Corrupted bifractal features in finite uncorrelated power-law distributed data

Felipe Olivares [*], Massimiliano Zanin

Instituto de Física Interdisciplinar y Sistemas Complejos CSIC-UIB, Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

ARTICLE INFO

ABSTRACT

Multifractal Detrended Fluctuation Analysis stands out as one of the most reliable methods for unveiling multifractal properties, specially when real-world time series are under analysis. However, little is known about how several aspects, like artefacts during the data acquisition process, affect its results. In this work we have numerically investigated the performance of Multifractal Detrended Fluctuation Analysis applied to synthetic finite uncorrelated data following a power-law distribution in the presence of additive noise, and periodic and randomly-placed outliers. We have found that, on one hand, spurious multifractality is observed as a result of data finiteness, while additive noise leads to an underestimation of the exponents h_q for $q < 0$ even for low noise levels. On the other hand, additive periodic and randomly-located outliers result in a corrupted inverse multifractality around $q = 0$. Moreover, the presence of randomly-placed outliers corrupts the entire multifractal spectrum, in a way proportional to their density. As an application, the multifractal properties of the time intervals between successive aircraft landings at three major European airports are investigated.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Originated from the concept of multifractal sets [1–3], multifractality, i.e. when the scaling becomes a local property rather than a global one, has been used for characterising the local scaling behaviour in empirical signals coming from numerous areas of science, as for instance air traffic flow [4], financial data [5–9], brain electrical activity [10], human heartbeat dynamics [11], extreme events in atmospheric turbulence [12], river discharge and precipitation [13], seismic records [14,15] wind speed [16], musical signals [17], narrative texts [18,19] and avian influenza [20]. As evidenced, many experimental measurements require multiple scaling exponents for a proper characterisation of the underlying complex dynamics. The multifractal nature originates from the presence of nonlinear correlations for small and large fluctuations [21]. Nevertheless, a multifractal behaviour can also be observed due to broad probability distributions of the data [9,21] or to finite-size linearly long-term correlated time series [22]. When analysing real-world data, it is common to face an interplay between these three sources of multifractality. Nevertheless, multifractality originated by nonlinear correlations is normally considered the only real one, placing linear correlations in short sequences and the heavy tailed distributed fluctuations as ingredients yielding "spurious" multifractality [9]. Particularly, power-law distributed data or distributions with tails obeying a power-law give rise to the most simple case of multifractality—bi-fractality [21,23]— that is usually observed in systems exhibiting phase transitions [24] and Levy processes [23], as for instance financial markets [9].

---

* Corresponding author.
 E-mail address: felipe@ifisc.uib-csic.es (F. Olivares).

Among all multifractal approaches available in the literature [21,25–30], the most frequently used techniques to quantify the multi-scaling nature in synthetic and empirical data sequences are the Wavelet Transform Modulus Maxima method (WTMM) [26] and the Multifractal Detrended Fluctuation Analysis (MF-DFA) [21], the latter one being a generalisation of the classical DFA method [31]. Despite the fact that both methodologies can remove unwanted polynomial trends in the data, MF-DFA one stands out over the WTMM approach for being more accurate with shorter signals. Furthermore, it is more reliable in properly detecting monofractal and bifractal nature [24], and easier to implement [32,33]. In general, MF-DFA is recommended when the multifractal properties of the signal under analysis are unknown *a priori* [24]. Specifically, MF-DFA generalises the classical Hurst exponent to a local exponent h to unravel the heterogeneous scaling in the data. In such a manner, it identifies and quantifies local scaling behaviours—multiple scaling exponents.

The identification and quantification of multifractal features in empirical data is a challenging task due to the unavoidable contamination with artefacts inherent to the data acquisition process [34–38]. It is worth mentioning here that Ludescher et al. [36] have studied the performance of the MF-DFA method for characterising mono and multifractal synthetic sequences under the influence of additive noise, short-term memory and periodicities. Particularly, they found that multifractal analysis can easily be biased by those artefacts; for instance, multifractality is underestimated for multifractal signals with additive periodic trends. Later, following the same idea, Gulich et al. [37] have studied the spurious multifractality generated by the influence of additive coloured noises in multifractal time series. Still further, it has been shown that a spurious broad multifractal spectrum is obtained when monofractal records having local isolated singularities are considered, such as records with outliers due to measurement errors [38].

The aim of this work is to address how bifractal properties manifest in finite uncorrelated data obeying a power-law distribution, in the presence of some artefacts usually related to empirical recordings, by implementing MF-DFA methodology. More specifically, we focus on (i) finite size effects, since experimental data sets are constrained by data availability and stationarity, (ii) the influence of additive noise that is inherent to any measurement, and (iii) the presence of periodic and random outliers, both intrinsic to the underlying dynamics or due to errors in the data acquisition process. Additionally, our findings are applied to the study of real-world time series describing the time between successive landings at three major European airports.

2. Multifractal detrended fluctuation analysis (MF-DFA)

The MF-DFA algorithm has been described in [21], and a detailed implementation can be found in [32,33]. Yet, for the sake of completeness, its main elements are described below.

Given a time series $\mathcal{X}_t = \{x_t, t = 1, \ldots, M\}$, with M being the number of observations, the cumulated time series $Y(i) = \sum_{t=1}^{i}(x_t - \langle x \rangle)$ is considered, where $\langle x \rangle$ stands for mean value. This profile is divided into $\lfloor N/s \rfloor$[1] non-overlapping windows of equal length s. A local polynomial fit $y_{\nu,m}(i)$ of degree m is fitted to the profile of each window $\nu = 1, \ldots, \lfloor N/s \rfloor$. The degree of the polynomial can be varied to eliminate constant ($m = 0$), linear ($m = 1$), quadratic ($m = 2$), or higher order trends of the profile. The variance of the detrended time series is aftward evaluated by averaging over all data points i in each segment ν,

$$F_m^2(s) = \frac{1}{s} \sum_{i=1}^{s} \left\{ Y[(\nu - 1)s + i] - y_{\nu,m}(i) \right\}^2, \tag{1}$$

for $\nu = 1, \ldots, \lfloor N/s \rfloor$. Aiming at analysing the influence of fluctuations of different magnitudes and at different time scales, the generalised qth order fluctuation function is defined by

$$F_q(s) = \left\{ \frac{1}{\lfloor N/s \rfloor} \sum_{\nu=1}^{\lfloor N/s \rfloor} [F_m^2(\nu, s)]^{q/2} \right\}^{1/q}. \tag{2}$$

When $q = 0$, a logarithmic averaging procedure has to be employed because of the divergent exponent

$$F_0(s) \equiv \exp \left\{ \frac{1}{2\lfloor N/s \rfloor} \sum_{\nu=1}^{\lfloor N/s \rfloor} \ln[F_m^2(\nu, s)] \right\}. \tag{3}$$

For $q = 2$ the classical fractal DFA algorithm is retrieved [31]. Generally, for long-term power-law correlated data, it holds that

$$F_q(s) \sim s^{h_q}, \tag{4}$$

inside a certain range of s. The scaling exponents h_q are usually known as generalised Hurst exponents, and allow accounting for heterogeneous scaling. Ideally, for a monofractal time series, h_q is independent of q and equal to the Hurst exponent, H. Its value quantifies the degree of correlation in the data: if $H = 0.5$ the time series is uncorrelated, while

---

[1] $\lfloor c \rfloor$ denotes the largest integer less than or equal to c.

