Self-Similarity Index Estimation via Wavelets for Locally Self-Similar Processes

Yazhen Wang   Joseph E. Cavanaugh   Changyong Song

Department of Statistics, University of Missouri, Columbia, MO 65211

Abstract

Many naturally occurring phenomena can be effectively modeled using self-similar processes. In such applications, accurate estimation of the scaling exponent is vital, since it is this index which characterizes the nature of the self-similarity. Although estimation of the scaling exponent has been extensively studied, previous work has assumed that this parameter is constant. Such an assumption may be unrealistic in settings where it is evident that the nature of the self-similarity changes as the phenomenon evolves. For such applications, the scaling exponent must be allowed to vary as a function of time, and a procedure must be available which provides a statistical characterization of this progression. In what follows, we propose and describe such a procedure. Our method utilizes wavelets to construct local estimates of time-varying scaling exponents for locally self-similar processes. We establish a consistency result for these estimates. We investigate the effectiveness of our procedure in a simulation study, and demonstrate its applicability in the analyses of a hydrological and a geophysical time series, each of which exhibit locally self-similar behavior.

Keywords: fractals; local stationarity; long memory; long-range dependence; time series.

1. Introduction

A self-similar process is loosely defined as a stochastic process which generates a sample path that retains the same general appearance regardless of the distance from which it is observed. Self-similarity is a pervasive characteristic in naturally occurring phenomena. As a result, self-similar processes have been used to successfully model data arising in a variety of different scientific fields, including hydrology, geophysics, biology, medicine, and economics.

Self-similar processes were introduced to statisticians through the work of B. B. Mandelbrot (Mandelbrot and van Ness, 1968; Mandelbrot and Wallis, 1968, 1969). A stochastic process $Y(t)$ is formally defined as self-similar if $\mathcal{L}(Y(ct)) = \mathcal{L}(c^H Y(t))$ for $c > 0$, where $\mathcal{L}$ denotes the finite joint distribution of $Y(t)$. The constant H, called the self-similarity parameter or scaling exponent, dictates the nature of the self-similarity.

In most applications, it is assumed that $Y(t)$ has finite second moments, and that its associated increment process $X(t) = Y(t) - Y(t-1)$ is stationary. Under these assumptions, H may be taken to lie over the interval $(0, 1)$, and the value of H may be used in describing the autocorrelation structure of the increment process (see Beran, 1994, pp. 52–53). For $H \in (1/2, 1)$, $X(t)$ is characterized by serial correlations which decay slowly, and therefore exhibits long-range dependence or long memory. For $H \in (0, 1/2)$, $X(t)$ is characterized by serial correlations which decay rapidly and sum to zero. Such a process often exhibits sample paths where high and low values tend to consecutively alternate, a property known as anti-dependence or anti-persistence. For $H = 1/2$, $X(t)$ is serially uncorrelated. The estimation of H has been extensively studied, predominantly in the context of long memory where it is assumed that $H \in (1/2, 1)$ (Geweke and Porter-Hudak, 1983; Chen et al., 1995; Robinson, 1995; Hall et al., 1997; Kent and Wood, 1997).

In certain modeling applications, treating the self-similarity parameter H as a constant seems justifiable. Yet for many phenomena which exhibit self-similar behavior, the nature of the self-similarity changes as the phenomenon evolves. To sufficiently model such data, the parameter H must be allowed to vary as a function of time. Gançlaves and Flandrin (1993) and Flandrin and Gançlaves (1994) propose a class of processes which are locally self-similar with time-dependent scaling exponents and discuss their applications. For such processes, the local scaling exponent function $H(t)$ conveys important, even decisive, information regarding the behavior of the process. It is therefore desirable to develop a procedure for estimating $H(t)$. In what follows, we propose, describe, and investigate such a procedure.

The statistical characterization of $H(t)$ will have direct application in many scientific areas where the accurate depiction of self-similarity is an important practical issue. For instance in hydrology, where an investigation of Nile River depth measurements by Hurst (1951) led to pioneering work on long memory, self-similarity features in flow records can be used to describe, predict, and regulate the movement and the distribution of water (see Mandelbrot and Wallis, 1968; Percival and Guttorp, 1994). In medicine, beat-to-beat fluctuations in heart rates exhibit self-similarity patterns which can be used to distinguish between healthy patients and patients with severe congestive heart failure (see Peng et al., 1995a, 1995b). In genetics, nucleotide sequences in coding regions of DNA chains possess different self-similarity characteristics from sequences in noncoding regions. These characteristics can therefore be used to locate coding regions in long DNA sequences (see Peng et al., 1992; Buldyrev et al., 1993; Ossadnik et al., 1994).

Because the scaling exponent of a locally self-similar process is time dependent, the process as well as its associated increment process are, in general, non-stationary. Wavelets have been proven effective at handling both non-stationarity and self-similarity (Farge et al., 1993; Hwang and Mallat, 1993), and thus serve a central role in our estimation procedure for $H(t)$. For a locally self-similar process, we establish an approximate local log-linear relationship between the square of the continuous wavelet transformation and the scale for the transformation. The strength of the association is dictated by $H(t)$. This local log-linear relationship leads naturally to a least-squares procedure for estimating $H(t)$. In practice, the procedure can be implemented using appropriate values of the discrete wavelet transformation of the sample data.

Our exposition is organized as follows. Section 2 defines and briefly discusses locally self-similar processes. Section 3 develops the estimation algorithm for $H(t)$. This section concludes with the statement of a consistency result which is established in the Appendix. Section 4 discusses technical aspects regarding the implementation of the estimation algorithm, presents simulations designed to check the effectiveness of the procedure, and features applications where the procedure is used to analyze a hydrological and a geophysical time series which exhibit locally self-similar behavior.

2. Locally Self-Similar Processes

Let $Y(t)$ represent a mean-zero stochastic process with covariance function

$$\Gamma_t(u_1, u_2) = \mathrm{Cov}\{Y(t+u_1), Y(t+u_2)\}.$$

The process $Y(t)$ is said to be locally self-similar if

$$\Gamma_t(u_1, u_2) = \Gamma_t(0,0) - C(t)\,|u_1 - u_2|^{2H(t)}\{1 + o(1)\}, \text{ as } |u_1| + |u_2| \to 0, \tag{2.1}$$

where $C(t)$ is a non-negative function of t, and $H(t)$ represents the local scaling exponent function (or scaling function for short). From (2.1), it can easily be seen that the variogram of $Y(t)$ satisfies

$$E\left[\{Y(t+u_1) - Y(t+u_2)\}^2\right] = 2\,C(t)\,|u_1 - u_2|^{2H(t)}\{1 + o(1)\}, \text{ as } |u_1| + |u_2| \to 0.$$

The relation (2.1) demonstrates that for small u_1 and u_2, the correlation between $Y(t+u_1)$ and $Y(t+u_2)$ is diminished in accordance with the factor $|u_1 - u_2|^{2H(t)}$. For fixed $|u_1 - u_2|$ near zero, this factor grows as $H(t)$ becomes smaller. Consequently, as $H(t)$ decreases from 1 to 0, the local autocorrelation of the process $Y(t)$ decreases, causing a typical sample path to appear increasingly rough and erratic.

Now let $X(t) = Y(t) - Y(t-1)$ denote the increment process of $Y(t)$, with covariance function

$$\gamma_t(u_1, u_2) = \mathrm{Cov}\{X(t+u_1), X(t+u_2)\}.$$

The Winger-Ville distribution of $X(t)$, which functions as a local version of the power spectrum, is defined as

$$f_t(\lambda) = \frac{1}{2\pi}\int \gamma_t(\tau/2, -\tau/2)\, e^{-i\lambda\tau} d\tau.$$

It can be shown that

$$f_t(\lambda) = K(t)\,\lambda^{1-2H(t)}\,\{1 + o(1)\}, \quad \text{as } \lambda \to 0, \tag{2.2}$$

where $K(t)$ is a non-negative function of t. When $H(t) \in (1/2, 1)$, (2.2) implies that $f_t(\lambda)$ is unbounded at $\lambda = 0$. This spectral characteristic is indicative of long-range dependence.

