Robustness of estimators of long-range dependence and self-similarity under non-Gaussianity

BY CHRISTIAN L. E. FRANZKE1,*, TIMOTHY GRAVES2, NICHOLAS W. WATKINS1, ROBERT B. GRAMACY2,3 AND CECILIA HUGHES1

1 British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK
2 Statistical Laboratory, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK
3 Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637-1610, USA

Long-range dependence (LRD) and non-Gaussianity are ubiquitous in many natural systems such as ecosystems, biological systems and climate. However, it is not always appreciated that the two phenomena may occur together in natural systems and that self-similarity in a system can be a superposition of both phenomena. These features, which are common in complex systems, impact the attribution of trends and the occurrence and clustering of extremes. The risk assessment of systems with these properties will lead to different outcomes (e.g. return periods) than the more common assumption of independence of extremes. Two paradigmatic models are discussed that can simultaneously account for LRD and non-Gaussianity: autoregressive fractional integrated moving average (ARFIMA) and linear fractional stable motion (LFSM). Statistical properties of estimators for LRD and self-similarity are critically assessed. It is found that the most popular estimators can be biased in the presence of important features of many natural systems like trends and multiplicative noise. Also the LRD and non-Gaussianity of two typical natural time series are discussed.

Keywords: memory; paradigmatic models; multiplicative noise

1. Introduction

Meteorology has been both a seedbed and a testbed for many advances in the mathematics of complex systems; the former being seen in its contribution of the Lorenz model to nonlinear science [1], while the latter is exemplified by ensemble forecasting. This parallel evolution of theory and practice has identified important problems, e.g. the need to model the effect of extra, noise-like, degrees of freedom.

*chan1@bas.ac.uk

One contribution of 13 to a Theme Issue ‘Climate predictions: the influence of nonlinearity and randomness’.
on deterministic low-dimensional dynamics. Other well-established paradigms such as Hasselmann’s stochastic model [2] have highlighted the importance of red noise to mathematical climatology.

However, since the pioneering work of the late Benoit Mandelbrot, increasing attention has been paid to two self-similar (or ‘fractal’) aspects of time series: long-range dependence (LRD) in time, and the spatial counterpart of LRD, heavy-tailed probability distributions in amplitude. First identified in hydrology, and since studied in research areas as diverse as biology, telecommunications, social networks, econometrics and the climate system, LRD is characterized by its low-frequency singular behaviour, the so-called $1/f$ power spectrum. When present in a signal, the correlations captured by LRD will both hamper the identification and the attribution of deterministic trends [3], and impede the quantification of their significance. As was the case in the original context of hydrology, the presence of LRD in climate time series has been intensely debated, and it is still sometimes ascribed to transient effects, calibration issues or other forms of non-stationarity [4,5].

LRD is also of practical significance. By making systems subject to longer-lived fluctuations [6], it changes the information available to make predictions about the state of a complex system. It impacts the occurrence and clustering of extremes [7–9], which are important for risk assessments and mitigation strategies, e.g. insurance pricing and flood defence. The traditional assumption is that extremes are independent events. But there is growing evidence for clustering of extreme extra-tropical storms [10] and precipitation events (e.g. 2007 United Kingdom floods). This clustering of extremes will lead to higher return periods of extreme events than the more common assumption of independence. Recently, detrended fluctuation analysis (DFA), a tool originally developed to detect LRD, has found application in the prediction of dangerous bifurcations in dynamical systems, such as climate ‘tipping points’ [11–13]. Our contribution in this paper is focused on both (i) the problem of accurately estimating LRD in the presence of other signal elements and (ii) the complicating effects of heavy-tailed amplitude distributions, when they are present.

A process is long-range-dependent when the prediction of its next state depends on the whole of its past. An imprint of this dependence structure is that the covariance $r(k) = \text{Cov}(X(k), X(0))$ decays slowly, as $k \to \infty$, so that

$$
\sum_{k=0}^{\infty} |r(k)| = \infty.
$$

This slow decay of the covariances means that the values of the process $X$ are strongly dependent over long periods of time. This contrasts with the more familiar short-range-dependent system, where $\sum_{k=0}^{\infty} |r(k)| = C < \infty$. In a short-range-dependent process, the next state only depends on the current state and the recent past. The archetype of a short-range-dependent process is a first-order Markov process, where the next state depends only on the present state and is conditionally independent of past states. See Beran [6] for more details.