$H > 0.5$ indicates long-term correlations or persistence. On the other hand, antipersistent behaviour is characterised by $H < 0.5$.

A bi- or multifractal structure is observed when the scaling behaviours of small and large fluctuations are different. In this case h_q decreases with q, and the main Hurst exponent can be estimated from the second moment ($h_2 = H$). More specifically, the generalised Hurst exponent with negative order q describes the scalings of small fluctuations, because the segments ν with small variance dominate the average F_q. On the contrary, for positive orders q, the windows ν with large variance have a stronger influence, and thus h_q focuses on large fluctuations. The strength of the multifractality present in the data is usually defined as the spread of the generalised Hurst exponent [22]. As small fluctuations are characterised by larger scaling exponents than those associated with large fluctuations, the multifractality degree can be quantified by

$$\Delta h_q \equiv h(q_{min}) - h(q_{max}), \tag{5}$$

where q_{min} and q_{max} are respectively the minimal and maximal value of the moment q considered in the analysis. Normally $q_{min} = -q_{max}$ is used. Another way to quantify multifractal features is by measuring the width of the singularity spectrum [21,39]. For further details about this methodology and its implementation see [32]. Specifically for a MATLAB implementation we recommend Ref. [33].

3. Numerical analysis

3.1. A model of bi-fractal spectra

We consider finite uncorrelated time series with power-law distribution function:

$$P(x) = x_{min} x^{-(\alpha+1)} \quad \text{for } x_{min} \leq x < \infty, \tag{6}$$

with $\alpha > 0$, and x_{min} corresponding to the smallest value of x for which the power law holds [40]. In the present study we set $P(x) = 0$ for $x < x_{min}$ [31]. Yet, empirical data normally do not follow a power-law over their entire range; in fact, the distribution usually deviates from the power-law from below the minimum value x_{min} [40]. For $\alpha \leq 2$, time series distributed according to Eq. (6) exhibit multifractal scaling behaviours on all scales [21]. It was shown in [31] that the generalised Hurst exponent for data distributed according to Eq. (6) can be expressed as

$$h_q = \begin{cases} 1/\alpha & q \leq \alpha, \\ 1/q & q > \alpha, \end{cases} \tag{7}$$

which describes a bi-fractal nature, i.e. a monofractal behaviour ($h_q = $ constant) for $q \leq \alpha$, while a multifractal nature for $q > \alpha$ [21].

3.2. Synthetic data generation

For generating numerical uncorrelated sequences distributed according to a power-law we consider a transformation procedure. Being r_i random real numbers uniformly distributed in the interval $[0, 1]$, these are transformed according to $r_i \rightarrow x_i = x_{min} r_i^{-1/\alpha}$, to obtain random power-law distributed real numbers x_i in the range $[x_{min}, \infty)$ [21]. We have generated a set of one hundred independent realisations of length $N = 10^n$ with $n = 3, 4, 5, 6$, different values of $\alpha \in [0.5, 2]$, and $x_{min} = 1$. The MF-DFA method considers 30 time scales $s \in [10, N/10]$ equally distributed in a logarithmic scale. We set the moments $q \in [-10, 10]$ with a step equal to 0.1. Results shown in the present work were obtained by using a detrending polynomial of second order ($m = 2$). Quantities averaged over one hundred realisations are reported hereafter.

3.3. Finite-size effects

We firstly investigate the validity of the theoretical model for h_q, given by Eq. (7), for short time series. Fig. 1(a) shows the generalised fluctuation function for $\alpha = 1$ and $N = 10^6$. Only integer values of the moment q are depicted. A good scaling behaviour is observed over the whole range of temporal scales considered and for all moments q. The slope of the $F_q(s)$ in the log–log scale decreases with q. These results qualitatively hold for all $\alpha \in [0.5, 2]$. The slopes h_q are depicted in Fig. 1(b) as a function of q, as estimated by a linear least-squares fit and for different time series length. On one hand, we found a good agreement with the numerical simulation for $q > \alpha = 1$, even for short time series. On the other hand, for $q < \alpha$ the theoretical monofractal behaviour is corrupted, mimicking a multifractal behaviour, i.e. h_q decreasing with q. Fig. 2(a)–(b) illustrate this spurious multifractality for different values of α and q. For $q = \alpha$, the exponent h_q is underestimated in comparison with the theoretical prediction, for all lengths N considered in this analysis. Yet, for negative moments q, h_q approaches from above as the time series length increases—see Fig. 2(b). Note that the convergence to the scaling $h_q = 1/\alpha$ with N is slower as α grows.

Furthermore, with the purpose of characterising the multifractal behaviour for $q > \alpha$, we define

$$h_q \sim q^{-\mu}, \tag{8}$$

where the exponent $\mu \equiv \mu(N, \alpha)$ accounts for deviations from the theoretical prediction ($\mu = 1$). Fig. 2(c) shows the estimated exponents μ as a function of α. We found that, independently of the length N, the theoretical expected value is recovered up to $\alpha = 1$. Yet, the exponent μ decreases for $\alpha \in [1, 2]$.
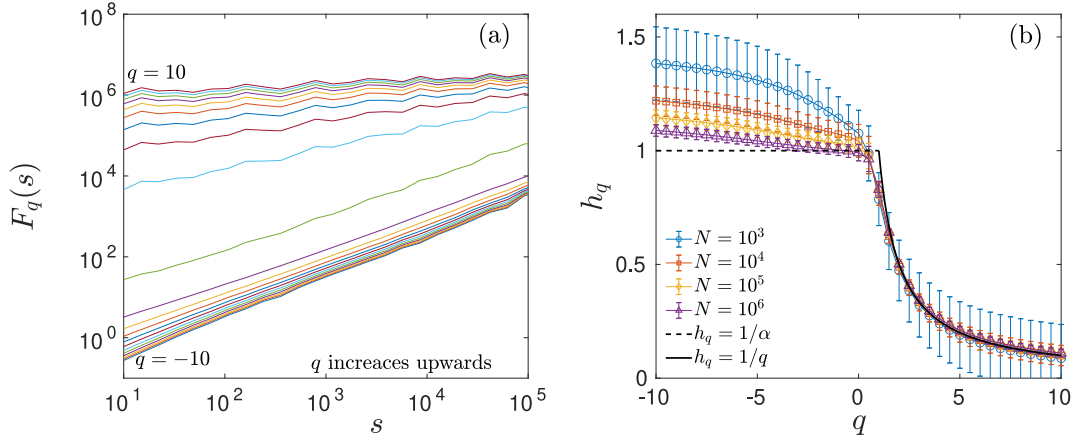
Fig. 1. (a) Generalised Fluctuation function as a function of the time scale s with $q \in [-10, 10]$ for $\alpha = 1$ and $N = 10^6$ (only integer values of q are depicted for sake of clarity). (b) Generalised Hurst exponent h_q a function of q for the same value of α and for different values of the time series length N. Dashed and solid lines describe the theoretical curve for power-law distributed data with $\alpha = 1$ at long scales for $q \leq \alpha$ and $q > \alpha$ respectively.
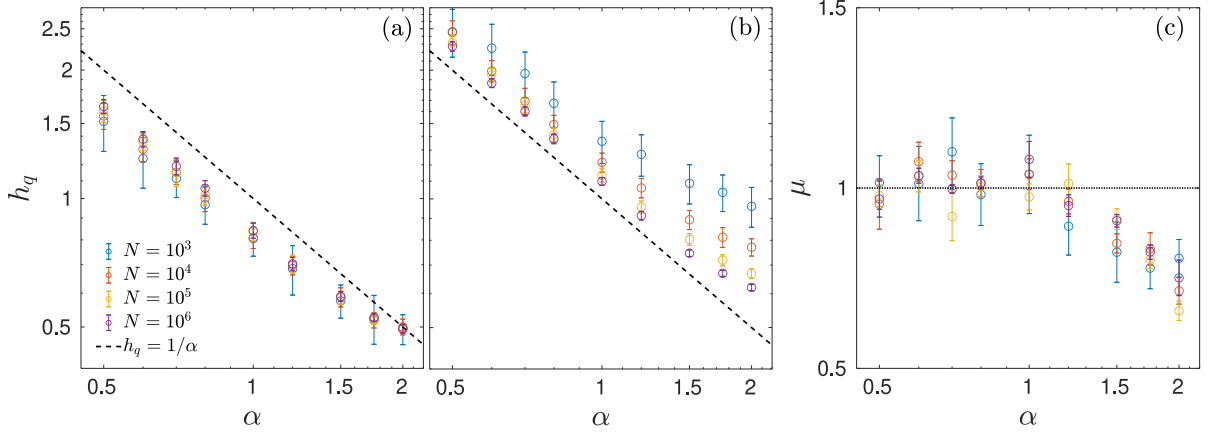

