Statistical Science
2012, Vol. 27, No. 2, 247–277
DOI: 10.1214/11-STS370

Estimators of Fractal Dimension: Assessing the Roughness of Time Series and Spatial Data

Tilmann Gneiting, Hana Ševčíková and Donald B. Percival

Abstract. The fractal or Hausdorff dimension is a measure of roughness (or smoothness) for time series and spatial data. The graph of a smooth, differentiable surface indexed in $\mathbb{R}^d$ has topological and fractal dimension d. If the surface is nondifferentiable and rough, the fractal dimension takes values between the topological dimension, d, and $d+1$. We review and assess estimators of fractal dimension by their large sample behavior under infill asymptotics, in extensive finite sample simulation studies, and in a data example on arctic sea-ice profiles. For time series or line transect data, box-count, Hall–Wood, semi-periodogram, discrete cosine transform and wavelet estimators are studied along with variation estimators with power indices 2 (variogram) and 1 (madogram), all implemented in the `R` package `fractaldim`. Considering both efficiency and robustness, we recommend the use of the madogram estimator, which can be interpreted as a statistically more efficient version of the Hall–Wood estimator. For two-dimensional lattice data, we propose robust transect estimators that use the median of variation estimates along rows and columns. Generally, the link between power variations of index $p > 0$ for stochastic processes, and the Hausdorff dimension of their sample paths, appears to be particularly robust and inclusive when $p = 1$.

Key words and phrases: Box-count, Gaussian process, Hausdorff dimension, madogram, power variation, robustness, sea-ice thickness, smoothness, variogram, variation estimator.

Lies, damn lies, and dimension estimates
Lenny Smith [(2007), p. 115]

Tilmann Gneiting is Professor, Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany (e-mail: t.gneiting@uni-heidelberg.de). Hana Ševčíková is Senior Research Scientist, Center for Statistics and the Social Sciences, University of Washington, Box 354322, Seattle, Washington 98195-4322, USA (e-mail: hanas@uw.edu). Donald B. Percival is Principal Mathematician, Applied Physics Laboratory; Professor, Department of Statistics, University of Washington, Box 355640, Seattle, Washington 98195-5640, USA (e-mail: dbp@apl.washington.edu).

1. INTRODUCTION

Fractal-based analyses of time series, transects, and natural or man-made surfaces have found extensive applications in almost all scientific disciplines (Mandelbrot, 1982). While much of the literature ties fractal properties to statistical self-similarity, no such link is necessary. Rather, we adopt the argument of Bruno and Raspa (1989), Davies and Hall (1999) and Gneiting and Schlather (2004) that the fractal or Hausdorff dimension quantifies the roughness or smoothness of time series and spatial data in the limit as the observational scale becomes infinitesimally fine. In practice, measurements can only be taken at a finite range of scales, and usable estimates of the fractal dimension depend on the availability of observations at suffi-

ciently fine temporal or spatial resolution (Malcai et al., 1997; Halley et al., 2004).

We follow common practice in defining the fractal dimension of a point set $X \subset \mathbb{R}^d$ to be the classical Hausdorff dimension (Hausdorff, 1919; Falconer, 1990). For $\varepsilon > 0$, an ε-cover of X is a finite or countable collection $\{B_i : i = 1, 2, \ldots\}$ of balls $B_i \subset \mathbb{R}^d$ of diameter $|B_i|$ less than or equal to ε that covers X. With

$$H^\delta(X) = \lim_{\varepsilon \to 0} \inf \left\{ \sum_{i=1}^{\infty} |B_i|^\delta : \{B_i : i = 1, 2, \ldots\} \text{ is an } \varepsilon\text{-cover of } X \right\}$$

denoting the δ-dimensional Hausdorff measure of X, there exists a unique nonnegative value D such that $H^\delta(X) = \infty$ if $\delta < D$ and $H^\delta(X) = 0$ if $\delta > D$. This value D is the Hausdorff dimension of the point set X. Under weak regularity conditions, the Hausdorff dimension coincides with the box-count dimension,

$$D_{\mathrm{BC}} = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}, \tag{1}$$

where $N(\varepsilon)$ denotes the smallest number of cubes of width ε in $\mathbb{R}^d$ which can cover X, and also with other natural and/or time-honored notions of dimension (Falconer, 1990).

In this paper we restrict attention to the case in which the point set

$$X = \{(t, X_t) \in \mathbb{R}^d \times \mathbb{R} : t \in \mathrm{T} \subset \mathbb{R}^d\} \subset \mathbb{R}^{d+1}$$

is the graph of time series or spatial data observed at a finite set $\mathrm{T} \subset \mathbb{R}^d$ of typically regularly spaced times or locations. The fractal dimension then refers to the properties of the curve ($d = 1$) or surface ($d \geq 2$) that arises in the continuum limit as the data are observed at an infinitesimally dense subset of the temporal or spatial domain, which without loss of generality can be assumed to be the unit interval or unit cube. In time series analysis and spatial statistics, this limiting scenario is referred to as infill asymptotics (Hall and Wood, 1993; Dahlhaus, 1997; Stein, 1999).

If the limit curve or limit surface is smooth and differentiable, its fractal dimension, D, equals its topological dimension, d. For a rough and nondifferentiable curve or surface, the fractal dimension may exceed the topological dimension. For example, suppose that $\{X_t : t \in \mathbb{R}^d\}$ is a Gaussian process with stationary increments, whose variogram or structure function,

$$\gamma_2(t) = \tfrac{1}{2}\mathbb{E}(X_u - X_{u+t})^2, \tag{2}$$

satisfies

$$\gamma_2(t) = |c_2 t|^\alpha + \mathcal{O}(|t|^{\alpha+\beta}) \quad \text{as } t \to 0, \tag{3}$$

where $\alpha \in (0, 2]$, $\beta \geq 0$, and $c_2 > 0$, and $|\cdot|$ denotes the Euclidean norm. Then the graph of a sample path has fractal dimension

$$D = d + 1 - \frac{\alpha}{2} \tag{4}$$

almost surely (Orey, 1970; Adler 1981). This relationship links the fractal dimension of the sample paths to the behavior of the variogram or structure function at the coordinate origin, and can be extended to broad classes of potentially anisotropic and nonstationary processes, as well as some non-Gaussian processes (Adler, 1981; Hall and Roy, 1994; Xue and Xiao, 2011). It allows us to think of fractal dimension as a second-order property of a Gaussian stochastic process, in addition to being a roughness measure for a realized curve or surface. Accordingly, we refer to the index α in the asymptotic relationship (3) as the fractal index.

Table 1 provides examples of Gaussian processes that exhibit this asymptotic behavior. Fractional Brownian motion is a nonstationary, statistically self-similar process that is defined in terms of the variogram (Mandelbrot and Van Ness, 1968). The other entries in the table refer to stationary processes with covariance function $\sigma(t) = \mathrm{cov}(X_u, X_{t+u})$, which relates to the variogram as

$$\gamma_2(t) = \sigma(0) - \sigma(t), \quad t \in \mathbb{R}^d.$$

Key examples include the Matérn family (Matérn, 1986; Guttorp and Gneiting, 2006), used by Goff and Jordan (1988) to parameterize the fractal dimension of oceanic features; the Cauchy class, introduced by Gneiting and Schlather (2004) to illustrate local and global properties of random functions; and the Dagum family (Berg, Mateu and Porcu, 2008). In simulation settings, the powered exponential class (Yaglom, 1987) is a convenient default example. The exponent β in the asymptotic relationship (3) equals $\beta = 2 - \alpha$ for the Matérn class, $\beta = \alpha$ for the powered exponential and Cauchy families, and $\beta = \tau$ for the Dagum family. Generally, the smaller the value of β, the harder the estimation of the fractal index, α, and the more pronounced the finite sample bias of estimators of the fractal index or fractal dimension.

As an illustration for time series or line transect data, Figure 1 displays Gaussian sample paths from the powered exponential family with the fractal index, α, ranging from 1.9 to 0.2, and the fractal dimension, D, ex-

TABLE 1

Some parametric classes of variograms and covariance functions for a Gaussian process $\{X_t : t \in \mathbb{R}^d\}$. *The covariance functions have been normalized such that* $\sigma(0) = 1$. *Here,* α *is the fractal index,* $c > 0$ *is a range parameter, and* K_ν *is a modified Bessel function of the second kind of order* ν. *The Matérn and Dagum families allow for less restrictive assumptions on the parameters than stated here*

Class	**Variogram or covariance**	**Parameters**
Fractional Brownian motion	$\gamma_2(t) = \lvert ct\rvert^\alpha$	$\alpha \in (0, 2]$
Matérn	$\sigma(t) = \frac{2^{(\alpha/2)-1}}{\Gamma(\alpha/2)} \lvert ct\rvert^{\alpha/2} K_{\alpha/2}(\lvert ct\rvert)$	$\alpha \in (0, 2]$
Powered exponential	$\sigma(t) = \exp(-\lvert ct\rvert^\alpha)$	$\alpha \in (0, 2]$
Cauchy	$\sigma(t) = (1 + \lvert ct\rvert^\alpha)^{-\tau/\alpha}$	$\alpha \in (0, 2]; \tau > 0$
Dagum	$\sigma(t) = 1 - (\frac{\lvert ct\rvert^\tau}{1+\lvert ct\rvert^\tau})^{\alpha/\tau}$	$\tau \in (0, 2]; \alpha \in (0, \tau)$

tending from 1.05 to 1.9. The smaller α, the rougher the sample path, and the larger the fractal dimension, with the lower limit, $D = 1$, being associated with a smooth curve, and the upper limit, $D = 2$, corresponding to a space-filling, exceedingly rough graph. Figure 2 shows a realization from the nonstationary Gaussian Matérn model of Anderes and Stein (2011), in a case in which the fractal index and the fractal dimension vary linearly along the unit interval. To illustrate the visual effects of the measurement scale, both the original sample path of size 10,000 and an equidistantly thinned version of size 1,000 are shown.

Turning to spatial data, Figure 3 shows Gaussian sample surfaces from the powered exponential class with the fractal index, α, being equal to 1.5, 1.0 and 0.2. The surfaces are increasingly rough with the frac-

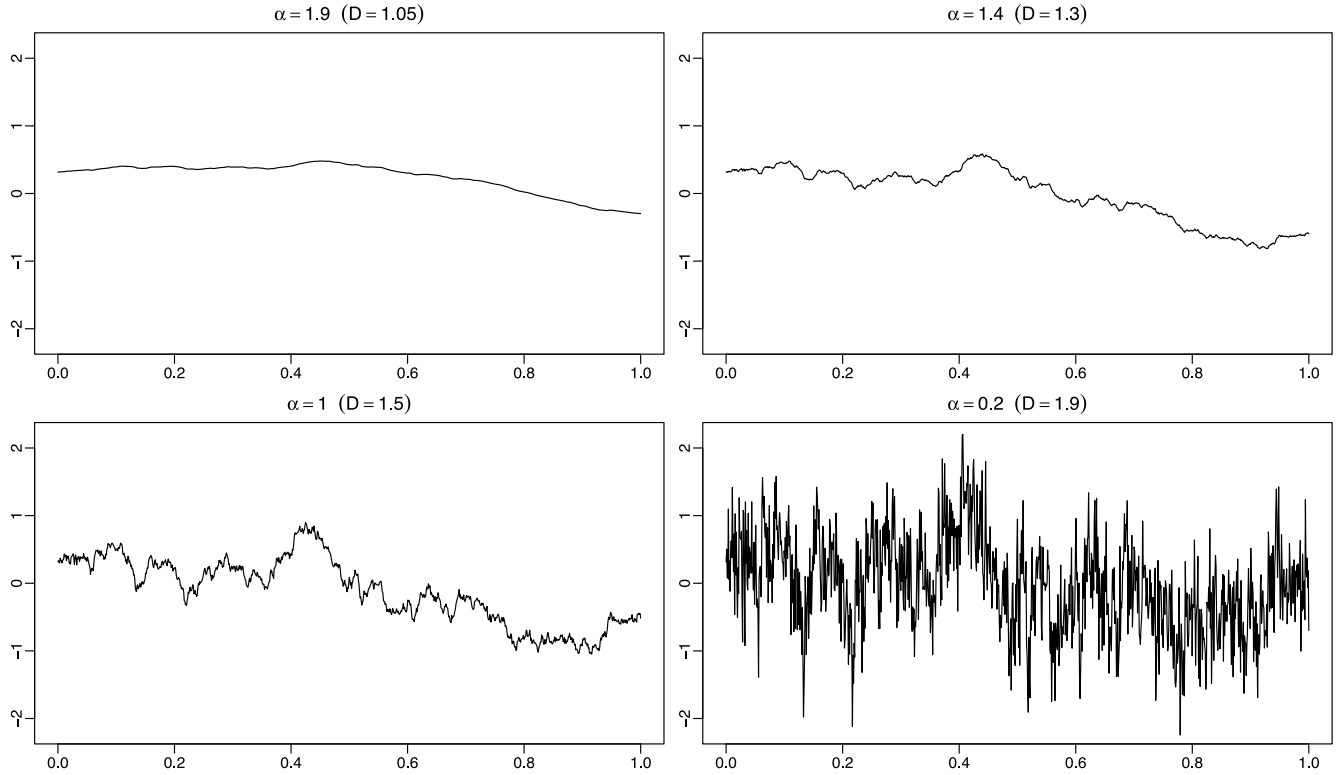

FIG. 1. *Sample paths of stationary Gaussian processes with powered exponential covariance function,* $\sigma(t) = \exp(-|t|^\alpha)$, *and fractal index,* α, *equal to* 1.9, 1.4, 1.0 *and* 0.2. *The corresponding values of the fractal dimension,* D, *are* 1.05, 1.3, 1.5 *and* 1.9, *respectively. The simulation domain is a grid over the unit interval with spacing* 1/1,024.

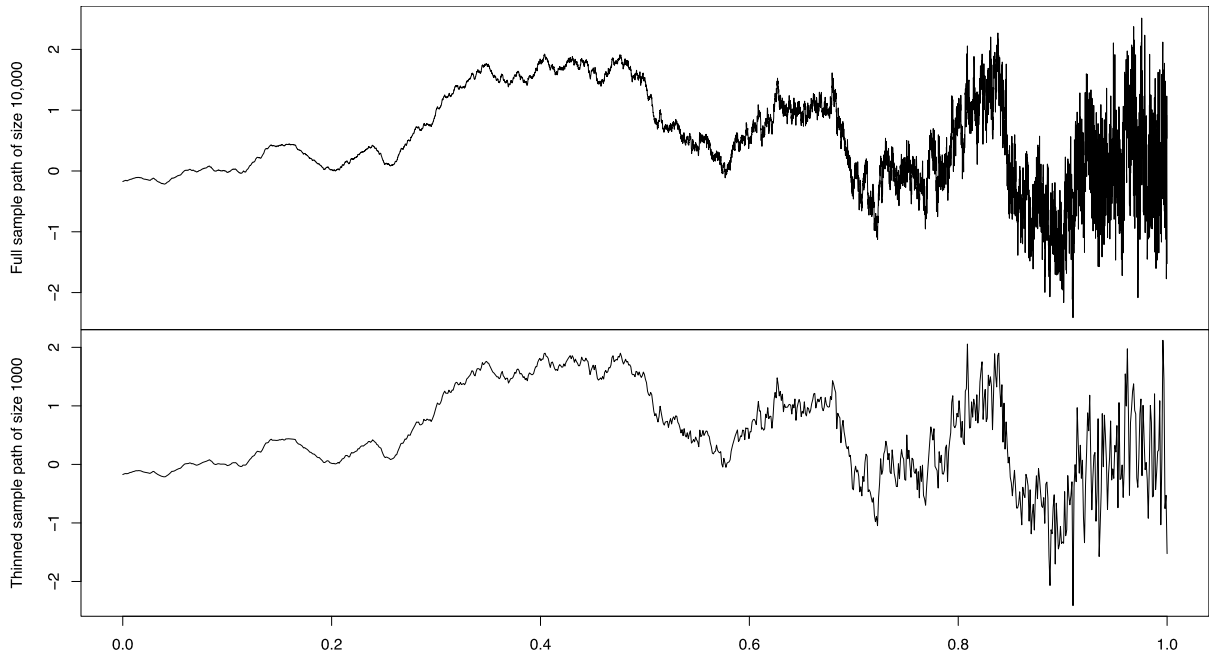

FIG. 2. *A sample path from the nonstationary Gaussian Matérn process of Anderes and Stein* (2011), *where the fractal dimension*, D, *grows linearly from* $D = 1$ *at time* $t = 0$ *to* $D = 2$ *at time* $t = 1$. *To illustrate the visual effects of the temporal resolution, both the original sample path with grid spacing* $1/10{,}000$ (*top panel*) *and an equidistantly thinned version with grid spacing* $1/1{,}000$ (*bottom panel*) *are shown. The nonstationary covariance is given by equation* (10) *of Anderes and Stein* (2011) *with* $\sigma^2 = 1$, $\rho = 1/2$, *and linearly varying local smoothness parameter,* $\nu_t = 1 - t$.

tal dimension, D, being equal to 2.25, 2.5 and 2.9, respectively.

A wealth of applications requires the characterization of the roughness or smoothness of time series, line transect or spatial data, with Burrough (1981) and Malcai et al. (1997) summarizing an impressive range of experimental results. For example, fractal dimensions have been studied for geographic profiles and surfaces, such as the underside of sea ice (Rothrock and Thorndike, 1980), the topography of the sea floor (Goff and Jordan, 1988), Martial surface (Orosei et al., 2003) and terrestrial features (Weissel, Pratson and Malinverno, 1994; Turcotte, 1997; Gagnon, Lovejoy and Schertzer, 2006). Further references can be found in Molz, Rajaram and Lu (2004) for applications in subsurface hydrology and in Halley et al. (2004) for applications in ecology. Not surprisingly then, estimators for the fractal dimension have been proposed and

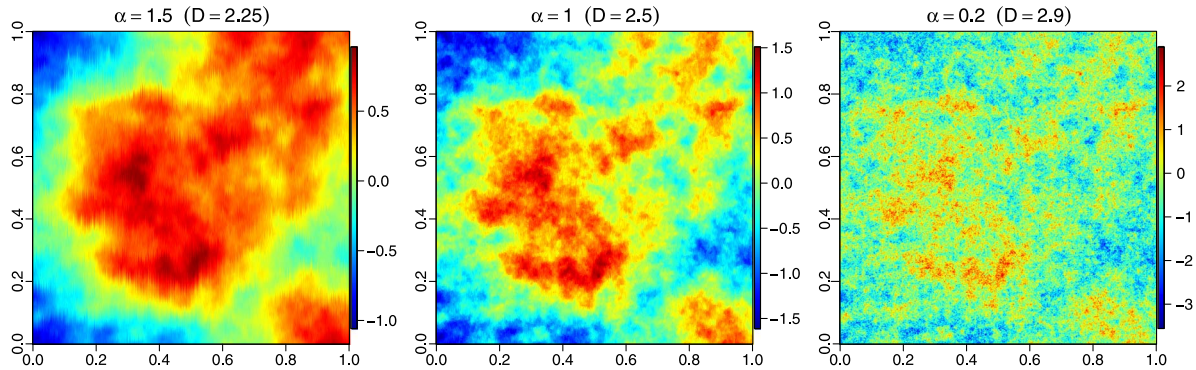

FIG. 3. *Sample surfaces of stationary spatial Gaussian processes with the powered exponential covariance function,* $\sigma(t) = \exp(-|t|^\alpha)$, *and fractal index,* α, *equal to* 1.5, 1.0 *and* 0.2. *The corresponding values of the fractal dimension,* D, *are* 2.25, 2.5 *and* 2.9. *The simulation domain is a grid on the unit square in* $\mathbb{R}^2$ *with spacing* $1/512$ *along each coordinate.*

widely used in various literatures, including physics, engineering, the earth sciences, and statistics, with the works of Burrough (1981), Goff and Jordan (1988), Bruno and Raspa (1989), Dubuc et al. (1989a, 1989b), Jakeman and Jordan (1990), Theiler (1990), Klinkenberg and Goodchild (1992), Schepers, van Beek and Bassingthwaighte (1992), Hall and Wood (1993), Constantine and Hall (1994), Kent and Wood (1997), Davies and Hall (1999), Chan and Wood (2000), Zhu and Stein (2002), Chan and Wood (2004) and Bez and Bertrand (2011) being examples. Our objectives in this paper are to survey the literature across disciplines, review and assess the various types of estimators, and provide recommendations for practitioners, along with novel directions for theoretical work.

