



Are Hurst exponents estimated from short or irregular time series meaningful?

Sergei Katsev, Ivan L'Heureux*

Ottawa-Carleton Institute for Physics, University of Ottawa, 150 Louis Pasteur, Ottawa, Ont., Canada K1N 6N5

Received 19 June 2002; received in revised form 4 March 2003; accepted 10 March 2003

Abstract

We show that several time series analysis methods that are often used for detecting self-affine fractal scaling and determining Hurst exponents in data sets may lead to spurious results when applied to short discretized data series. We show that irregularities in the series, such as jumps or spikes (as are often found in geophysical data) may lead to spurious scaling and consequently to an incorrect determination of the Hurst exponent. We also illustrate the statistical error in measuring Hurst exponent in series where self-affine fractal scaling does exist. Users should be aware of these caveats when interpreting the results of short time series analysis.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Self-affine fractal scaling; Power spectrum; Jumps; Spikes; Discontinuities

1. Introduction

Time and space series analysis methods have become widespread and valuable tools in studying geological data. However, many of the common methods, while accurate in the theoretical limit of continuous long series, may lead to significant errors when applied to short or irregular discretized sequences of real experimental data. In this paper, we outline some caveats that should be avoided in the analysis of such series. We demonstrate that the uncertainty in the Hurst exponent values measured from short data sets (less than 500 points) is usually too large for most practical purposes. We also show that the presence of large inhomogeneities in the series, such as large-magnitude abrupt changes in the series variable (jumps or spikes), may lead to spurious results in detecting fractal scaling and calculating Hurst exponents. To our knowledge, these caveats have not been reported previously. Our results suggest

that many of the published results of such series analysis, in order to be meaningful, need to be revised and substantiated by additional evidences.

2. Measurement error on the Hurst exponent

The Hurst exponent is commonly used as a measure of the geometric (fractal) scaling in the data series (Turcotte, 1997). If the series $x(t)$ is a self-affine fractal then $x(bt)$ is statistically equivalent to $b^H x(t)$, where H is the Hurst exponent. The Hurst exponent is frequently calculated for the experimentally obtained data sets to characterize noisy data series. It is also used in characterizing stochastic processes. For example, Brownian noise is a self-affine fractal with a Hurst exponent $H = 0.5$; white noise may be considered as having $H = 0$, and an exactly self-similar process would be characterized by $H = 1$.

In many cases, when dealing with geological and geophysical data, the Hurst exponent is calculated from a series that consists of a short discrete set of values. One of the commonly used methods for calculating the Hurst exponent is described below. Here, we specify the

*Corresponding author. Tel.: +1-613-562-5800; fax: +1-613-562-5190.

E-mail addresses: skatsev@physics.uottawa.ca (S. Katsev), ilheureu@physics.uottawa.ca (I. L'Heureux).

measurement error for the H value, i.e. we determine the range within which a Hurst exponent value calculated from such short series may deviate from its actual value. For relatively long series (more than 1000 points), most of the methods cannot reliably generate the Hurst exponent values with uncertainties smaller than 0.1, over the full range of their values (e.g. Meakin, 1998 and references therein). Slightly better results (with an uncertainty of about 0.05) can be achieved for H close to 0.5 for very large, high-quality data sets.

Here, we illustrate that the uncertainty in the Hurst exponent values measured from short series, such as many of the available geophysical data sets, is typically much larger. We use a Brownian noise as an exemplary process with a known Hurst exponent of $H = 0.5$. Theoretically, this value may be obtained in the limit of an infinitely long series or when a statistical average is taken over a very large number of noise realizations.

To measure the Hurst exponent, we can use a simple method that is frequently employed for that purpose in geosciences (Meakin, 1998; Holten et al., 1997; Hoskin, 2000). First, the so-called series width $w(l)$ is calculated as a function of the time scale l :

$$w(l) = \langle \langle (x - \langle x \rangle_l)^2 \rangle_l \rangle^{1/2}, \quad (1)$$

where $\langle x \rangle_l$ denotes the average over the part of the series of the length l . The outer brackets indicate averaging over all possible averaging windows of width l along the series. If the series exhibits fractal scaling over some range, a straight line is obtained in that range when the function $w(l)$ is plotted on a log–log plot. The slope of this line is typically used as a measure of the Hurst exponent.

In all our simulations described below (except for a few explicitly mentioned cases), the series $x(t)$ contained 256 data points.