Fig. 2. (a) Generalised Hurst exponent for $q = \alpha$ as a function of α for different values of N. (b) same as (a) but with h_q evaluated in $q = -10$. The dashed black lines in both previous panels represent the theoretical expected power law $h_q = 1/\alpha$. (c) Exponent μ as a function of α for different values of N.

3.4. Influence of additive noise

We next study the influence of additive noise. For that, we define a noisy finite sequence as follows:

$$y_i = x_i + A\eta_i, \tag{9}$$

where x_i is a finite random time series distributed according to Eq. (6), η_i is a Gaussian white noise with zero mean and unit variance, and A_n is the amplitude of the noise contamination defined as a multiple of the standard deviation of the original sequence $A = A_n \sigma_{x_i}$ ($A_n \in \mathbb{R}$). For generating the numerical time series we set the length $N = 10^5$.

The generalised Hurst exponents are depicted in Fig. 3, as a function of the moment q, for $\alpha = (0.5, 1, 2)$, and considering several intensities of additive noise. It stands out that, for $q < 2$, the exponents h_q are considerable underestimated, even for small amount of noise—negative values of the moment q describes small fluctuations, which are closer to the noise level. Of course, for very large noise levels, as for instance for $A_n = 10$, the exponent h_q becomes constant and equal to 0.5, since the monofractal noise dominates the dynamics. This behaviour was previously observed in numerical simulations of noisy multifractal cascades [36,37]. Interestingly enough, additive noise offsets the spurious multifractal behaviour due to finite size effects as $\alpha \to 2$ (see Fig. 3(c)), since $h_{q<\alpha} \to 0.5$. These two biases compete against each other and could eventually cancel out, e.g. when having short sequences in a noisy environment. Nevertheless, the multifractal degree is, without a doubt, underestimated for $\alpha < 1$.
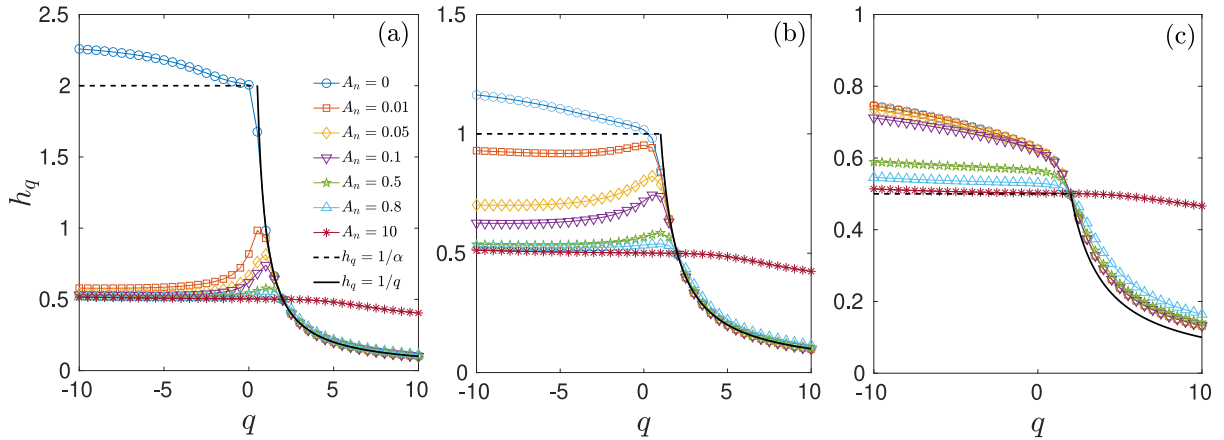
Fig. 3. Generalised Hurst exponent as a function of q for different values of the noise intensity A_n for (a) $\alpha = 0.5$, (b) $\alpha = 1$ and (c) $\alpha = 2$. No error bars are depicted for the sake of clarity.

3.5. Influence of additive periodic outliers

We consider the influence of periodic outliers by defining:

$$y_i = x_i + I\,\delta_\tau(i), \quad \text{with } \delta_\tau(i) = \sum_{i=\infty}^{-\infty} \delta(t - i\tau), \tag{10}$$

where the original sequence x_i is perturbed by a periodic spike characterised by the intensity I, defined in terms of the standard deviation of x_i, $I = A_p \sigma_{x_i} + \eta_i$ ($A_p \in \mathbb{R}$ and η_i is a Gaussian white noise with zero mean and unit variance), and with δ being the Dirac delta function. For generating the numerical time series we set the length $N = 10^5$ and the period $\tau = 720$, which can represent, for instance, monthly spikes in hourly sampled data.

Fig. 4 shows $F_q(s)$ as a function of the time scale s, for $\alpha = 1$ and for intensities of the periodic outliers (a) $A_p = 1$ and (b) $A_p = 100$. A crossover at $s_c \sim \tau$ is observed. This spurious crossover is generated by the presence of periodic trends [41]; and in this case it can be attributed to the added periodic outliers. The larger the intensity of the outlier event, the larger the perturbation of the generalised fluctuation functions for $s > \tau$. However, we observe that it is still feasible to estimate of the slope h_q if we restrict the fitting procedure to short time scales only, e.g. to $s < 300$ data points. To illustrate, the estimated slopes h_q are depicted in Fig. 5(a). We found that the multifractal nature is corrupted around $q = 0$. The exponent h_q is overestimated as A_p increases, since large events interfere mostly with moments in the range $[-1, 1]$. In fact, a good<30 1 linear scaling in the log–log scale of $F_q(s)$ with s is not observed, as most clearly ev9+306idenced for $q = 0$, purple dashed line in Fig. 4(b). Therefore, with strictness, those exponents are not valid as a measure of the fractal scaling. In synthesis, a spurious inverse multifractality is observed up to $q = 0$. Yet, for $q > \alpha$, the decreasing behaviour of h_q with q remains the same. Consequently, the multifractal degree estimated by Eq. (5) is not dramatically affected. Same conclusions can be drawn for other values of α.