The remainder of the paper is organized as follows. In Section 2 we describe estimators of fractal dimension for time series and line transect data, including box-count, Hall–Wood, variogram, madogram, power variation, semi-periodogram and wavelet-based techniques. Then in Section 3 we assess and compare the estimators. Considering both efficiency and robustness, we concur with Bruno and Raspa (1989) and Bez and Bertrand (2011) and recommend the use of the madogram estimator, that is, the variation estimator with power index $p = 1$, which can be interpreted as a statistically efficient version of the Hall–Wood estimator. An underlying motivation is that for an intrinsically stationary process with variogram of order $p > 0$ of the form

$$\begin{aligned}\gamma_p(t) &= \tfrac{1}{2}\mathbb{E}|X_u - X_{u+t}|^p \\ &= |c_p t|^{\alpha p/2} + \mathcal{O}\big(|t|^{(\alpha+\beta)p/2}\big) \quad \text{as } t \to 0,\end{aligned} \tag{5}$$

the relationship (4) between the fractal index, α, and the fractal dimension, D, appears to be more robust and inclusive when $p = 1$, as compared to the default case, in which $p = 2$.

Section 4 discusses ways in which estimators for the time series or line transect case can be adapted to spatial data observed over a regular lattice in $\mathbb{R}^2$, and Section 5 evaluates these proposals. Our preferred estimators in this setting are simple, robust transect estimators that employ the median of variation estimates with power index $p = 1$ along individual rows and columns. A data example on arctic sea-ice profiles is given in Section 6. The paper ends in Section 7, where we make a call for new directions in theoretical and methodological work that addresses both probabilists and statisticians. Furthermore, we hint at inference for nonstationary or multifractional processes, where the fractal dimension of a sample path may vary temporally or spatially. All computations in the paper use the `fractaldim` package (Ševčíková, Gneiting and Percival, 2011), which implements our proposals in `R` (Ihaka and Gentleman, 1996).

2. ESTIMATING THE FRACTAL DIMENSION OF TIME SERIES AND LINE TRANSECT DATA

Spurred and inspired by the now classical essay of Mandelbrot (1982), a large number of methods have been developed for estimating fractal dimension. By the early 1990s a sizable, mostly heuristic literature on the estimation of fractal dimension for time series and line transect data had accumulated in the physical, engineering and earth sciences, where various reviews are available (Dubuc et al., 1989a; Klinkenberg and Goodchild, 1992; Schepers, van Beek and Bassingthwaighte, 1992; Gallant et al., 1994; Klinkenberg, 1994; Schmittbuhl, Vilotte and Roux, 1995). These developments prompted the statistical community to introduce new methodology, along with asymptotic theory for box-count (Hall and Wood, 1993), variogram (Constantine and Hall, 1994; Kent and Wood, 1997), level crossing (Feuerverger, Hall and Wood, 1994) and spectral (Chan, Hall and Poskitt, 1995) estimators, among others.

Essentially all methods follow a common scheme, in that:

(a) a certain numerical property, say Q, of the time series or line transect data is computed as a function of "scale," say ε;

(b) an asymptotic power law $Q(\varepsilon) \propto \varepsilon^b$ as the scale $\varepsilon \to 0$ becomes infinitesimally small is derived or postulated; where

(c) the scaling exponent, b, is a linear function of the fractal dimension, D;

(d) and thus D is estimated by linear regression of $\log Q(\varepsilon)$ on $\log \varepsilon$, with emphasis on the smallest observed values of the scale ε.

Table 2 shows the data property, the measure of scale and the scaling law for various methods. For techniques working in the spectral domain, the scaling law applies as the frequency grows to infinity, equivalent to the scale becoming infinitesimally small in the time domain.

In the balance of this section, we describe the most popular estimators of fractal dimension in the equally spaced time series or line transect setting. Without loss of generality, we may assume that the observation domain is the unit interval. In the case of $n_s = n + 1$

TABLE 2
Some extant methods for estimating the fractal dimension of time series and line transect data

Method	Property	Scale	Scaling law	Regime
Box-count	$N(\varepsilon)$: number of boxes	ε: box width	$N(\varepsilon) \propto \varepsilon^{-D}$	$\varepsilon \to 0$
Divider	$L(\varepsilon)$: length of curve	ε: step size	$L(\varepsilon) \propto \varepsilon^{-1-D}$	$\varepsilon \to 0$
Level crossing	$M(h)$: number of crossings	h: bandwidth	$M(h) \propto h^{1-D}$	$h \to 0$
Variogram	$\gamma_2(t)$: variogram	t: lag	$\gamma_2(t) \propto t^{4-2D}$	$t \to 0$
Madogram	$\gamma_1(t)$: madogram	t: lag	$\gamma_1(t) \propto t^{2-D}$	$t \to 0$
Spectral	$f(\omega)$: spectral density	ω: frequency	$f(\omega) \propto \omega^{2D-5}$	$\omega \to \infty$
Wavelet	$\nu^2(\tau)$: wavelet variance	τ: scale	$\nu^2(\tau) \propto \tau^{4-2D}$	$\tau \to 0$

equally spaced observations, the data graph is the point set

$$\left\{(t, X_t) : t = \frac{i}{n}, i = 0, 1, \ldots, n\right\} \subset \mathbb{R}^2. \tag{6}$$

The relevant asymptotic regime then is infill asymptotics, in which the number of observations grows to infinity, whereas the underlying domain, namely the unit interval, remains fixed. For convenience in what follows, we refer to both n and n_s as the sample size.

2.1 Box-Count Estimator

The popular box-count estimator is motivated by the scaling law (1) that defines the box-count dimension. The basic idea is simple, in that the time series graph is initially covered by a single box. The box is divided into four quadrants, and the number of cells required to cover the curve is counted. Then each subsequent quadrant is divided into four subquadrants, and one continues doing so until the box width equals the resolution of the data, keeping track of the number of quadrants required to cover the graph at each step. If $N(\varepsilon)$ denotes the number of boxes required at width or scale ε, the box-count estimator equals the slope in an ordinary least squares regression fit of $\log N(\varepsilon)$ on $\log \varepsilon$. Similarly to Mandelbrot's (1967) divider technique, the method can be used to quantify the fractal dimension of any planar point set, rather than just equally spaced time series or line transect data.

In our setting of a data graph of the form (6), where, for simplicity, we assume that $n = 2^K$ is a power of 2, the box-count algorithm can be summarized as follows. Let $u = \max_{0 \le j \le n} X_{j/n} - \min_{0 \le j \le n} X_{j/n}$ denote the range of the data. Consider scales $\varepsilon_k = 2^{k-K}$ where $k = 0, 1, \ldots, K$. At the largest scale, $\varepsilon_K = 1$, the graph (6) can be covered by a single box of width 1 and height u, which we now call the bounding box. At scale ε_k the bounding box can be tiled by 4^{K-k} boxes of width 2^{k-K} and height $u2^{k-K}$ each, and we define $N(\varepsilon_k)$ to be the number of such boxes that intersect the linearly interpolated data graph. Figure 4 provides an illustration on two of the datasets in Figure 1, for which $n = 1024$ and $K = 10$. For example, the upper left plot considers $k = 8$, where $\varepsilon_8 = 2^{-2}$ and $N(\varepsilon_8) = 11$, and the middle left plot looks at $k = 5$, where $\varepsilon_5 = 2^{-5}$ and $N(\varepsilon_5) = 207$. The naive box-count estimator then uses the slope in an ordinary least squares regression fit of $\log N(\varepsilon)$ on $\log \varepsilon$, that is,

$$\widehat{D}_{\mathrm{BC}} = -\left\{\sum_{k=0}^{K} (s_k - \bar{s}) \log N(\varepsilon_k)\right\} \left\{\sum_{k=0}^{K} (s_k - \bar{s})^2\right\}^{-1},$$

where $s_k = \log \varepsilon_k$ and $\bar{s}$ is the mean of $s_0, s_1, \ldots, s_K$. In our illustrating example, this leads to the estimates shown in the lower row of Figure 4.

Several authors identified problems with the naive estimator that includes all scales in the regression fit of $\log N(\varepsilon)$ on $\log \varepsilon$, and proposed modifications that address these issues (Dubuc et al., 1989a; Liebovitch and Toth, 1989; Block, von Bloh and Schellnhuber, 1990; Taylor and Taylor, 1991). Indeed, one always has $N(\varepsilon_0) \ge n$ and $N(\varepsilon_K) = 1$, which suggests that the counts at both the smallest and the largest scales ought to be discarded. Liebovitch and Toth (1989) proposed to exclude the smallest scales ε_k for which $N(\varepsilon_k) > n/5$, as well as the two largest scales, from the regression fit. We adopt this proposal in our standard version of the box-count estimator, as illustrated in Figure 5. The restriction on the scales improves the statistical and computational efficiency of the estimator. However, it is in the limit as $\varepsilon \to 0$ that the underlying scaling law (1) operates, and thus it is unfortunate that information at the very smallest scales is discarded.

A natural variant of the box-count estimator uses scales $\varepsilon_l = l/n$, rather than scales $\varepsilon_k = 2^k/n$ at the

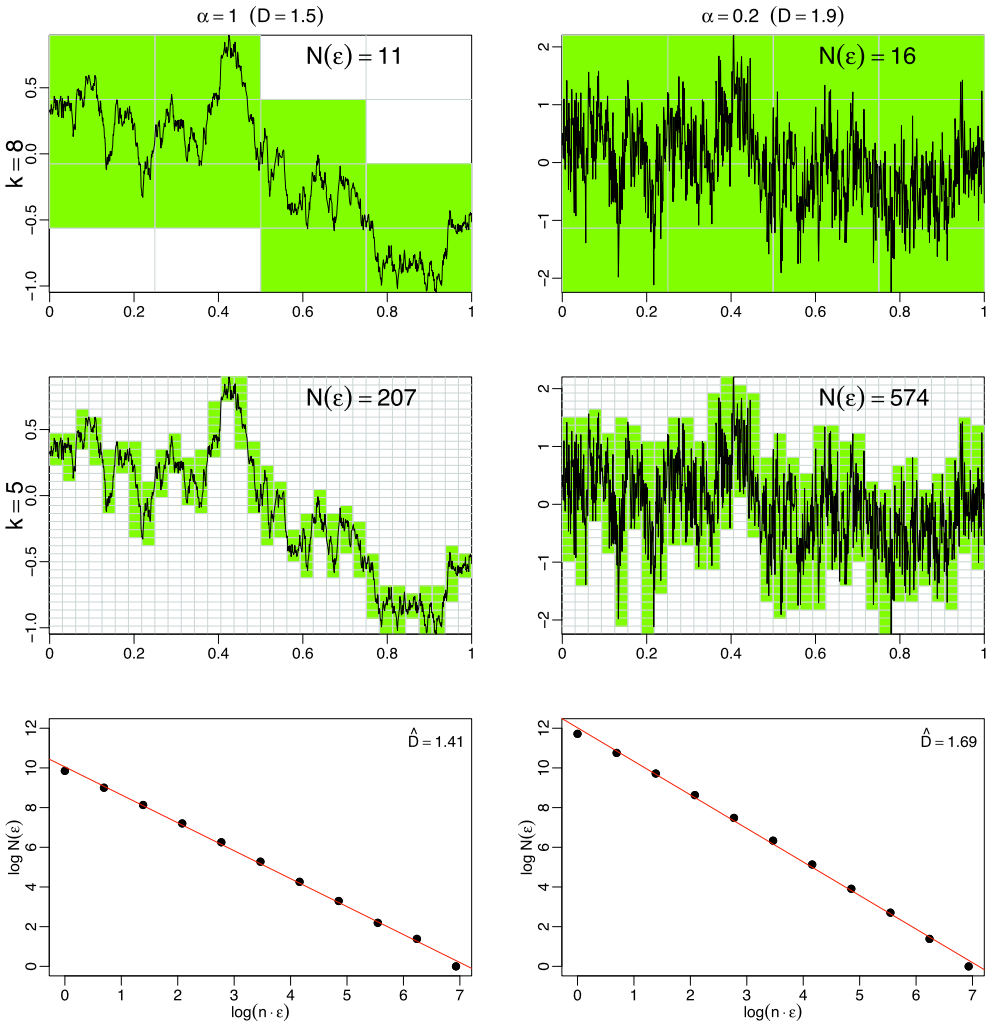

FIG. 4. *Illustration of the box-count algorithm and naive box-count estimates for the two datasets in the lower row of Figure 1. See the text for details.*

powers of 2 only. In the next section we discuss a related estimator that takes up this idea, addresses the aforementioned limitations, and is tailored to time series data of the form (6).

2.2 Hall–Wood Estimator

Hall and Wood (1993) introduced a version of the box-count estimator that operates at the smallest ob-

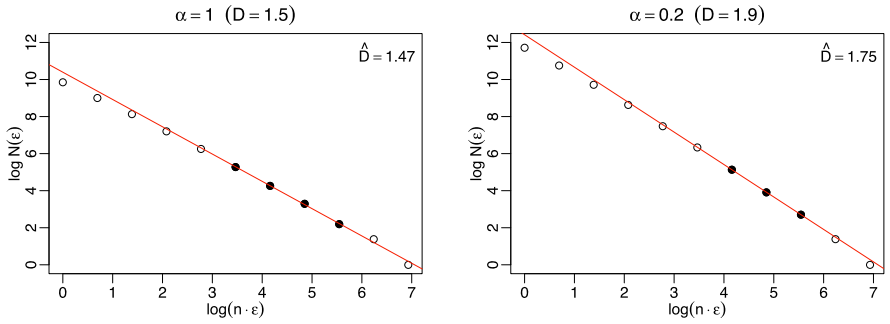

FIG. 5. *Log-log regression for the standard version of the box-count estimator and the datasets in the lower row of Figure 1. Only the points marked with filled circles are used when fitting the regression line.*

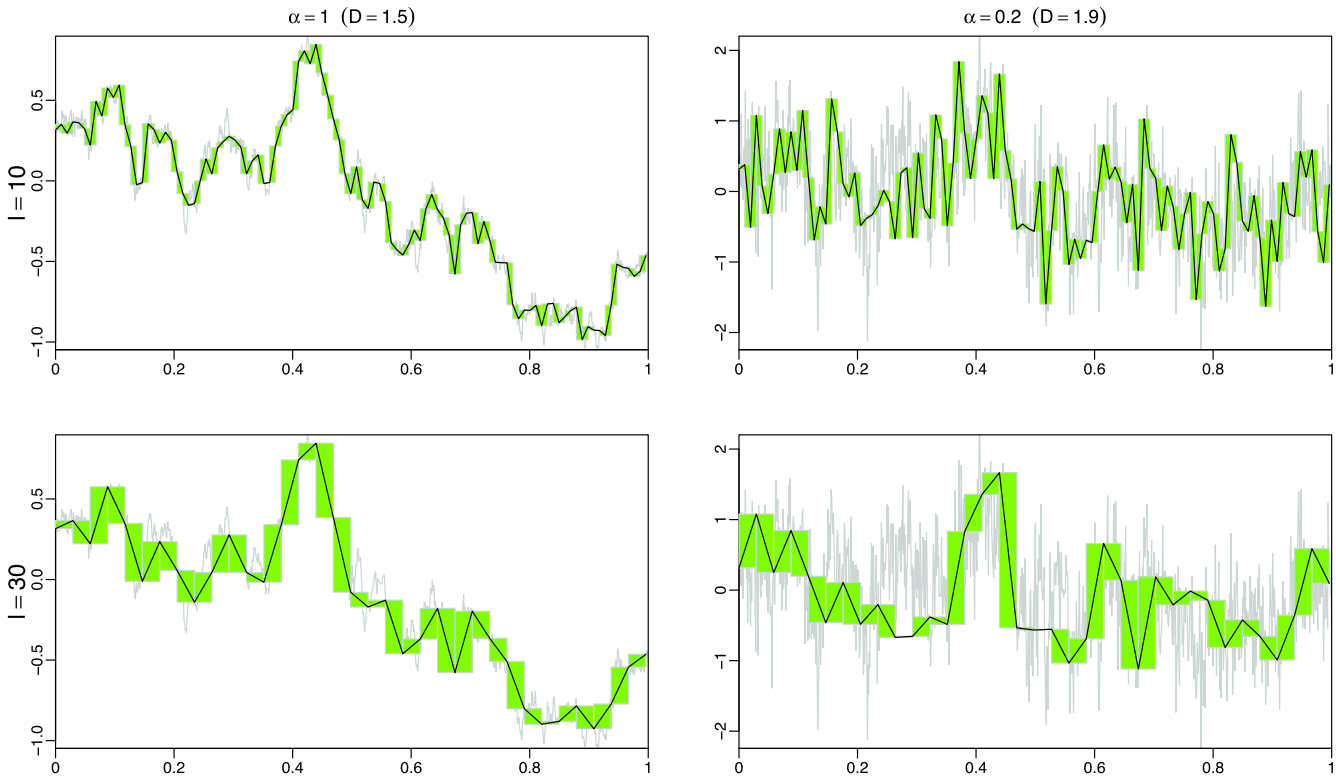

FIG. 6. *Illustration of the Hall–Wood algorithm for the datasets in the lower row of Figure 1. The quantity* $\widehat{A}(l/n)$ *is computed as the sum of the colored areas, where* $n = 1{,}024$, $l = 10$ *(upper row) and* $l = 30$ *(lower row).*

served scales and avoids the need for rules of thumb in its implementation.