We calculated the Hurst exponent values for 200 Brownian noise realizations. When the obtained H values are binned in a histogram, a rather symmetric bell-shaped distribution is obtained (inset in Fig. 1). The distribution of H values exhibits a maximum around $H = 0.5$ and occupies the range from 0.2 to 0.8. For longer series, the width of the distribution decreases thus improving the accuracy of measuring the Hurst exponent value (Fig. 1).

When $w(l)$ is calculated for individual realizations, a fractal scaling typically exists in a limited region of intermediate scales. For small scales, the $w(l)$ graph typically curves down as a result of averaging over a large number of small-sized data sets while for large l the number of data points is insufficient to obtain an accurate result, although methods exist to correct for this effect (Meakin, 1998). For short series, the region of the intermediate scales where fractal scaling exists may not be properly defined for all realizations and the exact

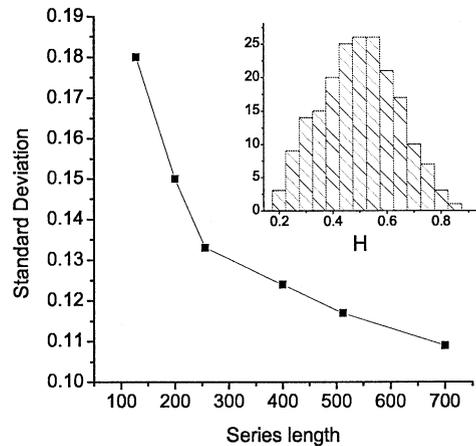


Fig. 1. Distribution of Hurst exponent values. Inset shows typical distribution of Hurst exponent values measured from different realizations of Brownian noise process. Graph shows standard deviation of such distributions as a function of series length.

shape of the distribution depends on the particular choice of the scaling region that is used to calculate H .

A similar bell-shaped type of distribution for the measured Hurst exponent values is obtained when a so-called bounded Brownian noise is used, whereby a Brownian noise process is confined to a specific range of values. When the process variable reaches the boundary of this range, it is reflected off this boundary. This noise type is frequently used to simulate fluctuations in the values of physical parameters that are naturally limited to a certain range, e.g. a percentage of an impurity in a crystal (Holten et al., 1997). Such restriction of the noise process introduces anti-persistence to the series. In agreement with the fact that anti-persistent behavior is characterized by $H < 0.5$, our tests indicate that the restriction of the process variable range decreases the distribution's mean value but does not significantly affect its variance.

Thus, in the case of Brownian noise, the values of H calculated according to Eq. (1) from a single short data set may deviate from its theoretical value by as much as ≈ 0.3 (or 60%), as measured by the range of the H values in the distribution in Fig. 1. Because of such a wide range, special care should be taken when interpreting the scaling parameters obtained from experimental data sets.

3. Impact of 'jumps' and 'spikes' on measured fractal scaling

Many experimentally obtained series are substantially inhomogeneous. Quite often, the regions where the fluctuations in the series variable are relatively small are

separated by transition regions where the value of the variable changes substantially over a very short interval (e.g. Pollok et al., 2001). Such transitions between the different ranges of the series variable will be referred to here as ‘jumps’. Another irregularity, which is commonly found in data sets, are ‘spikes’, i.e. unusually large or small single values of the series variable. In spite of these irregularities, the series are frequently hypothesized to be generated by a single stochastic or chaotic process and fractal scaling is often reported in such series (e.g. Cortini and Anastasio, 2001; Pollok et al., 2001). In many cases, when fractal scaling is observed, it is taken as a characteristic of the underlying pattern-forming process and corresponding conclusions are drawn about the pattern formation mechanism. Here, we show that, in the case of short series, jumps and spikes in the series may lead to the spurious detection of fractal scaling.

Suppose the series is a step function (inset in Fig. 2) and the location of the step is chosen randomly. Although such series clearly does not possess any fractal scaling, the function width $w(l)$ calculated according to Eq. (1) exhibits a power law-type dependence on the scale l typical of fractal scaling (Fig. 2). Spurious scaling arises because, in the process of calculating $w(l)$, the step is ‘swept’ more times by wider averaging windows than by the narrower ones so the calculated series width increases with the scale.

