(1,1) model. We can see power-law multifractal behavior in the nonunique scaling exponents shown in (a) and (b), however (c) does not show good power-law scaling, let alone multifractality.

First, we integrate the time series $|r_t|$ to generate a profile $y(t) \equiv \sum_{k=1}^{t} |r_t|$, which we can visualize as a one-dimensional random walk. We then divide the series $y(t)$ into $N_s$ segments of size $s$. We use non overlapping segments, so the product $sN_s$ gives the total length of the complete time series. For each segment $v$, we apply linear regression (generalizable to polynomial regression) to calculate a best fitting "trend" $y_v(i)$, where $i = 1, 2, \ldots, s$. For each segment $v$ we can thus define a mean square fluctuation

$$F_2(v, s) \equiv \frac{1}{s} \sum_{i=1}^{s} |y((v-1)s + i) - y_v(i)|^2,$$

which corresponds to a measure of the second moment for segment $v$. We then define a $q$th order fluctuation function

$$F_q(s) \equiv \frac{1}{N_s} \sum_{v=1}^{N_s} F_2(v, s)^{q/2}.$$

Positive values of $q$ weight large fluctuations while negative moments weight small fluctuations. The fact that $F_2(v, s)$ always remains positive avoids the divergence of the negative moments. The scaling of the fluctuation for moment $q$ follows

$$F_q(s) \sim s^{qh(q)}, \tag{11}$$

where $h(q)$ represents a generalized Hurst exponent, with the traditional Hurst exponent given by $H = h(2)$. Monofractal time series have a unique Hurst exponent $h(q) = H$, while for multifractal series the value of $h(q)$ depends on $q$.

We can relate $h(q)$ to the exponents $\tau(q)$ conventionally used to quantify the scaling of the partition function

$$Z_q(s) \sim s^{\tau(q)}$$

in the standard textbook multifractal formalism [53] via

$$\tau(q) = qh(q) - 1. \tag{12}$$

We can obtain this relation by noting that for any positive, normalized and stationary time series $x(t)$, we can define a "box probability"

$$p_s(v) \equiv \sum_{k=(v-1)s+1}^{vs} x(t)$$

for box size $s$, given the normalization $\sum_{k=1}^{sN_s} x(k) = 1$. The partition function is defined as

$$Z_q(s) \equiv \sum_{v=1}^{N_s} |p_s(v)|^q.$$

The relation given by Eq. (12) can be seen by noting that $Z_q(s) \sim N_s F_q(s)$. The advantage of methods such as WTMM and MF-DFA arises due to not being limited to stationary, positive, and normalized time series. Fig. 4(a) shows that except for the GARCH series, $\tau(q)$ is not linear, indicating multifractality. The GARCH series has approximately linear behavior, indicative of monofractality. Note that uncorrelated white noise has Hurst exponent $h(q) = \frac{1}{2}$ and $\tau(q) = q/2 - 1$, so plotting $\tau - q/2$ will give a horizontal (i.e., constant) line. We have therefore plotted $\tau(q) - q/2$ instead of $\tau(q)$ for easier visualization.

The Legendre transform of Eq. (12) gives the multifractal singularity spectrum $f(\alpha_s)$:

$$\alpha_s = \mathrm{d}\tau(q)/\mathrm{d}q, \tag{13}$$

$$f(\alpha_s) = q\alpha_s - \tau(q). \tag{14}$$

Here, $f$ gives the dimension of the subset of the series characterized by the singularity strength $\alpha_s$. The literature also refers to $\alpha_s$ as the Hölder exponent, which represents a measure of the number of continuous derivatives that the underlying signal possesses. The singularity spectrum of monofractal time series will show a unique Hölder exponent, while a multifractal series will show a range of values for $\alpha_s$. Fig. 4(b) shows that