To motivate their proposal, let $A(\varepsilon)$ denote the total area of the boxes at scale ε that intersect with the linearly interpolated data graph (6). There are $N(\varepsilon)$ such boxes, and so $A(\varepsilon) \propto N(\varepsilon)\varepsilon^2$, which leads us to a reformulation of definition (1), namely,

$$D_{\mathrm{BC}} = 2 - \lim_{\varepsilon \to 0} \frac{\log A(\varepsilon)}{\log(\varepsilon)}. \tag{7}$$

At scale $\varepsilon_l = l/n$, where $l = 1, 2, \ldots$, an estimator of $A(l/n)$ is

$$\widehat{A}(l/n) = \frac{l}{n} \sum_{i=1}^{\lfloor n/l \rfloor} \left| X_{il/n} - X_{(i-1)l/n} \right|, \tag{8}$$

where $\lfloor n/l \rfloor$ denotes the integer part of n/l. Figure 6 suggests a natural interpretation of this quantity, in that it approximates $A(l/n)$, with all features at scales less than l/n being ignored. For an alternative, and potentially preferable, interpretation in terms of power variations, see Section 2.4.

The Hall–Wood estimator with design parameter $m = 1$, as used in the numerical experiments of Hall and Wood (1993), is based on an ordinary least squares regression fit of $\log \widehat{A}(l/n)$ on $\log(l/n)$:

$$\widehat{D}_{\mathrm{HW}} = 2 - \left\{ \sum_{l=1}^{L} (s_l - \bar{s}) \log \widehat{A}(l/n) \right\} \left\{ \sum_{l=1}^{L} (s_l - \bar{s})^2 \right\}^{-1}, \tag{9}$$

where $L \geq 2$, $s_l = \log(l/n)$ and $\bar{s} = \frac{1}{L}\sum_{l=1}^{L} s_l$. Hall and Wood (1993) recommended the use of $L = 2$ to minimize bias, which is unsurprising, in view of the limit in (7) being taken as $\varepsilon \to 0$. Using $L = 2$ yields our standard implementation of the Hall–Wood estimator, namely,

$$\widehat{D}_{\mathrm{HW}} = 2 - \frac{\log \widehat{A}(2/n) - \log \widehat{A}(1/n)}{\log 2}. \tag{10}$$

Figure 7 shows the corresponding log–log plots and regression fits in our illustrating example.

2.3 Variogram Estimator

Owing to its intuitive appeal and ease of implementation, the variogram estimator has been very popular. A prominent early application is that of Burrough (1981). The first asymptotic study under the infill scenario was published in the physics literature (Jakeman

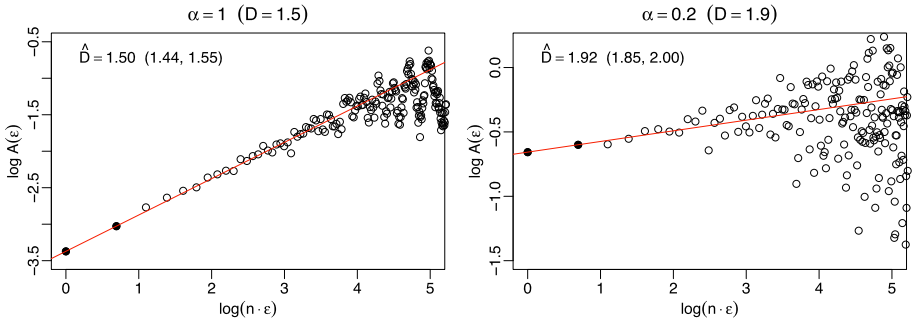

FIG. 7. *Log–log regression for the Hall–Wood estimator* (10) *and the datasets in the lower row of Figure* 1. *Only the two points marked with filled circles are used when fitting the regression line.*

and Jordan, 1990), followed by key contributions of Constantine and Hall (1994), Kent and Wood (1997), Davies and Hall (1999) and Chan and Wood (2004) in statistical journals.

Recall that the variogram or structure function $\gamma(t)$ of a stochastic process $\{X_t : t \in \mathbb{R}\}$ with stationary increments is defined in (2) as one-half times the expectation of the square of an increment at lag t. The classical method of moments estimator for $\gamma(t)$ at lag $t = l/n$ from time series or line transect data (6) is

$$\widehat{V}_2(l/n) = \frac{1}{2(n-l)} \sum_{i=l}^{n} \bigl(X_{i/n} - X_{(i-l)/n}\bigr)^2. \tag{11}$$

In view of the relationship (4) between the fractal index, defined in (3), and the fractal dimension, D, a regression fit of $\log \widehat{V}(t)$ on $\log t$ yields the variogram estimator,

$$\widehat{D}_{\mathrm{V};2} = 2 - \frac{1}{2}\left\{\sum_{l=1}^{L}(s_l - \bar{s}) \log \widehat{V}_2(l/n)\right\} \cdot \left\{\sum_{l=1}^{L}(s_l - \bar{s})^2\right\}^{-1}, \tag{12}$$

where $L \geq 2$, $s_l = \log(l/n)$ and $\bar{s}$ is the mean of $s_1, \ldots, s_L$. Figure 8 illustrates the log–log regression for the datasets in the lower row of Figure 1. In addition to the corresponding point estimate, we provide a 90% central interval estimate, using the parametric bootstrap method as proposed by Davies and Hall (1999).

Constantine and Hall (1994) argued that the bias of the variogram estimator increases with the cut-off L in the log–log regression, and Davies and Hall (1999) showed numerically that the mean squared error (MSE) of the estimator for a Gaussian process with powered exponential covariance is minimized when $L = 2$. Zhu and Stein (2002) argued similarly in a spatial setting. These results are unsurprising and have intuitive support from the fact that the behavior of the theoretical variogram in an infinitesimal neighborhood of zero determines the fractal dimension of the Gaussian sample paths. We thus choose $L = 2$ in our implementation, resulting in the estimate

$$\widehat{D}_{\mathrm{V};2} = 2 - \frac{1}{2}\frac{\log \widehat{V}_2(2/n) - \log \widehat{V}_2(1/n)}{\log 2}. \tag{13}$$

As Kent and Wood (1997) suggested, generalized least squares rather than ordinary least squares regression

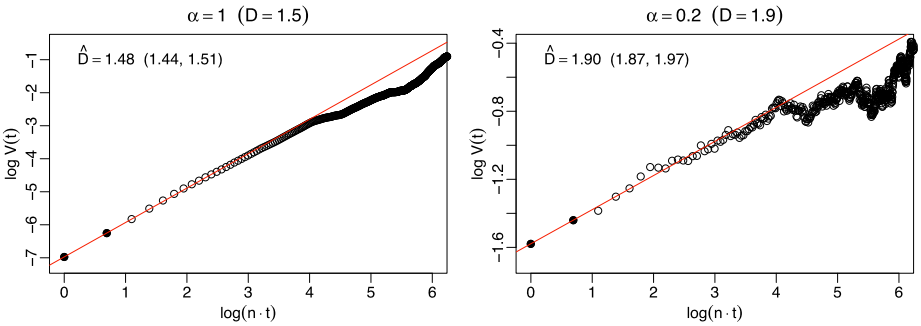

FIG. 8. *Log-log regression for the variogram estimator* (13) *and the datasets in the lower row of Figure* 1. *Only the two points marked with filled circles are used when fitting the regression line.*

could be employed, though the methods coincide when $L = 2$.

It is well known that the method of moments estimator (11) is nonrobust. It is therefore tempting to replace it by the highly robust variogram estimator proposed by Genton (1998), which is based on the robust estimator of scale of Rousseeuw and Croux (1993). We implemented an estimator of the fractal dimension that uses (12) with $L = 2$, but with the method of moments estimate (11) replaced by Genton's highly robust variogram estimator. In a simulation setting, this estimator works well. However, it breaks down frequently when applied to real data, where it yields very limited, discrete sets of possible values for the estimate only. Upon further investigation, this stems from the ubiquitous discreteness of real-world data, under which the Rousseeuw and Croux (1993) estimator can fail, in ways just described. While discreteness is a general issue when estimating fractal dimension, the problem is exacerbated by the use of this estimator. In this light, the next section investigates another approach to more robust estimators of fractal dimension.

2.4 Variation Estimators

We now discuss a generalization of the variogram estimator that is based on the variogram of order p of a stochastic process with stationary increments, namely,

$$\gamma_p(t) = \tfrac{1}{2}\mathbb{E}|X_u - X_{t+u}|^p. \tag{14}$$

When $p = 2$, we recover the variogram (2), when $p = 1$ the madogram, and when $p = 1/2$ the rodogram (Bruno and Raspa, 1989; Emery, 2005; Bez and Bertrand 2011). Standard arguments show that a Gaussian process with a variogram of the form (3) admits analogous expansions of the variogram of order p, in that

$$\gamma_p(t) = |c_p t|^{\alpha p/2} + \mathcal{O}\big(|t|^{(\alpha+\beta)p/2}\big) \quad \text{as } t \to 0, \tag{15}$$

with fixed values of the fractal index, $\alpha \in (0, 2]$, $\beta > 0$, and a constant $c_p > 0$ that satisfies

$$c_p = \left(\frac{2^{p-1}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\right)^{2/(\alpha p)} c_2.$$

The fractal index, α, of the Gaussian process and the Hausdorff dimension, D, of its sample paths then admit the linear relationship (4).

A natural generalization of the method of moments variogram estimator (11) for time series or line transect data of the form (6) is the power variation of order p, namely,

$$\widehat{V}_p(l/n) = \frac{1}{2(n-l)}\sum_{i=l}^{n}|X_{i/n} - X_{(i-l)/n}|^p. \tag{16}$$

We then define the variation estimator of order p for the fractal dimension as

$$\widehat{D}_{\mathrm{V};p} = 2 - \frac{1}{p}\left\{\sum_{l=1}^{L}(s_l - \bar{s})\log\widehat{V}_p(l/n)\right\} \cdot \left\{\sum_{l=1}^{L}(s_l - \bar{s})^2\right\}^{-1}, \tag{17}$$

where $L \geq 2$, $s_l = \log(l/n)$ and $\bar{s}$ is the mean of $s_l, \ldots, s_L$. This definition nests the variogram, madogram and rodogram estimators, which arise when $p = 2$, 1 and $1/2$, respectively. The general case, $p > 0$, has been studied by Coeurjolly (2001, 2008).

For the same reasons as before, and supported by simulation experiments, we let $L = 2$ in our implementation, so that

$$\widehat{D}_{\mathrm{V};p} = 2 - \frac{1}{p}\frac{\log\widehat{V}_p(2/n) - \log\widehat{V}_p(1/n)}{\log 2}. \tag{18}$$

As an illustration, Figure 9 shows the log–log regression fit for the variation estimator of order $p = 1$ and our example data. For instances of the use of the madogram estimator in the applied literature, see Weissel, Pratson and Malinverno (1994) and Zaiser et al. (2004).

A natural question then is for the choice of the power index $p > 0$. With the estimator depending on the relationship (4) between the fractal index, α, in the expansion (15) and the Hausdorff dimension, D, of the sample path, it is critically important to assess its validity when the assumption of Gaussianity is violated. In the standard case in which $p = 2$, Hall and Roy (1994) showed that, while the relationship (4) extends to some non-Gaussian processes, it fails easily. For example, it applies to marginally power-transformed Gaussian fields if and only if the transformation power exceeds $1/2$. Other counterexamples can be found in the work of Bruno and Raspa (1989) and Scheuerer (2010).

Bruno and Raspa (1989) and Bez and Bertrand (2011) applied the Lipschitz–Hölder heuristics of Mandelbrot [(1977), page 304] to argue that the relationship (4) is universal when $p = 1$. While we agree that the relationship is particularly robust and inclusive when $p = 1$, the Lipschitz–Hölder heuristics, which connects the Lipschitz exponent of a sample path to its Hausdorff dimension, is tied to continuity. Thus, it can fail if the sample paths are sufficiently irregular. For instance, the sample paths of a binary stochastic process, which attains the values 0 and 1 only, have Hausdorff dimension $D = 1$. As the corresponding variogram (14) is

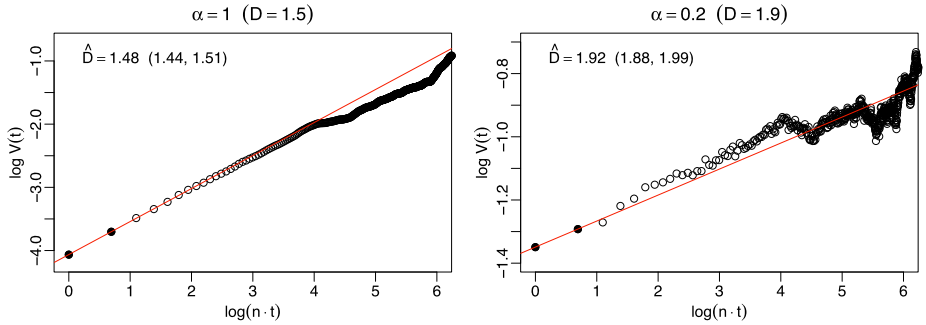

FIG. 9. *Log–log regression for the madogram estimator* [*variation estimator* (18) *with power index* $p = 1$] *and the datasets in the lower row of Figure* 1. *Only the two points marked with filled circles are used when fitting the regression line.*

independent of its order, we may refer to the common version as γ. If the expected number of sample path jumps per unit time is finite, then γ grows linearly at the coordinate origin (Masry, 1972). In this case, the relationship (4) holds if, and only if, the common variogram, γ, is understood to be of order $p = 1$. However, there are binary processes whose variogram behaves like $\gamma(t) = \mathcal{O}(|t|^{\gamma})$ as $t \to 0$, where $0 < \gamma < 1$, and then the relationship fails. Notwithstanding these examples, the argument of Bruno and Raspa (1989) and Bez and Bertrand (2011) is persuasive, and we maintain that the relationship (4) is particularly inclusive when $p = 1$. A natural conjecture is that if $p = 1$, then the relationship is valid if the process is ergodic (in a suitable sense) and the expected number of sample path discontinuities per unit time is finite. Furthermore, it is worth noting that there are processes for which the madogram exists and the foregoing is satisfied, while second moments do not exist (Ehm, 1981).

The following interesting connection between the Hall–Wood estimator and the madogram estimator also suggests a special role of the power index $p = 1$. For l a positive integer and $j = 0, 1, \ldots, l-1$, let

$$\widehat{A}^{(j)}(l/n) = \frac{l}{n} \sum_{i=1}^{\lfloor (n-j)/l \rfloor} |X_{(il+j)/n} - X_{(il+j-l)/n}|.$$

Then $\widehat{A}(l/n) = \widehat{A}^{(0)}(l/n)$ and

$$\widehat{V}_1(l/n) = \frac{1}{2} \frac{n}{n-1} \frac{1}{l} \sum_{j=0}^{l-1} \widehat{A}^{(j)}(l/n)$$

is, up to inessential constants, the mean of l distinct copies of $\widehat{A}(l/n)$. A comparison of the general forms (9) and (17), or the standard forms (10) and (18), of the Hall–Wood estimator with the madogram estimator suggests that the latter is a statistically more efficient version of the Hall–Wood estimator. A similar, more tedious calculation can be used to argue heuristically that the box-count estimator has a bias, typically leading to lower estimates of the fractal dimension than the Hall–Wood and variation estimators. For a confirmation in simulation studies, see Section 3.2.

Here we restrict attention to a small initial study that assesses the efficiency and outlier resistance of variation estimators. Figures 10 and 11 show the root mean squared error (RMSE) of the variation estimator (18) from Gaussian sample paths in dependence on the power index, p. The curves are computed from 1,000 independent realizations with sample size $n = 1{,}024$, correspond to fixed values of the fractal index, α, and have their minima marked. Figure 10 concerns the ideal Gaussian model, where the estimator performs best for power indices of about $p = 2$, corresponding to the variogram estimator, similarly to the observations of Coeurjolly [(2001), page 1417]. Figure 11 shows that the RMSE can deteriorate considerably under a single additive outlier, with the effect being stronger for smoother sample paths, that is, higher values of the fractal index. The smaller the power index, the more outlier resistant the variation estimator.

A possible variant of the variation estimator uses p-moments of higher increments, as proposed by Kent and Wood (1997) and Istas and Lang (1997) in the case $p = 2$. For example, one could turn to second differences, rather than first differences, and base a log–log regression on

$$\widehat{V}_p^{(2)}(l/n) = \frac{1}{2(n-2l)} \cdot \sum_{i=l}^{n-l} |X_{(i+l)/n} - 2X_{i/n} + X_{(i-l)/n}|^p, \tag{19}$$

rather than the standard variation (16). Also, if more than two points are used in the regression, the generalized least squares method could be used in lieu of

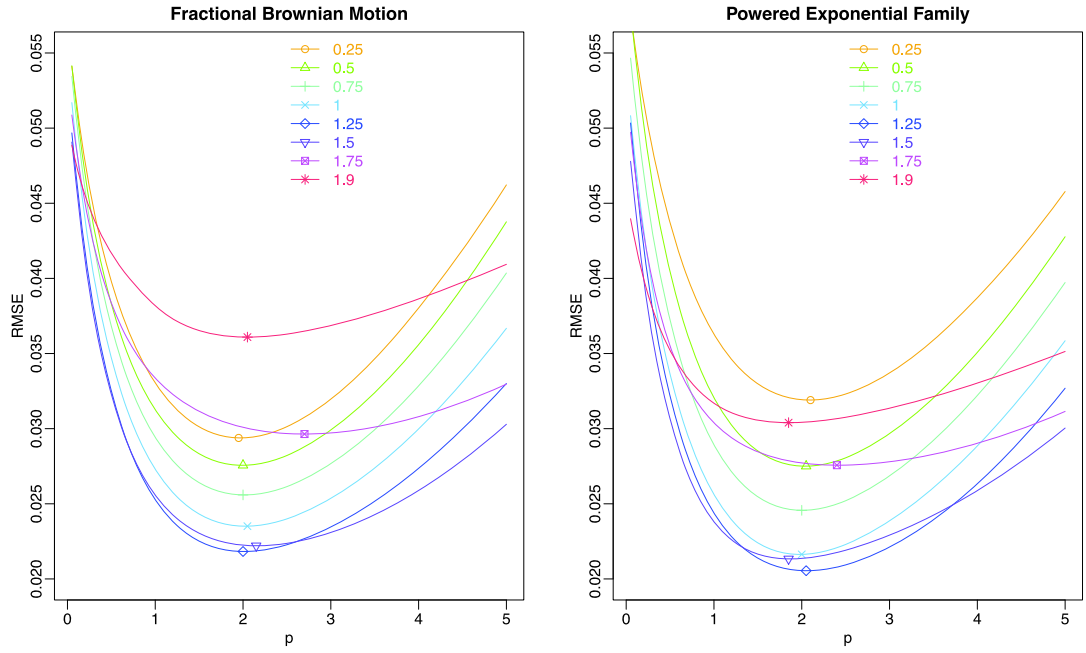

FIG. 10. *Root mean squared error (RMSE) of the variation estimator (18) as a function of the power exponent, p, for Gaussian fractional Brownian motion (left) and Gaussian processes with powered exponential covariance (right). The scale parameter used is $c = 1$, and each RMSE is computed from* 1,000 *Monte Carlo replicates of Gaussian sample paths under the sampling scheme (6), where $n = 1{,}024$. The curves correspond to fixed values of the fractal index, α, and have their minima marked.*

the ordinary least squares technique. However, there is no clear advantage in doing so in applied settings, in which the corresponding covariance structure is unknown and needs to be estimated as well.

