ESTIMATION OF FRACTAL INDEX AND FRACTAL DIMENSION OF A GAUSSIAN PROCESS BY COUNTING THE NUMBER OF LEVEL CROSSINGS

By Andrey Feuerverger, Peter Hall and Andrew T. A. Wood

University of Toronto, CSIRO Division of Mathematics and Statistics and Australian National University

First version received November 1992

Abstract. The fractal index α and fractal dimension D of a Gaussian process are characteristics that describe the smoothness of the process. In principle, smoother processes have fewer crossings of a given level, and so level crossings might be employed to estimate α or D. However, the number of crossings of a level by a non-differentiable Gaussian process is either zero or infinity, with probability one, so that level crossings are not directly usable. Crossing counts may be rendered finite by smoothing the process. Therefore, we consider estimators that are based on comparing the sizes of the average numbers of crossings for a small, bounded number of different values of the smoothing bandwidth. The averaging here is over values of the level. Strikingly, we show that such estimators are consistent, as the size of the smoothing bandwidths shrinks to zero, if and only if the weight function in the definition of 'average' is constant. In this important case we derive the asymptotic bias and variance of the estimators, assuming only a non-parametric description of covariance, and describe the estimators' numerical properties. We also introduce a novel approach to generating Gaussian process data on a very fine grid.

Keywords. Bias; count; fractal dimension; fractal index; Gaussian process; level crossing; smoothing; upcrossing; variance.

1. INTRODUCTION

The smoothness of a stationary Gaussian process may be characterized by the behaviour of its variogram in the neighbourhood of the origin. Roughly speaking, if the variogram at points distant t apart converges to zero like $|t|^\alpha$ as $t \to 0$, then the process has $\alpha/2$ derivatives, measured in terms of Lipschitz-like behaviour. The value of α is sometimes termed the fractal index of the process, and indeed the fractal dimension D of sample paths is equal to $2 - \alpha/2$ (see, for example, Adler, 1981, Chapter 8). Thus, the parameter α is of intrinsic interest and importance. It has been used extensively as an index of roughness; see Berry and Hannay (1978), Coster and Chermant (1983), Dubuc *et al.* (1989), Mandelbrot *et al.* (1984), Thomas and Thomas (1988), Taylor and Taylor (1991) and the references therein. In the present paper we consider the problem of estimating α by counting the level crossings of Gaussian processes.

The value of α can never exceed 2, and should α be less than 2 then the

0143-9782/94/06 587–606 JOURNAL OF TIME SERIES ANALYSIS Vol. 15, No. 6

process can be taken to be continuous but not differentiable (see, for example, Leadbetter *et al.*, 1983, and Yaglom, 1987). In this instance the number of times that the process crosses any given level, over a specified interval, is either zero or infinity. Thus, counts of crossings for the original process are not meaningful. However, this problem disappears if we smooth the process by forming a suitable moving average or convolution over a narrow window, of width proportional to h say. The smoothed process is then differentiable, and the number of crossings is finite with probability one. If we were to specify the level in advance then we might be unlucky and select a level which the process did not cross. Thus, we might look instead at the average number of crossings of a large number of levels, the average being with respect to a weight function w, which may be almost arbitrary. We show that, strikingly, the case where the weight function is constant has special and important properties. This is the only case where α can be estimated consistently by examining the average number of level crossings for a finite number – as few as 2 – of bandwidths h.

Having derived this result we focus attention on the case of the constant w. Here it transpires that the average number of crossings of the smoothed process over an interval is proportional to the total variation of the process on that interval. We propose an estimation procedure based on (log-log) regression of average number of crossings versus bandwidth. We also compute the asymptotic bias and variance of our estimator of α, as $h \to 0$, and thereby derive the rate of convergence. (It should be stressed that we are considering this problem in a non-parametric setting – see Equation (1.1) below – rather than a parametric context. Thus, convergence rates are particularly interesting.) Finally we present a simulation study which confirms our theoretical analysis. That work describes a novel way of generating Gaussian processes on a very fine grid, with predetermined values of α.

There is no loss of generality in assuming that the underlying stationary Gaussian process $X = X(t)$ has zero mean and unit variance. We shall suppose in addition that the variogram $v(t) = E\{X(0) - X(t)\}^2$ satisfies $v(t) \sim 2c|t|^\alpha$ as $t \to 0$, where $c > 0$. Equivalently, the covariance function $\gamma(t) = E\{X(0)X(t)\}$ has the property

$$\gamma(t) = 1 - c|t|^\alpha + o(|t|^\alpha) \tag{1.1}$$

as $t \to 0$. We suppose that γ has two derivatives on $(0, \infty)$, bounded on (ε, ∞), for $\varepsilon > 0$, and mimicking property (1.1) near the origin:

$$\gamma''(t) = -\alpha(\alpha - 1)c|t|^{\alpha-2} + o(|t|^{\alpha-2}) \tag{1.2}$$

as $t \to 0$.

Our theoretical development of this problem is confined to the case where the process is observed over only a fixed interval, which we take without loss of generality to be $(0, 1)$. In particular, our asymptotic theory does not rely on the length T, say, of the recording interval diverging to infinity. Rather, the size h of the smoothing parameter tends to zero. We could have developed a theoretical account for the case of increasingly large T, but that

would have distracted attention from the very important fact that consistant estimation of α is possible using only a very small trace of the process X.

In practice, X would usually not be recorded in the continuum, but at a sequence of discrete lattice points. Often the recording device itself imposes a degree of smoothing on the data, through some sort of inbuilt filter. Practical choice of the bandwidth h is determined by the level of gridding or filtering of the raw data, which ideally should be an order of magnitude less than the value of h.

For processes other than Gaussian ones the classical relationship $D = 2 - \alpha/2$ between fractal dimension and fractal index does not necessarily hold. In particular, if $X = |Z|^\upsilon$ and Z is a stationary Gaussian process then the relationship is valid if and only if $\upsilon \geq 1/2$. See Hall and Roy (1994) for further discussion. Variants of our results may be derived for classes of non-Gaussian processes by taking them to be functions of Gaussian ones. However, the proofs lack the elegance and simplicity of those given here.

2. DEFINITION OF ESTIMATOR, AND BASIC PROPERTIES

Let X denote a stationary Gaussian process whose covariance function γ satisfies (1.1). Then the number of crossings of any level by the process X, over any finite interval, has infinite mean. See, for example, Leadbetter *et al.* (1983, p. 216 ff). To render the number of crossings finite we might smooth X using a moving average or linear filter, generating a new process Y given by

$$Y(t) = h^{-1} \int K\left(\frac{u}{h}\right) X(t + u)\, du \qquad -\infty < t < \infty.$$

Here, $h > 0$ denotes bandwidth, and we require that the kernel K be differentiable, be compactly supported and satisfy $\int (|K| + |K'|)\, du < \infty$, $\int K\, du = 1$ and $\int uK(u)\, du = 0$. By letting $h \to 0$ we recover the process X:

$$\lim_{h \to 0} Y(t) = X(t) \qquad -\infty < t < \infty$$

with probability one. Although the sample paths of the original process X are not differentiable, those of Y are, and in fact

$$0 < \lambda \equiv E\{Y'(0)\}^2 = -\int\int K(u_1) K(u_2) \gamma''\{h(u_1 - u_2)\}\, du_1\, du_2 < \infty; \quad (2.1)$$

see the Appendix. Therefore, by Rice's formula (see Theorem 7.3.2 of Leadbetter *et al.*), we have that the random function

$$N(u) = \#\{t \in (0, 1): Y(t) = u\} \qquad -\infty < u < \infty$$

is well defined and finite with probability one. We shall base our estimator of α on the behaviour of this quantity as $h \to 0$. Almost identical results may be

proved for upcrossings or downcrossings, rather than the crossings considered here.