If the increment process $X(t)$ is stationary, $f_t(\lambda)$ is independent of t and reduces to the ordinary power spectrum of $X(t)$. In this instance, $K(t)$ and $H(t)$ are constant, and (2.2) becomes

$$f(\lambda) = K\lambda^{1-2H}\{1 + o(1)\}, \quad \text{as } \lambda \to 0. \tag{2.3}$$

As mentioned in the introduction, estimation of constant H in (2.3) has been considered by many authors, mostly in the context of long-memory time series where $H \in (1/2, 1)$.

We now present and discuss examples of processes which exhibit locally self-similar behavior. First, consider the process defined for real $t \geq 0$ by the stochastic integral

$$Y(t) = \int_{-\infty}^{0} \left\{(t-u)^{H(t)-1/2} - (-u)^{H(t)-1/2}\right\} dB(u) + \int_{0}^{t} (t-u)^{H(t)-1/2}\, dB(u), \tag{2.4}$$

where $B(t)$ is standard Brownian motion, and $H(t) \in (0, 1)$ represents the scaling function. This process is an extension of fractional Brownian motion, FBM, which allows for the self-similarity parameter to vary over time. For this reason, the process is called generalized fractional Brownian motion, GFBM.

A simple, explicit form for the covariance function of $Y(t)$ can be derived without placing any conditions on $H(t)$. For smooth $H(t)$, the covariance function satisfies (2.1), which establishes that $Y(t)$ is locally self-similar.

The increment process of FBM, called fractional Gaussian noise, is stationary with a spectrum which satisfies (2.3). The increment process of GFBM is non-stationary (unless $H(t)$ is constant) with a spectrum which satisfies (2.2). GFBM relaxes many homogeneous restrictions of FBM. The latter may exhibit sample paths of varying degrees of roughness which are characterized by varying modes of self-similarity. Due to this general behavior, the process can be used to model complicated natural phenomena (Mandelbrot, 1983; Gançlaves and Flandrin, 1993; Flandrin and Gançlaves, 1994).

Next, consider the process defined for non-negative integer t as

$$(1 - \Phi(\mathrm{B}))(1 - \mathrm{B})^{H(t)-1/2} X(t) = (1 - \Theta(\mathrm{B}))\epsilon(t), \tag{2.5}$$

where B is a backshift operator defined by $\mathrm{B}X(t) = X(t-1)$, $\Phi(\mathrm{B})$ and $\Theta(\mathrm{B})$ are polynomials in B having characteristic roots outside the unit circle, $\epsilon(t)$ is Gaussian white noise, and $H(t) \in (0,1)$ represents the scaling function. This process extends the definition of a fractional autoregressive integrated moving-average or ARFIMA process by again allowing for the self-similarity parameter to evolve over time. We refer to this process as a generalized ARFIMA or GARFIMA process.

An ARFIMA process is stationary with a spectrum which satisfies (2.3). A GARFIMA process is non-stationary (unless $H(t)$ is constant) with a spectrum which satisfies (2.2). It can be shown that the normalized partial sums of an ARFIMA process have the same limiting distribution as a globally self-similar process (see Beran, 1994, pp. 48–50). Analogously, it can be shown that the normalized partial sums of a GARFIMA process have the same limiting distribution as a locally self-similar process; i.e., a process having a covariance function which satisfies (2.1). Thus, an ARFIMA process can be regarded as the increment process for a globally self-similar process, and a GARFIMA process can be regarded as the increment process for a locally self-similar process.

ARFIMA processes have been extensively studied in the context of long-memory time series (see Beran, 1994). Though not previously considered in the literature, GARFIMA processes have even greater potential for widespread applicability, since they provide for the modeling of non-stationary time series which exhibit irregular patterns of roughness or variability. Utilizing GARFIMA processes in time series applications would require identification, estimation, and forecasting algorithms for GARFIMA models. Such algorithms have been previously proposed and studied for ARFIMA models (see, for instance, Brockwell and Davis, 1994), and could likely be adapted to the GARFIMA setting given a procedure for estimating $H(t)$ such as the one we propose here.

A special case of (2.5) is defined by

$$(1-\mathrm{B})^{H(t)-1/2}X(t) = \epsilon(t). \tag{2.6}$$

We refer to such a process as generalized fractionally integrated noise, GFIN, since it represents an extension of fractionally integrated noise, FIN.

3. Estimation of the Self-Similarity Index via the Wavelet Transformation

We now develop the procedure used to estimate the scaling function $H(t)$.

Let ψ denote the Daubechies mother wavelet, and let $TY(a,t)$ denote the wavelet transformation of the locally self-similar process $Y(t)$ corresponding to the scale a and the location t. We can write

$$TY(a,t) = a^{-1/2} \int \psi\left(\frac{u-t}{a}\right) Y(u)\, du = a^{1/2} \int \psi(x)\, Y(t+a\,x)\, dx. \tag{3.1}$$

Using (3.1), (2.1), and the first vanishing moment of ψ, we obtain

$$\begin{aligned} E\left(|TY(a,t)|^2\right) &= \frac{1}{a} \int\int \psi\left(\frac{u-t}{a}\right) \psi\left(\frac{v-t}{a}\right) E\{Y(u)\,Y(v)\}\, du\, dv \\ &= a \int\int \psi(x)\,\psi(y)\, E\{Y(t+a\,x)\,Y(t+a\,y)\}\, dx\, dy \\ &= a \int\int \psi(x)\,\psi(y)\,\Gamma_t(a\,x, a\,y)\, dx\, dy \\ &\sim a \int\int \psi(x)\,\psi(y)\left\{\Gamma_t(0,0) - C(t)\,|a\,x - a\,y|^{2\,H(t)}\right\} dx\, dy \\ &= C_1\, a^{1+2\,H(t)}, \qquad \text{as } a \to 0, \end{aligned} \tag{3.2}$$

where

$$C_1 = -C(t) \int\int |x-y|^{2\,H(t)} \psi(x)\,\psi(y)\, dx\, dy.$$

Now let

$$y_t(a) = \log\left(|TY(a,t)|^2\right),$$

$$C_2 = E\left[\log\left\{|TY(a,t)|^2 \big/ E\left(|TY(a,t)|^2\right)\right\}\right],$$

and

$$\varepsilon_t(a) = \log\left\{|TY(a,t)|^2 \big/ E\left(|TY(a,t)|^2\right)\right\} - C_2.$$

Then clearly,

$$y_t(a) = C_2 + \log\left\{E\left(|TY(a,t)|^2\right)\right\} + \varepsilon_t(a). \tag{3.3}$$

Note that (3.2) and (3.3) imply the approximate regression model

$$y_t(a) \approx c + \{2\,H(t) + 1\}\, \log\, a + \varepsilon_t(a), \qquad \text{for small scale } a, \tag{3.4}$$

where $c = \log C_1 + C_2$. This, in turn, suggests that we use the method of least squares to estimate $H(t)$. The general procedure is outlined as follows. More specific details are provided in Subsection 4.1.

First, we select a sequence of small scales $a_1 > \cdots > a_k$, say $a_j = 2^{-j}$ where $j = 1, \ldots, k$. Next, we set $x_j = \log a_j$ and $y_j = y_t(a_j)$ for each j. Finally, treating (x_j, y_j), $j = 1, \cdots, k$, as a set of bivariate data, we evaluate the least-squares estimate of $H(t)$ in (3.4) via

$$\hat{H}(t) = \left\{ \frac{\sum(x_j - \bar{x})\,(y_j - \bar{y})}{\sum(x_j - \bar{x})^2} - 1 \right\} /2, \tag{3.5}$$

where $\bar{x} = \sum x_j/k$, $\bar{y} = \sum y_j/k$.

In the Appendix, we establish the following consistency result for the estimator $\hat{H}(t)$.