2.5 Spectral and Wavelet Estimators

We now consider the semi-periodogram estimator of Chan, Hall and Poskitt (1995), which operates in the

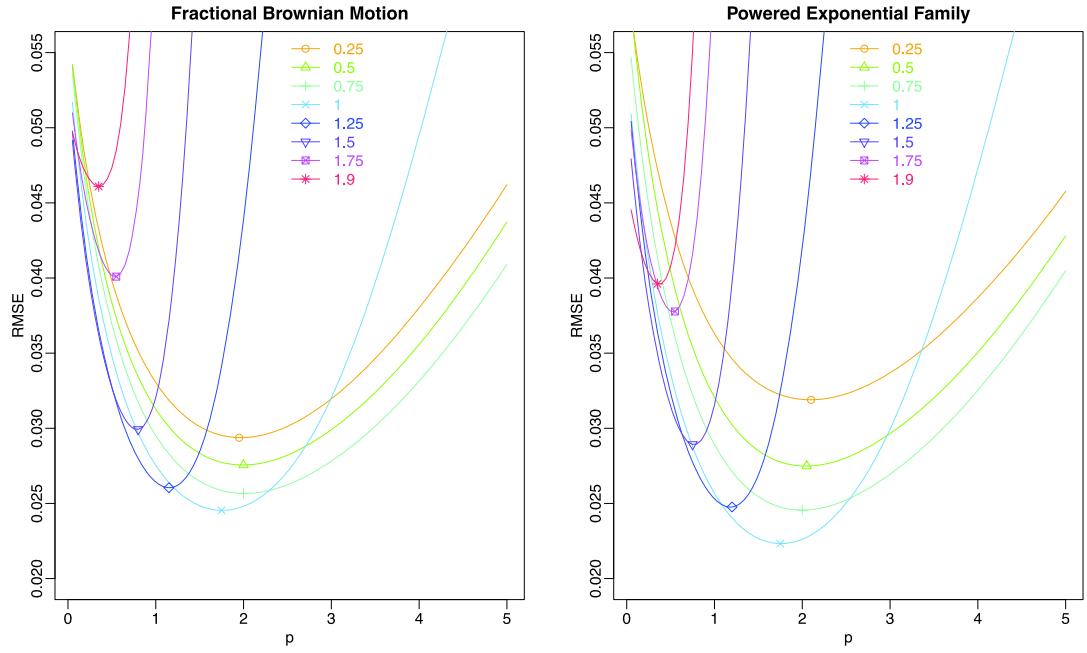

FIG. 11. *Same as Figure 10, except that in each sample path a randomly placed observation is contaminated by an additive Gaussian outlier with standard deviation* 0.1.

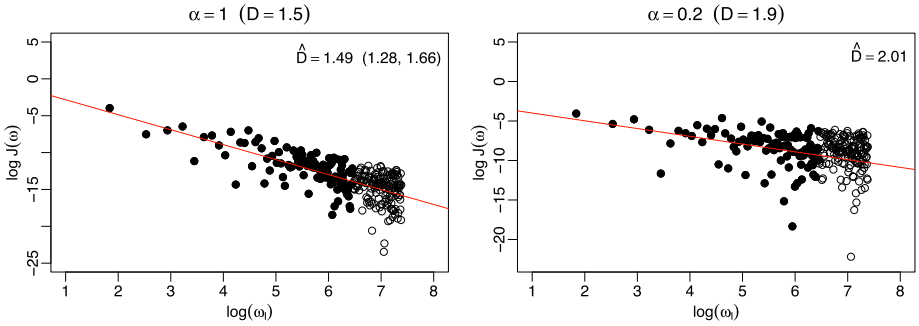

FIG. 12. *Log-log regression for the semi-periodogram estimator and the datasets in the lower row of Figure 1. Only the points marked with filled circles are used when fitting the regression line.*

frequency domain and is closely related to the spectral estimator of Dubuc et al. (1989a). The basis for this estimator is the well-known fact that the spectral density function for a stationary stochastic process with a second-order variogram of the form (3) decays like $|\omega|^{-\alpha-1}$ as frequency $|\omega| \to \infty$ (Stein, 1999). For a stationary Gaussian process $\{X_t : t \in [0,1]\}$, Chan, Hall and Poskitt (1995) defined

$$B(\omega) = 2\int_0^1 X_t \cos(\omega[2t-1])\,\mathrm{d}t$$

and called

$$J(\omega) = B(\omega)^2$$

the semi-periodogram. Under weak regularity conditions, the expected value of the semi-periodogram decays in the same way as the spectral density function. Suppose now that we have $n_s = 2m+1$ observations X_t at times $t = i/(2m)$, where $i = 0, 1, \ldots, 2m$. In this setting, Chan, Hall and Poskitt (1995) approximated $B(\omega)$ by

$$\widehat{B}(\omega) = \frac{1}{m}\left[\frac{X_0 + X_1}{2} + \sum_{i=1}^{2m-1} X_{i/(2m)} \cos\left(\omega\frac{i-m}{m}\right)\right]$$

and the semi-periodogram $J(\omega)$ by

$$\widehat{J}(\omega) = \widehat{B}(\omega)^2.$$

The semi-periodogram estimator of the fractal dimension D is

$$\widehat{D}_{\mathrm{P}} = \frac{5}{2} + \frac{1}{2}\left\{\sum_{l=1}^{L}(s_l - \bar{s})\log\widehat{J}(\omega_l)\right\} \cdot \left\{\sum_{l=1}^{L}(s_l - \bar{s})^2\right\}^{-1}, \tag{20}$$

where $\omega_l = 2\pi l$, $s_l = \log\omega_l$ and $\bar{s}$ is the mean of $s_1, \ldots, s_L$. The highest unaliased frequency (i.e., the Nyquist frequency) is πm, which is reflected by the fact that $\widehat{B}(\pi m + \delta) = \widehat{B}(\pi m - \delta)$ for any δ. This suggests setting $L = \lfloor m/2 \rfloor$, but Chan, Hall and Poskitt (1995) recommended using $L = \lfloor \min\{m/2, n_s^{2/3}\} \rfloor$, which is less than $\lfloor m/2 \rfloor$ for $m \geq 34$. As m grows, this rule thus has the effect of eliminating $\widehat{J}(\omega)$ at high frequencies in forming $\widehat{D}_{\mathrm{P}}$, which at first seems counterintuitive, given that the underlying scaling law applies as ω increases. However, as ω approaches the Nyquist frequency, aliasing causes the expectation of $\widehat{J}(\omega)$ to deviate markedly from the decay rate of $\omega^{-\alpha-1}$, leading to the need to eliminate high-frequency terms when forming $\widehat{D}_{\mathrm{P}}$. Figure 12 shows an example of the log–log regression for the semi-periodogram estimator for datasets from Figure 1, where $n_s = 1{,}025$ and hence $L = 101 < m/2 = 256$.

The definition of $\widehat{B}(\omega)$ is similar in spirit to the so-called type-II discrete cosine transform (DCT-II); see Ahmed, Natarajan and Rao (1974) and Strang (1999) for background. Davies (2001) noted that this transform has some attractive properties when used as a basis for spectral analysis, so it is of interest to explore the DCT-II as a substitute for the semi-periodogram in estimating fractal dimension. Given time series data of the form (6), taking the definition of the DCT-II given by Gonzalez and Woods (2007) and adjusting it for our convention for the samples $X_{i/(2m)}$ yields

$$\widetilde{B}(\omega) = \left(\frac{2}{2m+1}\right)^{1/2} \sum_{i=0}^{2m} X_{i/(2m)} \cos\left(\omega\frac{2i+1}{4m}\right)$$

and $\widetilde{J}(\omega) = \widetilde{B}(\omega)^2$. Here the Nyquist frequency is $2\pi m$, as can be seen by noting that $\widetilde{J}(2\pi m + \delta) = \widetilde{J}(2\pi m - \delta)$ for any δ. If we now let $\omega_l = 2\pi l m/(2m+1)$ with $s_l = \log\omega_l$ and $\bar{s}$ being the mean of the s_l's as before, the log–log regression estimator (20) can be applied with $\widehat{J}(\omega_l)$ replaced by $\widetilde{J}(\omega_l)$ and $L =$

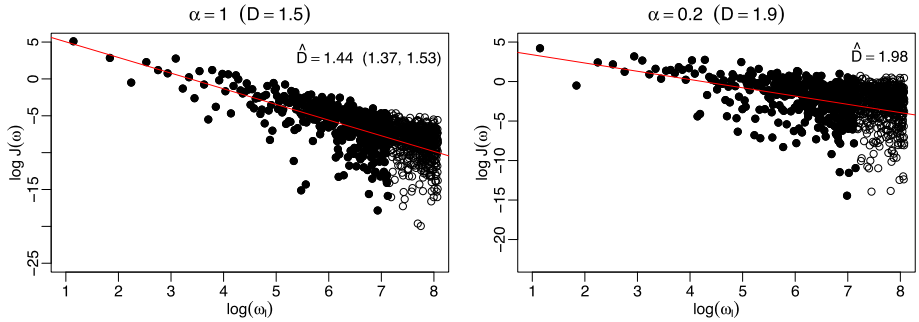

FIG. 13. *Log–log regression for the DCT-II estimator and the datasets in the lower row of Figure 1. Only the points marked with filled circles are used when fitting the regression line.*

$\lfloor \min\{2m, 4n_s^{2/3}\} \rfloor$. The DCT-II estimator uses approximately four times more points in the log–log regression than does the semi-periodogram estimator. For example, in Figure 13 we have $L = 406$, whereas we had $L = 101$ in Figure 12.

The semi-periodogram estimator also serves as motivation for a similar wavelet estimator, which is an adaptation of a weighted least squares estimator for the long memory parameter of a fractionally differenced process (Percival and Walden, 2000, Section 9.5). Given a time series of length n_s, we compute its maximal overlap discrete wavelet transform (MODWT) out to level $J_0 = \lfloor \log_2(n_s) \rfloor$ using reflection boundary conditions; this can be done using the function `modwt` in the `R` package `wavelets` (Aldrich, 2010). This MODWT yields J_0 vectors of wavelet coefficients $\widetilde{\mathbf{W}}_j$, $j = 1, \ldots, J_0$, each of which contains $2n_s$ coefficients. The coefficients in the jth vector are associated with the scale $\tau_j = 2^{j-1}$. The average of these coefficients squared, that is, $\|\widetilde{\mathbf{W}}_j\|^2/2n_s$, provides an estimator of the wavelet variance $\nu^2(\tau_j)$. This variance varies approximately as τ_j^α for large τ_j [Percival and Walden, 2000, equation (297b)], where α is the fractal index, from which the fractal dimension can be deduced; see Table 2. The scale τ_j corresponds to the band of frequencies $(\pi/2^j, \pi/2^{j-1}]$.

The information that is captured by the semi-periodogram at high frequencies is thus captured in the wavelet variance at small scales τ_j. Since Chan, Hall and Poskitt (1995) eliminated certain high frequencies in their semi-periodogram estimator, this suggests using just the wavelet variances indexed by $j = J_0, \ldots, J_1$, where $J_0 = \max\{1, \lfloor \log_2(n_s)/3 - 1 \rfloor\}$. Because the variance of the wavelet variance estimators depends upon τ_j, we replace the ordinary least squares estimator of the slope that is the basis for equation (20) with a weighted least squares estimator, say $\hat{\alpha}_{\mathrm{WL}}$ [Percival and Walden, 2000, equation (376c)]. The corresponding estimator of the fractal dimension D is $\widehat{D}_{\mathrm{WL}} = 2 - \frac{1}{2}\hat{\alpha}_{\mathrm{WL}}$. Figure 14 gives an example of the wavelet estimator of D. Note that the estimator of D in the right-hand plot is $\widehat{D}_{\mathrm{WL}} = 2.05$, which is greater than the upper limit $D = 2$ for the Hausdorff dimension of a curve. Sampling variability can cause this to happen on occasion with the other estimators also. In the simulation experiments reported below, the wavelet estimator proved to perform comparably to the DCT-II

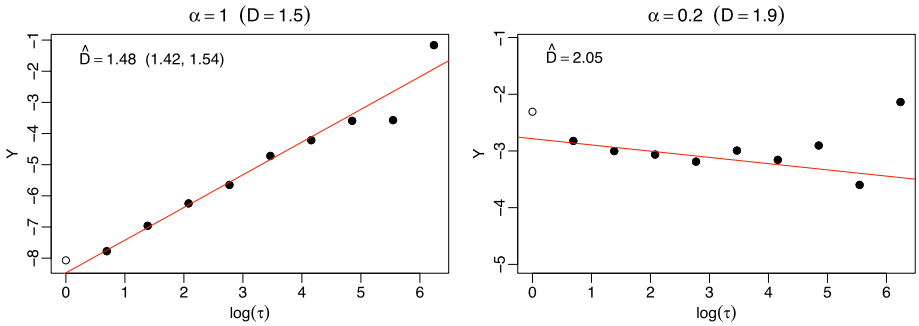

FIG. 14. *Log–log regression for the wavelet estimator and the datasets in the lower row of Figure 1. Only the points marked with filled circles are used when fitting the regression line.*

estimator, so we have chosen to drop the former and report only on the latter in what follows.

3. PERFORMANCE ASSESSMENT: TIME SERIES AND LINE TRANSECT DATA

We now turn to an evaluation of the various types of estimators, where we consider the large sample behavior under infill asymptotics and report on a finite sample simulation study that assesses both efficiency and robustness.

3.1 Asymptotic Theory

As noted, the fractal dimension refers to the properties of a graph in a hypothetical limiting process that might exist if the scale of measurement were to become infinitely fine. Hence, estimators of fractal dimension are studied under infill asymptotics (Hall and Wood, 1993; Stein, 1999), in which the number of observations grows to infinity, whereas the underlying domain, namely the unit interval, remains fixed. We assume that time series or line transect data of the form (6) arise from a Gaussian process $\{X_t : t \in [0,1]\}$ with a second-order structure of the type (3), where we let n grow without bounds. Typically, the literature assumes stationarity, so that the process $\{X_t : t \in [0,1]\}$ has a covariance function of the form

$$\sigma(t) = \sigma(0) - |ct|^{\alpha} + \mathcal{O}(|t|^{\alpha+\beta}) \quad \text{as } t \to 0,$$

where $\alpha \in (0,2)$, $\beta \geq 0$ and $c > 0$. The behavior of the bias, variance and mean squared error (MSE) of the estimators, and the corresponding types of limit distributions, then depend on the fractal index α and on β. Typically, the corresponding asymptotic results carry over to Gaussian processes with stationary increments and a variogram or structure function of the form (3).

For any Hall–Wood or variogram estimator $\widehat{D}$ of the form (9) or (12) with a fixed value of the design parameter L, the key results of Hall and Wood (1993) and Constantine and Hall (1994) are that

$$\text{MSE}(\widehat{D}) = \begin{cases} \mathcal{O}(n^{-1}) + \mathcal{O}(n^{-2\beta}), & \text{if } 0 < \alpha < \frac{3}{2}, \\ \mathcal{O}(n^{-1}\log n) + \mathcal{O}(n^{-2\beta}), & \text{if } \alpha = \frac{3}{2}, \\ \mathcal{O}(n^{2\alpha-4}) + \mathcal{O}(n^{-2\beta}), & \text{if } \frac{3}{2} < \alpha < 2, \end{cases} \tag{21}$$

where in each case the first term corresponds to the variance, and the second term to the squared bias. If $\alpha \leq 3/2$, then $\widehat{D}$ has a normal limit; if $\alpha > 3/2$, the limit is a Rosenblatt distribution as described by Taqqu (1975). In a recent far-reaching paper, Coeurjolly (2008) showed that in the Gaussian case and for the variation estimator (17) with general power index $p > 0$, the asymptotic behavior is still described by (21). Furthermore, the convergence rates are retained if the arithmetic mean in the definition of the power variation (16) is replaced by a trimmed mean, or by a convex combination of sample quantiles. While some of these results carry over to certain specific non-Gaussian processes (Chan and Wood, 2004; Achard and Coeurjolly, 2010), the limiting distribution theory is considerably richer then, and a general non-Gaussian theory remains lacking.

It is interesting to observe the change in the asymptotic rate of convergence at $\alpha = 3/2$ for all these types of estimators. However, Kent and Wood (1997) showed that the variogram estimator achieves an MSE of order

$$\text{MSE}(\widehat{D}) = \mathcal{O}(n^{-1}) + \mathcal{O}(n^{-2\beta}) \tag{22}$$

for all $\alpha \in (0,2)$ if the design parameter satisfies $L \geq 3$ and the generalized least squares technique, rather than the ordinary least squares method, is used in the log–log regression fit, and/or second differences of the form (19) are used. Similarly, Coeurjolly (2008) demonstrated that the asymptotic rate of convergence for the variation estimator with general power index $p > 0$ can be improved if second differences or related special types of increments are used. In finite sample simulation studies for the variogram estimator ($p = 2$), Kent and Wood (1997) did not find a clear-cut advantage in using generalized least squares and/or second differences, and our own experience with variation estimators of diverse power indices is similar. As Kent and Wood (1997) argued, the likely cause is that, the closer α is to 2, the larger n must be before the asymptotic regime is reached. This behavior is in marked contrast to the case of spatial lattice data, to be discussed below.

Chan, Hall and Poskitt (1995) developed asymptotic theory for the semi-periodogram estimator, but the MSE decays at best at rate $\mathcal{O}(n^{-1/4})$ in their results. The asymptotic scenario for the level crossing estimator in Feuerverger, Hall and Wood (1994) involves a bandwidth parameter and thus is not directly comparable.