Of course, $N(u) = 0$ if $|u|$ is too large, and so working with $N(u)$ alone could result in vacuous conclusions. We therefore suggest calculating a version of N that is 'averaged', in a sense, over all possible u's. To this end, let $w \geqslant 0$ denote a weight function satisfying

$$w(u) \leqslant D_1(1 + |u|)^{D_2} \qquad -\infty < u < \infty$$

for arbitrary constants D_1, $D_2 > 0$. This growth condition ensures that

$$M = \int N(u)w(u)\,du \tag{2.2}$$

is well defined and finite with probability one, and in fact has all moments finite. (Unqualified integrals are over $(-\infty, \infty)$.) It may also be shown that

$$M = \int_0^1 |Y'(t)| w\{Y(t)\}\,dt \tag{2.3}$$

and from this formula one may prove that, as $h \to 0$,

$$E(M) \sim \lambda^{1/2} E[w\{X(0)\}]E|Z_0| \tag{2.4}$$

where Z_0 denotes a standard normal random variable and λ is as in (2.1). Proofs of these three assertions will be given later in this section. It may further be proved from (1.2) and (2.1), on integrating by parts, that as $h \to 0$

$$\lambda \sim C_1 h^{\alpha-2} \tag{2.5}$$

where

$$C_1 = c \iint K'(u_1)K'(u_2)|u_1 - u_2|^{\alpha}\,du_1\,du_2$$

and c, α are as in (1.1). Therefore,

$$E(M) \sim C_2 h^{(\alpha-2)/2} \tag{2.6}$$

where $C_2 = C_1^{1/2} E[w\{X(0)\}]E|Z_0|$.

Formula (2.6) offers the possibility of estimating $\alpha/2 - 1$, and hence α and $D = 2 - \alpha/2$, from the slope of a linear regression of $\log\{E(M)\}$ on $\log h$ for small h. Since $E(M)$ is not known we shall have to replace it by M in any practical estimator. For this approach to be feasible, i.e. consistent, using only a bounded number of different bandwidths, it is essential that the measure of variation

$$\frac{\operatorname{var} M}{(EM)^2} \to 0 \tag{2.7}$$

as $h \to 0$. We shall show shortly that this result holds if and only if $w \equiv \text{const}$. Without loss of generality, $w \equiv 1$.

To confirm consistency of the estimator based on only a finite number of

bandwidths, note that by (2.7), and a Taylor expansion argument,

$$\log\{M(h)\} \equiv \log\{EM(h)\} + \log\left\{1 + \frac{M(h) - EM(h)}{EM(h)}\right\}$$
$$= \log\{EM(h)\} + o_p(1)$$

and in view of (2.6),

$$\log\{EM(h)\} = \tfrac{1}{2}(\alpha - 2)\log h + \log C_2 + o(1).$$

Therefore

$$\log\{M(h)\} = \tfrac{1}{2}(\alpha - 2)\log h + \log C_2 + o_p(1).$$

Hence for each fixed $k > 1$, the regression estimator defined by

$$\hat{\alpha} = \hat{\alpha}(h_1, \ldots, h_k) = 2\left[\sum_{j=1}^{k}(x_j - \bar{x})\log\{M(h_j)\}\right]\left\{\sum_{j=1}^{k}(x_j - \bar{x})^2\right\}^{-1} + 2 \quad (2.8)$$

where $x_j = \log h_j$ and $\bar{x} = k^{-1}\sum x_j$, is consistent for α provided only that $h_1, \ldots, h_k \to 0$ and $\sum(x_j - \bar{x})^2$ or, equivalently, $\sum\sum|\log h_i - \log h_j|$, is bounded away from zero. Here, consistency means that for some $\varepsilon > 0$ and any $\eta > 0$,

$$\lim_{h_1,\ldots,h_k \to 0, \sum\sum|\log h_i - \log h_j| > \varepsilon} P(|\hat{\alpha} - \alpha| > \eta) \to 0.$$

When $w \equiv 1$ we see from (2.3) that

$$M = M(h) = \int |Y'(t)|\, dt$$

representing the total variation of the process Y over the interval $(0, 1)$. We were initially surprised, as might be the reader, that the crucial condition (2.7) holds only if w is constant. It is instructive to consider first an intuitive explanation for this result. The starting point is formula (2.3), which integral we might approximate by a series as follows:

$$M \approx n^{-1}\sum\left|Y'\left(\frac{i}{n}\right)\right| w\left\{Y\left(\frac{i}{n}\right)\right\} \quad (2.9)$$

Now, the Gaussian random variables $Y'(t)$, $0 < t < 1$, are asymptotically independent as $h \to 0$, since for any fixed $t_1 \neq t_2$ the correlation

$$\operatorname{corrln}\{Y'(t_1), Y'(t_2)\} \sim \frac{-\gamma''(t_1 - t_2)}{\lambda} \to 0$$

as $h \to 0$, by (2.5). Therefore, if w is constant then the series at (2.9) is approximately an average of independent random variables, from which it is to be expected by the law of large numbers that (2.7) will hold. However, should w not be constant, then since

$$\operatorname{corrln}\{Y(t_1), Y(t_2)\} \to \gamma(t_1 - t_2)$$

as $h \to 0$, the approximate independence of the summands in (2.9) is no longer valid, and in fact (2.7) will fail.

We close this section by sketching proofs of the following: all moments of M are finite, (2.3) and (2.4) hold and $w \equiv \text{const}$ is necessary and sufficient for (2.7). To show that M has finite moments, we shall first assume the equivalence of (2.2) and (2.3) and work with the latter. Given pathwise continuity of $Y(t)$ and $Y'(t)$ (see (A1) in the Appendix) we have, from (2.3),

$$|M|^r \leqslant \sup_{t\in[0,1]} |Y'(t)|^r \times D_1^r\{1 + \sup_{t\in[0,1]} |Y(t)|\}^{rD_2}$$

under the growth condition for w. The finiteness of $E|M|^r$ for each $r > 0$ now follows from the Cauchy–Schwarz inequality, and the fact that the random variables $\sup_{t\in[0,1]} |Y(t)|$ and $\sup_{t\in[0,1]} |Y'(t)|$ have lighter-than-exponential tails (see Section 12.2 of Leadbette *et al.*, 1983). Note that the existence of the first two moments of M is required in (2.7).