LRD of a system is characterized by the parameter $d$ in the statistics literature and sometimes by Mandelbrot’s Joseph parameter $J$, for example, in the physics literature. Temporal LRD has been detected in the water level of the river Nile [14,15]. It was observed that the range of values grows as $\tau^J$, where $J$ refers to
the Joseph exponent [16, p. 157] and \( \tau \) to the time period under consideration. The growth of range was anomalously large compared with that in the familiar paradigms of random walks and Brownian motion, which have both played a central role in understanding diffusion. In the random walk, the variance grows as the square root of time \( t^{1/2} \) [17]. The subsequent theoretical explanation of the difference between the observed anomalous diffusion and the Brownian motion paradigm came first through the use of a self-similar process, in particular via the study of fractional Brownian motion (fBm) [18]. Here, self-similarity means that the statistical properties on all scales are similar. This property is controlled by the self-similarity parameter \( H \). A stochastic process \( S \) is self-similar when a rescaling of time by a factor \( \lambda \) leads to a rescaling of the amplitude of the process \( S \) by \( \lambda^H \). Thus, a process is said to be \( H \)-self-similar when the following relation holds [19]:

\[
S(\lambda t) \sim \lambda^H S(t),
\]

(1.2)

where \( \sim \) denotes equality in distribution. It is not always appreciated, and is sometimes confused in the literature, that two different aspects of a time series can contribute to its self-similarity \( H \). The first is LRD, which is synonymous with persistence, referred to by Mandelbrot & Wallis [18] as the Joseph effect. Persistent systems exhibit longevity by having a tendency to maintain the way they have been recently. Examples are heatwaves and drought conditions. The second source of self-similarity identified by Mandelbrot, the so-called Noah effect, is the property that change in a system can be rather large and can occur very abruptly, i.e. time series drawn from systems can exhibit sharp discontinuities, e.g. earthquakes.

It may at first seem odd that the two phenomena occur simultaneously in natural systems, because they pull in opposite directions. On reflection, taken together, however, we see that the two effects capture the facts that coherent structures in nature are real and that they can emerge, change or even vanish very quickly. The archetypal model of the Noah effect is non-Gaussian jumps whose complementary cumulative distribution function decays with a power law in size \( s \), \( p(s) \sim s^{-\alpha} \). Here, \( \alpha \) denotes the stability exponent (referred to as tail exponent in the statistics literature). Thus, the self-similarity parameter is

\[
H = J - \frac{1}{2} + \frac{1}{\alpha} = d + \frac{1}{\alpha}, \quad H \in (0, 1).
\]

(1.3)

The distinction between \( H \) and \( J = d + \frac{1}{2} \) is important because not all observed time series have Gaussian fluctuations. As such, one may find that popular diagnostics, as shown below, may be insensitive to heavy tails, measuring \( J \) and not \( H \). However, in many situations, both heavy tails and the clustering of extremes caused by LRD are very important and both contribute to the value of \( H \). This makes it necessary to be able to measure both \( H \) and \( J \) independently.

Many estimators for the LRD exponent \( d \) have thus been developed, and are widely used in the respective literature. Much of what has been rigorously established about the estimators, however, is for a particular LRD Gaussian model, fBm, which is also \( H \) self-similar and has a particularly simple relation between \( H \) and \( J \), i.e. \( H = J + \frac{1}{2} - \frac{1}{2} = J \) (\( \alpha = 2 \) for Gaussian increments). Most observed time series depart from the assumptions of fBm in some way or another,
so it is important to critically evaluate how sensitive the estimators are to deviations from fractional Gaussian noise (fGn). fBm is a random walk with long-range-correlated increments and the value of $J$ in such $H$ self-similar walks is directly connected to the presence of LRD in their increments.