3.2 Simulation Study: Gaussian Processes

We now turn to a simulation study, in which we confirm and complement the foregoing asymptotic results in a Gaussian setting. In doing so, exact simulation is critically important (Chan and Wood, 2000; Zhu and Stein, 2002), and we use the circulant embedding approach (Dietrich and Newsam 1993; Wood and Chan, 1994; Stein, 2002; Gneiting et al., 2006)

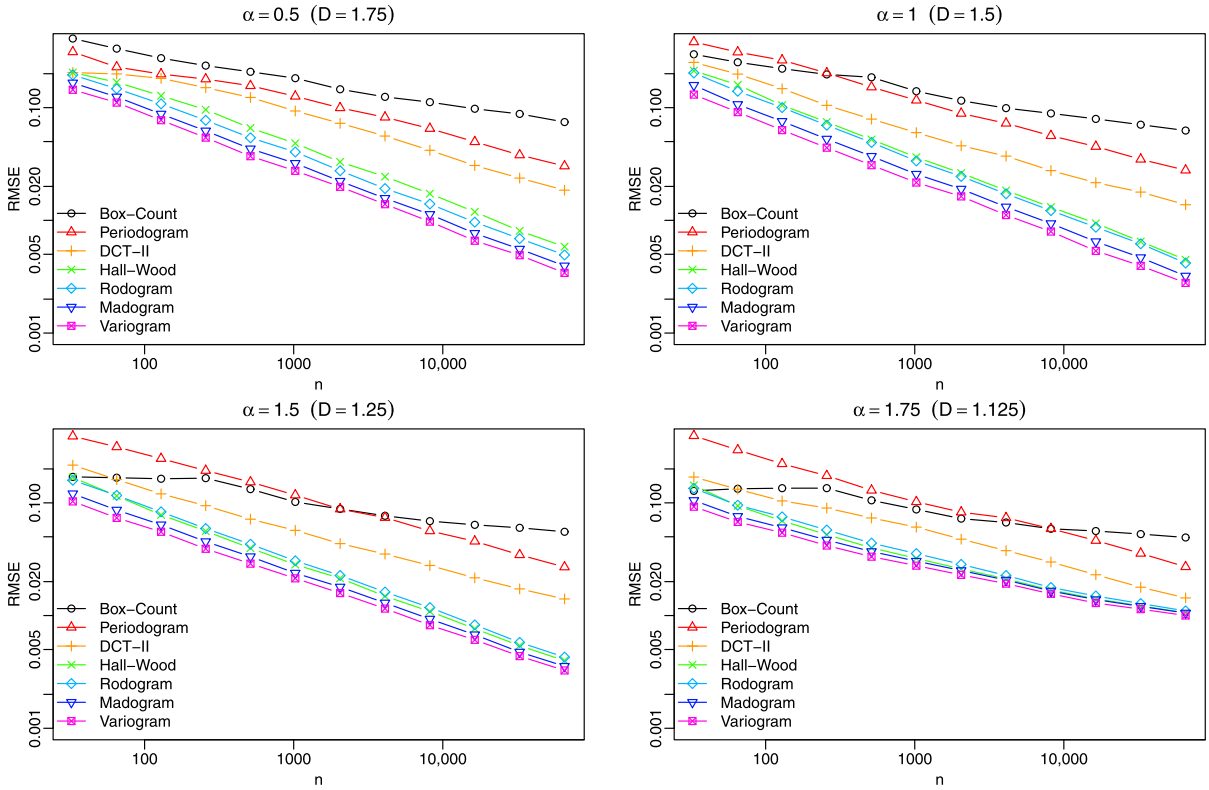

FIG. 15. *Root mean squared error (RMSE) of estimators of fractal dimension in dependence on the sample size, n, computed from Gaussian sample paths of the form (6) with powered exponential covariance function, $\sigma(t) = \exp(-|t|^\alpha)$. For each combination of α and n, the number of Monte Carlo replicates is* 1,000.

as implemented in the `R` package `RandomFields` (Schlather, 2001) to generate Gaussian sample paths, using the function `GaussRF`. The circulant embedding technique relies on the fast Fourier transform and is both exact and fast.

Figure 15 shows log–log plots for the root mean squared error (RMSE) of the various types of estimators in their dependence on the sample size n, computed from 1,000 independent trajectories of the form (6) from the corresponding stationary Gaussian process with a powered exponential covariance function, $\sigma(t) = \exp(-|t|^\alpha)$. The graphs are approximately linear, and their slopes show good agreement with the asymptotic laws in (21). Furthermore, they confirm the aforementioned observation of Kent and Wood (1997) that large values of the fractal index, α, require large sample sizes to reach the asymptotic regime. The variogram estimator generally shows the lowest MSE, followed by the madogram and rodogram, and then the Hall–Wood, DCT-II, periodogram and box-count estimators. This ranking is retained under Gaussian processes with covariance functions from the Matérn and Cauchy families, as well as for fractional Brownian motion, for all values of the fractal index α and all sufficiently large sample sizes, n. The use of variation estimators based on second differences, as defined in (19), typically does not yield lower RMSEs (results not shown).

Figures 16 and 17 show box- and scatterplots for the same types of estimators and the same class of Gaussian processes, where the sample size is $n = 1{,}024$. Three groups of estimators can be distinguished, the first comprising the variogram and other two variation estimators along with the Hall–Wood estimator, the second the spectral estimators, and the third the box-count estimator. The most efficient estimator is the variogram estimator, closely followed by the madogram estimator. As we have argued theoretically before, the madogram estimator is a more efficient version of the Hall–Wood estimator, in that the estimators are strongly correlated, but the former is less dispersed. While the spectral estimators are less competitive, showing substantially higher dispersion than the variation estimators, the DCT-II estimator improves considerably on the periodogram estimator. The box-

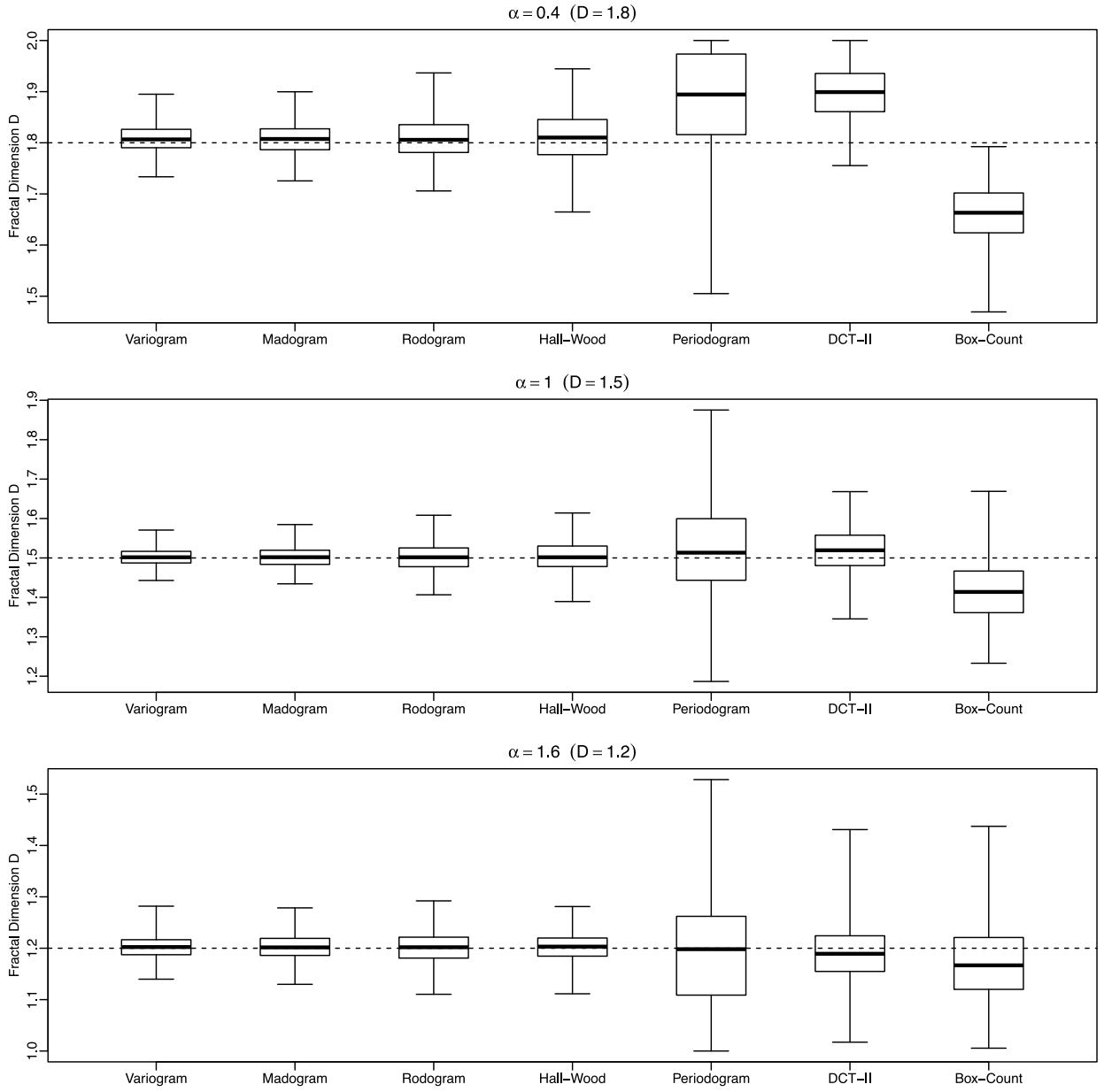

FIG. 16. *Boxplots for estimates of the fractal dimension from Gaussian sample paths of the form* (6), *where* $n = 1{,}024$, *with powered exponential covariance function*, $\sigma(t) = \exp(-|t|^{\alpha})$, *and the fractal index*, α, *being equal to* 0.4, 1.0 *and* 1.6, *respectively. The corresponding true values of the fractal dimension*, D, *namely* 1.8, 1.5 *and* 1.2, *are shown as dashed lines. The number of Monte Carlo replicates is* 500.

count estimator generally shows a bias, with the estimates being too low.

Figure 18 illustrates these results in a further experiment, in which we consider a Gaussian sample path with the exponential covariance function, $\sigma(t) = \exp(-|t|)$, and estimate the fractal dimension along sliding blocks of size 1,024. The corresponding estimates are plotted at the midpoint of the sliding block. It is clearly seen that the variogram and other variation estimators are the least dispersed, and that the madogram estimator is a more efficient version of the Hall–Wood estimator. The spectral estimators are the most dispersed, with the DCT-II estimator outperforming the periodogram estimator, and the box-count estimator is biased.

Thus far in this section, we have studied the efficiency of the estimators under an ideal Gaussian process assumption. However, robustness against deviations from Gaussianity is a critically important requirement on any practically useful estimator of the fractal

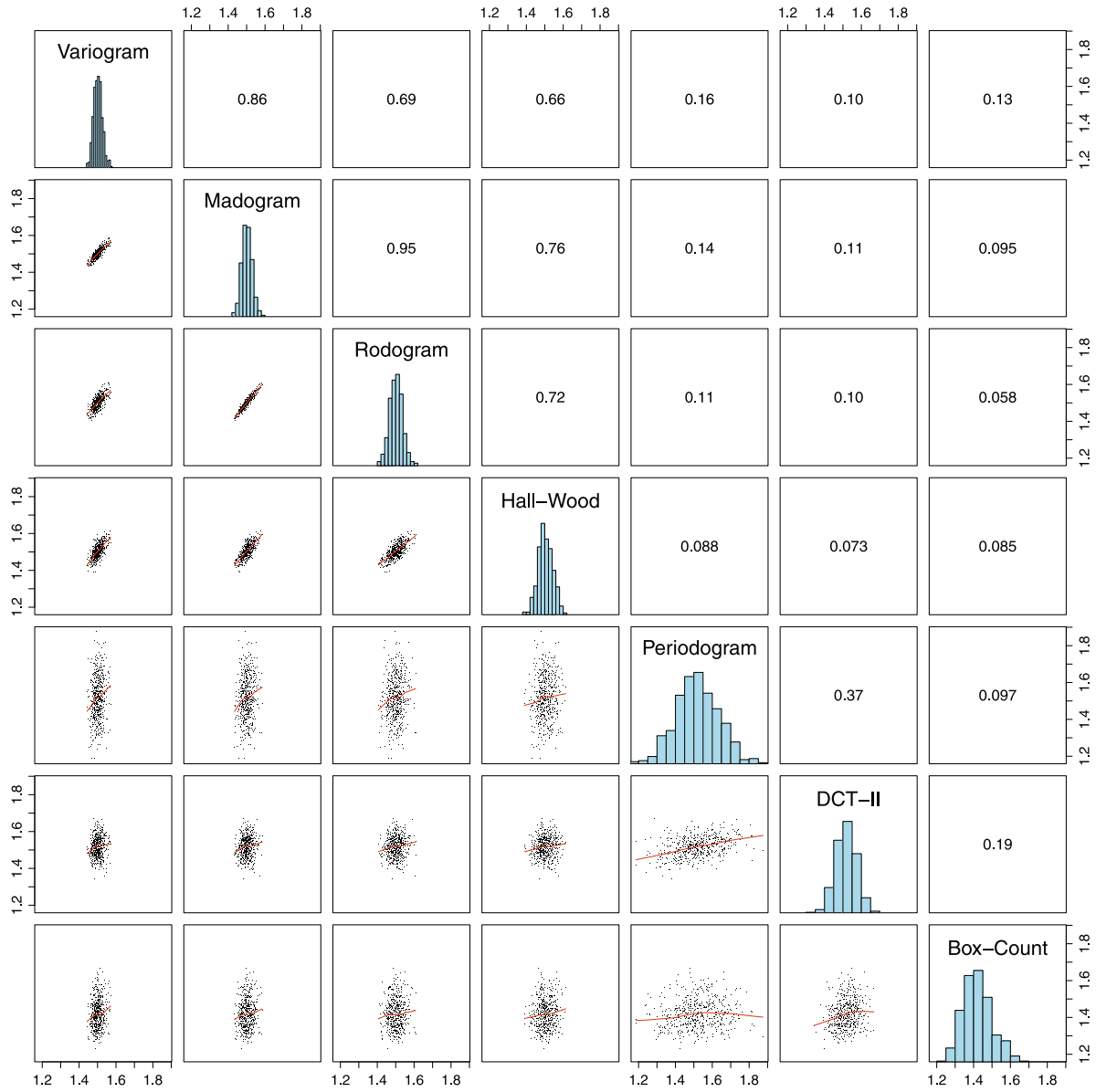

FIG. 17. *Scatterplot matrix for estimates of the fractal dimension from Gaussian sample paths of the form* (6) *with exponential covariance function,* $\sigma(t) = \exp(-|t|)$, *and sample size* $n = 1{,}024$. *The true value of the fractal dimension is* $D = 1.5$. *The panels along the diagonal show histograms of the estimates, and those above the diagonal pairwise Pearson correlation coefficients. The number of Monte Carlo replicates is* 500.

dimension. In this light, we now expand our simulation study, and consider a situation in which Gaussian sample paths are contaminated by additive outliers. Specifically, given a sample path of the form (6), we let i be discrete uniform on $\{0, 1, \ldots, n\}$ and replace $X_{i/n}$ by $X_{i/n} + y$, where y is normal with mean zero and standard deviation 0.1 and independent of i. This process is repeated to obtain the desired number of outliers.

Figure 19 shows RMSEs from such an experiment, using sample size $n = 1{,}024$, five additive outliers, and the powered exponential covariance function, $\sigma(t) = \exp(-|t|^\alpha)$, with values of the fractal index α that nearly span the full range from 0 to 2. Not surprisingly, the results resemble those in Figures 10 and 11, which considered variation estimators and the case of a single outlier only, and echo the findings of Achard and Coeurjolly (2010). Amongst the variation estimators considered here, the most outlier resistant is the rodogram estimator ($p = 1/2$), but the box-count estimator, which performs poorly overall when there are

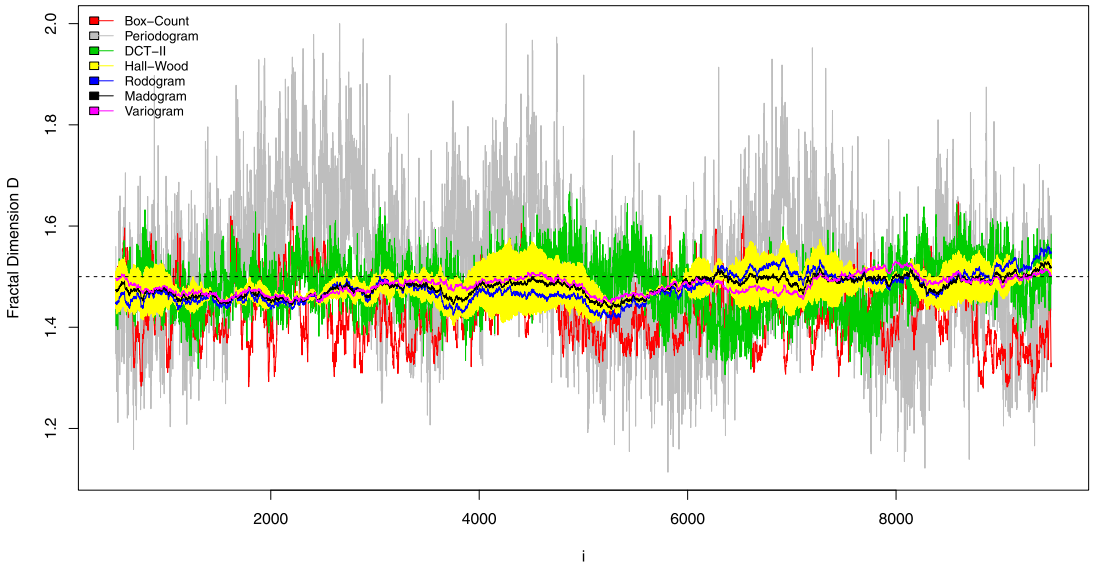

FIG. 18. *Estimates of the fractal dimension plotted at the midpoints of a sliding estimation window of size* 1,025, *for a Gaussian sample path of the form* (6), *where* $n = 10{,}000$, *with exponential covariance function,* $\sigma(t) = \exp(-|t|)$. *The true fractal dimension,* $D = 1.5$, *is marked by the dashed line. The label on the horizontal axis,* i, *indicates the midpoint of the sliding estimation block, at* $i/10{,}000$.

no outliers, becomes competitive at the highest α considered (1.9).

3.3 Discussion

The foregoing results and arguments lead us to a recommendation for practitioners, in that we join Bruno and Raspa (1989) and Bez and Bertrand (2011) and call for the use of the madogram estimator, that is, the variation estimator with power index $p = 1$. The madogram estimator can be interpreted as a statistically superior version of the Hall–Wood estimator, is simultaneously more outlier resistant and more efficient than

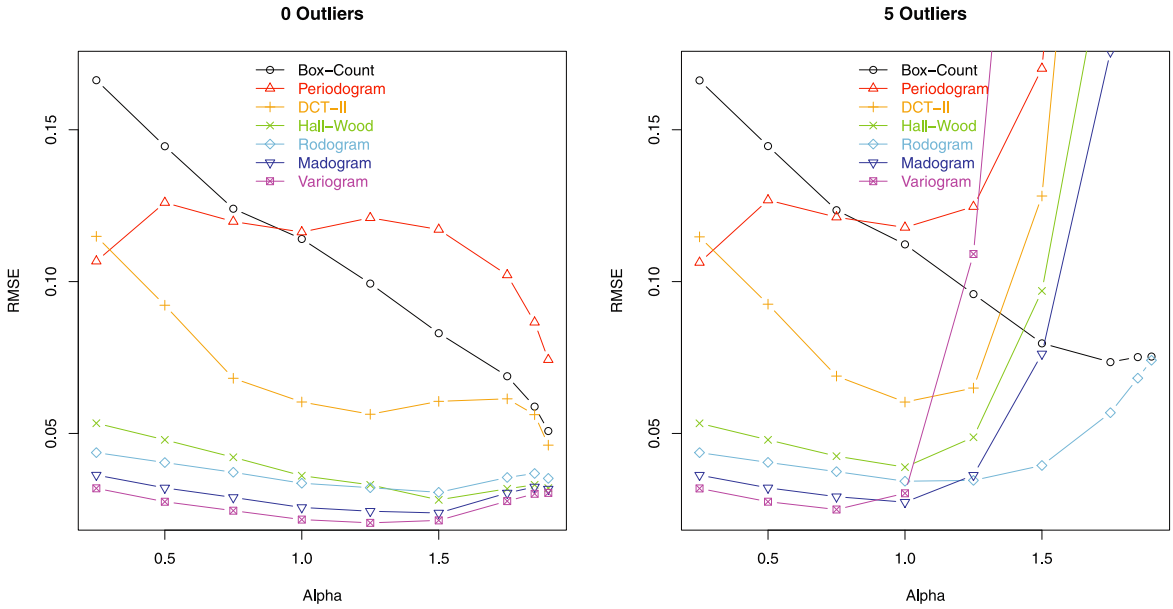

FIG. 19. *Root mean squared error (RMSE) of estimators of fractal dimension in dependence on the fractal index,* α, *for Gaussian sample paths of the form* (6) *with sample size* $n = 1{,}024$ *and powered exponential covariance function,* $\sigma(t) = \exp(-|t|^{\alpha})$. *The panel on the left corresponds to the ideal Gaussian process setting; the panel on the right to a situation with five additive outliers in each sample path. The number of Monte Carlo replicates is* 1,000.

many of its competitors, and has strong intuitive appeal. Importantly, the critical relationship (4) between the fractal index of a stochastic process whose variogram of order p shows a behavior of the form (5) at the origin, and the fractal dimension of its sample paths, is apparently valid for a larger class of non-Gaussian processes when $p = 1$ than for certain other choices of p (in particular, $p = 2$), thereby justifying the use of the madogram estimator for both Gaussian and non-Gaussian stochastic processes. Its resistance to outliers can be enhanced further if the arithmetic mean in the definition of the power variation (16) is replaced by a trimmed mean, as proposed by Coeurjolly (2008).