Theorem. *Suppose that Y(t) is Gaussian with a covariance function which satisfies (2.1). Then $\hat{H}(t)$ is a consistent estimator of $H(t)$.*

As mentioned in the introduction, if a globally self-similar process $Y(t)$ has stationary increments and finite second moments, H may be taken to lie over the interval $(0, 1)$, since values of H outside of this range lead to processes which are not of practical interest. However, in applying our estimation procedure for $H(t)$ to real data, estimates of $H(t)$ which are either negative or exceed 1 will occasionally occur. Such estimates may arise for a variety of reasons. For instance, if the procedure is applied to the first difference of a locally self-similar process, estimates of $H(t)$ between -1 and 0 generally result; if the procedure is applied to the partial sums of a locally self-similar process, estimates of $H(t)$ between $+1$ and $+2$ generally result. Thus, estimates of $H(t)$ outside the interval (0,1) may imply that the underlying series (or subseries) must be suitably differenced or accumulated in order to obtain a series which can be regarded as locally self-similar with locally stationary increments and finite second moments. Of course, such estimates may also occur if the underlying series (or subseries) reflects the dynamics of a process which lies outside the realm of locally self-similar processes. This issue, however, lies outside the scope of the present paper.

In the next section, we discuss the implementation of our estimation procedure and investigate its effectiveness in a simulation study. We also use our procedure to characterize $H(t)$ for two real time series which exhibit locally self-similar behavior.

4. Simulations and Applications

4.1 Algorithm Description

In practice, we generally observe a process $Y(t)$ at discrete, equally-spaced time points. For convenience, we will assume that the time points are scaled to lie over the interval [0,1), and that the sample size is a power of 2; say, $t_i = (i-1)/n$, where $i = 1, \ldots, n = 2^J$. The resulting data are analyzed using the discrete wavelet transformation, a discretized version of the continuous wavelet transformation utilized in the development of the methodology of Section 3.

Let the vector $Y = [Y(t_1), \ldots, Y(t_n)]'$ represent a sample of n measurements on $Y(t)$. The discrete wavelet transformation of Y can be written in the form $\mathcal{W}Y$, where $\mathcal{W}$ is an $n \times n$ orthogonal matrix which depends on both the wavelet and the boundary adjustment (Cohen et al., 1993; Daubechies, 1994). Fast algorithms with complexity of order n are available for performing both the transformation and the inversion of the transformation which results in the reconstruction of the original data.

To label the $n = 2^J$ coefficients of the transformation, we index $n-1$ of the coefficients dyadically:

$$y_{j,k}; \quad k = 0, \ldots, 2^j - 1; \quad j = 0, \ldots, J-1.$$

We then label the remaining coefficient $y_{-1,0}$. The coefficient $y_{j,k}$ is referred to as the discrete wavelet transformation of Y at level j and location $k2^{-j}$. For $k = 0, \ldots, 2^j - 1$ and $j = 0, \ldots, J-1$, the quantity $y_{j,k}$ can be viewed as an approximation to $TY(2^{-j}, k2^{-j})$, the continuous wavelet transformation $TY(a,t)$ evaluated at scale $a = 2^{-j}$ and location $t = k2^{-j}$ (Donoho and Johnstone, 1994).

In what follows, we discuss technical aspects of the algorithm used in our simulations and applications to estimate $H(t)$.

We partition the sampling interval $[0,1)$ into 2^l nonoverlapping subintervals of equal length, where l is an integer chosen such that $1 \le l \le (J-1)$. The 2^l subintervals are of the form

$$I_m = [(m-1)2^{-l}, m2^{-l}); \quad m = 1, \ldots, 2^l.$$

For each I_m, $m = 1, \ldots, 2^l$, an estimate $\hat{H}(t)$ is constructed. One may envision each $\hat{H}(t)$ as estimating the average value of the scaling function $H(t)$ over the corresponding subinterval I_m. The appropriate time index for the $\hat{H}(t)$ associated with I_m might be regarded as the midpoint of I_m, namely $2^{-l-1}(2m-1)$.

For a given I_m, we compute $\hat{H}(t)$ by pooling together values of $y_{j,k}$ where the level j is "large," say no less than some positive integer J' ($J' \leq (J-1)$), and where the location $k2^{-j}$ is within the subinterval I_m. For each $m = 1, \ldots, 2^l$, we define the bivariate collections of data

$$(X_m, Y_m) = \left[\left\{\log\left(2^{-j}\right), \log\left(|y_{j,k}|^2\right)\right\} \;\; \middle| \;\; k2^{-j} \in I_m; \;\; 0 \leq k \leq 2^j - 1, \;\; J' \leq j \leq (J-1)\right].$$

We then fit a least-squares line to (X_m, Y_m), treating the X_m as the regressor measurements and the Y_m as the predictor measurements. The estimate $\hat{H}(t)$ is found by adjusting the estimate of the slope in the least-squares fit by first subtracting 1 and then dividing by 2, as indicated in (3.5).

As a final, optional stage in our procedure, we construct a curve from the collection of estimates $\hat{H}(t)$ by employing local polynomial smoothing. This curve then serves to approximate the shape of the scaling function $H(t)$.

For applications where the sample size is not a power of two, we augment the data in order to achieve such a sample size, and we adjust the results accordingly. (See Subsection 4.4 for an example.) In the discrete wavelet transformation, we utilize the least asymmetric wavelet with the symmetric boundary condition and 8 vanishing moments (see Daubechies, 1992).

4.2 Simulations

We test the performance of our algorithm in six simulation sets based on the GFBM process (2.4) and the partial sums of the GFIN process (2.6). We employ three different specifications of the scaling function $H(t)$. We simulate the processes so that the time index t is confined to the interval $[0, 1)$.

To simulate realizations of the GFBM process (2.4), consider setting

$$B_n(t_i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{i} \epsilon(t_k); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \tag{4.1}$$

where the $\epsilon(t_k)$ are variates of a Gaussian white noise process. For large n, the values of $B_n(t_i)$ can be treated as realizations of Brownian motion, since the normalized partial sum in (4.1) approximates the stochastic integral which defines such a process. The second of the two stochastic integrals which comprise (2.4) can then be approximated by its discretized analogue

$$\begin{aligned} Y_{2,n}(t_i) &= \sum_{k=1}^{i} (t_i - t_k)^{H(t_i)-1/2} \{B_n(t_k) - B_n(t_{k-1})\} \qquad \{B(t_0) \equiv 0\} \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^{i} (t_i - t_k)^{H(t_i)-1/2} \epsilon(t_k); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n. \end{aligned}$$

Similarly, the first of the two stochastic integrals in (2.4) can be approximated by a sum of the form

$$\begin{aligned} Y_{1,N}(t_i) &= \frac{1}{\sqrt{N}} \sum_{k=-N}^{0} \left\{ (t_i - u_k)^{H(t_i)-1/2} - (-u_k)^{H(t_i)-1/2} \right\} \epsilon(u_k); \\ & \qquad u_k = \left(\frac{k}{N}\right) K; \; k = 0, \ldots, -N; \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \end{aligned}$$

where the $\epsilon(u_k)$ are variates of a Gaussian white noise process, and the integer $N > 0$ and constant $K > 0$ are chosen such that $(u_{-N}, u_{-N+1}], (u_{-N+1}, u_{-N+2}], \ldots, (u_{-1}, u_0]$ provides a sufficiently fine partition of a suitably large interval $(-K, 0]$. The sequence

$$Y_n(t_i) = Y_{1,N}(t_i) + Y_{2,n}(t_i); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \tag{4.2}$$

can then be treated as realizations of the process (2.4). We remark that the contribution of $Y_{1,N}(t_i)$ to $Y_n(t_i)$ is negligible for the purpose at hand.

To simulate realizations of the GFIN process defined by (2.6), we use the fact that (2.6) can be represented as an infinite moving average of the form

$$X(t_i) = \sum_{k=0}^{\infty} a(t_i, k) \epsilon(t_{i-k}), \tag{4.3}$$

where the $\epsilon(t_{i-k})$ are variates of a Gaussian white noise process and the coefficients $a(t_i, k)$ are defined by

$$a(t_i, k) = \frac{\Gamma(k + \{H(t_i) - 1/2\})}{\Gamma(k+1)\Gamma(H(t_i) - 1/2)}$$

(see Beran, 1994, p. 65). The infinite sum in (4.3) can be approximated by the finite sum

$$X_n(t_i) = \sum_{k=0}^{N} a(t_i, k)\epsilon(t_{i-k}); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \tag{4.4}$$

provided that N is chosen to be a suitably large integer. The sequence (4.4) can then be treated as realizations of the process (4.3), or equivalently, (2.6). Accordingly, for large n, the normalized partial sums

$$Y_n(t_i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{i} X_n(t_k); \quad t_i = \frac{(i-1)}{n}; \; i = 1, \ldots, n; \tag{4.5}$$

can be regarded as realizations of a self-similar process with a covariance function of the form (2.1). These sums are therefore amenable to our algorithm.