The maximum value reached by h_q gives account, in some way, of the intensity of the spikes added to the sequence, as can be seen in Fig. 5(b). As α increases, the overestimation of h_q for low A_p is compensated. Undeniably, when the intensity of the event is sufficiently high, the estimated exponents h_q are always overvalued, independently of α. A similar behaviour is observed when different periodicities are considered. Fig. 6(a) summarises the results for $\alpha = 1$ and different values of τ. The same range of temporal scales ($s \in [20, 100]$) for all τ were considered to estimate h_q. On one hand, as the period increases, for a fixed intensity A_p, the inverse corrupted multifractality around $q = 0$ is magnified. On the other hand, independently of the period, the maximum value of h_q around $q = 0$ linearly scales with the logarithm of A_p (for large intensities), as can be seen in Fig. 6(b). As pointed out before, the observation of the crossover at s_c arises due to additive periodic outliers, and the scale s_c can approximately determine their period. Fig. 5(c) compares s_c with the real period τ. The estimated values s_x obtained from the observed crossover scales are in a good agreement with the periods τ considered in corrupted the synthetic sequences (Eq. (10)).

3.6. Influence of randomly located outliers

Lastly, we examine the effects when randomly located outliers perturb the original sequence x_i. Fig. 7(a)–(b) summarises the results for sequences of length $N = 10^5$, characterised by $\alpha = 1$ and perturbed with randomly located outliers with an intensity $A_p = 100$. We found that the linear scaling of the function F_q with s is corrupted in a qualitatively similar way as though periodic spikes were perturbing the sequence, as can be seen in Fig. 7(a). It is observed that even for a
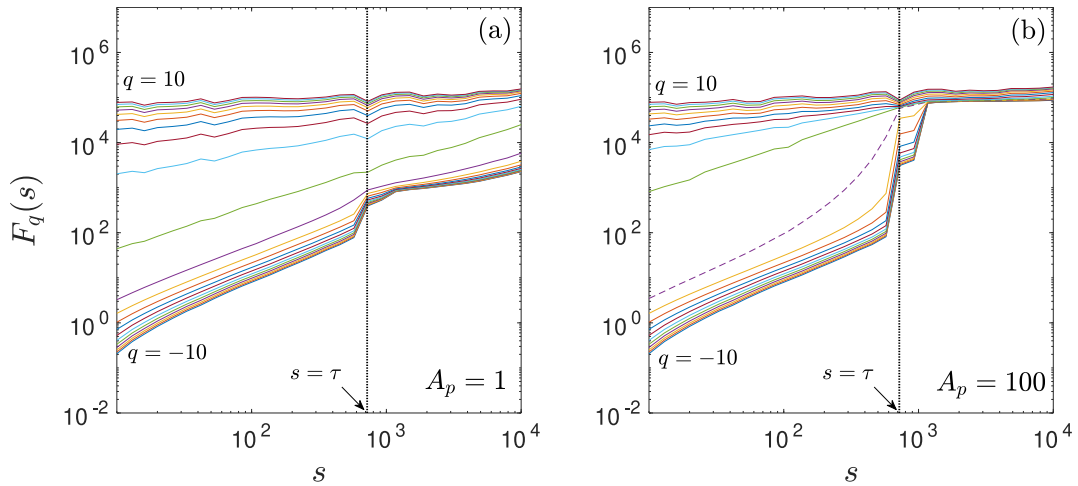
Fig. 4. Generalised fluctuation function as a function of the time scale s with $q \in [-10, 10]$ for $\alpha = 1$, $N = 10^5$, for (a) $A_p = 1$ and (b) $A_p = 100$. In (b) purple dashed line represents $q = 0$. Only integer values of q are depicted for the sake of clarity.
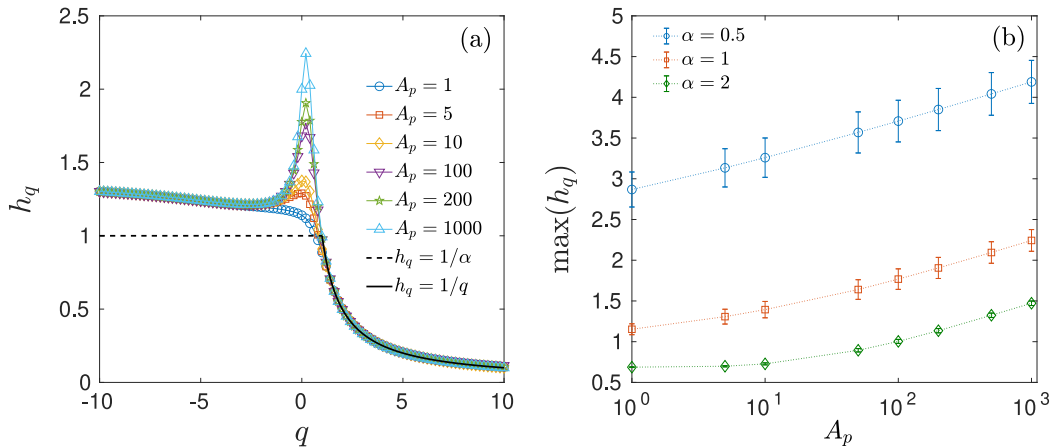

Fig. 5. (a) Generalised Hurst exponent as a function of q, estimated in the range $s \in [20, 300]$, for $\alpha = 1$, $\tau = 720$ and different values of A_p; and (b) maximum value of h_q around $q = 0$ as a function of A_p. In panel (a) no error bars are depicted for the sake of clarity.
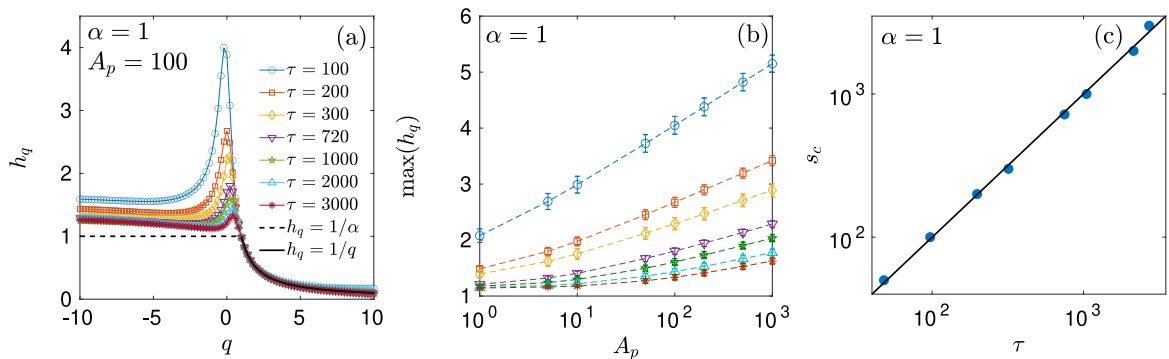