In this study, we critically re-evaluate four popular methods for measuring $d$ and two for measuring $H$: variable bandwidth (VB) estimator [20], wavelets (WLs) [21], DFA [22] and the rescaled range analysis ($R/S$) [14,16]; exact Whittle estimator [23,24] and semi-parametric estimators (in this study we will focus on the Geweke–Porter-Hudak (GPH) estimator [25]; semi-parametric estimators are also described in Bardet et al. [26], Robinson [27,28] and Hurvich et al. [29]). This is a not a complete set of estimators for the self-similarity or LRD parameter (other estimators that are well established in the statistics community include the fractional exponential (FEXP) model [30]). Here, we want to focus on some of the most popular estimators in the physics and geosciences communities ($R/S$ [31–35], DFA [7,35–37], GPH [3,38,39], VB [40] and WL [21,41]).

We show scatter plots that compare the imposed parameters for a number of trials with the values detected by the above methods. These have the advantage of visually indicating which methods measure the complete self-similarity exponent $H$, and which, instead, measure LRD via $d$. For evaluating the performance of these estimators, we employ two well-established paradigmatic time-series models: a self-similar process (linear fractional stable motion (LFSM) [42]) and an asymptotically self-similar process (autoregressive fractional integrated moving average (ARFIMA) [6]).

In most natural systems, trends are common properties, so we systematically examine the performance of the various estimators when a linear trend is superposed on each time series. This is an important issue because in practice it is not always easy to remove trends. That they typically manifest themselves in higher moments presents further challenges still. Trends are the hardest to deal with because long-range-dependent processes can produce apparent trends over rather long periods of time [3]. Therefore, it can be difficult to decide if a trend is due to external forcing or due to finite time-series length.

In §2, we present the two paradigmatic time-series models in more detail and discuss various special cases. In §3, we introduce the estimators of LRD and self-similarity. We test their utility in §4. In §5, we discuss the LRD and self-similarity properties of two exemplary time series from nature, and in §6, we provide conclusions.

## 2. Paradigmatic models of natural time series

### (a) Natural time-series examples

The most prominent and widely used paradigmatic LRD time-series model, especially in the physics literature, is fGn. However, most natural time series are not strictly Gaussian and many are actually highly non-Gaussian. To illustrate this point, we discuss two time series from nature. In figure 1a, we display monthly mean temperature from the Antarctic station Faraday–Vernadsky [3,43]. It is easy to see that the raw time series is highly non-Gaussian. This is confirmed by the probability distribution functions (PDFs) plotted in figure 1b. There is also a striking seasonal signal visible in the skewness of the PDFs. However, it must be
noted that polar temperature time series are much more non-Gaussian than mid-latitude ones, which are typically nearly Gaussian. Our second exemplary time series is the auroral electrojet (AE) index (figure 1c [44,45]). The AE index is derived from 1 min resolution time series from 12 high-latitude magnetometers. Reflecting the intermittent nature of the ionospheric and solar wind processes that influence it, the AE index is seen to be spiky and strongly non-Gaussian (figure 1d). Clearly, many natural time series are non-Gaussian.

(b) Paradigmatic models

While fGn and fBm are Gaussian, they are still useful idealized paradigmatic models for understanding many observed phenomena. A paradigmatic model is an idealized framework that captures some properties of observed time series, though not all. Most paradigmatic models allow analytical work, although at the expense of an over-idealization of the physical or biological phenomena. There is a need for better paradigmatic models of observed phenomena that allow simultaneously non-Gaussian statistics and LRD. Such models are needed, for example, for hypothesis testing and time-series simulation. Thus, we briefly discuss two paradigmatic time-series models that exhibit LRD and non-Gaussianity. The widely used fGn and fBm are special cases in the Gaussian limit of these more general models, thereby retaining some of the analytical tractability of these models.

In this study, we use two classes of processes with symmetric $\alpha$-stable (SaS) distributions. They are the LFSM and the ARFIMA model with SaS innovations. Stoev & Taqqu [46] present efficient methods for the simulation of these processes using the fast Fourier transform.