Analytically, this effect may be illustrated as follows. Suppose the series $x(t)$ represented by a step-function starts at $t = 0$ and has the total length T . Suppose also that the location of the jump in the series is chosen such that $x = 0$ for $t < a$ and $x = 1$ for $t > a$, where a is some

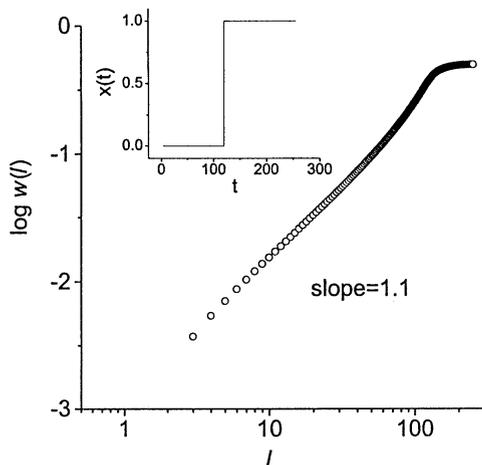


Fig. 2. Spurious scaling due to jump in series. Series width $w(l)$ was calculated according to Eq. (1) from series constructed by step function (shown in inset). Slope is measured by fitting a straight line in the region of l values from 10 to 60.

value between 0 and T . Let us calculate the series width $w(l)$ according to Eq. (1).

If s denotes the starting point of an averaging window of length l then the variance $\langle x^2 \rangle_l - \langle x \rangle_l^2$ and the mean $\langle x \rangle_l$ that are taken inside this window are calculated from

$$\langle x \rangle_l = \frac{1}{l} \int_s^{s+l} x dt, \quad \langle x^2 \rangle_l = \frac{1}{l} \int_s^{s+l} x^2 dt. \quad (2)$$

Let us assume, for simplicity, that the jump is located far enough from either end of the series and let us consider the spatial scales l such that $l < a$ and $l < T - a$. If the jump in the step function is not covered by the averaging window, i.e. $0 < s < a - l$ or $a < s < (T - l)$, then the value of the series $x(t)$ is constant everywhere within the window and therefore $x - \langle x \rangle_l = 0$.

For $a - l < s < a$, integration over the interval from a to $s + l$, in which $x(t)$ is different from zero, gives

$$\langle x \rangle_l = (s + l - a)/l \quad (3)$$

and thus

$$\begin{aligned} \langle (x - \langle x \rangle_l)^2 \rangle_l &= \frac{1}{l} \int_s^{s+l} \left[x(t) - \frac{1}{l}(s + l - a) \right]^2 dt \\ &= \frac{1}{l} \int_s^a \left[0 - \frac{1}{l}(s + l - a) \right]^2 dt \\ &\quad + \frac{1}{l} \int_a^{s+l} \left[1 - \frac{1}{l}(s + l - a) \right]^2 dt \\ &= \frac{1}{l^2} (s + l - a)(a - s). \end{aligned} \quad (4)$$

Averaging over all possible windows of width l is done by integrating over the starting point s and gives

$$\begin{aligned} w(l) &= \frac{1}{T - l} \int_0^{T-l} [\langle (x - \langle x \rangle_l)^2 \rangle_l]^{1/2} ds \\ &= \frac{1}{T - l} \int_{a-l}^a \left[\frac{(s + l - a)(a - s)}{l^2} \right]^{1/2} ds \\ &= \frac{\pi}{8} \frac{l}{T - l}. \end{aligned} \quad (5)$$

For $l \ll T$, this dependence on l results in an approximately straight line on a log–log plot of $w(l)$ and a spurious Hurst exponent of $H = 1$. Slightly different values of the Hurst exponent may be obtained if the jump is located close to either end of the series.

The value $H = 1$ is obtained here as a result of applying a scaling exponent determination technique to an object, which is not, strictly speaking, a self-affine fractal. The same value is obtained by other techniques, such as a box-counting method. However, when such objects, e.g. jumps in the series variable, appear in experimental data, they lead to a false fractal scaling detection.

In order to illustrate this statement, we have performed a simulation by adding jumps to a series generated by a white noise (Fig. 3). White noise may be

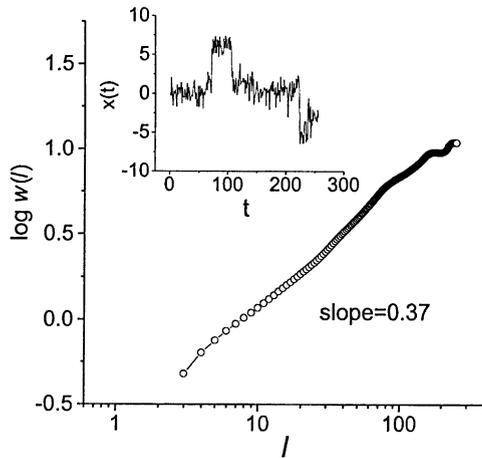


Fig. 3. Spurious scaling is obtained when jumps are added to series generated by white noise process. Four jumps of random magnitude and sign (up or down) were added at random locations in series. Resulting series is shown in inset and graph shows corresponding series width $w(l)$.