An heuristic proof of (2.3) runs as follows. Observe that

$$\begin{aligned} N(u) &= \lim_{n\to\infty} \sum_i I\left\{ Y\left(\frac{i}{n}\right) < u < Y\left(\frac{i+1}{n}\right) \text{ or } Y\left(\frac{i}{n}\right) > u > Y\left(\frac{i+1}{n}\right)\right\} \\ &= \lim_{n\to\infty} \sum_i I\left\{ Y\left(\frac{i}{n}\right) < u < Y\left(\frac{i}{n}\right) + n^{-1}Y'\left(\frac{i}{n}\right) \right. \\ &\qquad \left. \text{or } Y\left(\frac{i}{n}\right) > u > Y\left(\frac{i}{n}\right) + n^{-1}Y'\left(\frac{i}{n}\right)\right\} \end{aligned}$$

whence

$$\begin{aligned} M &= \int N(u)w(u)\,du \\ &= \lim_{n\to\infty} \sum_i \int I\left\{ Y\left(\frac{i}{n}\right) < u < Y\left(\frac{i}{n}\right) + n^{-1}Y'\left(\frac{i}{n}\right) \right. \\ &\qquad \left. \text{or } Y\left(\frac{i}{n}\right) > u > Y\left(\frac{i}{n}\right) + n^{-1}Y'\left(\frac{i}{n}\right)\right\} w(u)\,du \\ &= \lim_{n\to\infty} \sum_i I\left\{ Y'\left(\frac{i}{n}\right) > 0\right\} \int_{Y(i/n)}^{Y(i/n)+n^{-1}Y'(i/n)} w(u)\,du \\ &\quad + I\left\{ Y'\left(\frac{i}{n}\right) < 0\right\} \int_{Y(i/n)+n^{-1}Y'(i/n)}^{Y(i/n)} w(u)\,du \\ &= \lim_{n\to\infty} \sum_i n^{-1}\left| Y'\left(\frac{i}{n}\right)\right| w\left\{ Y\left(\frac{i}{n}\right)\right\} \\ &= \int_0^1 |Y'(t)| w\{Y(t)\}\,dt. \end{aligned}$$

A rigorous proof is longer but similar.

Formula (2.3) implies that

$$\operatorname{var} M = \int_0^1\int_0^1 \operatorname{cov}(U_1, U_2)\, dt_1\, dt_2$$

where $U(t) = w\{Y(t)\}|Y'(t)|$ and $U_i = U(t_i)$. (This result may also be derived directly from (2.2) using the arguments of Cramér and Leadbetter (1967, p. 202 ff).) To check our claim that condition (2.7) is equivalent to $w \equiv \text{const}$, put $\lambda = \operatorname{var}\{Y'(t)\}$ and $Z(t) = \lambda^{-1/2}Y'(t)$ and observe that

$$\lambda^{-1} \operatorname{var} M = \int_0^1\int_0^1 \sigma(t_1, t_2)\, dt_1\, dt_2$$

where

$$\sigma(t_1, t_2) = \operatorname{cov}[w\{Y(t_1)\}|Z(t_1)|, w\{Y(t_2)\}|Z(t_2)|].$$

As $h \to 0$, and for $t_1 \neq t_2$, the bivariate joint limiting distribution of $|Z(t_1)|$ and $|Z(t_2)|$ conditional on $Y(t_1)$ and $Y(t_2)$ is that of $|Z_1|$ and $|Z_2|$, where Z_1 and Z_2 are independent standard normal random variables. Arguing thus, rigorously from the formula for the Gaussian joint density of $Y(t_1)$, $Y(t_2)$, $Z(t_1)$ and $Z(t_2)$, we may prove that for each fixed $t_1 \neq t_2$,

$$\sigma(t_1, t_2) \to \operatorname{cov}[w\{X(t_1)\}, w\{X(t_2)\}](E|Z_0|)^2$$

as $h \to 0$. In fact, for any t_1 and t_2,

$$\begin{aligned} |\sigma(t_1, t_2)| &\leq E[w\{Y(0)\}^2 Z(0)^2] \\ &\leq (E[w\{Y(0)\}^4]E\{Z(0)\}^4)^{1/2} \\ &\to (E[w\{X(0)\}^4]E(Z_0^4))^{1/2} \end{aligned}$$

as $h \to 0$. Therefore, by the dominated convergence theorem,

$$\begin{aligned} \lambda^{-1} \operatorname{var} M &\to \int_0^1\int_0^1 \operatorname{cov}[w\{X(t_1)\}, w\{X(t_2)\}]\, dt_1\, dt_2 (E|Z_0|)^2 \\ &= \operatorname{var}\left[\int_0^1 w\{X(t)\}\, dt\right](E|Z_0|)^2. \end{aligned} \tag{2.10}$$

Next, using (2.3), observe that

$$\begin{aligned} \lambda^{-1/2} E(M) &= \int_0^1 E[w\{Y(t)\}|Z'(t)|]\, dt \\ &= E[w\{Y(0)\}|Z'(0)|] \\ &\to E[w\{X(0)\}]E|Z_0| \end{aligned} \tag{2.11}$$

which incidentally verifies (2.4). By (2.10) and (2.11),

$$\frac{\operatorname{var} M}{(EM)^2} \to \frac{\operatorname{var}\left[\int_0^1 w\{X(t)\}\, dt\right]}{[Ew\{X(0)\}]^2}.$$

The right-hand side equals zero if and only if w is constant almost everywhere!

3. BIAS AND VARIANCE OF THE ESTIMATOR

Define $\hat{\alpha}$ as in (2.8), with k fixed and bandwidths $h_1, \ldots, h_k$ converging to zero. For the sake of definiteness we shall take $h_j = a_j h$, where $a_1, \ldots, a_k$ denote fixed positive constants, and $h \to 0$. In this instance the quantities

$$y_j = x_j - \bar{x} = \log a_j - k^{-1} \sum \log a_i$$

do not depend on h, and this simplifies notation. As suggested in Section 2, we take $w \equiv 1$ when defining M, so that $M = M(h) = \int N$. Thus,

$$\hat{\alpha} = 2\left(\sum_{j=1}^{k} y_j^2\right)^{-1}\left[\sum_{j=1}^{k} y_j \log\{M(a_j h)\}\right] + 2. \tag{3.1}$$

To elucidate the first-order asymptotic behaviour of this estimator, put $m(h) = EM(h)$ and Taylor-expand in (3.1), obtaining

$$\hat{\alpha} = A + B + O_p[m(h)^{-2} \operatorname{var}\{M(h)\}] \tag{3.2}$$

where

$$A = 2\left(\sum_{j=1}^{k} y_j^2\right)^{-1} \sum_{j=1}^{k} \frac{y_j\{M(a_j h) - m(a_j h)\}}{m(a_j h)}$$

denotes the primary contribution to the error about the mean of $\hat{\alpha}$, and

$$B = 2 + 2\left(\sum_{j=1}^{k} y_j^2\right)^{-1} \sum_{j=1}^{k} y_j \log\{m(a_j h)\}$$

is the main contribution to the mean. Note that $E(A) = 0$ and that B is non-random.