Phil. Trans. R. Soc. A (2012)
(i) **Linear fractional stable motion**

LFSM is a model that exhibits LRD and heavy tails at the same time. It is an extension to the simpler Brownian random walk model. It links individual heavy-tailed jumps by means of a retarded memory kernel. It can be represented by a stochastic process\(^1\) \(X_{H,\alpha} = \{X_{H,\alpha}(t), t \in \mathbb{R}\}\), which is defined by the stochastic integral

\[
X_{H,\alpha}(t) = C_{H,\alpha}^{-1} \int_{\mathbb{R}} \left( (t - s)^{H - (1/\alpha)} - (-s)^{H - (1/\alpha)} \right) dL_{\alpha}(s), \quad (2.1)
\]

where \(0 < H < 1\), \(\alpha \in (0,2)\) and where \(L_{\alpha} = \{L_{\alpha}(s), s \in \mathbb{R}\}\) is a standard symmetric \(\alpha\)-stable (SaS) Lévy process. The process \(X_{H,\alpha}\) is called LFSM and is self-similar with self-similarity parameter \(H\). Thus, it satisfies relation (1.2) and has stationary increments. The parameter \(\alpha\) controls the tails of the distribution of an SaS random variable \(\xi\), that is,

\[
\mathbb{P}(|\xi| \geq x) \sim x^{-\alpha}, \quad \text{as } x \to \infty. \quad (2.2)
\]

The greater the value of \(\alpha\), the lower the probability of extreme fluctuations of the SaS process \(X_{H,\alpha}\). We recommend the introductions by Taqqu [47] and Mercik et al. [48] for detailed expositions.

(ii) **Fractional Brownian motion**

In the Gaussian case, we have that \(\alpha = 2\) and the LFSM process (2.1) reduces to fBm. In this case, the self-similarity parameter \(H\) will always equal the Joseph exponent \(J\), thus \(H = J = d + \frac{1}{2}\).

(iii) **Ordinary Lévy motion**

In the memory-less case, we have \(d = 0\) and \(0 < \alpha < 2\). In this case, equation (2.2) holds and the tails of the distribution decay according to a power law and consequently \(L_{\alpha}\) has infinite variance and is called ordinary Lévy motion. This encapsulates the Noah effect. In this case, the self-similarity exponent is \(H = 1/\alpha\).

(iv) **Autoregressive fractional integrated moving average**

Another paradigmatic model exhibiting LRD with heavy-tailed fluctuations is ARFIMA, which has the added ability of exhibiting short-range-dependent behaviour. This model, denoted ARFIMA\((p, d, q)\), \(p, q \in \mathbb{N}\), extends the usual ARIMA\((p, d, q)\) models, in which \(d\) takes integer values [49]. An ARIMA model is written as

\[
\Phi(B)(1 - B)^d X_t = \Psi(B) \epsilon_t, \quad (2.3)
\]

where \(B\) denotes the back shift operator defined by \(BX_t = X_{t-1}\), \(B^2X_t = X_{t-2}\), \ldots. The polynomials \(\Phi\) and \(\Psi\) are defined as \(\Phi(x) = 1 - \sum_{j=1}^{p} a_j x^j\) and \(\Psi(x) = 1 + \sum_{j=1}^{q} b_j x^j\), where \(p\) and \(q\) are integers. The innovations \(\epsilon_t, (t = 1, 2, \ldots)\) are usually independent and identically distributed (iid) normal variables with zero

\(^1\)This is the traditional way of defining LFSM, though \(H\) has been defined in (1.3) as the dependent and \(d\) and \(\alpha\) as the independent variables.
expectation and variance $\sigma^2$, but can also be $\alpha$-stable distributed. The widely used autoregressive process of first-order (AR(1)) is a special case with $p = 1$, $d = 0$ and $q = 0$. Typically $d$ is an integer. In the case that $d$ is a fractional real number, $X(t)$ is an ARFIMA($p$, $d$, $q$) process with $-\frac{1}{2} < d < \frac{1}{2}$ and exhibits LRD. In the Gaussian case, ARFIMA is stationary. An ARFIMA($0$, $d$, $0$) is asymptotically equivalent to fBm [47]. An ARFIMA($p$, $d$, $q$) with $p > 0$ and $q > 0$ is long-range-dependent but not self-similar. However, it is asymptotically self-similar for long time scales.

3. Estimators of long-range dependence and self-similarity

Here, we briefly describe the four estimators of the self-similarity (SS) and LRD parameters used in this study.