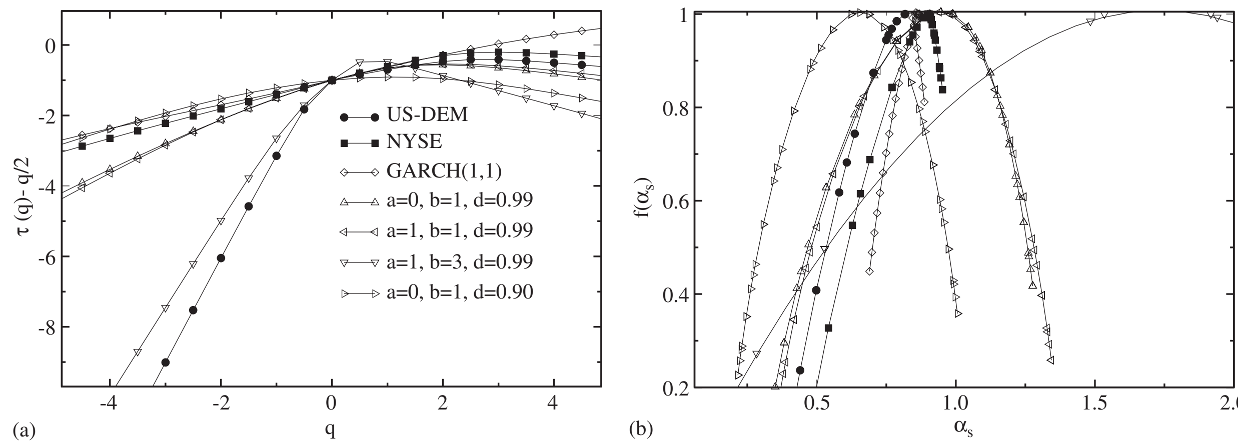


Fig. 4. (a) The multifractal scaling exponent $\tau(q) = qh(q) - 1$ for the absolute returns. For easier visualization, we have plotted $\tau - q/2$. (b) The multifractal singularity spectrum $f(\alpha_s)$. The proposed model, the NYSE and DEM/USD time series each have nonuniversal multifractal behavior. We see the multifractality in the nonlinearity of $\tau(q)$ and in the wide $f(\alpha_s)$, except for the GARCH series which has almost linear $\tau(q)$ and a narrow $f(\alpha_s)$, indicating monofractality.

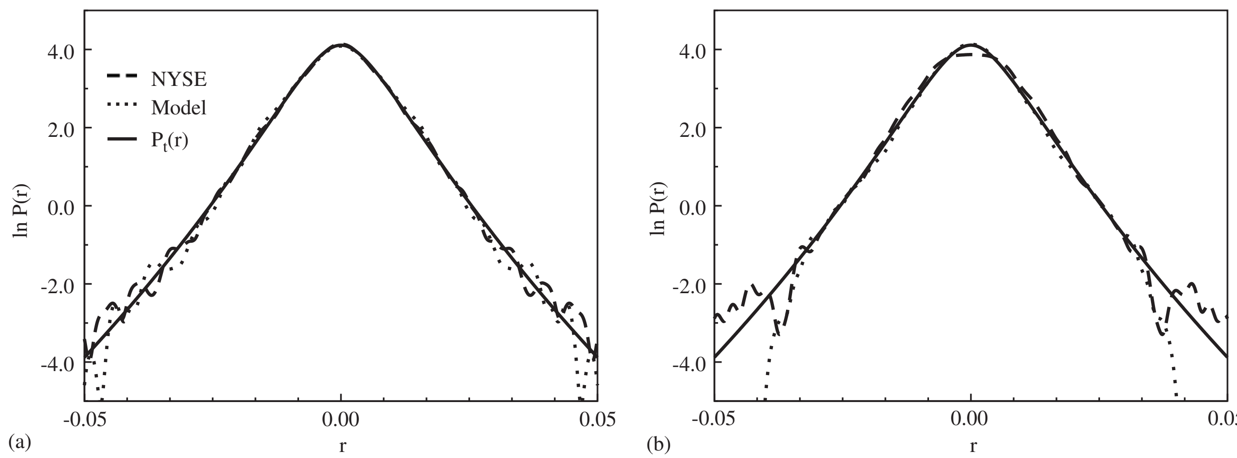


Fig. 5. Symmetrized PDF $P(r)$ of the returns $r_\tau$ measured over lags (a) $\tau = 1$ d and (b) $\tau = 25$ d (about one business month), shown for the artificial data (dashed line), the NYSE data (long dashed) and compared with the "theoretical" distribution $P_t(r)$ (solid line)). For $\tau = 1$ d all three distributions are almost identical, while for $\tau = 25$ we find surprisingly good agreement except for the expected minor discrepancy that appears for actual data for small returns.

except for the GARCH series, we obtain non universal multifractal singularity spectra. The GARCH model appears nearly monofractal, with a narrow $f(\alpha_s)$. Note that sometimes $\tau$ and $f(\alpha_s)$ appear defined differently (e.g., what we define as $f(\alpha_s)$ here would equal $-f(\alpha) + 1$ in the terminology of Refs. [55,56]). We use the definitions contained in Refs. [53,57].