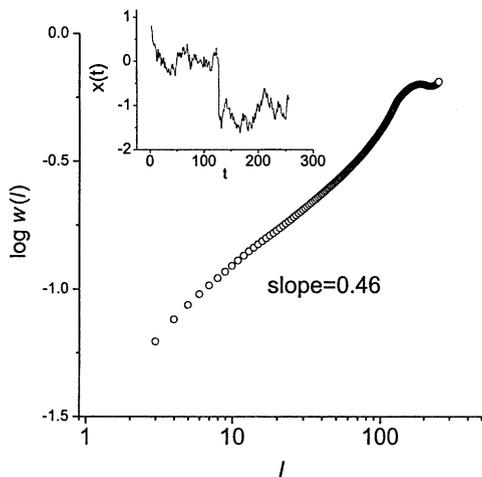


Fig. 4. Series width $w(l)$ computed from series generated by Ornstein-Uhlenbeck noise with noise correlation time $\tau = 10$ to which one jump was added. Slope is measured by fitting a straight line in the region of l values from 10 to 50.

regarded to exhibit scaling with $H = 0$ (Turcotte, 1997). Adding jumps, which are larger in magnitude than typical noise fluctuations, produces a spurious fractal scaling. As seen in Fig. 3, the jumps produce a false fractal scaling even if their magnitude is not very large so that it is hard to distinguish them from the white noise fluctuations. Spurious scaling regions may also be obtained, for example, when jumps are added to a series produced by a non-scaling Ornstein-Uhlenbeck noise (Fig. 4).

Spikes in the series also lead to spurious scaling (Fig. 5). Analytical calculations similar to the ones

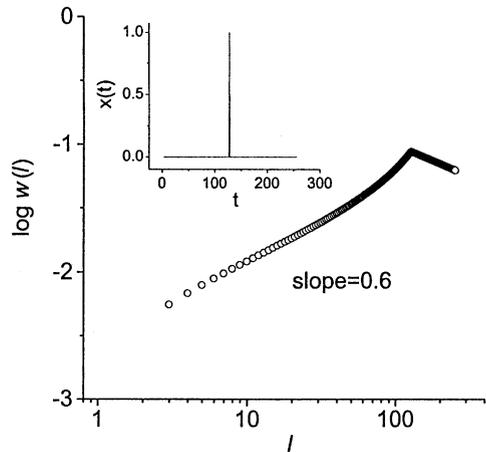


Fig. 5. Series width $w(l)$ computed from series containing one spike. Difference between obtained value of slope and its analytically calculated value of 0.6 is due to approximations taken in deriving Eq. (6). Slope is measured by fitting a straight line in the region of l values from 10 to 50.

performed above for the case of a single jump indicate that a single spike in the series leads to the series width $w(l)$ of the form

$$w(l) = \frac{\sqrt{l-1}}{T-l} \tag{6}$$

and consequently, for $1 \ll l \ll T$, to a spurious Hurst exponent $H = 0.5$.

Besides measuring the series width according to Eq. (1), a power spectrum method is often employed to detect fractal scaling in the series and to determine the value of the Hurst exponent (e.g. Halden and Hawthorne, 1993). For a series which exhibits such scaling, the power spectrum $S(\omega)$ (usually obtained by performing a Fourier transform on the series) follows the power law:

$$S \propto \omega^{-\beta}, \tag{7}$$

where ω is the frequency and the exponent β is related to the Hurst exponent H by $H = (\beta - 1)/2$ (Meakin, 1998). Thus, for Brownian noise, $\beta = 2$.

Consider an infinitely long series, which consists of one jump described by a step function $\Theta(t)$ defined as $\Theta(t) = 0$ for $t \leq 0$ and $\Theta(t) = 1$ for $t > 0$. The Fourier transform of this function yields

$$Y(\omega) = \int_{-\infty}^{+\infty} \Theta(t) e^{-i\omega t} dt = \frac{1}{i\omega} \tag{8}$$

and the power spectrum is $S(\omega) = |Y|^2 \propto \omega^{-2}$. Consequently, a spurious fractal scaling may be obtained with the scaling exponent $\beta=2$, which is the same value as for the Brownian noise (Fig. 6)! Applying this method to a noisy series of finite length with one jump typically

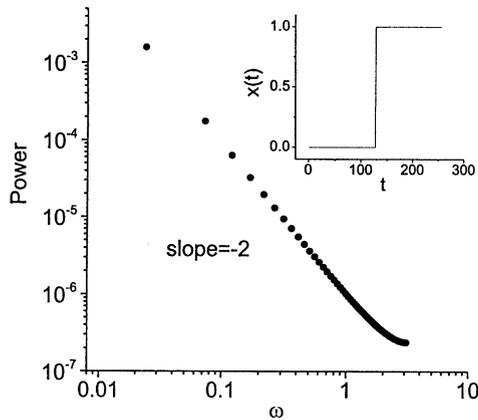


Fig. 6. Spurious power spectrum scaling due to jump in series.