Fig. 6. (a) Generalised Hurst exponent as a function of q for $\alpha = 1$, $A_p = 100$ and different values of τ (no error bars are depicted for the sake of clarity), (b) maximum value of h_q around $q = 0$ as a function of A_p; and (c) the crossover scale s_c as a function of the period τ.
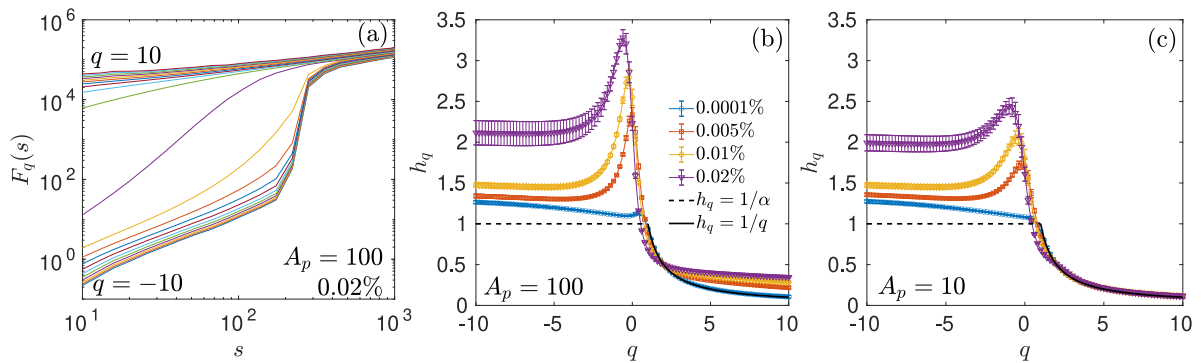
Fig. 7. (a) Generalised fluctuation function as a function of the time scale s with $q \in [-10, 10]$ for $\alpha = 1$, $N = 10^5$, and 0.02% of the data being random outliers with an intensity $A_p = 100$ (only integer values of q are depicted for the sake of clarity). Generalised Hurst exponent as a function of q for $\alpha = 1$, and different percentages of randomly located events, for (b) $A_p = 100$ and (c) $A_p = 10$.

Table 1
Main characteristics of the considered airports. First column: Airports names. Second column: night inactivity period in hours. Third column: average number of flights per day. Fourth column: time series lengths.

Airports	Night inactivity period (hr.)	Averaged #flights/day	N
Frankfurt	6	740	451534
Heathrow	6.5	531	323766
Tegel	7	79	48112

small amount of perturbing events, e.g. 0.02% of the data, the multifractal spectrum is dramatically squeezed upwards at large temporal scales.

Fig. 7(b) contrasts the exponents h_q estimated from sequences with different percentages of the data being randomly located outliers. For negative moments, there exists an overestimation of the exponents h_q, proportional to the amount of events in the data. On the other hand, for negative values of q, the theoretical decay $\sim q^{-1}$ is not longer valid. Moreover, not even the model $\sim q^{-\mu}$ can be fitted. Consequently, the entire multifractal spectrum is corrupted, and thus, the multifractal degree estimated by Eq. (5). In contrast, the effect on the exponent estimation when having perturbing outliers with low intensity $A_p = 10$, is limited to $q < 0$, as observed in Fig. 7(c). Not least, the peak of h_q appreciated around $q \in 0$, accounts for the presence of outliers as aforementioned.

4. Application to landing return interval dynamics

In order to show how these results can impact real-world analyses, we study time series describing aircraft landing dynamics. More specifically, we estimate the time between consecutive landings – landing return interval – starting on May 1 of 2018 and ending on December 31 of 2019 (609 days) from Frankfurt, Heathrow and Tegel airports. Estimated landing times were extracted from ADS-B position reports, obtained from the OpenSky Network (https://opensky-network. org) [42]. ADS-B (Automatic Dependent Surveillance - Broadcast) is a technology allowing aircraft to continuously send radio messages, stating their position and other information of relevance [43,44]; these messages are then received by ground stations, and integrated into coherent reports. For each airport, flights have been identified as performing a landing when the last known position was within a radius of 3 nautical miles from the center of the airport, and the last reported altitude below 500 m. Note that there exists observational noise as a result of the uncertainty associated with the time of landing.

Table 1 lists, in the second column, the average number of flights per day. Additionally, being these airports close to populated areas, operations are not allowed at night (except for emergencies and other specified exceptions); the third column then reports the duration of this night inactivity period, as extracted from the corresponding Jeppesen's airport charts. Depending of the airport activity, different length N of the landing return interval are obtained. The fourth column in Table 1 shows N for each airport, being Tegel airport the one with the least number of landings per day and with the longest inactivity period.

As a representative time series, in Fig. 8(a) we illustrate a small portion of the landing return interval sequence from Frankfurt airport. Note that the vertical axis is in logarithmic scale for better visualisation of the outliers, corresponding to the inactivity time during the night. Fig. 8(b) shows the histograms, by using logarithmic binning [45], of the return interval sequences from the three airports. It is observed that a power-law scaling applies only for times larger than $t_{\min} \sim 2$ and less than 100 min. We can therefore say that the distributions follow a power-law upon a certain range [40]. For comparison purposes, we have depicted a power-law scaling of t^{-2} and t^{-3}—see dotted and dashed black lines in
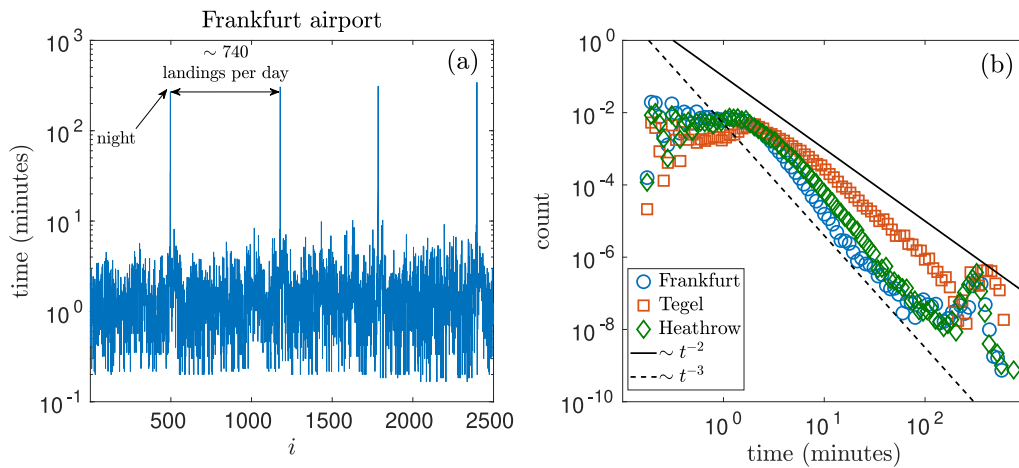
Fig. 8. (a) A portion of the landing return intervals from Frankfurt airport in semilog scale. (b) Histogram of the landing return interval for Frankfurt (solid blue circles), Tegel (solid red squares) and Heathrow (solid green diamonds) airports. Dotted and dashed black lines indicate power law t^{-2} and t^{-3} respectively.