### 2.5. Fat tails and time scaling in the PDF

Having clearly established the multifractal properties of the absolute return time series, we now turn our attention to the PDF. We consider here the PDF of returns for the two actual data sets (long dashed lines in Figs. 5 and 6) and the kinematic model time series (dashed lines). We expect that the model is able to reproduce the PDF and its scaling behavior. As parameters in the model we use the same two sets we already

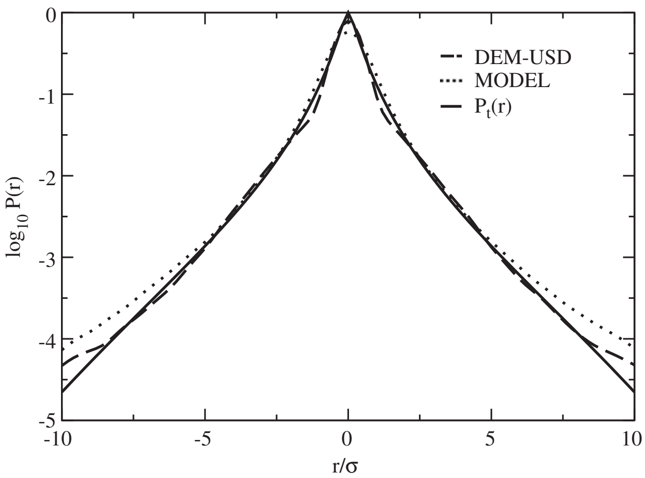


Fig. 6. Symmetrized PDF $P(r)$ of the returns $r_\tau$ shown for the high frequency DEM/USD series, the artificial data, and the theoretical distribution $P_t(r)$. We see that a Gaussian with log-normally distributed variance provides a good description of the absolute return distribution up until 10 standard deviations.

used for autocorrelations. We also compare the PDF of actual and artificial data with a "theoretical" distribution (solid line), which we introduce to obtain a better understanding of the form of the actual data and model PDF. The proposed "theoretical" distribution is a Gaussian $P_{\mathrm{t}}$ with log normally distributed local variance:

$$P_{\mathrm{t}}(r) = \int \rho(\sigma) \frac{\exp\left(\frac{-r^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}\, \mathrm{d}\sigma, \tag{15}$$

$$\rho(\sigma) = \frac{1}{\sqrt{2\pi}s\sigma} \exp\left[-\frac{1}{2}\left(\frac{\log\sigma - m}{s}\right)^2\right]. \tag{16}$$

Indeed, the resulting PDF is the convolution of Gaussian PDF with different (lognormally distributed) volatilities. We obtain two sets of constants for the two data sets, $s = 0.41$ and $m = -0.34$ for the NYSE series, and $s = 0.78$ and $m = -0.35$ for the DEM/USD series. In order to estimate the PDF of returns from the two data sets using the standard histogram method, we have used the equivalent smooth interpolation given by

$$P(r) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\Delta} \exp\left\{\frac{-(r-r_i)^2}{2\Delta^2}\right\}$$

with a smoothing window of $\Delta = 0.001$ (Fig. 5(a) for the NYSE index and $\Delta = 0.4$ for the DEM/USD rates (Fig. 6). The $r_i$ in this expression can be either from the actual or artificial data and $N$ is their number. The solid lines in Figs. 5(a) and 6 suggest that this distribution is among the better candidates for describing the distribution of price variations. The agreement found is unexpectedly strong.

We next study the PDF time scaling typically observed in financial times series. For both actual and artificial data we estimate (using the same smoothing window) the distribution for monthly (25 business days) returns (Fig. 5(b)) of the NYSE index

$$\frac{\sum_{i=1}^{25} r_{i+t}}{(25)^\delta} = \frac{1}{(25)^\delta}\log\frac{S_{t+25}}{S_t},$$

using the measured value $\delta = 0.53$ of the scaling exponent. Note that the solid line curve is exactly the same in the two figures, therefore both actual and artificial data exhibit almost perfect time scale invariance of the return distributions (see also Ref. [34]). The lack of agreement for small monthly returns is due to short-range

(1 d) correlations in the signs of the returns that are neglected in our model, but that could easily be incorporated [34]. Compared to other models of financial returns, this model shows remarkably good agreement with actual data.