To describe the variance of A, and hence that of $\hat{\alpha}$, define

$$L = L(h) = \int_0^\infty \left\{\iint K'(u_1) K'(u_2) |t + u_1 - u_2|^\alpha du_1 du_2\right\}^2 dt$$

for $0 < \alpha < 3/2$ (for α in this range, $L < \infty$), put

$$H = H(h) = \begin{cases} h^{2(2-\alpha)} & \text{if } \alpha > 3/2 \\ h \log h^{-1} & \text{if } \alpha = 3/2 \\ h & \text{if } \alpha < 3/2 \end{cases}$$

and

$$C = \begin{cases} (3/C_1)^2 \int_0^1 (1-t)\{\gamma''(t)\}^2 dt & \text{if } \alpha > 3/2 \\ (9c/4C_1)^2 & \text{if } \alpha = 3/2 \\ (3c/C_1)^2 L & \text{if } \alpha < 3/2 \end{cases}$$

where C_1 is as in (2.5). In this notation it may be proved that

$$m(h)^{-2}\operatorname{var}\{M(h)\} \sim CH \tag{3.3}$$

as $h \to 0$, and similarly, for a constant $\Delta = \Delta(a_1, \ldots, a_k)$, that

$$\operatorname{var}(A) \sim \Delta H. \tag{3.4}$$

We shall shortly derive formula (3.3). The extension to (3.4) may be established by a notationally more tedious, but mechanically very similar, argument. We shall also prove that if condition (1.1) is specialized to

$$\gamma(t) = 1 - c|t|^\alpha - d|t|^\beta + o(|t|^\beta) \tag{3.5}$$

where β, α are constants and $\alpha < \beta < \infty$, then

$$\log\{m(h)\} = \frac{1}{2}\log\left(\frac{2C_1}{\pi}\right) + \frac{1}{2}(\alpha - 2)\log h + C_1^{-1}C_3 h^{\beta-\alpha} + o(h^{\beta-\alpha}) \tag{3.6}$$

where

$$C_3 = \iint K'(u_1)K'(u_2)|u_1 - u_2|^\beta du_1 du_2.$$

It follows that

$$B = \alpha + C_4 h^{\beta-\alpha} + o(h^{\beta-\alpha}) \tag{3.7}$$

where

$$C_4 = 2C_1^{-1}C_3\left(\sum_{j=1}^k y_j^2\right)^{-1}\sum_{j=1}^k a_j^{\beta-\alpha}.$$

Combining (3.2), (3.4) and (3.7) we may deduce that

$$\hat{\alpha} - \alpha = C_4 h^{\beta-\alpha} + \Delta^{1/2}H^{1/2}V + o_p(h^{\beta-\alpha} + H^{1/2})$$

where the random variable V has unit variance. Thus, $\hat{\alpha}$ is consistent for α with asymptotic bias of size $h^{\beta-\alpha}$ and variance of size H. The error of $\hat{\alpha}$ about its mean will thus converge to zero more slowly when α approaches 2. This is a common feature of estimators of fractal dimension; see, for example, Constantine and Hall (1994) and Hall and Wood (1993).

We close by proving (3.3) and (3.6). Let V_1, V_2 have a bivariate normal distribution with zero mean, unit variance and correlation coefficient ρ. Then as $\rho \to 0$,

$$\operatorname{cov}(|V_1|, |V_2|) = 9\pi^{-1}\rho^2 + O(\rho^4). \tag{3.8}$$

Let r denote the covariance function of the process Y, and put $\lambda = E\{Y'(0)^2\} = -r''(0)$. It is shown in (A5) of the Appendix that

$$\operatorname{cov}\{Y'(0), Y'(t)\} = -r''(t) = -\iint K(u_1)K(u_2)\gamma''\{t + h(u_1 - u_2)\}\, du_1\, du_2.$$

In this notation, and in view of (3.8),

$$\begin{aligned}\operatorname{var} M &= \int_0^1\int_0^1 \operatorname{cov}\{|Y'(t_1)|, |Y'(t_2)|\}\,dt_1\,dt_2 \\ &= 9\pi^{-1}\{EY'(0)^2\}^{-1}\int_0^1\int_0^1 r''(t_1 - t_2)^2\,dt_1\,dt_2 \\ &\quad + O\left[\{EY'(0)^2\}^{-2}\int_0^1\int_0^1 r''(t_1 - t_2)^4\,dt_1\,dt_2\right] \\ &= 18\pi^{-1}\lambda^{-1}\int_0^1 (1 - t)\{r''(t)\}^2\,dt + O\left\{\lambda^{-2}\int_0^1 (1 - t)r''(t)^4\,dt\right\}. \qquad (3.9)\end{aligned}$$

Now when $3/2 < \alpha < 2$,

$$\begin{aligned}J &\equiv \int_0^1 (1 - t)\{r''(t)\}^2\,dt \\ &= h^{-4}\int_0^1 (1 - t)\left[\iint K'(u_1)K'(u_2)\gamma\{t + h(u_1 - u_2)\}\,du_1\,du_2\right]^2 dt \\ &\sim h^{-4}\int_0^1 (1 - t)\left[\iint K'(u_1)K'(u_2)\{\gamma(t) + h(u_1 - u_2)\gamma'(t) \right. \\ &\qquad \left. + \frac{1}{2}h^2(u_1 - u_2)^2\gamma''(t)\right\}\,du_1\,du_2\Bigg]^2 dt \\ &= \int_0^1 (1 - t)\gamma''(t)^2\,dt. \qquad (3.10)\end{aligned}$$

In the case $0 < \alpha \leq 3/2$, let $0 < \delta \leq 1$ and write

$$\begin{aligned}J &= \int_0^{1/h} h^{-3}(1 - ht)\left[\iint K'(u_1)K'(u_2)\gamma\{h(t + u_1 - u_2)\}\,du_1\,du_2\right]^2 dt \\ &= \int_0^{\delta/h} + \int_{\delta/h}^{1/h} \\ &= J_{1\delta} + J_{2\delta} \qquad (3.11)\end{aligned}$$

say. Given any $\varepsilon > 0$ we may choose $\delta > 0$ so small that for all sufficiently small h, and some $\theta = \theta(h) \in (-\varepsilon, \varepsilon)$, we have

$$J_{1\delta} = (1 + \theta)c^2h^{2\alpha - 3}\int_0^{\delta/h}\left\{\iint K'(u_1)K'(u_2)|t + u_1 - u_2|^\alpha\,du_1\,du_2\right\}^2 dt. \qquad (3.12)$$

Also, when $0 < \alpha < 2$, $J_{2\delta} = 0(1)$. If $\alpha = 3/2$,

$$\int_0^{\delta/h}\left\{\iint K'(u_1)K'(u_2)|t + u_1 - u_2|^{3/2}\,du_1\,du_2\right\}^2 dt$$

$$\sim \int_1^{a/h} t^3\left\{\iint K'(u_1)K'(u_2)|1 + t^{-1}(u_1 - u_2)|^{3/2}\,du_1\,du_2\right\}^2 dt$$

$$\sim \left\{\frac{1}{2}\frac{3}{2}\left(\frac{3}{2} - 1\right)\right\}^2\left(\int_1^{\delta/h} t^{-1}dt\right)\left\{\iint K'(u_1)K'(u_2)(u_1 - u_2)^2\,du_1\,du_2\right\}^2$$

$$\sim \frac{9}{16}\log h^{-1}$$

and so

$$J \sim \frac{9c^2}{16}\log h^{-1}. \tag{3.13}$$

If $0 < \alpha < 3/2$ then we may show from (3.11) and (3.12) that

$$J \sim c^2 L h^{2\alpha-3}. \tag{3.14}$$

Similar arguments may be used to approximate the integral in the second, remainder term in (3.9). Arguing thus it may be shown that the remainder term in (3.9) is of smaller order than the first term. Hence, combining (3.9), (3.10), (3.13) and (3.14), we obtain

$$\operatorname{var} M \sim 18\pi^{-1}\lambda^{-1} \times \begin{cases} \int_0^1 (1-t)\gamma''(t)^2 dt & \text{if } \alpha > 3/2 \\ (9c^2/16)\log h^{-1} & \text{if } \alpha = 3/2 \\ c^2 L h^{2\alpha-3} & \text{if } \alpha < 3/2 \end{cases}$$

We know from (2.4) that $EM \sim (2/\pi)^{1/2}\lambda^{1/2}$, and so

$$\frac{\operatorname{var} M}{(EM)^2} \sim 9C_1^2 \times \begin{cases} h^{2(2-\alpha)}\int_0^1 (1-t)\gamma''(t)^2 & \text{if } \alpha > 3/2 \\ (9c^2/16)h\log h^{-1} & \text{if } \alpha = 3/2 \\ c^2 L h & \text{if } \alpha < 3/2 \end{cases}$$

which proves (3.3).