Table 2
Multifractal properties of the considered airports. First column: Airports names. Second column: α value estimated by a linear fitting from the histogram in the range time $\in [2, 100]$. Third column: Hurst exponents $h_2 = H$. Fourth column: μ exponent estimated in the range $q \in [2, 5]$. SD stands for the error from the fitting procedure.

Airports	$\alpha \pm$ SD	$H \pm$ SD	$\mu \pm$ SD
Frankfurt	2.3 ± 0.1	0.53 ± 0.01	0.90 ± 0.005
Heathrow	2 ± 0.1	0.58 ± 0.02	0.76 ± 0.010
Tegel	1.2 ± 0.2	0.53 ± 0.02	0.56 ± 0.017

Fig. 8(b), respectively. By a simple linear least-squares fit in the range (2, 100) minutes we have estimated the power-law parameter α, reported in the second column in Table 2. We found a correlation between the total number of flights per day and the distribution parameter α. The more flights, the shorter the separation time between them, and consequently, the lower the probability of finding two flights widely separated in time. For times larger than approximately 100 min, a peak is observed, corresponding to the night inactivity period. Bearing in mind the noisy environment inherent to the data acquisition procedure, and furthermore, that inactivity times may be considered as outliers of the power-law distribution, these empirical sequences are ideal to study how these artefacts may corrupt their multiscaling properties.

We have analysed the multifractal properties of the landing return intervals with a second order polynomial. Sixty temporal scales $s \in [10, 10^4]$ equally distributed in the logarithmic scale were considered and $q \in [-5, 5]$ with a step equal to 0.1. Fig. 9 shows the F_q as a function of s for the landing return intervals for each airport. The time between the closing and opening hours squeezes upwards the scaling for a windows size $s > 1000$ for Frankfurt and Heathrow airports. On the other hand, since Tegel has less flights per day, the scaling is corrupted at smaller time scales. These results show that the night inactivity period is having the role of a periodic outlier.

By setting a fitting range $s \in [15, 100]$ for the three airports, we have estimated h_q for all values of q. These results are depicted in Fig. 10 (blue open circles). Slightly linear correlations are found, which are quantified by the classical Hurst exponent—see the intersection between solid vertical black line and h_q in Fig. 10. The second column of Table 2 summarises the values of $H = h_2$ for each airport. The generalised exponents h_q are overestimated in the range $q \in [0, 1]$, most probably because of the long return times, i.e. of the nights. Out of this region, the evolution of h_q behaves as one could expect from the probability distribution of the data, that is, bi-fractal. We have estimated the exponent μ in the range $q \in [2, 5]$—see last column in Table 2. These results are in accordance with those obtained by using synthetic data (compare with Fig. 2(c)).

In order to clarify the origin of the observed multifractality, we have contrasted the evolution of h_q estimated from the original data, with what was obtained from surrogate data. By shuffling the data points, all temporal correlations are removed, and only the multifractal nature due to the distribution probability remains. Open red squares in Fig. 10 show these results. One can observe that the multifractality is slightly affected for Frankfurt and Heathrow airports, indicating that the corrupted bi-fractality is mostly originated by the probability distribution. On the contrary, the result for Tegel airport shows some differences for negative moments. We hypothesise that this could be a product of the presence of nonlinearities in the dynamics. Certainly, a Hurst exponent equal to 0.5 is estimated for the shuffled versions of all the three sequences. Finally, by considering a Fourier Transform surrogate approach [9] to retain all temporal correlations but
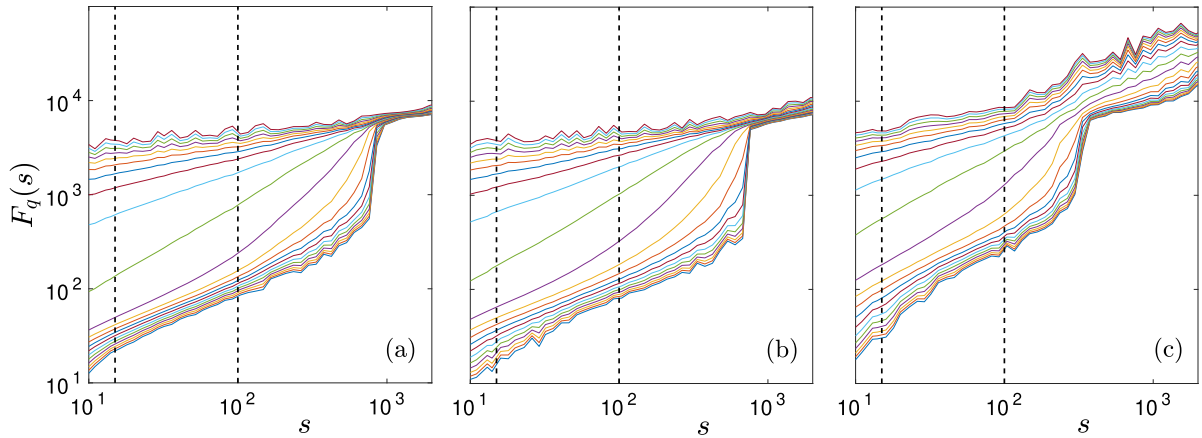
Fig. 9. F_q as a function of s for the landing return interval for (a) Frankfurt, (b) Heathrow, and (c) Tegel airports. Black dashed lines indicate the temporal range to estimate the exponents h_q. Only integer values of the moments q used to estimate the generalised fluctuation function are depicted.
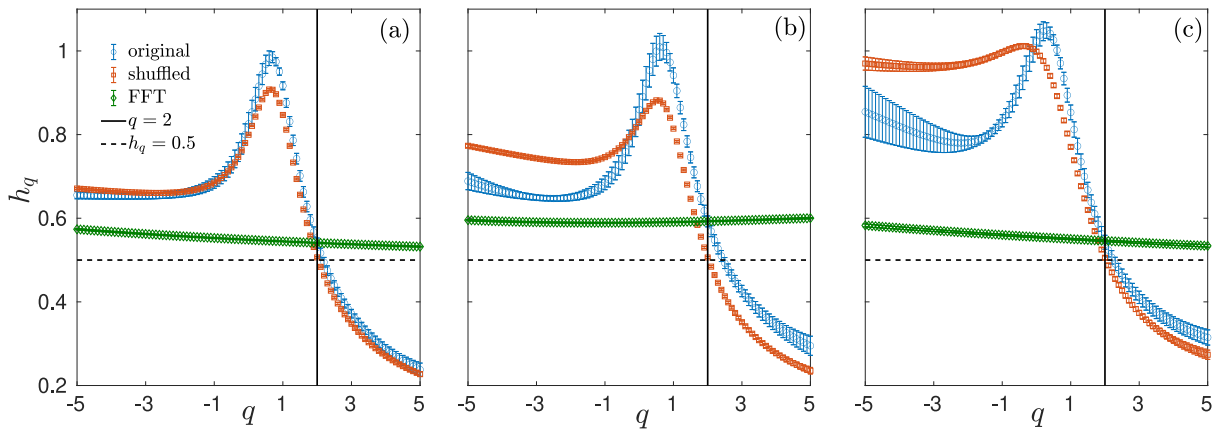

Fig. 10. Generalised Husrt exponents as a function of q for (a) Frankfurt, (b) Heathrow and (c) Tegel airports. Error bars stand for the error from the fitting procedure for the original sequences. Solid black lines indicate the second moment $q = 2$, and therefore the classical Hurst exponent $h_2 = H$. Dashed black lines show $h = 0.5$ (totally uncorrelated dynamics).

changing the probability distribution to a Gaussian one by phase randomisation [46], we found that the exponents h_q are constant in the range $q \in [-5, 5]$ and close to the Hurst exponent estimated from the original sequences—see open green diamonds in Fig. 10. These results confirm the origin of the bi-fractality.