## 3. Discussion

Our results show that the proposed kinematic model can reproduce the non universal multifractality and the PDF of the absolute returns. The model is able to generate a wide range of multifractal behavior. Nevertheless, note that we are not interested in obtaining perfectly identical multifractal properties for the proposed model and the actual data. Indeed, the observed non universality would invalidate such attempts.

A second important difference relates to how the proposed model generates volatility similarly to a multiplicative process, in spite of being an additive process. The similarity with multiplicative processes arises due to the log-normally distributed variance (Eq. (16)), which enters the model via the exponential term in Eq. (2). The additivity enters the model in the moving average of Eq. (2). This ingredient in the model is desirable in order to account for the approximately lognormal PDF of volatilities observed in actual financial time series.

Finally, we note that these results are obtained by means of the proposed Markov process and that the multifractal volatility correlations are not in the input, but are generated a posteriori by the model. This kinematic model is thus different from others (see for example Ref. [58]) where correlations are already present in the input ingredients, i.e., it is assumed a priori that increments are long-range correlated. Furthermore, the model presented here also seems to represent an advance due to the much improved agreement with actual data. Indeed, we find good agreement for up to 10 standard deviations in the PDF (Fig. 6). How strong this agreement is becomes particularly evident if one considers that the two data sets analyzed significantly differ in nature. One consists of a daily stock market index and the other of high frequency currency exchange rate data sampled up to every 2 s. An interesting question for future study is how to modify the model to account for the existence of return-volatility asymmetric correlations (i.e., the "leverage effect") [30,33,59]. Another interesting question concerns the implications for efficiency of a Markov model of financial returns.

## Acknowledgements

We thank the anonymous referees for very helpful suggestions. We thank Y. Ashkenazy and Michele Pasquini for discussions and the Brazilian research agencies CAPES, CNPq and FAPEAL. One of us (F.P.) acknowledges the financial support of Cofin MIUR 2002 prot. 2002027798_005.

## References

[1] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, 2000.
[2] J.-P. Bouchaud, A. Matacz, M. Potters, Phys. Rev. Lett. 87 (2001) 228701.
[3] B. Podobnik, P.Ch. Ivanov, Y. Lee, A. Chessa, H.E. Stanley, Europhys. Lett. 50 (6) (2000) 711.
[4] M. Pasquini, M. Serva, Euro. Phys. J. B 16 (2000) 195.
[5] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Phys. Rev. E. 60 (1999) 1390.
[6] P. Gopikrishnan, V. Plerou, L.A. NunesAmaral, M. Meyer, H.E. Stanley, Phys. Rev. E 60 (1999) 5305.
[7] M. Pasquini, M. Serva, Econ. Lett. 65 (1999) 275.
[8] R.N. Mantegna, H.E. Stanley, Nature 383 (1996) 587.
[9] R.N. Mantegna, H.E. Stanley, Nature 376 (1995) 46.
[10] L. Bachelier, Ann. Sci. École Norm. Sup. 17 (1900) 21.
[11] Newton Da Costa Jr., Silvia Nunes, Paulo Ceretta, Sergio Da Silva, Econ. Bull. 7 (3) (2005) 1–9.
[12] Z. Ding, C.W.J. Granger, R.F. Engle, J. Empirical Finance 1 (1993) 83.
[13] N. Crato, P. De Lima, Econ. Lett. 45 (1994) 281.
[14] J.-P. Bouchaud, D. Sornette, J. Phys. I France 4 (1994) 863.
[15] M. Mevy, S. Solomon, Int. J. Mod. Phys. C 7 (1996) 595.
[16] R.T. Baillie, J. Econometrics 73 (1996) 5.
[17] S. Gallucio, G. Caldarelli, M. Marsili, Y.-C. Zhang, Physica A 245 (1997) 423.