To establish (3.6), observe that since $E(M) = \lambda^{1/2}(2/\pi)^{1/2}$ and

$$\lambda = -h^{-2}\iint K'(u_1)K'(u_2)\gamma\{h(u_1 - u_2)\}\,du_1\,du_2$$

then under condition (3.5),

$$E(M) = \left(\frac{2}{\pi}\right)^{1/2}\{C_1 h^{\alpha-2} + C_3 h^{\beta-2} + o(h^{\beta-2})\}^{1/2}$$

whence follows the desired result.

4. THE SIMULATION METHOD

Suppose that we wish to generate a random vector

$$U = \left[X(0), X\left(\frac{1}{n}\right), \ldots, X\left(\frac{n-1}{n}\right)\right]^{\mathrm{T}}$$

from a zero-mean, stationary Gaussian process $\{X(t)\}$ with given covariance function γ. Then $U \sim N_n(0, G)$, where

$$G = \begin{bmatrix} \gamma(0) & \gamma(1/n) & \dots & \gamma\{(n-1)/n\} \\ \gamma(1/n) & \gamma(0) & \dots & \gamma\{(n-2)/n\} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma\{(n-1)/n\} & \gamma\{(n-2)/n\} & \dots & \gamma(0) \end{bmatrix} \tag{4.1}$$

One possibility is to base a simulation method on the Cholesky factorization of G, but in the present setting this has the disadvantage that storage requirements are $O(n^2)$. An alternative method is described below. It is based on properties of circulant matrices and the fast fourier transform (FFT). See Brockwell and Davis (1987) for details of both.

The simulation scheme may be outlined as follows.

STEP 1. Embed G in a circulant covariance matrix $C(m \times m)$, where $m = 2^g$ for some integer g, and $m \geqslant 2n$.

STEP 2. Use the FFT twice, as indicated below, to generate a random vector $Y = (Y_0, Y_1, \dots, Y_{m-1})^{\mathrm{T}} \sim \mathrm{N}_m(0, C)$. Then, with appropriate construction of C in Step 1, $U = (Y_0, Y_1, \dots, Y_{n-1})^{\mathrm{T}} \sim \mathrm{N}_n(0, G)$.

In Step 1, we take C to be the circulant matrix

$$C = \begin{bmatrix} C_0 & C_1 & \dots & C_{m-1} \\ C_{m-1} & C_0 & \dots & C_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ C_1 & C_2 & \dots & C_0 \end{bmatrix}$$

where

$$C_j = \begin{cases} \gamma\left(\dfrac{j}{n}\right) & \text{if } 0 \leqslant j \leqslant \dfrac{m}{2} \\ \gamma\left(\dfrac{m-j}{n}\right) & \text{if } \dfrac{m}{2} < j \leqslant m-1. \end{cases} \tag{4.2}$$

Note that if $m \geqslant 2n$, then the $n \times n$ submatrix in the top left-hand corner of C is equal to G in (4.1).

Since C is a circulant matrix, we may use Proposition 4.5.1 of Brockwell and Davis (1987) and write $C = Q\Lambda Q^*$, where $\Lambda = \mathrm{diag}\{\lambda_0, \lambda_1, \dots, \lambda_{m-1}\}$ is the diagonal matrix of eigenvalues of C, the λ_k are given by (4.6) below, Q is the matrix whose columns consist of the (left) eigenvectors of C, and is given by

$$Q = \{q_{jk}: 0 \leqslant j, k \leqslant m-1\} \qquad \text{where } q_{jk} = m^{-1/2} \exp\left(\frac{-2\pi ijk}{m}\right)$$

and Q^* is the conjugate transpose of Q.

It is easily established that if C is real and symmetric, then (i) Λ is real and (ii) if, in addition,

$$\lambda_j \geq 0 \text{ for } 0 \leq j \leq m-1 \tag{4.3}$$

$Q\Lambda^{1/2}Q^*$ is real and symmetric, where $\Lambda^{1/2} = \text{diag}\{\lambda_0^{1/2}, \lambda_1^{1/2}, \ldots, \lambda_{m-1}^{1/2}\}$. It follows that if $Z = (Z_0, Z_1, \ldots, Z_{m-1})^T$ is a vector whose components are independent N(0, 1) random variables, and (4.3) holds, then $Q\Lambda^{1/2}Q^*Z$ is a real random vector with an $N_m(0, C)$ distribution.

To generate the complex normal random vector Q^*Z directly, we generate independent N(0, 1) random variables R_0, R_1, W_{jk}, $j = 1, 2$ and $k = 1, 2, \ldots, m/2 - 1$, and put

$$S_0 = R_0,\ T_0 \equiv 0,\ S_{m/2} = R_1,\ T_{m/2} \equiv 0,\ S_k = S_{m-k} = 2^{-1/2}W_{1k},$$
$$T_k = -T_{m-k} = 2^{-1/2}W_{2k},\ 1 \leq k \leq m/2 - 1. \tag{4.4}$$

Then, using the orthogonality properties of the relevant trigonometric functions, it is straightforward to establish that

$$Q^*Z = S + iT \text{ in distribution}$$

where $S = (S_0, S_1, \ldots, S_{m-1})^T$ and $T = (T_0, T_1, \ldots, T_{m-1})^T$.

Define the discrete Fourier transform of the sequence a of complex numbers $a_0, a_1, \ldots, a_{m-1}$ by

$$d_a(k) = \sum_{j=0}^{m-1} a_j \exp\left(\frac{-2\pi ijk}{m}\right) \qquad k = 0, 1, \ldots, m-1. \tag{4.5}$$

If m is of the form 2^g for some integer g, then (4.5) can be calculated very efficiently using the FFT algorithm. Using Proposition 4.5.1 of Brockwell and Davis (1987) again, it is seen that the eigenvalues of C are given by

$$\lambda_k = d_C(k) \qquad k = 0, 1, \ldots, m-1 \tag{4.6}$$

where the C-sequence $C_0, C_1, \ldots, C_{m-1}$ is defined as in (4.2). Also, assuming that (4.3) holds, note that $Q\Lambda^{1/2}Q^*Z$ is equal in distribution to $m^{1/2}Qa$, where $a = (a_0, a_1, \ldots, a_{m-1})^T$ and

$$a_j = \frac{\lambda_j^{1/2}(S_j + iT_j)}{m^{1/2}} \qquad j = 0, 1, \ldots, m-1. \tag{4.7}$$

Finally, observe that premultiplying the vector a by the matrix $m^{1/2}Q$ is precisely equivalent to calculating the discrete Fourier transform d_a in (4.5).