Lastly, we focus on the presence of linear correlations in the landing return intervals per day. We consider Frankfurt and Heathrow airports only, in order to get sequences long enough, and consequently a reliable estimation of H. Fig. 11(a) and (b) show a histogram of the Hurst exponents estimated for the 609 days of Frankfurt and Heathrow airports respectively. We found that the landing return interval dynamics is linearly correlated. By comparing this to the exponent H estimated from the whole data set (nights included), we observe that the presence of outliers underestimate the degree of temporal correlation—see dashed black lines in Fig. 11(a) and (b). To corroborate this finding, we have simulated linearly correlated data with $H = 0.6$ and 0.8 following a power-law distribution with $\alpha = 2$, corrupted by additive periodic outliers, with $\tau = 720$ and different amplitudes, as described in Section 3.5. For generating temporal correlated time series with the desired probability distribution, we have followed the iterative algorithm introduced by Schreiber and Schimitz [47]. In Fig. 11(c) is observed that the degree of correlation is underestimated due to the presence of additive periodic outliers. The larger the amplitude A_p, the closer is the estimated Hurst exponent to 0.5. Then, we conclude that the intermittency in the landing return interval dynamics due to nights (large and periodic return intervals) leads to a slightly decrease of the linear correlation. From an applied point of view, the dynamics of landing return intervals can be used as a way of characterising the underlying airport dynamics. Specifically, one may hypothesise that the time between successive landings may present correlations when the airport reaches a saturation state, i.e. when the landing of one aircraft is limited by those of the aircraft preceding it. Results here reported indicate that the calculation of a simple H may not be enough, as the presence of outliers and observational noise leads to an underestimation.
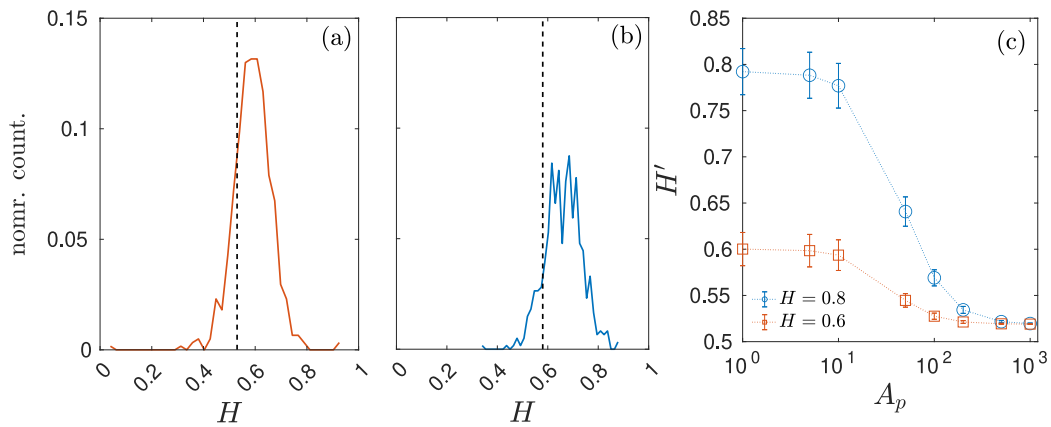
Fig. 11. Histogram of the estimated Hurst exponents H per day for (a) Frankfurt and (b) Heathrow airports. Dashed black lines indicate H estimated from the entire time series (nights included). (c) Averaged H, over 100 independent realisations, as a function of the amplitude of the outliers A_p added to a linear correlated sequence following a power-law distribution with $\alpha = 2$ ($N = 10^5$ and $\tau = 720$).

5. Conclusions

By investigating the performance of MF-DFA applied to finite uncorrelated power-law distributed data in the presence of additive noise and periodic and random-located outliers, we have evidenced that the original bi-fractal properties of the data are considerable corrupted by those artefacts. In particular, short time series overestimate exponents h_q, for negative moments and for $\alpha \in [0.5, 2]$, which leads to a spurious multifractal spectrum rather than a bi-fractal one. This result may be useful for analysing short time series, to avoid erroneous conclusions. When dealing with observational noise, we found that bi-fractality is corrupted for $q < 0$. However, for $\alpha = 2$ we found a noise-enhanced bi-fractality phenomenon, i.e. the presence of noise offsets the spurious multifractality originated by finite-size effects. This result is valuable for analysing experimental data characterised by $\alpha \sim 2$. The influence of additive periodic outliers corrupts the bi-fractal spectrum around $q = 1$, since the outliers affect the moments close to the mean value. This effect does not change the estimation of the multifractal degree, yet it generates an inverted behaviour of the exponent h_q with q. Finally, randomly-located outliers deform the entire bi-fractal spectrum, making the theoretical prediction not longer valid for all moments q. We believe that the present study gives a more solid characterisation of the bi-fractal nature in empirical time series by MF-DFA approach.

CRediT authorship contribution statement

Felipe Olivares: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Massimiliano Zanin:** Supervision, Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Felipe Olivares reports financial support was provided by European Research Council. Massimiliano Zanin reports financial support was provided by European Commission Seventh Framework Programme for Research and Technological Development IDEAS The European Research Council. Felipe Olivares reports equipment, drugs, or supplies was provided by State Agency of Research. Massimiliano Zanin reports equipment, drugs, or supplies was provided by State Agency of Research

Acknowledgements

References

[1] R. Benzi, G. Paladin, G. Parisi, A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, J. Phys. A: Math. Gen. 17 (18) (1984) 3521.

[2] A.B. Chhabra, C. Meneveau, R.V. Jensen, K. Sreenivasan, Direct determination of the f(α) singularity spectrum and its application to fully developed turbulence, Phys. Rev. A 40 (9) (1989) 5284.

[3] C. Meneveau, K. Sreenivasan, Measurement of f(α) from scaling of histograms, and applications to dynamical systems and fully developed turbulence, Phys. Lett. A 137 (3) (1989) 103–112.

[4] X. Zhang, H. Liu, Y. Zhao, X. Zhang, Multifractal detrended fluctuation analysis on air traffic flow time series: A single airport case, Physica A 531 (2019) 121790.

[5] D.O. Cajueiro, B.M. Tabak, Long-range dependence and multifractality in the term structure of LIBOR interest rates, Physica A 373 (2007) 603–614.

[6] H. Jin, J. Lu, Origins of the multifractality in Shanghai Stock Market, Nuovo Cimento B Serie 121 (9) (2006) 987–994.

[7] L. Zunino, B.M. Tabak, A. Figliola, D. Pérez, M. Garavaglia, O. Rosso, A multifractal approach for stock market inefficiency, Physica A 387 (26) (2008) 6558–6566.

[8] X. Zhang, L. Yang, Y. Zhu, Analysis of multifractal characterization of Bitcoin market based on multifractal detrended fluctuation analysis, Physica A 523 (2019) 973–983.