[18] N. Vandewalle, M. Ausloos, Physica A 246 (1997) 454.
[19] Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Physica A 245 (1997) 437.
[20] B. Podobnik, K. Matia, A. Chessa, P.Ch. Ivanov, Y. Lee, H.E. Stanley, Physica A 300 (2001) 300.
[21] R.F. Engle, A.J. Patton, Quant. Finance 1 (2001) 237.
[22] E. Bacry, J. Delour, J.F. Muzy, Phys. Rev. E 64 (2001) 026103.
[23] A.P. Kirman, G. Teyssiere, Microeconomic models for long-memory in the volatility of financial time series, in: Studies in Nonlinear Dynamics and Econometrics, vol. 5(4), Berkeley Electronic Press, 2002, p. 1083.
[24] R.F. Engle, Econometrica 50 (1982) 987.
[25] T. Bollerslev, J. Econometrics 31 (1986) 307–327.
[26] B.B. Mandelbrot, J. Bus. 36 (1963) 394.
[27] R.N. Mantegna, H.E. Stanley, Phys. Rev. Lett. 73 (1994) 2946.
[28] Ole E. Barndorff-Nielsen, Neil Shephard, J. Statistics 30 (2003) 277–295.
[29] R.T. Baillie, J. Econometrics 73 (1996) 5.
[30] J. Perell, J. Masoliver, Phys. Rev. E 67 (2003) 037102.
[31] J. Perell, J. Masoliver, N. Anento, Physica A 344 (2004) 134–137.
[32] G. Zumbach, Quant. Finance 4 (2004) 70.
[33] J. Perell, J. Masoliver, J.-P. Bouchaud, Appl. Math. Finance 11 (2004) 27–50.
[34] G.M. Viswanathan, U.L. Fulco, M.L. Lyra, M. Serva, Physica A 329 (2003) 273–280.
[35] U.L. Fulco, M.L. Lyra, F. Petroni, M. Serva, G.M. Viswanathan, Int. J. Mod. Phys. B 18 (2004) 681–689.
[36] R.F. Engle, A.J. Patton, Quant. Finance 1 (2001) 237–245.
[37] H.D. Jennings, P.Ch. Ivanov, A.M. Martins, P.C. Silva, G.M. Viswanathan, Physica A 336 (2004) 585–594.
[38] D. Panter, Modulation, Noise and Spectral Analysis, McGraw-Hill, New York, 1965.
[39] P.A. Samuelson, Ind. Manage. Rev. 6 (1965) 41–50.
[40] B.B. Mandelbrot, J. Bus. Univ. Chicago 39 (1966) 242.
[41] E.F. Fama, J. Finance 25 (1970) 383–417.
[42] F. Petroni, M. Serva, Euro. Phys. J. B 34 (2003) 495–500.
[43] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E 49 (1994) 1685.
[44] K. Hu, Plamen Ch. Ivanov, Zhi Chen, Pedro Carpena, H.E. Stanley, Phys. Rev. E 64 (2001) 011114.
[45] B.B. Mandelbrot, in: M. Rosenblatt, C. Van Atta (Eds.), Statistical Models and Turbulence, Springer, New York, 1972.
[46] B.B. Mandelbrot, J. Fluid Mech. 62 (1974) 331–358.
[47] B.B. Mandelbrot, Fractals and Scaling in Finance: Discontinuity, Concentration, Risk, Springer, New York, 1997.
[48] B.B. Mandelbrot, Sci. Am. 280 (1999) 70–73.
[49] B.B. Mandelbrot, A.J. Fisher, L.E. Calvet, Cowles Foundation Discussion Paper No. 1164, http://ssrn.com/abstract=78588.
[50] T. Lux, Quantitative Finance 1 (2001) 632–640.
[51] T. Lux, Int. J. Mod. Phys. 15 (2004) 481–491.
[52] A.L. Barabasi, T. Vicsek, Phys. Rev. A 44 (1991) 2730.
[53] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H.E. Stanley, Physica A 316 (2002) 87–114.
[54] J.F. Muzy, E. Bacry, A. Arneodo, Phys. Rev. Lett. 67 (1991) 3515.
[55] G. Paladin, A. Vulpiani, Phys. Rep. 156 (1987) 147–225.
[56] G. Paladin, A. Vulpiani, Phys. Rev. B 35 (1987) 2015.
[57] K. Matia, Y. Ashkenazy, H.E. Stanley, Europhys. Lett. 61 (2003) 422–428.
[58] E. Bacry, J. Delour, J.F. Muzy, Phys. Rev. E 64 (2001) 026103.
[59] J.-P. Bouchaud, A. Matacz, M. Potters, Phys. Rev. Lett. 87 (2001) 228701.