4. ESTIMATING THE FRACTAL DIMENSION OF SPATIAL DATA

We now turn to estimators of the fractal dimension of spatial data, as discussed by Dubuc et al. (1989b), Constantine and Hall (1994), Davies and Hall (1999), Chan and Wood (2000) and Zhu and Stein (2002), among other authors. Burrough (1981) noted a wealth of applications to landscape and other environmental data, with those of Rothrock and Thorndike (1980) on the underside of sea ice, and Goff and Jordan (1988) on the topography of the sea floor, being particularly interesting examples.

From a probabilistic point of view, a natural initial question is for the theoretical relationship between the fractal dimension of a surface indexed in $\mathbb{R}^2$, and the fractal dimension of its sections or line transects. Assuming stationarity of the spatial random field and additional (weak) regularity conditions, Hall and Davies (1995) showed that the fractal dimensions along line transects are all identical to one another, except that in one special direction the dimension may be less than in all others. Very general results that do not depend on the stochastic process setting are available from the fundamental work of Marstrand (1954). We join Davies and Hall (1999) in arguing that these results provide substantial support for the use of fractal dimension as a canonical measure of surface roughness. In particular, they allow us to estimate the fractal dimension of a surface by adding 1 to any estimate of the fractal dimension of the corresponding line transects.

Technically, we focus discussion on the situation in which a spatial stochastic process, indexed by the unit square in $\mathbb{R}^2$, is sampled on a regular lattice, to yield a surface graph of the form

$$\left\{(t, X_t) : t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{1}{n}\begin{pmatrix} i_1 \\ i_2 \end{pmatrix}, i_1 = 0, 1, \ldots, n, \; i_2 = 0, 1, \ldots, n\right\} \subset \mathbb{R}^3. \tag{23}$$

Before reviewing estimators of fractal dimension studied in the extant literature, and introducing new estimators, we propose a simple, unified notation. Specifically, for $k > 0$ we let

$$S(k) = \left\{(i_1, i_2, j_1, j_2) \in \{0, 1, \ldots, n\}^4 : \left|\begin{pmatrix} i_1 \\ i_2 \end{pmatrix} - \begin{pmatrix} j_1 \\ j_2 \end{pmatrix}\right| = k\right\},$$

and denote the cardinality of this set by $N(k)$. If $N(k) > 1$, we refer to k as a relevant distance. The estimators in the subsequent Sections 4.1 to 4.2 then take the form

$$\widehat{D} = 2 - \frac{1}{p}\left\{\sum_{k\in\mathcal{K}}(s_k - \bar{s}) \log \widehat{V}_p(k/n)\right\} \cdot \left\{\sum_{k\in\mathcal{K}}(s_k - \bar{s})^2\right\}^{-1}, \tag{24}$$

where $\mathcal{K}$ is a finite collection of relevant distances, $s_k = \log(k/n)$, $\bar{s}$ is the mean of $\{s_k : k \in \mathcal{K}\}$, and $\widehat{V}_p(k/n)$ is a certain variation with general power index $p > 0$.

Two-dimensional geometry allows for many options in the choice of the distance representatives and the variation, and we restrict attention to the most plausible and best performing estimators in the literature, all of which are based on power variations (Davies and Hall, 1999; Chan and Wood, 2000; Zhu and Stein, 2002). As concerns the set $\mathcal{K}$ of distance representatives, simulation experiments, experience in the line transect case, and the work of Chan and Wood (2000) and Zhu and Stein (2002) all suggest that a restriction to the smallest relevant distances only tends to lead to the best performance. In addition to minimizing the bias of the estimator, this strategy keeps the computational complexity low as well.

4.1 Isotropic Estimator

Davies and Hall (1999) considered an estimator based on the isotropic empirical variogram, which we now generalize. For a relevant distance k, consider the variation

$$\widehat{V}_{\mathrm{ISO};p}(k/n) = \frac{1}{2N(k)} \cdot \sum_{S(k)} |X_{i_1/n, i_2/n} - X_{j_1/n, j_2/n}|^p \tag{25}$$

with general power index $p > 0$. The isotropic estimator $\widehat{D}_{\mathrm{ISO};p}$ with power index p then is defined by (24) with the set $\mathcal{K} = \{1, \sqrt{2}, 2\}$ of distance representatives and the variation $\widehat{V}$ given by (25). Thus, we consider variations at horizontal and vertical spacings of one and two grid points ($k = 1$ and $k = 2$), and a diagonal spacing of a single grid point ($k = \sqrt{2}$), respectively.

4.2 Filter Estimator

Zhu and Stein (2002) studied a broad range of increment-based estimators, among which the "Filter 1" estimator shows good performance. We generalize by defining a filter estimator with general power index $p > 0$, rather than just $p = 2$ as in the work of Zhu and Stein (2002). Specifically, for a relevant distance $k > 0$ let

$$\begin{aligned}\widehat{V}_{\mathrm{F};p}(k/n) = {} & \frac{1}{2N(k)} \\ & \cdot \sum_{S(k)} |X_{i_1/n,i_2/n} \\ & - 2X_{(i_1+j_1)/(2n),(i_2+j_2)/(2n)} \\ & + X_{j_1/n,j_2/n}|^p.\end{aligned} \tag{26}$$

The filter estimator $\widehat{D}_{\mathrm{F};p}$ with power index p then is defined by (24) with the set $\mathcal{K} = \{2, 2\sqrt{2}, 4\}$ of distance representatives and the variation $\widehat{V}$ given by (26). This considers variations at horizontal and vertical spacings of one and two grid points ($k = 1$ and $k = 2$), and a diagonal spacing of a single grid point ($k = \sqrt{2}$), respectively. Hence, the filter estimator is the natural equivalent of the isotropic estimator $\widehat{D}_{\mathrm{ISO};p}$, but now using second differences, rather than first differences.

4.3 Square Increment Estimator

The square increment estimator is based on a proposal of Chan and Wood (2000), who restricted attention to quadratic variations. Here we define a square increment variation of general power index $p > 0$, namely,

$$\begin{aligned}\widehat{V}_{\mathrm{SI};p}(k/n) = {} & \frac{1}{2N(k)} \\ & \cdot \sum_{S(k)} |X_{i_1/n,i_2/n} - X_{i_1/n,j_2/n} \\ & - X_{j_1/n,i_2/n} + X_{j_1/n,j_2/n}|^p,\end{aligned} \tag{27}$$

where k is a relevant distance. The square increment estimator $\widehat{D}_{\mathrm{SI};p}$ then is defined by (24) with the set $\mathcal{K} = \{\sqrt{2}, 2\sqrt{2}\}$ of distance representatives, corresponding to squares that have side widths of one and two grid points, and the variation $\widehat{V}$ given by (27).

4.4 Transect Estimators

Finally, we consider two very simple estimators that are based on the variation estimator $\widehat{D}_{\mathrm{V};p}$ with general power index $p > 0$ of Section 2.4, or a variant that uses second differences, as defined in equation (19). In either case, a line transect estimate of the fractal dimension is computed for each row and each column in the grid. The transect-variation and transect-increment estimators $\widehat{D}_{\mathrm{TV};p}$ and $\widehat{D}_{\mathrm{TI};p}$ with power index p then add 1 to the median of the $2n$ corresponding line transect estimates. In the former case, the line transect estimates are based on first differences, in the latter on second differences.

5. PERFORMANCE ASSESSMENT: SPATIAL DATA

We now assess the estimators by their large sample behavior under infill asymptotic as well as in finite sample simulation studies, considering both efficiency and robustness.

5.1 Asymptotic Theory

Davies and Hall (1999), Chan and Wood (2000) and Zhu and Stein (2002) developed asymptotic theory for a very wide range of estimators of the form (24) that are based on variations with power index $p = 2$. Generally, their results apply under an infill asymptotic scenario for sample paths of the form (23) from an intrinsically stationary Gaussian spatial process with fractal index $\alpha \in (0, 2)$ and a variogram that behaves like (3) at the origin. For estimators that are based on variations corresponding to a first difference, the generic result is that

$$\begin{aligned}&\mathrm{MSE}(\widehat{D}) \\ &= \begin{cases} \mathcal{O}(n^{-2}) + \mathcal{O}(n^{-4\beta}), & \text{if } 0 < \alpha < 1, \\ \mathcal{O}(n^{-2}L(n)) + \mathcal{O}(n^{-4\beta}), & \text{if } \alpha = 1, \\ \mathcal{O}(n^{2\alpha-4}) + \mathcal{O}(n^{-4\beta}), & \text{if } 1 < \alpha < 2, \end{cases}\end{aligned} \tag{28}$$

where L is a function which is slowly varying at infinity. If $\alpha \leq 1$, then $\widehat{D}$ has a normal limit, while, if $\alpha > 1$, the limit is related to a Rosenblatt distribution, with some of these results carrying over to certain specific non-Gaussian processes (Chan and Wood, 2004). However, if the variations correspond to a second difference, such as in the cases of the filter and square increment estimators, and/or the generalized least squares techniques, rather than the ordinary least

squares method, is used, an improved asymptotic behavior, namely,

$$\text{MSE}(\widehat{D}) = \mathcal{O}(n^{-2}) + \mathcal{O}(n^{-4\beta}) \tag{29}$$

for all $0 < \alpha < 2$, is achieved, with an associated limit distribution that is normal. For regularity conditions and further detail, we refer to the original work of Davies and Hall (1999), Chan and Wood (2000), Zhu and Stein (2002) and Chan and Wood (2004), which is impressive. While these authors restricted attention to quadratic variations with power index $p = 2$ only, we conjecture, based on the work of Guyon and León (1989), Istas and Lang (1997) and Barndorff-Nielsen, Corcuera and Podolskij (2009) on the power variations of Gaussian processes, that analogous results continue to hold under a general power index $p > 0$, similar to the line transect case studied by Coeurjolly (2008).

As the amount of data in the sampling scheme (23) is about n^2, the asymptotic rates of convergence in (28) and (29) conform with those in the time series or line transect case, except that the transition from the classical to slower rates of convergence occurs already at $\alpha = 1$ (or $D = 3/2$), rather than at $\alpha = 3/2$ (or $D = 7/4$). Hence, these results suggest that there may be a higher benefit to using increments that are based on second differences in the spatial case than in the time series or line transect case.

5.2 Simulation Study: Gaussian Spatial Processes

In the subsequent simulation study, we use state-of-the-art implementations of the circulant embedding method (Stein, 2002; Gneiting et al., 2006) to generate Gaussian sample surfaces. Figure 20 shows the root mean squared error (RMSE) of the estimators versus the (square root of the) sample size, n, using data of the form (23) from stationary spatial processes with powered exponential covariance function, $\sigma(t) = \exp(-|t|^\alpha)$. The estimators use the traditional power index $p = 2$, and the number of Monte Carlo replicates is 1,000.

The asymptotic rates of convergence in (28) and (29) suggest linear graphs with slope -1 on the logarithmic scale for the filter and square increment estimators and all values of the fractal index, α. For the isotropic estimator, they suggest slope -1 for $\alpha \leq 1$, slope $-1/2$ for $\alpha = 3/2$, and slope $-1/4$ for $\alpha = 7/4$. Our empirical results are in good agreement with the theoretical slopes, and attest to the superior performance of

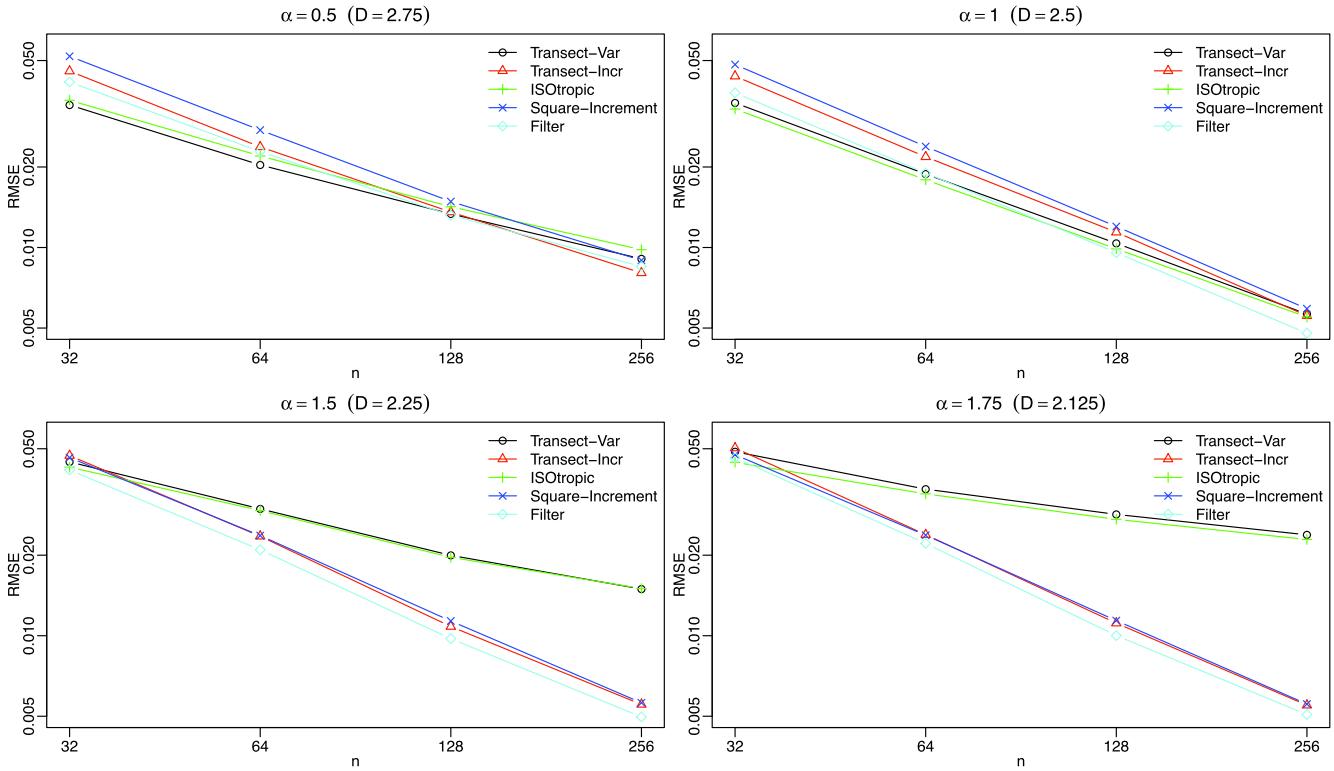

FIG. 20. *Root mean squared error (RMSE) for estimators of fractal dimension for data of the form (23) from spatial stochastic processes with powered exponential covariance, $\sigma(t) = \exp(-|t|^\alpha)$, versus the (square root of the) sample size, n. The power index used is $p = 2$, and for each combination of α and n, the number of Monte Carlo replicates is* 1,000.

the filter estimator in the ideal Gaussian process setting, as noted by Zhu and Stein (2002). Also, these and other simulation results lead us to conjecture that the transect-variation estimator behaves like (28), while the transect-increment estimator shares the favorable uniform asymptotic rate of convergence in (29).

In Figure 21 we consider estimators with general power index $p > 0$, but fix $n = 256$ in the spatial sampling scheme (23). Again, we use the powered exponential covariance model, and the number of Monte Carlo replicates is 1,000, each comprising a total of $(n+1)^2 = 66{,}049$ observations within the unit square. The left column shows the RMSE in the ideal Gaussian process setting, in which the filter, square increment and transect-increment estimators perform best. Furthermore, the efficiency of these estimators depends only very little on the choice of the power index.

Figure 21 also studies the behavior of the estimators in the presence of outliers. Specifically, the middle and right-hand columns show RMSEs in situations in which the Gaussian sample paths have been contaminated by 10 and 20 additive outliers, respectively, in ways essentially identical to those described in Section 3.2. Note that the vertical scale differs from row to row, with the largest RMSEs corresponding to the smoothest surfaces. The smoother the surface, that is, the larger the value of the fractal index, α, the more impact the outliers have on the estimators. Two of the estimators that dominate in the uncontaminated setting, namely the filter and the square increment estimators, are the least outlier resistant, even though the outlier fraction is very small in our study, at 0.015 and 0.03 percent, respectively. The most resistant estimators are the transect estimators.

These results allow for an interpretation in terms of breakdown points. While comprehensive formal definitions of breakdown points for dependent data have recently become available (Genton and Lucas, 2003), it suffices here to take a heuristic point of view, and define the breakdown point of an estimator as the minimal fraction of contaminated data that can ruin an estimator. Evidently, the isotropic, filter and square increment estimators have breakdown point zero. In contrast, the transect estimators have a positive breakdown point of about $1/(2\sqrt{m})$ under the spatial sampling scheme (23), where $m = n^2$ is the approximate amount of data. To see this, note that each outlier affects the individual estimate for at most two of the $2n$ transects, from which the median is formed. Thus, a transect estimator cannot be ruined, unless a fraction of at least $(n/2)/n^2$, or $1/(2\sqrt{m})$, of the data are contaminated.

5.3 Discussion

While confirming extant theoretical and simulation results for spatial data, which suggest the use of variations that are based on second differences, the results of this section lead us to two novel insights.

Thus far, the statistical literature has restricted attention to variation-based estimators with power index $p = 2$. A first observation is that the efficiency of these estimators depends little on the power index, $p > 0$. Outlier resistance and robustness arguments then suggest the use of smaller power indices, with Section 2.4 supporting the choice of $p = 1$.