We may summarize Step 2 of the procedure as follows. First, apply the FFT to the C-sequence in (4.2) to obtain the λ_j in (4.6), generate the S_j and T_j according to (4.4), and then calculate the a_j in (4.7). Second, apply the FFT to the a-sequence and put $U = \{d_a(0), d_a(1), \ldots, d_a(n-1)\}^T$. Then $U \sim N_n(0, G)$.

The only thing which can go wrong with the above procedure is that (4.3) is violated with the given $m = 2^g$, i.e. some of the λ_j are negative. We now give conditions which ensure that (4.3) is satisfied when m is sufficiently large.

PROPOSITION 4.1. Let

$$f(\omega) = \gamma(0) + 2\sum_{h=1}^{\infty} \gamma\left(\frac{h}{n}\right) \cos(2\pi h\omega) \qquad \omega \in [0, 1]$$

denote the spectral density of the process $\{X(t): t = 0, \pm n^{-1}, \pm 2n^{-1}, \ldots\}$. Suppose that $f(\omega)$ is strictly positive on $[0, 1]$, and that the covariance function is absolutely summable, i.e.

$$\sum_{h=0}^{\infty} \left|\gamma\left(\frac{h}{n}\right)\right| < \infty.$$

Then the matrix (4.2) is positive definite for all m sufficiently large. Note: $f(\omega)$ is a wrapping of the spectral density of the process $\{X(t): -\infty < t < \infty\}$.

PROOF OF PROPOSITION 4.1. Suppose that $f(\omega)$ has a minimum $\eta > 0$ on $[0, 1]$. Absolute summability of the covariance function implies we can choose an m_0 such that

$$\sum_{h=m/2}^{\infty} \left|\gamma\left(\frac{h}{n}\right)\right| \leqslant \frac{\eta}{4}$$

when $m \geqslant m_0$. Now the eigenvalues of C in (4.2) may be written

$$\lambda_m(j) = \gamma(0) + (-1)^{m/2}\gamma\left(\frac{m}{2n}\right) + \sum_{h=1}^{m/2-1} \gamma\left(\frac{h}{n}\right) \cos\left(\frac{2\pi hj}{m}\right)$$

when m is even; and when m is odd,

$$\lambda_m(j) = \gamma(0) + 2\sum_{h=1}^{(m-1)/2} \gamma\left(\frac{h}{n}\right) \cos\left(\frac{2\pi hj}{m}\right).$$

Thus, for $m \geqslant m_0$,

$$\max_{j=0,1,\ldots,m-1} \left|f\left(\frac{j}{m}\right) - \lambda_m(j)\right| \leqslant 2 \sum_{h=[(m-1)/2]+1}^{\infty} \left|\gamma\left(\frac{h}{n}\right)\right| \leqslant \frac{\eta}{2}$$

where [.] denotes integer part; and therefore

$$\min_{j=0,1,\ldots,m-1} \lambda_m(j) \geqslant \inf_{\omega\in[0,1]} f(\omega) - \frac{\eta}{2} = \frac{\eta}{2} > 0. \qquad \square$$

In Table I, minimum values of g are given for which (4.3) is satisfied with the covariance function given by

$$\gamma(t) = \exp(-c|t|^{\alpha}) \tag{4.8}$$

and $m = 2^g \geqslant 2n$.

TABLE I

VALUES OF g REQUIRED FOR POSITIVITY (SEE (4.3))

	n				
	100	250	500	1000	5000
$\alpha \leq 1.0$, all c	8	9	10	11	14
$c = 0.1$, $1.25 \leq \alpha \leq 1.75$	12	13	14	15	18
$c = 0.1$, $\alpha = 1.99$	12	13	14	15	17
$c = 1$, $\alpha = 1.25$	9	10	11	12	15
$c = 1$, $\alpha = 1.5$	9	11	12	13	16
$c = 1$, $1.75 \leq \alpha \leq 1.99$	10	11	12	13	16
$c = 10$, $1.25 \leq \alpha \leq 1.99$	8	9	10	11	14
$c = 10$, $\alpha = 1.99$	8	10	11	12	14
$c = 100$, all $1.25 \leq \alpha \leq 1.99$	8	9	10	11	14

Note: These results are for the covariance function given in (4.8).

5. A SIMULATION STUDY

We now describe the results of a simulation study. Random vectors $X = [X(0), X(1/n), \ldots, X\{(n-1)/n\}]^T$ were generated from the stationary Gaussian process with covariance function given by (4.8), for various values of n, α and c. The values chosen were $n = 500$, 1000, 5000; $c = 0.1$, 1, 10, 100; and $\alpha = 0.25$, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 1.99. All combinations were included except for the four given by $n = 5000$, $c = 0.1$ and $\alpha = 1.25$, 1.5, 1.75, 1.99. For each combination of n, α and c considered, 1000 random vectors $X = [X(0), X(1/n), \ldots, X\{(n-1)/n\}]^T$ were generated using the simulation method described in Section 4; and for each X generated, eight versions of the regression estimator described in Section 2 were calculated, as indicated below.

The kernel function $K(x)$ given by

$$k(x) = \begin{cases} 3(1 - x^2)/4 & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| > 1 \end{cases}$$

was used throughout. The smoothing bandwidth h was chosen to be of the form $h = r/n$, for $r = 5$, 10, 20, 40 and 80.

Once a value of the smoothing bandwidth h has been selected, $M(h)$ may be calculated using either (2.2) or (2.3). In the simulation study described here, M was always calculated using (2.3), but in some contexts it may be more convenient to use (2.2), depending on the form in which the data is received. For given $h = r/n$ and X, we approximated $Y'(u/n)$ by

$$Y'\left(\frac{u}{n}\right) \approx nr^{-1} \sum_{j=-r}^{r} K\left(\frac{j}{r}\right)\left\{X\left(\frac{u+j+1}{n}\right) - X\left(\frac{u+j}{n}\right)\right\}$$

$$u = 0, 1, \ldots, n-1$$

and then, using (2.3), M was approximated by

$$M(h) = M\left(\frac{r}{n}\right) = n^{-1}\sum_{u=0}^{n-1}\left|Y'\left(\frac{u}{n}\right)\right|. \tag{5.1}$$

Note that (5.1) depends not only on $X = [X(0), \ldots, X\{(n-1)/n\}]^{\mathrm{T}}$, but also on $X\{-(r+1)/n\}, \ldots, X(-1/n)$ and $X(1), X\{(n+1)/n\}, \ldots, X\{(n+r+1)/n\}$. Thus, a sequence of length a little longer than n was used, but for convenience we have taken n as the sample size index. Eight regression estimators were considered: the six 2-point regressions based on

$$\begin{aligned} E_1 &= \left\{M\left(\frac{5}{n}\right), M\left(\frac{10}{n}\right)\right\} & E_2 &= \left\{M\left(\frac{10}{n}\right), M\left(\frac{40}{n}\right)\right\} \\ E_3 &= \left\{M\left(\frac{5}{n}\right), M\left(\frac{40}{n}\right)\right\} & E_4 &= \left\{M\left(\frac{10}{n}\right), M\left(\frac{20}{n}\right)\right\} \\ E_5 &= \left\{M\left(\frac{20}{n}\right), M\left(\frac{40}{n}\right)\right\} & E_6 &= \left\{M\left(\frac{10}{n}\right), M\left(\frac{80}{n}\right)\right\} \end{aligned} \tag{5.2}$$

and the two 3-point regressions based on

$$E_7 = \left\{M\left(\frac{5}{n}\right), M\left(\frac{10}{n}\right), M\left(\frac{40}{n}\right)\right\} \qquad E_8 = \left\{M\left(\frac{10}{n}\right), M\left(\frac{20}{n}\right), M\left(\frac{80}{n}\right)\right\} \tag{5.3}$$

In each case, $\hat{\alpha}$ was estimated using (2.8), and then the fractal dimension was estimated by $\hat{D} = 2 - \hat{\alpha}/2$.