We generate three samples of the GFBM process (2.4) and three of the GFIN process (2.6) using the following scaling functions:

$$H_1(t) = \begin{cases} 0.3 & t \in (0, 0.5], \\ 0.7 & t \in (0.5, 1], \end{cases} \tag{4.6}$$

$$H_2(t) = 0.2 + 0.7t, \tag{4.7}$$

$$H_3(t) = 3(t - 0.5)^2 + 0.1. \tag{4.8}$$

In each instance, we consider a sample of size $n = 2^{12} = 4096$ ($J = 12$).

In the estimation algorithm, we partition the sampling interval $[0, 1)$ into $32 = 2^5$ ($l = 5$) subintervals, which results in 32 estimates of $H(t)$ corresponding to the time indices $t = 1/64, 3/64, \ldots, 63/64$. In pooling values of $y_{j,k}$ for each estimate, we restrict the level j to be no less than 4 ($J' = 4$).

For the simulation sets based on the GFBM processes, the sample paths and the estimated curves for $H(t)$ are displayed in Figures 1.1(a) through 1.3(b); for the sets based on the GFIN processes, the sample paths for the partial sums and the estimated curves for $H(t)$ are displayed in Figures 2.1(a) through 2.3(b). In constructing the estimated curves for the sets involving the step function $H_1(t)$, we do not smooth over the point of discontinuity.

We note that each estimated curve effectively approximates the general shape of the corresponding scaling function $H(t)$. Although the curves reflect a minor amount of negative bias in the $\hat{H}(t)$, they also reflect a relatively low degree of variability. These encouraging results suggest that our estimation procedure should result in an effective characterization of $H(t)$ in large-sample settings.

4.3 Application: Vertical Shear Measurements

Percival and Guttorp (1994) analyze a set of vertical ocean shear measurements. The data for the measurements are collected by dropping a probe into the ocean which records the water velocity every 0.1 meter as it descends. The "time" index is depth (in meters). The shear measurements (in 1/seconds) are obtained by taking a first difference of the velocity readings over 10 meter intervals, and then applying a low-pass filter to the differenced readings.

Vertical shear measurements display characteristics typical of self-similar processes; in particular, the increments of such a series often exhibit long-memory behavior. The data considered by Percival and Guttorp consist of 6875 values collected from a depth of 350.0 meters down to 1037.4 meters. The authors analyze 4096 of the values (chosen from the middle of the series) using wavelets and the Allan variance. Their justification for selecting the central part of the sample for their analysis is that "this subseries can be regarded as a portion of one realization of a process whose first backward difference is a stationary process" (p. 334). In other words, this part of the sample can be regarded as a sample path of a globally self-similar process with stationary increments.

We analyze the entire sample under the premise that the series can be treated as a realization of a locally self-similar process with an autocovariance function of the form (2.1). Our goal is to estimate the scaling function $H(t)$.

Rather than redefining the time index so that the sampling interval is $[0, 1)$, we retain the values of the original index (in meters). Since the size of the series is not a power of two, we extend its length to $2^{13} = 8192$ ($J = 13$) by repeating the last value in the series 1317 times. We then apply our estimation algorithm by partitioning the sampling interval into $32 = 2^5$ ($l = 5$) subintervals, which results in 32 estimates of $H(t)$. Only the first 27 estimates are relevant, since the remaining estimates pertain to the augmented portion of the series. In pooling values of $y_{j,k}$ for each estimate, we again restrict the level j to be no less than 4 ($J' = 4$).

The series is plotted in Figure 3(a), and the smoothed estimated curve for $H(t)$ in Figure 3(b). We note that the curve provides strong evidence that the self-similarity parameter $H(t)$ is not constant. The curve tends to vary over the range from 0.6 to 1.0, which reinforces the notion that the increments of the series exhibit long-memory behavior. However, the shape of the curve indicates that the nature of the long-range dependence changes considerably as a function of depth. To model $H(t)$ as a constant over the entire sampling interval would be to ignore the local aspect of self-similarity which appears to characterize the series.

We note that the estimated curve for $H(t)$ does appear fairly constant over the range of t corresponding to the portion of the sample analyzed by Percival and Guttorp (1994). This supports the notion that the central section of the series can be regarded as globally self-similar with stationary increments. The remaining sections, however, exhibit time-varying self-similarity patterns which can only be characterized by modeling the scaling exponent as a function of t and by estimating $H(t)$ accordingly.

4.4 Application: Regional Earthquake Seismogram Measurements

Figure 4(a) features a portion of a seismogram produced by a regional earthquake, recorded at the Norwegian seismic array NORESS on April 13, 1992. The local magnitude of the earthquake was measured to be 4.4 on the Richter scale. A total of $2^{13} = 8192$ readings are displayed in the figure. The time interval between observations is .025 seconds.

The seismogram for a regional earthquake is often divided into the P (primary) phase and the S (secondary) phase. In Figure 4(a), the P phase roughly corresponds to the first half of the seismogram, where the amplitude is relatively small and fairly constant. The S phase roughly corresponds to the second half of the seismogram, where the amplitude is relatively large and somewhat variable. For many regional earthquake seismograms (including the one considered here), higher frequencies are more prevalent in the P phase than in the S phase.

We analyze the portion of the seismogram featured in Figure 4(a) assuming that the series can be regarded as a realization of a locally self-similar process with an autocovariance function of the form (2.1). Again, our goal is to estimate the scaling function $H(t)$.

Rather than redefining the time index so that the sampling interval is $[0, 1)$, we retain the values of the original index. We apply our estimation algorithm by partitioning the sampling interval into $32 = 2^5$ ($l = 5$) subintervals, which results in 32 estimates of $H(t)$. In pooling values of $y_{j,k}$ for each estimate, we again restrict the level j to be no less than 4 ($J' = 4$).

The smoothed estimated curve for $H(t)$ is featured in Figure 4(b). The curve leads to several interesting conclusions regarding the behavior of the seismogram. First, as the time index increases, $\hat{H}(t)$ tends to increase from 0.2 to 1.0. Recall from Section 2 that as $H(t)$ decreases from 1 to 0, a typical sample path for a locally self-similar process will become increasingly rough and erratic. The fact that $\hat{H}(t)$ is smaller for the P phase than for the S phase is consistent with the notion that the P phase exhibits higher frequency behavior than the S phase. Second, the $\hat{H}(t)$ curve displays a local minimum which seems to correspond to the time at which the amplitude of the seismogram dramatically increases. If this is a general tendency which holds for seismograms from regional earthquakes, the scaling function $H(t)$ could possibly be used as a tool for delineating between the P phase and the S phase. Finally, the $\hat{H}(t)$ curve rises rapidly after the onset of the S phase. This indicates that the increments of the latter section of the series exhibit long-memory behavior. Moreover, as the S phase progresses, the serial correlation of the increments tends to increase.

Delineating between seismograms resulting from regional earthquakes and regional underground explosions is an important problem in geophysics (see Shumway, 1982, 1988). In future work, we hope to explore the potential application of self-similarity patterns, as summarized by the behavior of $H(t)$, in developing discriminants for this purpose.

5. Conclusion

We have proposed, discussed, and investigated an algorithm for estimating the time-varying self-similarity parameter $H(t)$ in a locally self-similar process. We show that our procedure leads to a consistent estimator of $H(t)$ at a given time point t. Our simulation results indicate that our method provides an accurate characterization of the progression of $H(t)$. Moreover, our applications illustrate that this characterization can be used to quantify time-dependent self-similarity patterns which arise in actual spatial and temporal series.