[9] R. Rak, D. Grech, Quantitative approach to multifractality induced by correlations and broad distribution of data, Physica A 508 (2018) 48–66.

[10] A.K. Maity, R. Pratihar, A. Mitra, S. Dey, V. Agrawal, S. Sanyal, A. Banerjee, R. Sengupta, D. Ghosh, Multifractal detrended fluctuation analysis of alpha and theta EEG rhythms with musical stimuli, Chaos Solitons Fractals 81 (2015) 52–67.

[11] P.C. Ivanov, L.A.N. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, Z.R. Struzik, H.E. Stanley, Multifractality in human heartbeat dynamics, Nature 399 (6735) (1999) 461–465.

[12] F. Olivares, G. Funes, D.G. Perez, High frequency multifractality in return intervals from fading induced by turbulence, Fractals 29 (2) (2021) 2150049–2151888.

[13] J.W. Kantelhardt, D. Rybski, S.A. Zschiegner, P. Braun, E. Koscielny-Bunde, V. Livina, S. Havlin, A. Bunde, Multifractality of river runoff and precipitation: Comparison of fluctuation analysis and wavelet methods, Physica A 330 (1–2) (2003) 240–245.

[14] L. Telesca, V. Lapenna, M. Macchiato, Multifractal fluctuations in seismic interspike series, Physica A 354 (2005) 629–640.

[15] L. Telesca, V. Lapenna, Measuring multifractality in seismic sequences, Tectonophysics 423 (1–4) (2006) 115–123.

[16] M. Laib, J. Golay, L. Telesca, M. Kanevski, Multifractal analysis of the time series of daily means of wind speed in complex regions, Chaos Solitons Fractals 109 (2018) 118–127.

[17] P. Oświecimka, J. Kwapień, I. Celińska, S. Drożdż, R. Rak, Computational approach to multifractal music, 2011, arXiv preprint arXiv:1106.2902.

[18] M. Ausloos, Generalized hurst exponent and multifractal function of original and translated texts mapped into frequency and length time series, Phys. Rev. E 86 (3) (2012) 031108.

[19] S. Drożdż, P. Oświecimka, A. Kulig, J. Kwapień, K. Bazarnik, I. Grabska-Gradzińska, J. Rybicki, M. Stanuszek, Quantifying origin and character of long-range correlations in narrative texts, Inform. Sci. 331 (2016) 32–44.

[20] Y. Leung, E. Ge, Z. Yu, Temporal scaling behavior of avian influenza A (H5N1): The multifractal detrended fluctuation analysis, Ann. Assoc. Am. Geogr. 101 (6) (2011) 1221–1240.

[21] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H.E. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, Physica A 316 (1–4) (2002) 87–114.

[22] D. Grech, G. Pamuła, On the multifractal effects generated by monofractal signals, Physica A 392 (23) (2013) 5845–5864.

[23] H. Nakao, Multi-scaling properties of truncated Lévy flights, Phys. Lett. A 266 (4–6) (2000) 282–289.

[24] P. Oświecimka, J. Kwapień, S. Drożdż, Wavelet versus detrended fluctuation analysis of multifractal structures, Phys. Rev. E 74 (1) (2006) 016103.

[25] A.-L. Barabási, P. Szépfalusy, T. Vicsek, Multifractal spectra of multi-affine functions, Physica A 178 (1) (1991) 17–28.

[26] J.-F. Muzy, E. Bacry, A. Arneodo, The multifractal formalism revisited with wavelets, Int. J. Bifurcation Chaos 4 (02) (1994) 245–302.

[27] W.-X. Zhou, et al., Multifractal detrended cross-correlation analysis for two nonstationary signals, Phys. Rev. E 77 (6) (2008) 066211.

[28] E. Serrano, A. Figliola, Wavelet leaders: A new method to estimate the multifractal singularity spectra, Physica A 388 (14) (2009) 2793–2805.

[29] H. Xiong, P. Shang, Weighted multifractal analysis of financial time series, Nonlinear Dynam. 87 (4) (2017) 2251–2266.

[30] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, D. Sornette, Multifractal analysis of financial markets: A review, Rep. Progr. Phys. 82 (12) (2019) 125901.

[31] J.W. Kantelhardt, E. Koscielny-Bunde, H.H. Rego, S. Havlin, A. Bunde, Detecting long-range correlations with detrended fluctuation analysis, Physica A 295 (3–4) (2001) 441–454.

[32] J.R. Thompson, J.R. Wilson, Multifractal detrended fluctuation analysis: Practical applications to financial time series, Math. Comput. Simulation 126 (2016) 63–88.

[33] E.A.F.E. Ihlen, Introduction to multifractal detrended fluctuation analysis in Matlab, Front. Physiol. 3 (2012) 141.

[34] S. Drożdż, J. Kwapień, P. Oświecimka, R. Rak, Quantitative features of multifractal subtleties in time series, Europhys. Lett. 88 (6) (2010) 60003.

[35] A. Turiel, C.J. Pérez-Vicente, J. Grazzini, Numerical methods for the estimation of multifractal singularity spectra on sampled data: A comparative study, J. Comput. Phys. 216 (1) (2006) 362–390.

[36] J. Ludescher, M.I. Bogachev, J.W. Kantelhardt, A.Y. Schumann, A. Bunde, On spurious and corrupted multifractality: The effects of additive noise, short-term memory and periodic trends, Physica A 390 (13) (2011) 2480–2490.

[37] D. Gulich, L. Zunino, The effects of observational correlated noises on multifractal detrended fluctuation analysis, Physica A 391 (16) (2012) 4100–4110.

[38] P. Oświecimka, S. Drożdż, M. Frasca, R. Gebarowski, N. Yoshimura, L. Zunino, L. Minati, Wavelet-based discrimination of isolated singularities masquerading as multifractals in detrended fluctuation analyses, Nonlinear Dynam. 100 (2) (2020) 1689–1704.

[39] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, Fractal measures and their singularities: The characterization of strange sets, Phys. Rev. A 33 (2) (1986) 1141.

[40] M.E. Newman, Power laws, Pareto distributions and Zipf's law, Contemp. Phys. 46 (5) (2005) 323–351.

[41] K. Hu, P.C. Ivanov, Z. Chen, P. Carpena, H.E. Stanley, Effect of trends on detrended fluctuation analysis, Phys. Rev. E 64 (1) (2001) 011114.

[42] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, M. Wilhelm, Bringing up OpenSky: A large-scale ADS-B sensor network for research, in: IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, IEEE, 2014, pp. 83–94.

[43] G. Williams, GPS for the sky: A survey of automatic dependent surveillance-broadcast (ADS-b) and its implementation in the United States, J. Air L. Com. 74 (2009) 473.

[44] R. Salcido, A. Kendall, Y. Zhao, Analysis of automatic dependent surveillance-broadcast data, in: 2017 AAAI Fall Symposium Series, 2017.

[45] Y.S. Virkar, Power-Law Distributions and Binned Empirical data (Ph.D. thesis), University of Colorado at Boulder, 2012.

[46] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data, Physica D 58 (1–4) (1992) 77–94.

[47] T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests, Phys. Rev. Lett. 77 (4) (1996) 635.