Whenever $\hat{D}$ was less than 1, it was reset to 1; and whenever $\hat{D}$ was greater than 2, it was reset to 2. Counts were kept of the number of times that resetting was necessary.

In Tables II, III and IV, a selection of the results of the simulation study is presented. In Table II, the eight regression estimators in (5.2) and (5.3) are compared when $\alpha = 1.5$ and $c = 1.0$. These values of α and c give results which are fairly representative of the whole study. The main finding is that

TABLE II

COMPARISON OF THE EIGHT REGRESSION ESTIMATORS OF FRACTAL DIMENSION D

	$n = 500$			$n = 1000$			$n = 5000$		
	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
E_1	0.0071	0.0759	0.0058	0.0017	0.0583	0.0034	−0.0084	0.0285	0.0009
E_2	0.0270	0.1152	0.0140	0.0158	0.0829	0.0071	−0.0003	0.0433	0.0019
E_3	0.0204	0.0933	0.0091	0.0111	0.0695	0.0050	−0.0030	0.0365	0.0013
E_4	0.0175	0.1059	0.0115	0.0111	0.0771	0.0061	−0.0018	0.0397	0.0016
E_5	0.0571	0.1606	0.0291	0.0287	0.1148	0.0140	0.0032	0.0578	0.0034
E_6	0.0439	0.1283	0.0184	0.0228	0.0937	0.0093	0.0015	0.0487	0.0024
E_7	0.0213	0.0961	0.0097	0.0118	0.0712	0.0052	−0.0026	0.0374	0.0014
E_8	0.0458	0.1323	0.0196	0.0236	0.0964	0.0098	0.0018	0.0499	0.0025

Note: The covariance function given in (4.8) was used with $c = 1.0$ and $\alpha = 1.5$.These results are for estimators of the quantity $D = 2 - \alpha/2 = 1.25$.

TABLE III

COMPARISON OF ESTIMATORS E_1 AND E_7: ROUGHER SAMPLE PATHS

	E_1				E_7			
	Bias	SD	MSE		Bias	SD	MSE	
$c = 1.0$, $\alpha = 0.25$								
$n = 500$	0.0954	0.0522	0.0118	(600)	0.0771	0.0631	0.0099	(444)
$n = 1000$	0.1018	0.0384	0.0118	(565)	0.0814	0.0490	0.0090	(352)
$n = 5000$	0.1130	0.0190	0.0131	(556)	0.0809	0.0301	0.0075	(115)
$c = 1.0$, $\alpha = 0.5$								
$n = 500$	0.0523	0.1014	0.0130	(38)	0.0577	0.0958	0.0125	(46)
$n = 1000$	0.0538	0.0736	0.0083	(13)	0.0458	0.0714	0.0072	(4)
$n = 5000$	0.0396	0.0332	0.0027		0.0273	0.0336	0.0019	
$c = 1.0$, $\alpha = 0.75$								
$n = 500$	0.0257	0.0981	0.0103	(1)	0.0444	0.1071	0.0135	(4)
$n = 1000$	0.0175	0.0679	0.0049		0.0181	0.0734	0.0057	
$n = 5000$	0.0105	0.0303	0.0010		0.0085	0.0333	0.0012	
$c = 1.0$, $\alpha = 1.0$								
$n = 500$	0.0161	0.0918	0.0087		0.0281	0.1001	0.0108	
$n = 1000$	0.0072	0.0651	0.0043		0.0129	0.0740	0.0056	
$n = 5000$	0.0010	0.0285	0.0008		0.0031	0.0332	0.0011	
$c = 10.0$, $\alpha = 0.5$								
$n = 500$	0.2100	0.0605	0.0478	(532)	0.2390	0.0293	0.0580	(786)
$n = 1000$	0.1904	0.0581	0.0396	(263)	0.2212	0.0412	0.0506	(500)
$n = 5000$	0.1100	0.0339	0.0132		0.1259	0.0340	0.0170	
$c = 10.0$, $\alpha = 0.75$								
$n = 500$	0.1305	0.1020	0.0274	(11)	0.2195	0.0981	0.0578	(85)
$n = 1000$	0.0832	0.0703	0.0119		0.1342	0.0726	0.0233	(2)
$n = 5000$	0.0307	0.0306	0.0019		0.0456	0.0334	0.0032	
$c = 10.0$, $\alpha = 1.0$								
$n = 500$	0.0670	0.0936	0.0133		0.1466	0.1002	0.0315	(1)
$n = 1000$	0.0341	0.0663	0.0056		0.0739	0.0756	0.0112	
$n = 5000$	0.0065	0.0288	0.0009		0.0166	0.0334	0.0014	

Notes: The covariance function given in (4.8) was used. These results are for estimators of the quantity $D = 2 - \alpha/2$. The numbers in parentheses indicate the number of times, out of 1000, that the estimator was reset to 2. The results for $c = 10.0$ and $\alpha = 0.25$ have been omitted because, in each case, resetting occurred ≥ 996 times.

estimator E_1 is clearly superior to the other estimators: in the great majority of cases, E_1 had the smallest bias and smallest standard deviation of all the estimators. It is interesting to note that E_1 is the estimator which involves the least smoothing, and also the estimator which estimates the derivative process $Y'(t)$ least accurately.

In Table III, E_1 is compared with E_7 over smaller values of α, corresponding to rougher sample paths. The estimator E_7 was chosen as it proved to be the better of the two 3-point estimators. The numbers in parentheses in Table III refer to the number of times, out of 1000, that $\hat{D}$ was reset to 2. Note that, in Table III, the smaller the value of α, the larger the bias of both E_1 and E_7, as predicted by (3.7).

TABLE IV

COMPARISON OF ESTIMATORS E_1 AND E_7: SMOOTHER SAMPLE PATHS

	E_1				E_7			
	Bias	SD	MSE		Bias	SD	MSE	
$c = 1.0$, $\alpha = 1.25$								
$n = 500$	0.0082	0.0852	0.0073		0.0252	0.1024	0.0111	
$n = 1000$	−0.0003	0.0604	0.0036		0.0046	0.0735	0.0054	
$n = 5000$	−0.0019	0.0270	0.0007		0.0028	0.0317	0.0010	
$c = 1.0$, $\alpha = 1.5$								
$n = 500$	0.0071	0.0759	0.0058		0.0213	0.0961	0.0097	
$n = 1000$	0.0017	0.0583	0.0034		0.0118	0.0712	0.0052	
$n = 5000$	−0.0084	0.0285	0.0009		−0.0026	0.0374	0.0014	
$c = 1.0$, $\alpha = 1.75$								
$n = 500$	0.0032	0.0702	0.0049	(2)	0.0184	0.0853	0.0076	(1)
	0.0031	0.0702	0.0049		0.0184	0.0853	0.0076	
$n = 1000$	−0.0019	0.0565	0.0032	(1)	0.0098	0.0700	0.0050	
	−0.0019	0.0565	0.0032		0.0098	0.0700	0.0050	
$n = 5000$	−0.0040	0.0336	0.0011		0.0044	0.0405	0.0017	
$c = 1.0$, $\alpha = 1.99$								
$n = 500$	0.0001	0.0152	0.0002	(768)	0.0040	0.0198	0.0004	(394)
	−0.0060	0.0176	0.0003		0.0030	0.0203	0.0004	
$n = 1000$	0.0007	0.0162	0.0003	(756)	0.0035	0.0198	0.0004	(536)
	−0.0053	0.0187	0.0004		0.0021	0.0205	0.0004	
$n = 5000$	−0.0006	0.0122	0.0001	(754)	0.0018	0.0151	0.0002	(561)
	−0.0065	0.0148	0.0003		0.0003	0.0158	0.0003	