A second and potentially very surprising insight is the superior performance of the transect-variation and transect-increment estimators. These estimators have positive breakdown points and outperform the traditional estimators even under minimal deviations from the ideal Gaussian process setting. Indeed, in practice, we would expect much larger deviations than in our simulation setting, in which the outliers had substantially lower variability than the process itself, and occurred at fractions of about 1 in 6,000 and 1 in 3,000 only. As the transect-increment estimator appears to share the favorable uniform rate of convergence (29) with the best performing extant estimators, this suggests the availability of an estimator, namely the transect-increment estimator with power index $p = 1$, that is both robust and efficient. We believe that these are highly promising prospects that deserve further study.

If spatial data are observed at irregularly scattered locations, rather than a regular grid, the only available estimator is a suitably modified version of the isotropic estimator (25), with $S(k)$ now representing pairs of stations that are approximately a distance k apart. In such cases, the use of the power index $p = 1$ again seems prudent.

6. DATA EXAMPLE: ARCTIC SEA-ICE PROFILES

In this penultimate section, we estimate the fractal dimension of arctic sea-ice profiles based upon measurements of sea-ice draft (93% of thickness). These profiles can be regarded as line transects through the underwater surface of sea ice because they were collected using upward-looking sonars by submarines traveling under the sea ice with no appreciable deviations from a single direction and depth. The data used here were sampled one meter apart at a resolution of one meter by the U.S. Navy in August 1998 in the Arctic Ocean, and are available online from the National

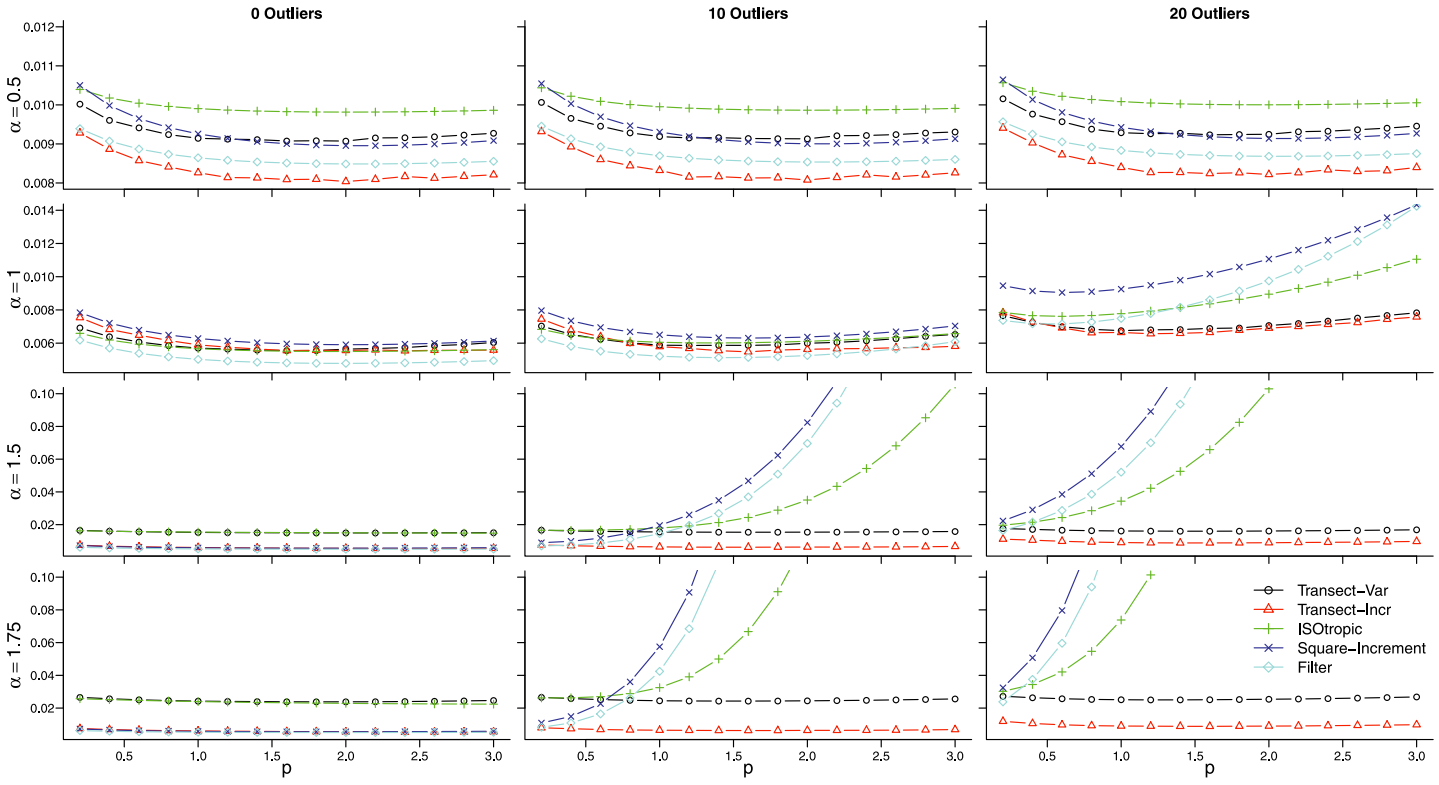

FIG. 21. *Root mean squared error of isotropic, filter, square increment, transect-variation and transect-increment estimators of fractal dimension in dependence on the power index, p. The underlying sample paths are of the form (23) with $n = 256$ from Gaussian spatial processes with powered exponential covariance, $\sigma(t) = \exp(-|t|^{\alpha})$. The columns correspond to situations with no outliers (left), 10 additive outliers (middle) and 20 additive outliers (right), respectively. The rows are for distinct values of the fractal index, namely $\alpha = 0.5, 1.0, 1.5$ and 1.75. The number of Monte Carlo replicates is* 1,000.

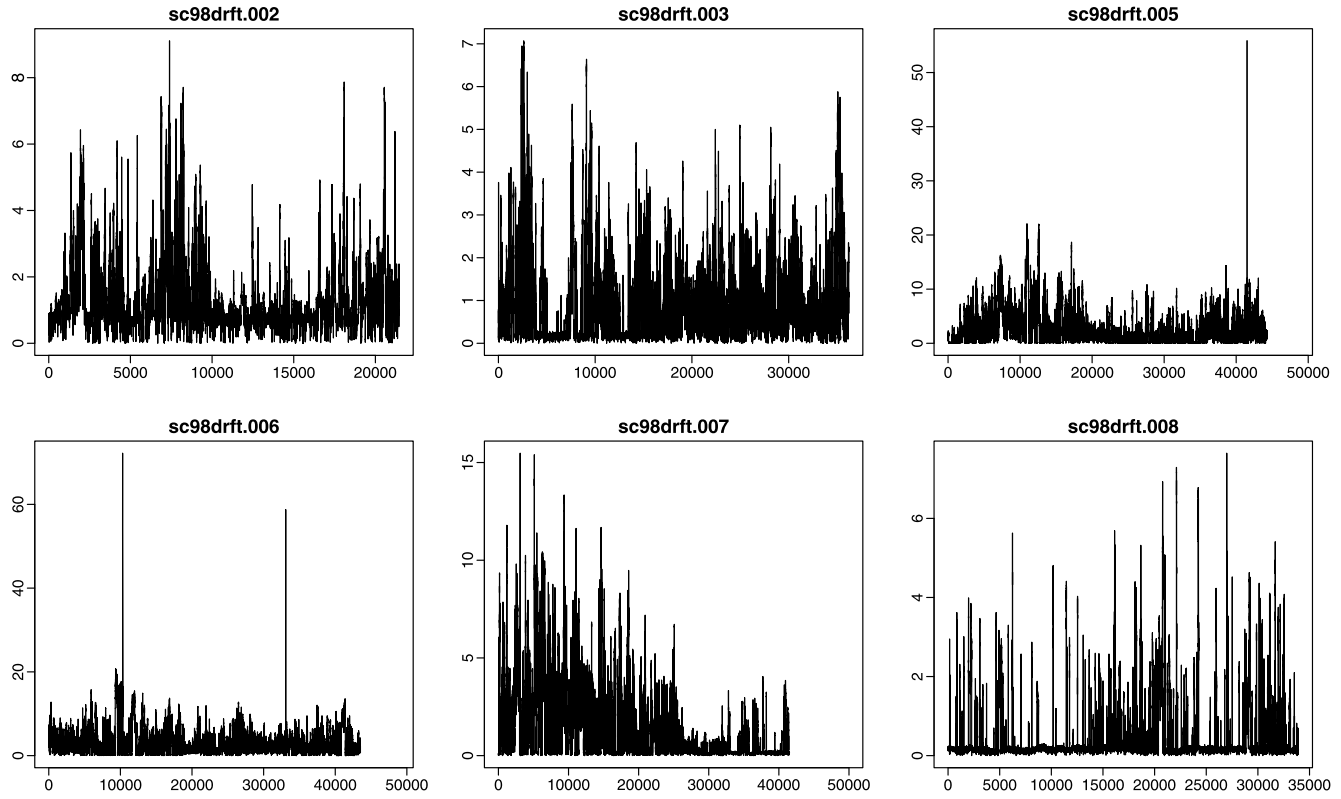

FIG. 22. *Arctic sea-ice profiles, with the panels showing files sc*98*drft*.002, 003 *and* 005–008, *respectively, and the labels being in meters. See the text for details.*

Snow and Ice Data Center at http://nsidc.org/data/docs/noaa/g01360_upward_looking_sonar. We examined six profiles of about 240 km total length (files sc98drft.002, 003 and 005–008). These profiles are illustrated in Figures 22 and 23 and show pronounced non-Gaussian features. Given the resolution, we join the extant literature in using fractal dimension to characterize the surface roughness of the macroscopic, topographic structure of sea ice, so that meters correspond to sufficiently small scales.

We use a sliding estimation window of width $n = 1{,}024$ meters, move these blocks along the profiles in increments of 10 meters, and estimate fractal dimension for each of them, using variation estimators with power indices $p = 2$ (variogram), 1 (madogram) and $1/2$ (rodogram) along with the Hall–Wood estimator. Thus, for each method there are 23,303 blocks in total, with Figure 23 showing examples of profiles and dimension estimates. Figures 24 and 25 show the corresponding boxplots, histograms and pairwise scatter diagrams, composited over all blocks. While all four methods result in similar estimates, the strongest correlation is, not surprisingly, between the Hall–Wood and madogram estimators, with the latter being our estimator of choice.

Overall, the Arctic sea-ice profiles have Hausdorff dimension of about 1.3 along line transects, and thus of about 2.3 spatially. These estimates are at the lower end of results reported in the literature. For instance, Bishop and Chellis (1989) argued that profiles of ice keels have fractal dimensions ranging from 1.2 to 1.7, while Connors, Levine and Shell (1990) reported estimates of about 1.4 and 1.6 for first year and multiyear ice segments, respectively. Further results and background information can be found in the works of Rothrock and Thorndike (1980) and Goff et al. (1995).

7. CONCLUDING REMARKS

In closing this review, we return to the opening quote of Lenny Smith [(2007), page 115] that precedes the abstract of our paper. Indeed, healthy skepticism about dimension estimates based on real-world data is well justified, in that estimates of fractal dimension depend on the availability of data at suitable scales and face issues of discreteness, and the effects of measurement error might also need to be disentangled. Notwithstand-

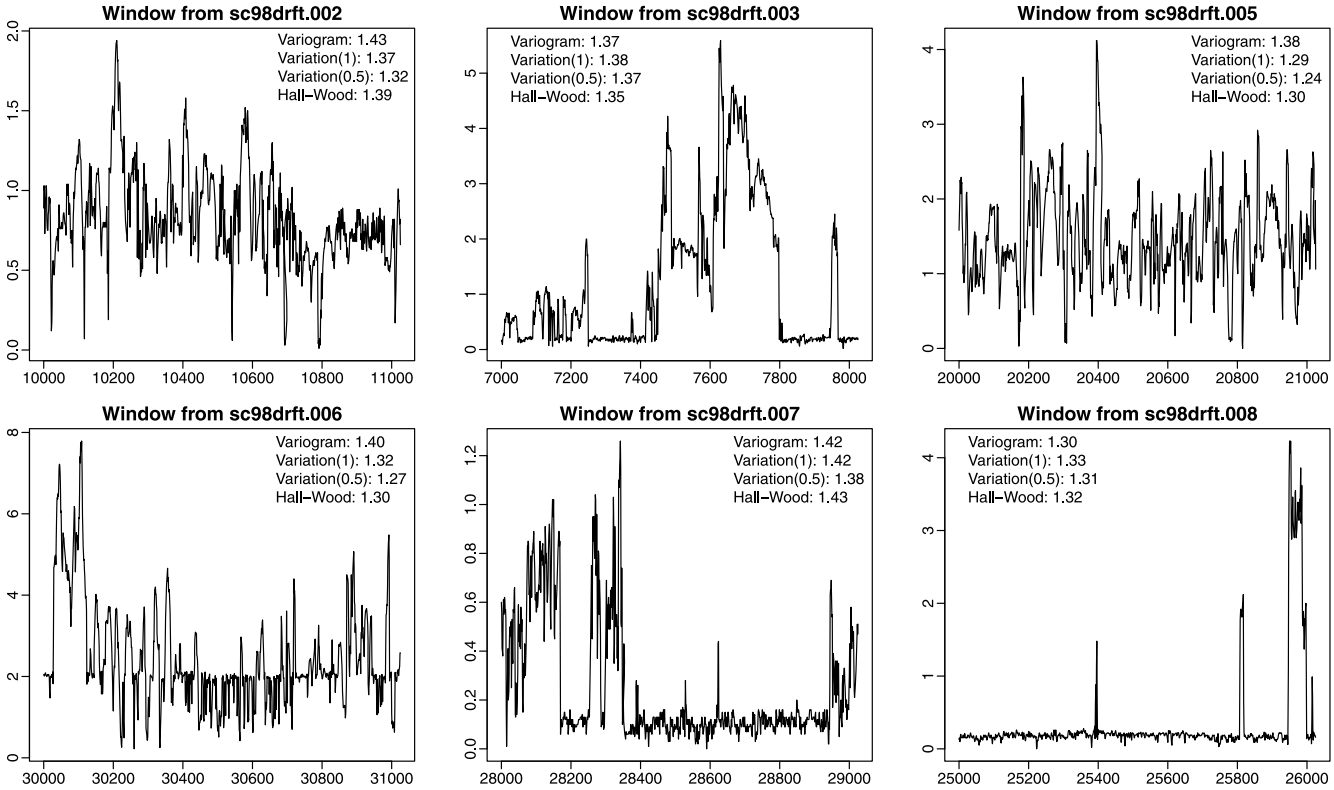

FIG. 23. *Selected blocks of width* 1,024 *meters from the sea-ice profiles in Figure* 22, *along with the corresponding estimates of fractal dimension.*

ing these issues, estimates of fractal dimension can serve as informative descriptors of surface roughness.

In the case of time series or line transect data, we recommend the use of the madogram estimator, that is, the variation estimator with power index $p = 1$. In the case of spatial lattice data, transect estimators based on madogram estimators along rows and columns show high promise. These recommendations echo observations of Bruno and Raspa (1989) and Bez and Bertrand (2011), who argued that the critical relationship (4) be-

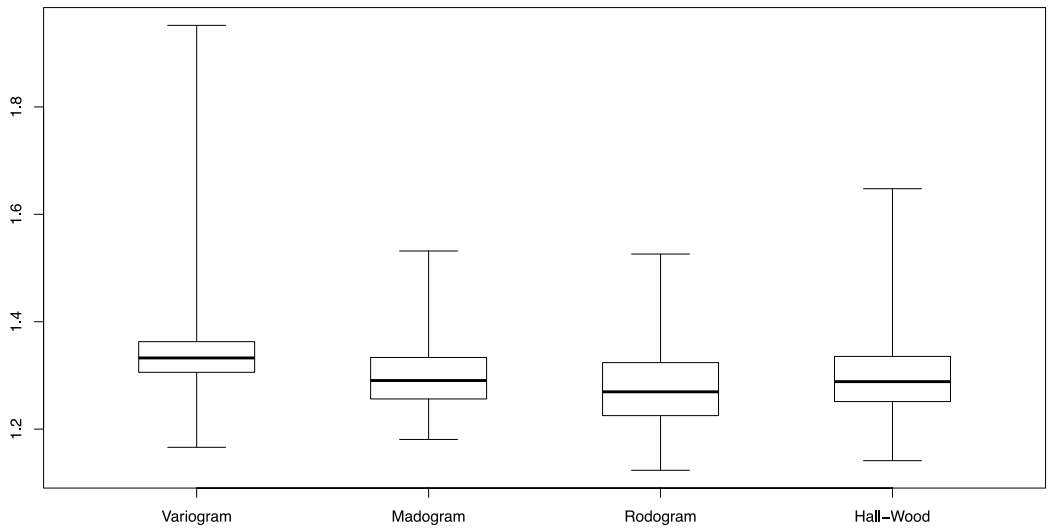

FIG. 24. *Boxplots for variation estimates with power indices* $p = 2$ (*variogram*), 1 (*madogram*) *and* 1/2 (*rodogram*) *along with Hall–Wood estimates of fractal dimension for the blocks in the ice profile data. See the text for details.*

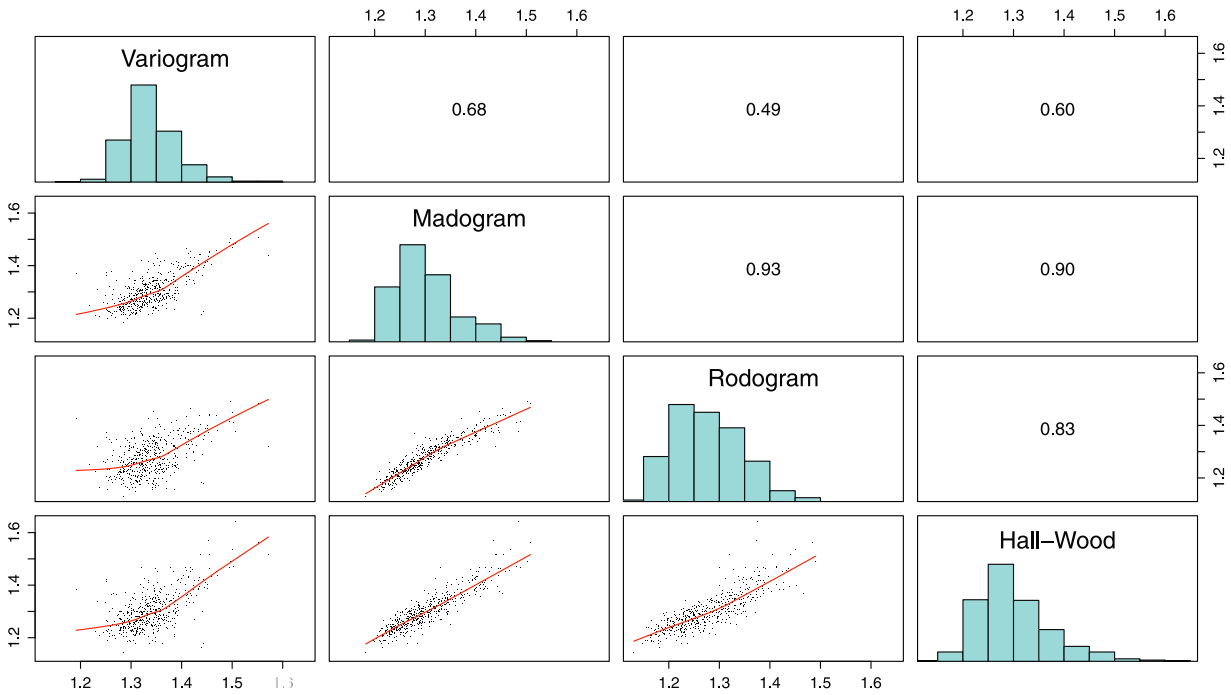

FIG. 25. *Scatterplot matrix for variation estimates with power indices* $p = 2$ (*variogram*), 1 (*madogram*) *and* $1/2$ (*rodogram*) *along with Hall–Wood estimates of fractal dimension for the blocks in the ice profile data. See the text for details.*

tween the fractal index of a stochastic process whose variogram of order $p > 0$ shows a behavior of the form (5) at the origin, and the fractal dimension of its sample paths, is particularly robust and inclusive when $p = 1$. We encourage work toward rigorous results in these directions, including both variograms and their equivalents for general types of line transect and spatial increments.