(a) Variable bandwidth

The VB method [20] is a technique for estimating the self-similarity exponent, $H$, from a time series $x(t)$. The time series of length $T$ is divided into windows or ‘bands’ of width $r$. The VB method can deploy two different algorithms to estimate $H$: (i) the standard deviation of the time series, $\sigma(r)$, is computed in each band; or (ii) the difference between the maximum and minimum value in each band, $\epsilon(r)$, is computed. Then $\sigma(r)$ and $\epsilon(r)$ are averaged over all the possible bands by varying the origin at fixed $r$,

$$
VB_w(r) = \frac{1}{L_r} \sum_{i=1}^{L_r} \sigma_i(r) \quad \text{and} \quad VB_\delta(r) = \frac{1}{L_r} \sum_{i=1}^{L_r} \epsilon_i(r),
$$

(3.1)

where $L_r$ is the number of windows of length $r$. This is repeated over a range of window sizes. Both quantities follow a power-law behaviour for self-similar time series [20] such that $VB_w(r) = r^H$ and $VB_\delta(r) = r^H$. Thus, the self-similarity exponent, $H$, is obtained from the slope of the corresponding log–log plot.

(b) Wavelets

A WL $\psi$ is a function with zero average and is normalized to one. A family of WLs is generated by scaling $\psi$ by a factor $s$ and translating it by $u$ ($\psi_{u,s}(t) = (1/\sqrt{s})\psi((t - u)/s)$). The WL transform allows one to construct a time–frequency representation of a signal, the WL spectrum. One can then infer the self-similarity parameter from the WL spectrum via ordinary least squares at large WL scales [21,50].

(c) Rescaled range

The rescaled range $R/S$ [15] is a technique for estimating the LRD parameter $d$ from a time series. The $R/S$ estimator is given by

$$
R/S(\tau) = \frac{\max x(t, \tau) - \min x(t, \tau)}{\sqrt{(1/\tau) \sum_{t=1}^{\tau} (\xi(t) - \langle \xi \rangle)^2}},
$$

(3.2)
where \( 1 \leq t \leq \tau \), \( x(t, \tau) = \sum_{u=1}^{t} x(u) - \langle x \rangle_{\tau} \), and \( \langle x \rangle_{\tau} = (1/\tau) \sum_{t=1}^{\tau} x(t) \). It scales like \( \tau^{J} \), where the value of \( J \) can be estimated from the slope of \( R/S \) from a log–log plot.

\( (d) \) Detrended fluctuation analysis

DFA [22] also estimates the LRD parameter \( d \) from a time series. In DFA, a profile is first computed by \( Y(i) = \sum_{t=1}^{N} x(t) \). The profile is cut into \( N_s \) non-overlapping segments of equal length \( s \) and then the local trend is subtracted for each segment \( v \) by a polynomial least-squares fit of the data. Linear (DFA1), quadratic (DFA2), cubic (DFA3) or higher-order polynomials can be used for detrending. In the \( n \)th-order DFA, trends of order \( n \) in the profile, and of order \( n - 1 \) in the original record, are eliminated. Next, the variance for each of the \( N_s \) segments is calculated by averaging over all data points \( i \) in the \( v \)th segment:

\[
F^2_s(v) = \langle Y^2_s(i) \rangle = \frac{1}{s} \sum_{i=1}^{s} Y^2_s[(v - 1)s + i].
\] (3.3)

Finally, the average over all segments is computed and the square root is applied to obtain the following fluctuation function:

\[
F(s) = \sqrt{\frac{1}{N_s} \sum_{v=1}^{N_s} F^2_s(v)}. \tag{3.4}
\]

For different detrending orders, \( n \), we obtain different fluctuation functions \( F(s) \), which are denoted by \( F^{(n)}(s) \). The fluctuation function scales according to \( F^{(n)}(s) \sim s^{\zeta} \), with \( d = \frac{1}{2} - \frac{1}{2} \). There are many variants of DFA, but use of standard DFA is recommended by Bashan et al. [37] if the functional form of a trend is not \textit{a priori} known.