Our algorithm can be conveniently implemented using the discrete wavelet transform of the sample data. Such a simple and direct procedure for statistically characterizing $H(t)$ expands the potential applications of locally self-similar processes in modeling complicated natural phenomena.

Appendix

Proof of Theorem in Section 3

The theorem can be easily established by proving

$$E\left\{\hat{H}(t)\right\} \to H(t) \qquad \text{and} \qquad \text{Var}\left\{\hat{H}(t)\right\} \to 0. \tag{A.1}$$

The result then follows by Chebyshev's inequality.

Recall from Section 3 that the bivariate data values which lead to $\hat{H}(t)$ are denoted by $x_j = \log(2^{-j})$ and $y_j = y_t(2^{-j})$ for $j = 1, \ldots, k$, where $y_t(a) = \log\left(|TY(a,t)|^2\right)$. Thus, $\hat{H}(t)$ is based on a bivariate sample of size k. We will prove (A.1) by demonstrating that as $k \to \infty$,

$$E\{\hat{H}(t)\} = H(t) + o(k^{-1}), \qquad \text{Var}\{\hat{H}(t)\} = O(k^{-3}). \tag{A.2}$$

To ease notation, we suppress the time index t in the specifications of $y_t(a)$, $\beta(t)$, $H(t)$, and $\hat{H}(t)$, and write these objects simply as $y(a)$, β, H, and $\hat{H}$, respectively.

Direct computations show that

$$\sum_{j=1}^{k}(x_j - \bar{x})^2 = (\log\ 2)^2 \left\{\sum_{j=1}^{k} j^2 - k\,(k+1)^2/4\right\} = (\log\ 2)^2\ k\,(k+1)\,(k-1)/12 \sim k^3, \tag{A.3}$$

and

$$\sum_{j=1}^{k} |x_j - \bar{x}| = (\log\ 2)\ \sum_{j=1}^{k} |j - (k+1)/2| \sim k^2. \tag{A.4}$$

By (3.2), we have as $a \to 0$,

$$E\left(|TY(a,t)|^2\right) = C_1\, a^{2H+1}\ \{1 + o(1)\},$$

and thus

$$\log\left\{E\left(|TY(a,t)|^2\right)\right\} = \log C_1 + (2\,H + 1)\ \log\ a + o(1).$$

The preceding relation and (3.3) together imply

$$E(y_j) = c + (2\,H + 1)\,x_j + o(1), \tag{A.5}$$

where $c = \log C_1 + C_2$. From (3.5) along with (A.3), (A.4), and (A.5), we obtain

$$\begin{aligned} E\{\hat{H}\} &= \frac{\sum(x_j - \bar{x})\, E(y_j - \bar{y})}{2\sum(x_j - \bar{x})^2} - \frac{1}{2} \\ &= H + o\left(\frac{\sum |x_j - \bar{x}|}{\sum(x_j - \bar{x})^2}\right) \\ &= H + o(k^{-1}). \end{aligned}$$

This establishes the first result of (A.2). We now derive the second result.

In what follows, we use K to denote constants which may depend on t but not on the scales a, and which may vary in definition from one usage to the next. Utilizing (3.1), the covariance property (2.1), and the first vanishing moment of the Daubechies wavelet ψ, we have

$$\begin{aligned} &\mathrm{Cov}\{TY(a_1,t), TY(a_2,t)\} \\ &= \frac{1}{(a_1\, a_2)^{1/2}} \int\int \psi\left(\frac{u-t}{a_1}\right)\, \psi\left(\frac{v-t}{a_2}\right)\, E\{Y(u)\, Y(v)\}\, du\, dv \\ &= (a_1\, a_2)^{1/2} \int\int \psi(x)\, \psi(y)\, E\{Y(t + a_1\, x)\, Y(t + a_2\, y)\}\, dx\, dy \\ &= (a_1\, a_2)^{1/2} \int\int \psi(x)\, \psi(y)\, \Gamma_t(a_1\, x, a_2\, y)\, dx\, dy \\ &\sim (a_1\, a_2)^{1/2} \int\int \psi(x)\, \psi(y)\, \left[\Gamma_t(0,0) - C(t)\, |a_1\, x - a_2\, y|^{2H}\right]\, dx\, dy \\ &= -C(t)\, (a_1\, a_2)^{1/2} \int\int \psi(x)\, \psi(y)\, |a_1\, x - a_2\, y|^{2H}\, dx\, dy \\ &= -C(t)\, a_1^{2H+1/2}\, a_2^{1/2} \int\int \psi(u + y\, a_2/a_1)\, \psi(y)\, |u|^{2H}\, du\, dy, \quad \text{as } a_1, a_2 \to 0. \qquad \text{(A.6)} \end{aligned}$$

For the Daubechies wavelet ψ with M vanishing moments, we take an M^{th}-order expansion of $\psi(u + y\, a_2/a_1)$ about u to establish that as $(a_2/a_1) \to 0$,

$$\psi(u + y\, a_2/a_1) = \psi(u) + y\psi'(u)(a_2/a_1) + \ldots + (1/M!) y^M \psi^{(M)}(u)(a_2/a_1)^M \{1 + o(1)\}.$$

This expansion allows us to show

$$\int\int \psi(u + y\, a_2/a_1)\, \psi(y)\, |u|^{2H}\, du\, dy \sim K(a_2/a_1)^M, \qquad \text{as } (a_2/a_1) \to 0.$$

Using the preceding in (A.6), we see that as $a_1, a_2 \to 0$ with $(a_2/a_1) \to 0$,

$$\mathrm{Cov}\{TY(a_1,t), TY(a_2,t)\} \sim K\, a_1^{2H+1/2}\, a_2^{1/2}\, (a_2/a_1)^M.$$

This result implies that for small a_1, a_2 chosen such that $a_2 < a_1$,

$$|\text{Corr}\{TY(a_1,t), TY(a_2,t)\}| \leq K\,(a_2/a_1)^{M-H}. \tag{A.7}$$

Now for a bivariate normal pair of random variables (W, Z) with $E(W) = E(Z) = 0$, it can be shown that

$$|\text{Corr}\{\log(W^2), \log(Z^2)\}| \leq [\text{Corr}(W,Z)]^2\,.$$

It therefore follows from (A.7) that for small $a_2 < a_1$,

$$|\text{Corr}\{y(a_1), y(a_2)\}| \;\leq\; [\text{Corr}\{TY(a_1,t), TY(a_2,t)\}]^2 \;\leq\; K\,(a_2/a_1)^{2M-2H}. \tag{A.8}$$

Furthermore, if W is normal with variance σ^2, the variance of $\log(W^2)$ does not depend of σ^2, as can be seen by writing $\log(W^2) = \log\{(W/\sigma)^2\} + \log\sigma^2$. Utilizing this fact along with (A.8), we have that for small $a_2 < a_1$,

$$|\text{Cov}\{y(a_1), y(a_2)\}| \leq K\,(a_2/a_1)^{2M-2H}.$$

From the preceding, we have

$$|\text{Cov}(y_i, y_j)| \leq K\,2^{|i-j|\,(2H-2M)}. \tag{A.9}$$

With (A.9), we can write

$$\begin{aligned}
|\sum_{ij} x_i\,x_j\,\text{Cov}(y_i,y_j)| &\leq K \sum_{ij} i\,j\,2^{|i-j|\,(2H-2M)} \\
&\leq K \sum_{i\leq j} i\,j\,2^{(j-i)\,(2H-2M)} \\
&= K \sum_{i=1}^{k} \left[i\,2^{-i\,(2H-2M)} \left\{ \sum_{j=i}^{k} j\,2^{j\,(2H-2M)} \right\} \right] \\
&\leq K \sum_{i=1}^{k} \left[i\,2^{-i\,(2H-2M)} \left\{ i\,2^{i\,(2H-2M)} \right\} \right] \\
&= K \sum_{i=1}^{k} i^2 \\
&\sim k^3.
\end{aligned} \tag{A.10}$$