Notes: The covariance function given in (4.8) was used. These results are for estimators of the quantity $D = 2 - \alpha/2$. The numbers in parentheses indicate the number of times, out of 1000, that the estimator was reset to 1. The rows immediately below those with numbers in parentheses give the corresponding results without resetting.

A rather surprising phenomenon is evident in Table IV: the standard deviations are small when $\alpha = 1.99$, and there is an apparent discrepancy with the theoretical result (3.4). Three points are worth noting.

(i) An independently conducted simulation study, not reported here, shows that the box-counting estimator (see Hall and Wood, 1993) and the variogram estimator (ee Constantine and Hall, 1994) exhibit a similar discrepancy. That is, when α is close to 2, standard deviations are surprisingly small. Thus, this phenomenon would appear not to be specific to the crossings estimators described here.

(ii) While it is true that the standard deviations when $\alpha = 1.99$ tend to be considerably smaller than the corresponding standard deviations when $\alpha < 1.99$, there is an important respect in which there is clear agreement between Table IV and the theory: the rate of decrease, as n increases, of the standard deviation is substantially slower when $\alpha = 1.99$ than when $\alpha < 1.99$.

(iii) The results of Table IV allow us to rule out the possibility that the discrepancy is due to the resetting procedure, even though in some cases resetting has a noticeable effect on bias and standard deviation.

Assuming that both the theory and numerical results are correct, which we believe to be the case, it seems reasonable to suggest that the discrepancy is due to constants and/or remainder terms playing unexpectedly influential rôles. However, further study of this phenomenon is required.

APPENDIX

We collect some relevant facts about the smoothed process $Y(t)$ and the derivative process $Y'(t)$ which appear in Sections 2 and 3. The derivations are elementary once the following result is noted: a stationary Gaussian process with a covariance function which satisfies (1.1) has a version whose sample paths are continuous with probability 1. In other words, there exists a version $X(t)$ for which

$$P\{X(t) \text{ continuous at each } t \in (-\infty, \infty)\} = 1 \tag{A1}$$

(see Cramer and Leadbetter, 1967, Section 9.2; and also Doob, 1953, p. 62). As a consequence of (A1), the process Y of Section 2, viewed as the convolution of the process X with the differentiable kernel K, is a smooth process which has everywhere-differentiable sample paths with probability 1.

Using (A1) and integration by parts, it is seen that

$$Y \equiv h^{-1}\int K\left(\frac{u}{h}\right)X(t+u)\,du = h^{-1}\int K\left(\frac{u-t}{h}\right)X(u)\,du \tag{A2}$$

and differentiating under the integral sign with respect to t, we obtain

$$Y'(t) = -h^{-2}\int K'\left(\frac{u-t}{h}\right)X(u)\,du = -h^{-1}\int K'(u_1)X(t+hu_1)\,du_1. \tag{A3}$$

The following identities are obtained directly from (A1)–(A3), Fubini's theorem and integration by parts:

$$\begin{aligned} r(s-t) &\equiv \operatorname{cov}\{Y(s), Y(t)\} \\ &= h^{-2}\iint K\left(\frac{u_1}{h}\right)K\left(\frac{u_2}{h}\right)\gamma\{t-s+(u_1-u_2)\}\,du_1\,du_2 \\ &= \iint K(u_1)K(u_2)\gamma\{t-s+h(u_1-u_2)\}\,du_1\,du_2 \end{aligned} \tag{A4}$$

$$\begin{aligned} \operatorname{cov}\{Y'(s), Y'(t)\} &= \iint K(u_1)K(u_2)\gamma''\{t-s+h(u_1-u_2)\}\,du_1\,du_2 \\ &= h^{-2}\iint K'(u_1)K'(u_2)\gamma\{t-s+h(u_1-u_2)\}\,du_1\,du_2 \end{aligned} \tag{A5}$$

and

$$\operatorname{cov}\{Y(s), Y'(t)\} = -h^{-1}\int K'(u_1)K(u_2)\gamma\{t-s+h(u_1-u_2)\}\,du_1\,du_2.$$

After differentiating the final term in (A4) with respect to s and then t, and then using (A5), we obtain

$$\operatorname{cov}\{Y'(s), Y'(t)\} = -r''(s-t). \tag{A6}$$

ACKNOWLEDGEMENTS

The work of the first author was partly supported by an Australian Research Council grant to the CMA, Australian National University, and by a grant from NSERC Canada. We are also grateful to the referee for very detailed and helpful comments.

REFERENCES

ADLER, R. J. (1981) *The Geometry of Random Fields*. New York: Wiley.

BERRY, M. V. and HANNAY, J. H. (1978) Topography of random surfaces. *Nature* 273, 573.

BROCKWELL, P. J. and DAVIS, R. A. (1987) *Time Series: Theory and Methods*. New York: Springer.

CONSTANTINE, A. G. and HALL, P. (1994) Characterising surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc. Ser. B.* 56, 97–113.

COSTER, M. and CHERMANT, J. C. (1983) Recent developments in quantitative fractography. *Int. Metals Rev.* 28, 234–38.

CRAMÉR, H. and LEADBETTER, M. R. (1967) *Stationary and related stochastic Processes*. New York: Wiley.

DOOB, J. L. (1953) *Stochastic Processes*. New York: Wiley.

DUBUC, B., QUINIOU, J. F., ROQUES-CARMES, C., TRICOT, C. and ZUCKER, S. W. (1989) Evaluating the fractal dimension of profiles. *Phys. Rev. A* 39, 1500–12.

HALL, P. and ROY, R. (1994) On the relationship between fractal dimension and fractal index for stationary stochastic processes. Forthcoming in Ann. Appl. Probab.

—— and WOOD, A. T. A. (1993) On the performance of box-counting estimators of fractal dimension. *Biometrika* 80, 246–52.

LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer.

MANDELBROT, B. B., PASSOJA, D. E. and PAULLAY, A. J. (1984) Fractal character of fracture surfaces of metals. *Nature* 308, 721–22.

TAYLOR, C. C. and TAYLOR, S. J. (1991) Estimating the dimension of a fractal. *J. Roy. Statist. Soc. Ser. B* 53, 353–64.

THOMAS, T. R. and THOMAS, A. P. (1988) Fractals and engineering surface roughness. *Surf. Topogr.* 1, 143–52.

YAGLOM, A. M. (1987) *Correlation Theory of Stationary and Related Random Functions I*. New York: Springer.