\( (e) \) Exact Whittle estimator

The exact Whittle estimator is a semi-parametric estimator [23,24]. This method assumes that the underlying model of LRD can be represented by \( (1 - B)^d X_t = \epsilon_t \), where \( \epsilon \) is iid noise. See Shimotsu & Phillips [23,24] for more details.

\( (f) \) Power spectral method

As a semi-parametric estimator, we use the power spectral method of Geweke & Porter-Hudak [25] and Hurvich et al. [51]. Spectral methods find \( d \) by estimating the spectral slope. The periodogram is used, which is an estimate of the spectral density of a finite-length time series and is given by

\[
\hat{S}(\lambda_j) = \frac{1}{N} \left| \sum_{t=1}^{N} X(t) e^{-i2\pi t\lambda_j} \right|^2, \quad j = 1, \ldots, [N/2], \tag{3.5}
\]
where \( \lambda_j = j/N \) is the frequency and the square brackets denote rounding towards zero. A series with LRD has a spectral density proportional to \(|\lambda|^{-2d}\) close to the origin. Since \( \hat{S}(\lambda) \) is an estimator of the spectral density, \( d \) is estimated by a regression of the logarithm of the periodogram versus the logarithm of the frequency \( \lambda \). Thus having calculated the spectral density estimate \( \hat{S}(\lambda) \), semi-parametric estimators fit a power law of the form \( f(\lambda, b, d) = b|\lambda|^d \), where \( b \) is the scaling factor.

4. Empirical tests

In order to examine the statistical properties of the LRD estimators, like bias, spread and outliers, we generate various test time series from the above paradigmatic models. Specifically, we generate ensembles of 1000 members with randomly selected parameters from uniform distributions: fGn with \( d \in (-0.5, 0.5) \), ordinary Lévy noise (oLn) with \( \alpha \in (1, 2) \), ARFIMA with Gaussian increments ARFIMA-G(1, \( d \), 1) and with ordinary Lévy increments ARFIMA-L(1, \( d \), 1) with the autoregressive coefficient \( a_1 \in \left(-\frac{1}{2}, \frac{1}{2}\right) \), and the moving average coefficient \( b_1 \in (0, 1), \ d \in \left(-\frac{1}{2}, \frac{1}{2}\right) \) and \( \alpha \in (1, 2) \). We also use various first-order autoregressive processes with parameters randomly chosen in \((-1, 1)\). The time-series length is always \( 2^{15} - M \) with \( M = 6000 \) (see Stoev & Taqqu [46] for an explanation of \( M \)), which is comparable to the length of most observed climatic and other natural time series, thus long enough for our purposes.

(a) Paradigmatic time series

First we examine how well the SS and LRD estimators work for the models of self-similarity. Most of the methods (e.g. DFA and \( R/S \)) require a regression fit in order to estimate the SS or LRD parameters. As shown by Chen et al. [36], non-stationarities and short-range dependences can cause crossovers in the fluctuation curves. Because of this we only regress on the long-range part of the fluctuation curves. Our results are robust to the particular choice of the cut-off.