Similarly, we can show

$$\bar{x}\,|\sum_{ij} x_i\,\text{Cov}(y_i,y_j)| \sim k^3, \qquad (\bar{x})^2\,|\sum_{ij} \text{Cov}(y_i,y_j)| \sim k^3. \tag{A.11}$$

By (A.10) and (A.11), we have

$$\begin{aligned}\sum_{ij}(x_i-\bar{x})\,(x_j-\bar{x})\,\mathrm{Cov}(y_i,y_j) &= \sum_{ij} x_i\,x_j\,\mathrm{Cov}(y_i,y_j)-\bar{x}\sum_{ij}(x_i+x_j)\,\mathrm{Cov}(y_i,y_j)\\ &\quad+(\bar{x})^2\sum_{ij}\mathrm{Cov}(y_i,y_j)\\ &\sim k^3,\end{aligned}$$

Finally, utilizing the preceding along with (3.5) and (A.3), we have

$$\mathrm{Var}(\hat{H})=\frac{\sum_{ij}(x_i-\bar{x})\,(x_j-\bar{x})\,\mathrm{Cov}(y_i,y_j)}{4\{\sum(x_i-\bar{x})^2\}^2}\sim k^{-3}.$$

Figure Captions

Figure 1.1(a). Sample path of GFBM process with scaling function $H_1(t)$ ($n = 4096$).

Figure 1.1(b). Smoothed estimates for $H_1(t)$ for GFBM sample path in Figure 1.1(a).

Figure 1.2(a). Sample path of GFBM process with scaling function $H_2(t)$ ($n = 4096$).

Figure 1.2(b). Smoothed estimates for $H_2(t)$ for GFBM sample path in Figure 1.2(a).

Figure 1.3(a). Sample path of GFBM process with scaling function $H_3(t)$ ($n = 4096$).

Figure 1.3(b). Smoothed estimates for $H_3(t)$ for GFBM sample path in Figure 1.3(b).

Figure 2.1(a). Partial sums for sample path of GFIN process with scaling function $H_1(t)$ ($n = 4096$).

Figure 2.1(b). Smoothed estimates for $H_1(t)$ for GFIN sample path in Figure 2.1(a).

Figure 2.2(a). Partial sums for sample path of GFIN process with scaling function $H_2(t)$ ($n = 4096$).

Figure 2.2(b). Smoothed estimates for $H_2(t)$ for GFIN sample path in Figure 2.2(a).

Figure 2.3(a). Partial sums for sample path of GFIN process with scaling function $H_3(t)$ ($n = 4096$).

Figure 2.3(b). Smoothed estimates for $H_3(t)$ for GFIN sample path in Figure 2.3(b).

Figure 3(a). Vertical shear series ($n = 6875$).

Figure 3(b). Smoothed estimates of $H(t)$ for vertical shear series in Figure 3(a).

Figure 4(a). Earthquake series ($n = 8192$).

Figure 4(b). Smoothed estimates of $H(t)$ for earthquake series in Figure 4(a).

Figure 1.1(a)

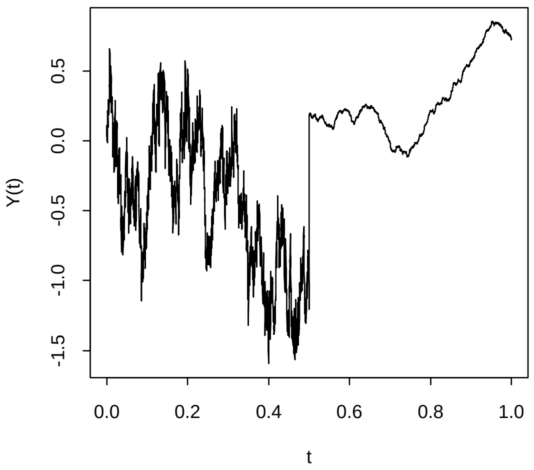


Figure 1.1(b)

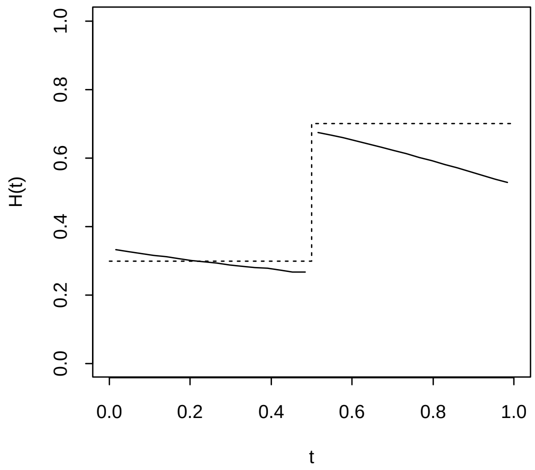


Figure 1.2(a)

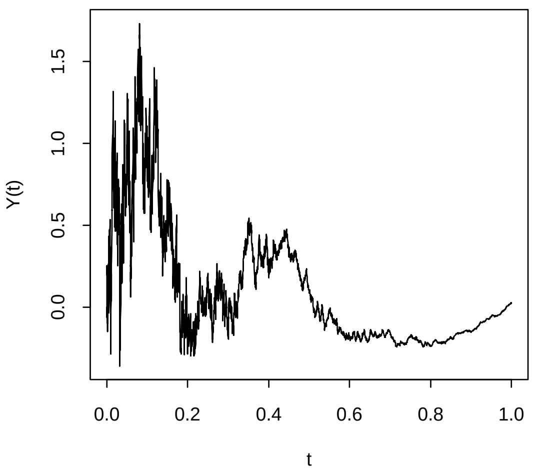


Figure 1.2(b)

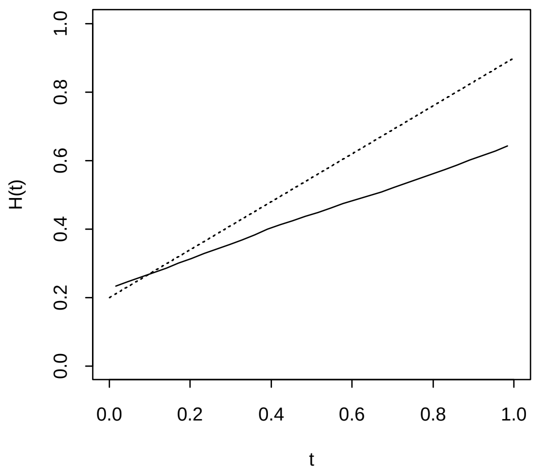


Figure 1.3(a)

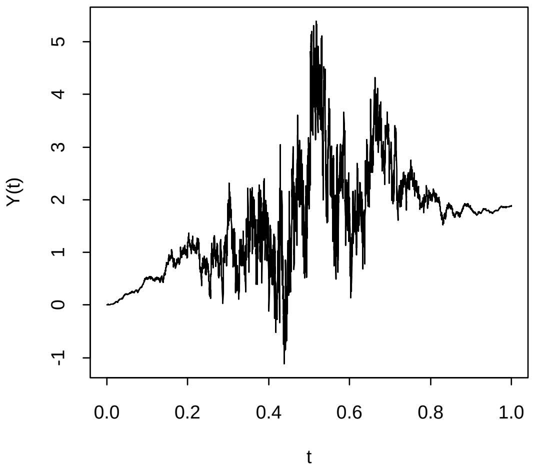


Figure 1.3(b)

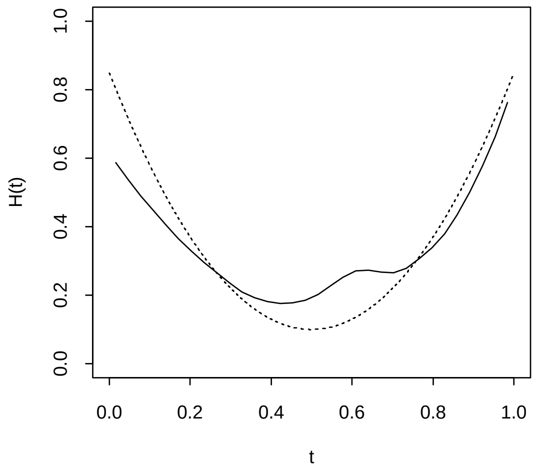


Figure 2.1(a)