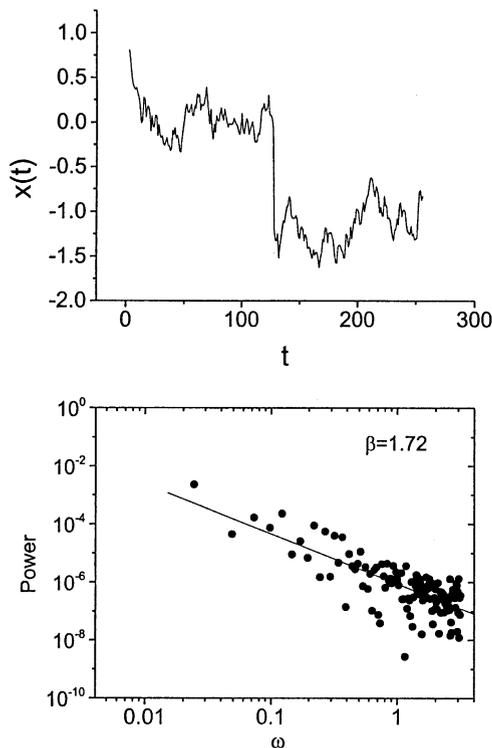


Fig. 7. Spurious power spectrum scaling in series produced by adding one jump to Ornstein–Uhlenbeck noise process with correlation time $\tau = 10$. Noise realization is same as in Fig. 4.

produces slightly smaller values of β . In Fig. 7, spurious power spectrum scaling is shown for an Ornstein–Uhlenbeck noise (Gardiner, 1983). Fitting a straight line to the power spectrum yields $\beta = 1.72$, which corresponds to $H = (\beta - 1)/2 = 0.36$.

The power spectrum of a single spike is approximately that of a delta function and is flat. Therefore, the

presence of spikes in the series produced by a scaling process may decrease the measured value of β . In general, detecting the scaling exponent by fitting a straight line to the series power spectrum is, in practice, insufficiently accurate for most short experimental series because of the large error in the fit.

4. Conclusion

We have outlined some of the caveats in interpreting the results of the time series analysis when the data set is short or contains substantial inhomogeneities, such as jumps or spikes in the series variable. Spurious fractal scaling was shown to arise when the scaling is estimated by calculating the width of the series or by estimating scaling in the power spectrum. Our results suggest that many results of the analysis of short experimental data sets, in order to be meaningful, need to be substantiated by additional evidence.

Acknowledgements

The authors wish to thank the Natural Science and Engineering Research Council of Canada and the Ministry of Training, Colleges and Universities of Ontario for financial support.

References

- Cortini, M., Anastasio, D., 2001. Chemical banding in volcanic minerals: a statistical phenomenological approach. *European Journal of Mineralogy* 13, 571–575.
- Gardiner, C.W., 1983. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer, Berlin, 442pp.
- Halden, N.M., Hawthorne, F.C., 1993. The fractal geometry of oscillatory zoning in crystals: application to zircon. *American Mineralogist* 78, 1113–1116.
- Holten, T., Jamtveit, B., Meakin, P., Cortini, M., Blundy, J., Austrheim, H., 1997. Statistical characteristics and origin of oscillatory zoning in crystals. *American Mineralogist* 82, 596–606.
- Hoskin, P.W.O., 2000. Patterns of chaos: fractal statistics and the oscillatory chemistry of zircon. *Geochimica et Cosmochimica Acta* 64 (11), 1905–1923.
- Meakin, P., 1998. *Fractals, Scaling and Growth Far from Equilibrium*. Cambridge University Press, Cambridge, 674pp.
- Pollok, K., Jamtveit, B., Putnis, A., 2001. Analytical transmission electron microscopy of oscillatory zoned grandite garnets. *Contributions to Mineralogy and Petrology* 141, 358–366.
- Turcotte, D.L., 1997. *Fractals and Chaos in Geology and Geophysics*, 2nd Edition. Cambridge University Press, Cambridge, 398pp.