As can be seen in figure 2, the WL and the VB methods are the only methods that are able to infer the self-similarity of oLn. While the WL method has no large estimation spread, the VB method exhibits large errors. All other estimators estimate \( d \) rather than \( H \), with \( R/S \) producing the largest estimation variance, while Whittle, DFA1 and GPH have considerably smaller estimation variances, but with the odd outlier (figure 2). All four estimators do a reasonably good job of inferring \( H \) from fGn, with VB\(_w\)(\( r \)) having the largest bias for \( H \) close to 0 and close to 1, and \( R/S \) also having some bias at small \( H \) with a relatively large estimation variance. WL, DFA1 and GPH produce very tight estimates, with WL and DFA1 only biased for small values, and Whittle and GPH working well over the entire range. For these paradigmatic time series, the two semi-parametric estimators, Whittle and GPH, give the best estimates. This picture changes once we allow for short-term dependence structures to contaminate the pure self-similar character of fGn and oLn. As figure 3 shows, all estimators work considerably worse for ARFIMA surrogate data, with large estimation spreads and many outliers. Again the Whittle estimator and GPH perform better than the other estimators.
Figure 2. Estimates of $H$ for walks derived from fractional Gaussian noise (fGn) and ordinary Lévy noise (oLn) for wavelet (WL) method and the variable bandwidth (VB) estimator (crosses denote VB$_d$ and circles VB$_w$). Estimates of $d$ for fGn and oLn for Whittle estimator, R/S, DFA1 and GPH. The solid line indicates the true self-similarity parameter $H$ ($d$ in the case of fGn or $1/\alpha$ for oLn).
Figure 3. Estimates of $H$ for walks derived from ARFIMA-G(1, $d$, 1) and ARFIMA-L(1, $d$, 1) with the WL method and VB estimator. Estimates of $d$ for ARFIMA-G(1, $d$, 1) and ARFIMA-L(1, $d$, 1) for the Whittle estimator, $R/S$, DFA1 and GPH. The solid line indicates the true self-similarity parameter $H$. 
Before we go on to examine how well the estimators work with superimposed trends, we test their accuracy on data generated from three basic short-range dependence models: independent white noise, AR(1) and ARFIMA(1,0,1). As figure 4a–c show, the Whittle estimator, GPH and DFA have the least bias. But it also has to be noted that all estimators have many outliers, suggesting that, given an individual time series, the estimates of \( d \) or \( H \) can be severely biased. The performance of higher-order DFA is very similar to DFA1 in the above test cases (not shown). These results are consistent with previous empirical studies [52].
(b) Trends

Another important test is to see how the presence of trends affects the various estimators. Here, we consider three cases: (1) a linear trend superimposed by realizations of fGn, oLn, ARFIMA-G and ARFIMA-L; (2) a linear trend only in the second half of the time series superimposed by realizations of fGn, oLn, ARFIMA-G and ARFIMA-L; and (3) a linear trend in the variance. Cases 2 and 3 are motivated by climate change where the time series may (case 2) or may not (case 1) include the pre-industrial era, and where climate change also influences the frequency and the strength of fluctuations (e.g. storms; case 3). Based on the evidence gathered so far, it is reasonable to expect that this can lead to bias any LRD estimate.

For cases 1 and 2, we assume the magnitude of the linear trend to be 1. The empirical tests reveal that the GPH estimator is slightly biased for fGn, ARFIMA-G and ARFIMA-L, and has a relatively large negative bias for oLn (figure 5). DFA is least biased for fGn data, but has considerable bias for oLn, ARFIMA-G and ARFIMA-L (figure 5) generated data. Both ARFIMA-G and ARFIMA-L estimators have a large number of outliers. Both $R/S$ and VB show qualitatively similar behaviour compared with GPH (not shown). Our results are qualitatively consistent with the study by Hu et al. [53].

Now, we examine how the estimators handle a trend in the variance (case 3). For this purpose, we use an AR(1) model with increasing variance:

$$x_{t+1} = F + \alpha x_t + \sigma \left(1 + \frac{t}{10000}\right) \zeta_t. \quad (4.1)$$

Here, again, we generate 1000 realizations by sampling values for $F$ from a uniform distribution $U(0, 2)$, using $\alpha \in U(-0.5, 0.5)$ and $\sigma \in U(0, 1)$. For this case, DFA is the least biased, and GPH and $R/S$ show considerable bias, while the VB estimators have a huge bias, although their estimates are very narrow (figure 4e).

(c) Correlated additive and multiplicative noise

Both LFSM and ARFIMA are additive noise models. In previous studies, it has been shown that multiplicative noise is important in natural systems [54,55]. This raises the question of what the effect of multiplicative noise would be on LRD estimators. To investigate, we use data generated from a process that has a so-called correlated additive and multiplicative (CAM) noise term. An example is the normal form for reduced climate models [54], which is given by the following stochastic differential equation (SDE):