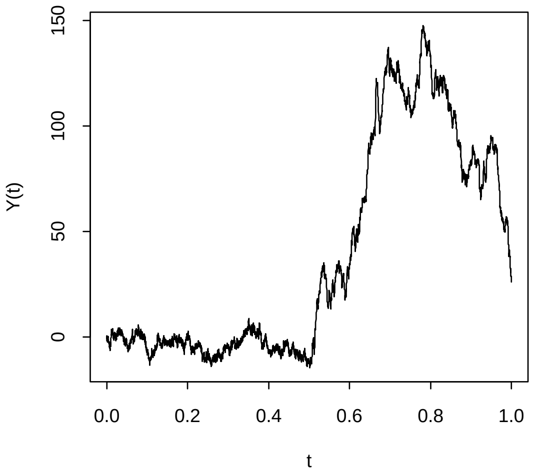


Figure 2.1(b)

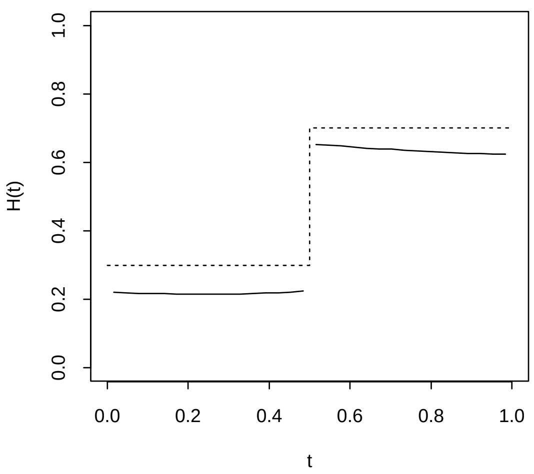


Figure 2.2(a)

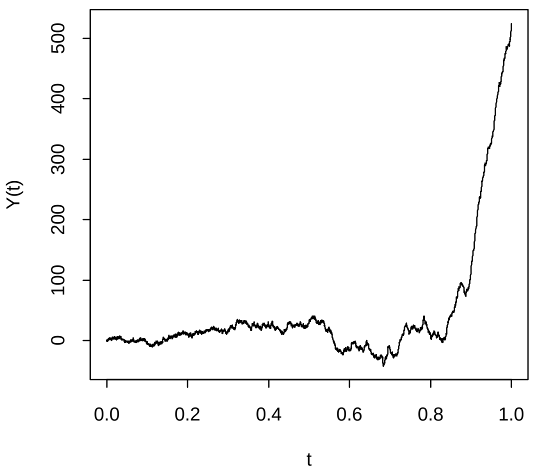


Figure 2.2(b)

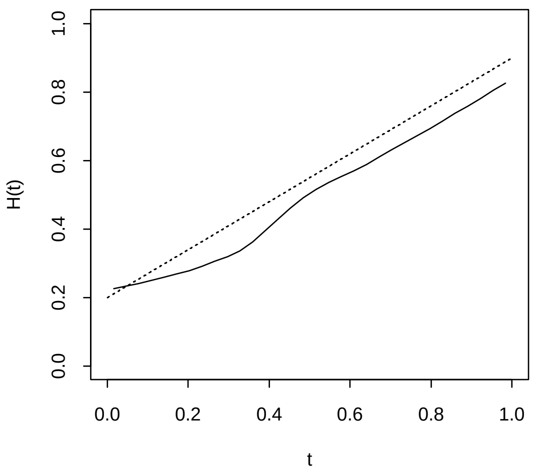


Figure 2.3(a)

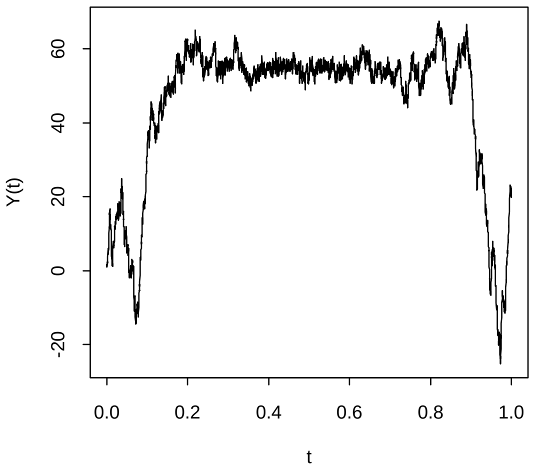


Figure 2.3(b)

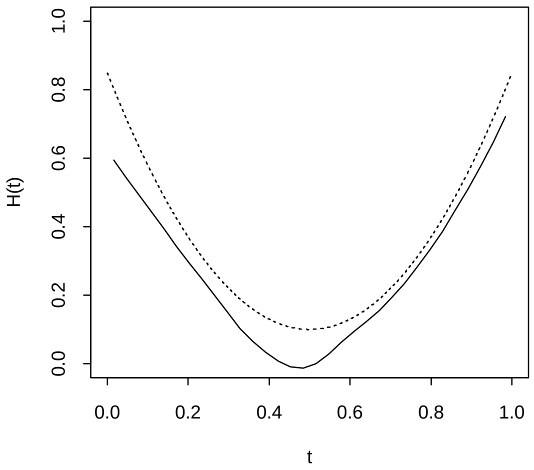


Figure 3(a)

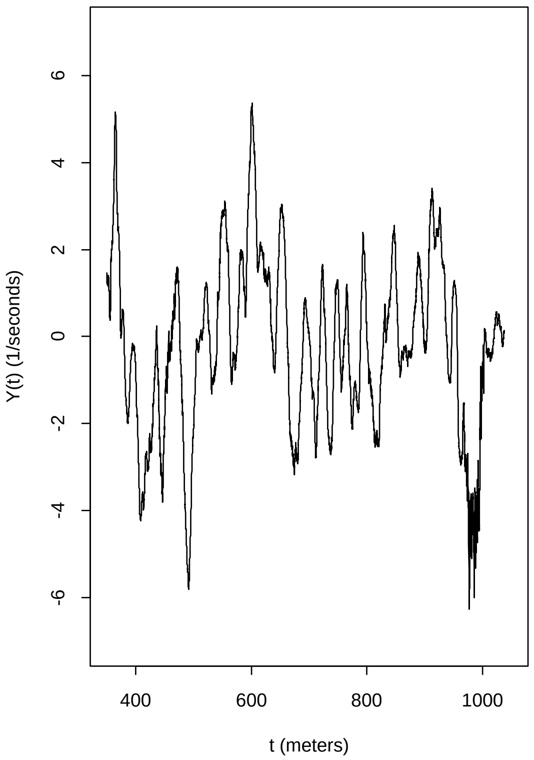


Figure 3(b)

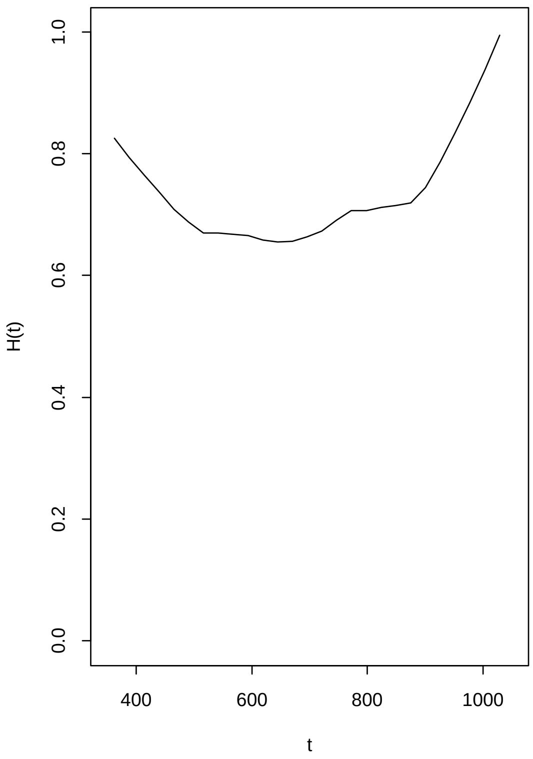


Figure 4(a)