$$dx = (F + ax + bx^2 - cx^3) \, dt + \sigma (L + Ix) \, dW. \quad (4.2)$$

As shown in Majda et al. [54], the PDF of equation (4.2) exhibits a power-law decay over a particular range, although its tail ultimately decays exponentially. We generate 1000 realizations by sampling random values for the SDE parameters from a uniform distribution while complying with the parameter relations as described in Majda et al. [54].
The GPH estimator is slightly biased towards positive values, while DFA and $R/S$ have larger biases. Observe that the WL estimator (as well as both VB estimators (not shown)) estimates self-similar behaviour (figure 4f). While the PDF of equation (4.2) decays in a power-law-like way over a given range, it is not self-similar because its ultimate decay is exponential [54]. This makes the estimates obtained by the WL and VB estimators questionable. Furthermore, $R/S$, DFA, GPH and the Whittle estimator again show an uncomfortably large number of outliers, again suggesting that the estimators may not be very reliable for this case, or that the signal is not characterized simply by $H$ and $d$. We note that Kantelhardt et al. [56] have studied the performance of more general multi-fractal DFA methods, which were found to extract the full range of scaling exponents in a particular multi-fractal test case.

5. Natural time-series examples

Now we return to the two natural time series from §2, and analyse their self-similarity and LRD characteristics. We have shown above that the two time series are non-Gaussian; are they also long-range-dependent or self-similar?

Phil. Trans. R. Soc. A (2012)
Applying the various estimation methods to the Faraday–Vernadsky temperature time series gives evidence for LRD but with a wide variety of values: Whittle $d = 0.24$, GPH $d = 0.28$, DFA2 $d = 0.43$, $R/S$ $d = 0.33$, WL $H = 0.53$, $VB_w H = 0.92$ and $VB_d H = 0.96$. While the three LRD estimators all provide evidence for LRD, they provide a rather large range of values of the LRD parameter, $d$. However, all of the estimates are between 0 and 0.5, suggesting a long-range-dependent but stationary process. The $H$ estimates are larger than the $d$ estimates; this might suggest self-similarity with $\alpha \approx 1.5$.

There is also evidence for LRD in the AE index, again with a wide variety of estimates from the various estimators: Whittle $d = 0.28$, GPH $d = 0.72$, DFA2 $d = 0.66$, $R/S$ $d = 0.4$, WL $H = 0.94$, $VB_w H = 0.97$ and $VB_d H = 0.99$. The estimates from GPH and DFA2 are larger than 0.5, suggesting a non-stationary time series, while the Whittle and $R/S$ estimates suggest a stationary long-range-dependent time series. Again, the $H$ estimates are larger than the $d$ estimates, providing some evidence for self-similarity. Other evidence has been presented that AE may not be a simple fractal [57]; recently work has focused on high-frequency non-stationary and lower-frequency $1/f$ properties [58].

While all LRD estimators agree that there is evidence for LRD in these two time series, they provide a rather large range of possible values, illustrating the problem of statistically robust LRD estimation in practice and the need for further investigation of the performance of statistical indicators in the presence of departures from fractality, including possible multi-fractality.

6. Conclusions

There are two contributions to self-similarity: (i) LRD and (ii) non-Gaussian jumps. This is not always appreciated in the various communities with interests in detecting self-similarity and LRD. We have shown that empirical estimators of LRD are at best biased for fractional Gaussian noise, but at worst not robust for processes that deviate from this idealized model. We have also shown that the empirical estimators are not very robust in the presence of trends and multiplicative noise.

Our results have several important implications for the modelling of natural time series. In our view, the ARFIMA model is a much better paradigmatic model of natural time series than fGn since it allows one to explicitly model short-range and long-range behaviour while also allowing for non-Gaussian increments.

Finally, there is a need for estimation procedures that can deal with multiplicative noise and trends in the variance. Such effects introduce sizeable biases and estimation uncertainty. As such, all estimations of LRD have to be taken with precaution. While it is true that all of the estimators we tested perform reasonably well for fractional Gaussian noise, once a time series is non-Gaussian or is non-stationary (in trend or volatility) the estimators can be problematic.

We thank M. Freeman and two anonymous reviewers for their comments on an earlier version of this manuscript. In particular, we are grateful to one of the reviewers for their emphasis on the Whittle and WL estimators. This study is part of the British Antarctic Survey Polar Science for Planet Earth Programme. It was funded by the Natural Environment Research Council. T.G. acknowledges generous support through an EPSRC studentship. R.B.G. was funded under EPSRC research grant EP/D065704/1.
References


Long-range dependence, non-Gaussianity