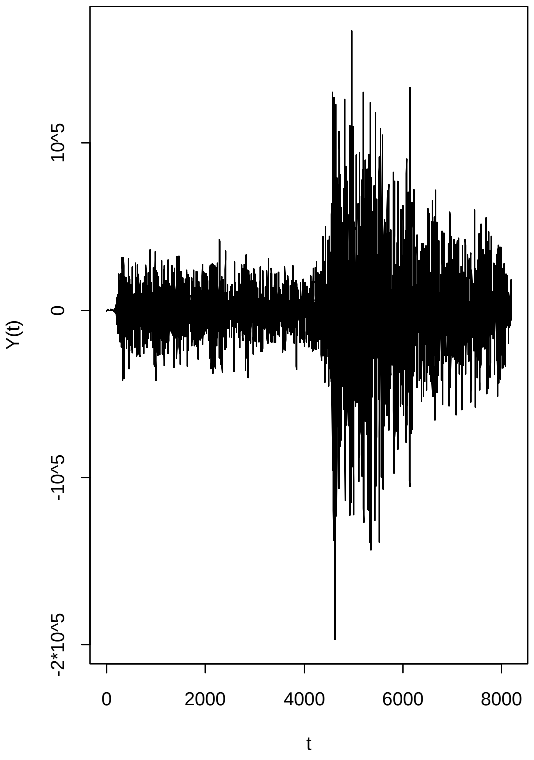
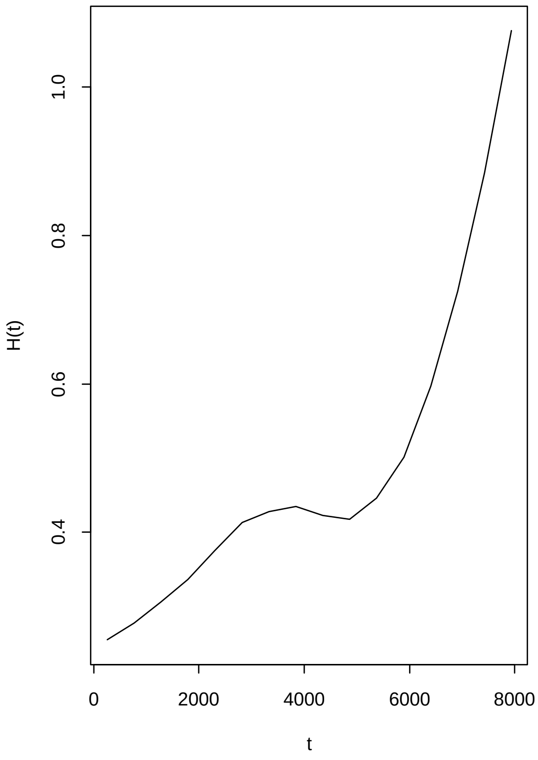


Figure 4(b)



References

Beran, J. (1994), *Statistics for Long Memory Processes*, New York: Chapman and Hall.

Brockwell, P. J., and Davis, R. A. (1991), *Time Series: Theory and Methods* (2nd ed.), New York: Springer-Verlag.

Buldyrev, S. V., Goldberger, A. L., Havlin, S., Peng, C-K., Stanley, H. E., Stanley, M. H. R., and Simons, M. (1993), "Fractal Landscapes and Molecular Evolution: Modeling the Moyosin Heavy Chain Gene Family," *Biophysical Journal*, 65, 2673–2679.

Chen, G., Hall, P., and Poskitt, D. S. (1995), "Periodogram-Based Estimators of Fractal Properties," *The Annals of Statistics*, 23, 1684–1711.

Cohen, A., Daubechies, I., Jawerth, B., and Vail, P. (1993), "Multiresolution Analysis, Wavelets and Fast Algorithms on an Interval," *Comptes Rendus des Séances de l'Académie des Sciences, Série I, Mathématique*, 316, 417–421.

Daubechies, I. (1992), *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia: SIAM.

Daubechies, I. (1994), "Two Recent Results on Wavelets: Wavelet Bases for the Interval, and Biorthogonal Wavelets Diagonalizing the Derivative Operator," in *Recent Advances in Wavelet Analysis*, eds. Schumaker and Webb, Boston: Academic Press, pp. 237–257.

Donoho, D. L., and Johnstone, I. M. (1994), "Ideal Spatial Adaptation by Wavelet Shrinkage," *Biometrika*, 81, 425–455.

Farge, M., Hunt, C. R., and Vassilicos, J. C. (1993), *Wavelets, Fractals and Fourier Transformations*, Oxford: Clarendon Press.

Flandrin, P., and Ganςlaves, P. (1994), "From Wavelets to Time-Scale Energy Distributions," in *Recent Advances in Wavelet Analysis*, eds. Schumaker and Webb, Boston: Academic Press, pp. 309–334.

Ganςlaves, P., and Flandrin, P. (1993), "Bilinear Time-Scale Analysis Applied to Local Scaling Exponents Estimation," in *Progress in Wavelet Analysis and Applications*, eds. Meyer and Roques, Paris: Frontieres, pp. 271–276.

Geweke, J., and Porter-Hudak, S. (1983), "The Estimation and Application of Long-Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–237.

Hall, P., Koul, H. L., and Turlach, B. A. (1997), "Note on Convergence Rates of Semiparametric Estimators of Dependence Index," *The Annals of Statistics*, 4, 1725–1739.

Hurst, H. E. (1951), "Long-Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, 116, 770–799.

Hwang, W. L., and Mallat, S. (1994), "Characterization of Self-Similar Multifractals with Wavelet Maxima," *Applied and Computational Harmonic Analysis*, 4, 316–328.

Kent, J. T., and Wood, A. T. A. (1997), "Estimating the Fractal Dimension of a Locally Self-Similar Gaussian Process by Using Increments," *Journal of the Royal Statistical Society B*, 3, 679–700.

Mandelbrot, B. B. (1983), *The Fractal Geometry of Nature* (updated and augmented ed.), New York: Freeman.

Mandelbrot, B. B., and van Ness, J. W. (1968), "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, 10, 422–437.

Mandelbrot, B. B., and Wallis, J. R. (1968), "Noah, Joseph and Operational Hydrology," *Water Resources Research*, 4, 909–918.

Mandelbrot, B. B., and Wallis, J. R. (1969), "Computer Experiments with Fractional Gaussian Noises," *Water Resources Research*, 5, 228–267.

Neuman, M. H., and von Sachs, R. (1997), "Wavelet Thresholding in Anisotropic Function Classes and Application to Adaptive Estimation of Evolutionary Spectra," *The Annals of Statistics*, 25, 38–76.

Ossadnik, S. M., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Mantegna, R. N., Peng, C-K., Simons, M., and Stanley, H. E. (1994), "Correlation Approach to Identify Coding Regions in DNA Sequences," *Biophysical Journal*, 67, 64–70.

Peng, C-K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., and Stanley, H. E. (1992), "Long-Range Correlation in Nucleotide Sequences," *Nature*, 356, 168–170.

Peng, C-K., Hausdorff, J. M., Mietus, J. E., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995a), "Fractals in Physiology Control: From Heart Beat to Gait," in *Fractals in Biology and Medicine*, eds. Nonnenmacher, Losa, and Weibel, Boston: Birkhauser Verlag.

Peng, C-K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995b), "Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series," *Chaos*, 5.

Percival, D. P., and Guttorp, P. (1994), "Long-Memory Processes, the Allen Variance and Wavelets," in *Wavelets in Geophysics*, eds. Foufoula-Georgiou and Kumar, New York: Academic Press, pp. 325-357.

Robinson, P. (1995), "Gaussian Semiparametric Estimation of Long Range Dependence," *The Annals of Statistics*, 23, 1630–1661.

Shumway, R. H. (1982), "Discriminant Analysis for Time Series," in *Handbook of Statistics 2*, eds. Krishnaiah and Kanal, Amsterdam: North Holland, pp. 1–46.

Shumway, R. H. (1988), *Applied Statistical Time Series Analysis*, New Jersey: Prentice-Hall.