Furthermore, we call for the development of large sample theory for variation estimators of general power index, including but not limited to our preferred choice of $p = 1$. While for time series this has been achieved in the far-reaching recent work of Coeurjolly (2008), the case of spatial data remains open. As noted in Section 5.1, we believe that many of the results in the extant default case $p = 2$ can be carried over to a general power index, $p > 0$. In doing so, the results of Guyon and León (1989), Istas and Lang (1997) and Barndorff-Nielsen, Corcuera and Podolskij (2009) on the power variations of Gaussian processes provide tools that can be applied in concert with the methodology put forth in an impressive strand of asymptotic literature for $p = 2$, which includes the work of Hall and Wood (1993), Constantine and Hall (1994), Kent and Wood (1997), Davies and Hall (1999), Chan and Wood (2000) and Zhu and Stein (2002), among others. Of particular interest is our conjecture in the spatial setting of Section 5.1, according to which the transect-variation approach allows for estimators of the fractal dimension that are simultaneously highly efficient and robust. This approach can be paired with the use of trimmed means or linear combinations of sample quantiles in defining power variations, as suggested by Coeurjolly (2008) for time series, or with the use of bipower and multipower variations (Barndorff-Nielsen et al. 2006; Barndorff-Nielsen, Corcuera and Podolskij, 2011). While the case of Gaussian processes appears to be challenging yet tractable, a general asymptotic theory for non-Gaussian processes remains elusive, despite the recent progress by Chan and Wood (2004) and Achard and Coeurjolly (2010).

To our knowledge, Bayesian methods for estimating fractal dimension have not been explored yet, except that the method of Handcock and Stein (1993) could be applied in a parametric context. Physical insight can drive the choice of the prior, and the development of Bayesian madogram estimators might be a promising option.

Multifractional Brownian motion (Peltier and Levy Vehel, 1995; Benassi, Jaffard and Roux, 1997; Herbin, 2006) is a generalization of the classical fractional Brownian motion, where the fractal dimension is allowed to vary along the sample paths. Similarly, the

nonstationary Gaussian random fields described by Anderes and Stein (2011) allow for location-dependent, local Hausdorff dimensions, which need to be estimated as functions, rather than a single number. In this setting, the estimators considered in our paper can be used as building blocks for the more complex estimators needed to handle these nonstationary processes, as studied by Benassi, Cohen and Istas (1998), Ayache and Lévy Véhel (2004) and Coeurjolly (2005), among others. For an applied perspective, see Gagnon, Lovejoy and Schertzer (2006).

To close on a practical note, we have developed an R package, called `fractaldim`, that implements the estimators of fractal dimension discussed in this paper (Ševčíková, Gneiting and Percival, 2011). It has the ability to compute estimates for a single dataset, or a series of estimates along sliding blocks, as in our data example. Multiple estimates can be conveniently bundled into a single call, and the default arguments correspond to the recommendations in this paper.

For example, to generate four log–log plots of the type shown in Figure 5, a few lines of code suffice:

```
par(mfrow=c(2,2))
series <- GaussRF(x=c(0,1,1/1024),
  model='stable', gridtriple=TRUE,
  param=c(mean=0, variance=1,
  scale=1, kappa=1))
methods <- c('hallwood',
  'periodogram', 'variogram',
  'madogram'))
D <- fd.estimate(series,
  method=methods, plot.loglog=T,
  plot.allpoints=T)
```

Here, the function `GaussRF` from the `RandomFields` package is used to generate a Gaussian sample path with exponential covariance and fractal index $\alpha = 1$. The desired estimation methods are passed to the estimation function, `fd.estimate`, as character strings as in our example. Alternatively, each method can be given as a list with a `name` element determining the method, and any additional elements specifying parameters. For example, an entry `list(name='variation', p.index=1)` is equivalent to `'madogram'`.

To obtain interval estimates with a parametric bootstrap method, as proposed by Davies and Hall (1999) and shown in many of our figures, the following call suffices, using the simulated `series` from above, `boot.it` bootstrap replicates, and the madogram estimator:

```
D <- fd.estimate(series,
  methods='madogram')
D.boot <- rep(NA, boot.it)
alpha <- 4 - 2*D$fd
for (i in 1:boot.it)
 {
 boot.series <- GaussRF(
   x=c(0,1,1/1024), model='stable',
   gridtriple=TRUE,
   param=c(mean=0, variance=1,
   scale=D$scale^(-1/alpha),
   kappa=alpha))
 D.boot[i] <- fd.estimate(
   boot.series,
   methods='madogram')$fd
 }
```

Confidence intervals can be obtained from the array `D.boot` in the usual way. Additional functions for variation estimators in the line transect case are available within the DVFBM package (Coeurjolly, 2009).

ACKNOWLEDGMENTS

We are grateful to two anonymous reviewers, Nicolas Bez, Werner Ehm, Marc Genton, Mark Podolskij, Roopesh Ranjan, Michael Scheuerer and Michael Zaiser for comments and discussions. Our research has been supported by the National Science Foundation under Awards DMS-0706745 and ARC-0529955, by the Alfried Krupp von Bohlen und Halbach Foundation and by Grant Number R01 HD054511 from the National Institute of Child Health and Human Development.

REFERENCES

ACHARD, S. and COEURJOLLY, J.-F. (2010). Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise. *Stat. Surv.* **4** 117–147. MR2658892

ADLER, R. J. (1981). *The Geometry of Random Fields*. Wiley, Chichester. MR0611857

AHMED, N., NATARAJAN, T. and RAO, K. R. (1974). Discrete cosine transform. *IEEE Trans. Computers* **C-23** 90–93. MR0356555

ALDRICH, E. (2010). Wavelets: A package of funtions for computing wavelet filters, wavelet transforms and multiresolution analyses, R package version 0.2-6.

ANDERES, E. B. and STEIN, M. L. (2011). Local likelihood estimation for nonstationary random fields. *J. Multivariate Anal.* **102** 506–520. MR2755012

AYACHE, A. and LÉVY VÉHEL, J. (2004). On the identification of the pointwise Hölder exponent of the generalized multifractional Brownian motion. *Stochastic Process. Appl.* **111** 119–156. MR2049572

BARNDORFF-NIELSEN, O. E., CORCUERA, J. M. and PODOLSKIJ, M. (2009). Power variation for Gaussian processes with stationary increments. *Stochastic Process. Appl.* **119** 1845–1865. MR2519347

BARNDORFF-NIELSEN, O. E., CORCUERA, J. M. and PODOLSKIJ, M. (2011). Multipower variation for Brownian semistationary processes. *Bernoulli* **17** 1159–1194.

BARNDORFF-NIELSEN, O. E., GRAVERSEN, S. E., JACOD, J., PODOLSKIJ, M. and SHEPHARD, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Calculus to Mathematical Finance* 33–68. Springer, Berlin. MR2233534

BENASSI, A., COHEN, S. and ISTAS, J. (1998). Identifying the multifractional function of a Gaussian process. *Statist. Probab. Lett.* **39** 337–345. MR1646220

BENASSI, A., JAFFARD, S. and ROUX, D. (1997). Elliptic Gaussian random processes. *Rev. Mat. Iberoam.* **13** 19–90. MR1462329

BERG, C., MATEU, J. and PORCU, E. (2008). The Dagum family of isotropic correlation functions. *Bernoulli* **14** 1134–1149. MR2543589

BEZ, N. and BERTRAND, S. (2011). The duality of fractals: Roughness and self-similarity. *Theoretical Ecology* **4** 371–383.

BISHOP, G. C. and CHELLIS, S. E. (1989). Fractal dimension: A descriptor of ice keel surface roughness. *Geophysical Research Letters* **16** 1007–1010.

BLOCK, A., VON BLOH, W. and SCHELLNHUBER, H. J. (1990). Efficient box-counting determination of generalized fractal dimensions. *Phys. Rev. A* (3) **42** 1869–1874. MR1068470

BRUNO, R. and RASPA, G. (1989). Geostatistical characterization of fractal models of surfaces. In *Geostatistics*, *Vol. 1* (M. Armstrong, ed.) 77–89. Kluwer, Dordrecht.

BURROUGH, P. A. (1981). Fractal dimensions of landscapes and other environmental data. *Nature* **294** 240–242.

CHAN, G., HALL, P. and POSKITT, D. S. (1995). Periodogram-based estimators of fractal properties. *Ann. Statist.* **23** 1684–1711. MR1370303

CHAN, G. and WOOD, A. T. A. (2000). Increment-based estimators of fractal dimension for two-dimensional surface data. *Statist. Sinica* **10** 343–376. MR1769748

CHAN, G. and WOOD, A. T. A. (2004). Estimation of fractal dimension for a class of non-Gaussian stationary processes and fields. *Ann. Statist.* **32** 1222–1260. MR2065204

COEURJOLLY, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process.* **4** 199–227. MR1856174

COEURJOLLY, J.-F. (2005). Identification of multifractional Brownian motion. *Bernoulli* **11** 987–1008. MR2188838

COEURJOLLY, J.-F. (2008). Hurst exponent estimation of locally self-similar Gaussian processes using sample quantiles. *Ann. Statist.* **36** 1404–1434. MR2418662

COEURJOLLY, J. F. (2009). dvfBm: Discrete variations of a fractional Brownian motion, R package version 1.0.

CONNORS, D. N., LEVINE, E. R. and SHELL, R. R. (1990). A small-scale under-ice morphology study in the high arctic. *Cold Regions Research and Engineering Laboratory Monograph* **90-1** 145–151.

CONSTANTINE, A. G. and HALL, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc. Ser. B* **56** 97–113. MR1257799

DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916

DAVIES, R. B. (2001). Integrated processes and the discrete cosine transform. *J. Appl. Probab.* **38A** 105–121. MR1915538

DAVIES, S. and HALL, P. (1999). Fractal analysis of surface roughness by using spatial data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **61** 3–37. MR1664088

DIETRICH, C. R. and NEWSAM, G. N. (1993). A fast and exact method for multidimensional Gaussian stochastic simulations. *Water Resources Research* **29** 2861–2869.

DUBUC, B., QUINIOU, J. F., ROQUES-CARMES, C., TRICOT, C. and ZUCKER, S. W. (1989a). Evaluating the fractal dimension of profiles. *Phys. Rev. A* (3) **39** 1500–1512. MR0983085

DUBUC, B., ZUCKER, S. W., TRICOT, C., QUINIOU, J. F. and WEHBI, D. (1989b). Evaluating the fractal dimension of surfaces. *Proc. Roy. Soc. London Ser. A* **425** 113–127. MR1019289

EHM, W. (1981). Sample function properties of multiparameter stable processes. *Z. Wahrsch. Verw. Gebiete* **56** 195–228. MR0618272

EMERY, X. (2005). Variograms of order ω: A tool to validate a bivariate distribution model. *Mathematical Geology* **37** 163–181.

FALCONER, K. (1990). *Fractal Geometry*: *Mathematical Foundations and Applications*. Wiley, Chichester. MR1102677

FEUERVERGER, A., HALL, P. and WOOD, A. T. A. (1994). Estimation of fractal index and fractal dimension of a Gaussian process by counting the number of level crossings. *J. Time Ser. Anal.* **15** 587–606. MR1312323

GAGNON, J. S., LOVEJOY, S. and SCHERTZER, D. (2006). Multifractal earth topography. *Nonlinear Processes in Geophysics* **13** 541–570.

GALLANT, J. C., MOORE, I. D., HUTCHINSON, M. F. and GESSLER, P. (1994). Estimating fractal dimension of profiles: A comparison of methods. *Math. Geol.* **26** 455–481.

GENTON, M. G. (1998). Highly robust variogram estimation. *Math. Geol.* **30** 213–221. MR1610687

GENTON, M. G. and LUCAS, A. (2003). Comprehensive definitions of breakdown points for independent and dependent observations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65** 81–94. MR1959094

GNEITING, T. and SCHLATHER, M. (2004). Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Rev.* **46** 269–282 (electronic). MR2114455

GNEITING, T., ŠEVČÍKOVÁ, H., PERCIVAL, D. B., SCHLATHER, M. and JIANG, Y. (2006). Fast and exact simulation of large Gaussian lattice systems in $\mathbb{R}^2$: Exploring the limits. *J. Comput. Graph. Statist.* **15** 483–501. MR2291260

GOFF, J. A. and JORDAN, T. H. (1988). Stochastic modeling of seafloor morphology — inversion of sea beam data for 2nd-order statistics. *Journal of Geophysical Research* **93** 13589–13608.

GOFF, J. A., STEWART, K. W., SINGH, H. and TANG, X. (1995). Quantitative analysis of sea ice draft. 2. Application of stochastic modeling to intersecting topographic profiles. *Journal of Geophysical Research* **100** 7005–7017.

GONZALEZ, R. C. and WOODS, R. E. (2007). *Digital Image Processing*, 3rd ed. Prentice-Hall, Englewood Cliffs, NJ.

GUTTORP, P. and GNEITING, T. (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika* **93** 989–995. MR2285084

GUYON, X. and LEÓN, J. (1989). Convergence en loi des H-variations d'un processus gaussien stationnaire sur $\mathbf{R}$. *Ann. Inst. H. Poincaré Probab. Stat.* **25** 265–282. MR1023952

HALL, P. and DAVIES, S. (1995). On direction-invariance of fractal dimension on a surface. *Appl. Phys. A* **60** 271–274.

HALL, P. and ROY, R. (1994). On the relationship between fractal dimension and fractal index for stationary stochastic processes. *Ann. Appl. Probab.* **4** 241–253. MR1258183

HALL, P. and WOOD, A. (1993). On the performance of box-counting estimators of fractal dimension. *Biometrika* **80** 246–252. MR1225230

HALLEY, J. M., HARTLEY, S., KALLIMANIS, A. S., KUNIN, W. E., LENNON, J. J. and SGARDELIS, S. P. (2004). Uses and abuses of fractal methodology in ecology. *Ecology Letters* **7** 254–271.

HANDCOCK, M. and STEIN, M. L. (1993). A Bayesian analysis of kriging. *Technometrics* **35** 403–410.

HAUSDORFF, F. (1919). Dimension und äußeres Maß. *Math. Ann.* **79** 157–179.

HERBIN, E. (2006). From N parameter fractional Brownian motions to N parameter multifractional Brownian motions. *Rocky Mountain J. Math.* **36** 1249–1284. MR2274895

IHAKA, R. and GENTLEMAN, R. (1996). *R*: A language for data analysis and graphics. *J. Comput. Graph. Statist.* **5** 299–314.

ISTAS, J. and LANG, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Stat.* **33** 407–436. MR1465796

JAKEMAN, E. and JORDAN, D. L. (1990). Statistical accuracy of measurements on Gaussian random fractals. *J. Phys. D: Appl. Phys.* **23** 397–405.

KENT, J. T. and WOOD, A. T. A. (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc. Ser. B* **59** 679–699. MR1452033

KLINKENBERG, B. (1994). A review of methods used to determine the fractal dimension of linear features. *Math. Geol.* **26** 23–46.

KLINKENBERG, B. and GOODCHILD, M. F. (1992). The fractal properties of topography: A comparison of methods. *Earth Surface Processes and Landforms* **17** 217–234.

LIEBOVITCH, L. S. and TOTH, T. (1989). A fast algorithm to determine fractal dimensions by box counting. *Phys. Lett. A* **141** 386–390. MR1030923

MALCAI, O., LIDAR, D. A., BIHAM, O. and AVNIR, D. (1997). Scaling range and cutoffs in empirical fractals. *Phys. Rev. E* **56** 2817–2828.

MANDELBROT, B. B. (1977). *Fractals*: *Form*, *Chance*, *and Dimension*, Revised ed. W. H. Freeman and Co., San Francisco, CA. MR0471493

MANDELBROT, B. B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman and Co., San Francisco, CA. MR0665254

MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR0242239

MARSTRAND, J. M. (1954). Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc.* (3) **4** 257–302. MR0063439

MASRY, E. (1972). On covariance functions of unit processes. *SIAM J. Appl. Math.* **23** 28–33. MR0309220

MATÉRN, B. (1986). *Spatial Variation*, 2nd ed. *Lecture Notes in Statistics* **36**. Springer, Berlin. MR0867886

MOLZ, F. J., RAJARAM, H. and LU, S. L. (2004). Stochastic fractal-based models of heterogeneity in subsurface hydrology: Origins, applications, limitations, and future research questions. *Reviews of Geophysics* **42** RG1002.

OREY, S. (1970). Gaussian sample functions and the Hausdorff dimension of level crossings. *Z. Wahrsch. Verw. Gebiete* **15** 249–256. MR0279882

OROSEI, R., BIANCHI, R., CORADINI, A., ESPINASSE, S., FEDERICO, C., FERRICCIONI, A. and GAVRISHIN, A. I. (2003). Self-affine behavior of Martian topography at kilometer scale from Mars Orbiter Laser Altimeter data. *Journal of Geophysical Research — Planets* **108** E4 8023.

PELTIER, R. F. and LEVY VEHEL, J. (1995). Multifractional Brownian motion: Definition and preliminary results. Technical report, INRIA Rocquencourt.

PERCIVAL, D. B. and WALDEN, A. T. (2000). *Wavelet Methods for Time Series Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics* **4**. Cambridge Univ. Press, Cambridge. MR1770693

ROTHROCK, D. A. and THORNDIKE, A. S. (1980). Geometric properties of the underside of sea ice. *Journal of Geophysical Research* **85** 3955–3963.

ROUSSEEUW, P. J. and CROUX, C. (1993). Alternatives to the median absolute deviation. *J. Amer. Statist. Assoc.* **88** 1273–1283. MR1245360

SCHEPERS, H. E., VAN BEEK, J. H. G. M. and BASSINGTHWAIGHTE, J. B. (1992). Four methods to estimate the fractal dimension from self-affine signals. *IEEE Engineering in Medicine and Biology* **11** 57–64.

SCHEUERER, M. (2010). Regularity of the sample paths of a general second order random field. *Stochastic Process. Appl.* **120** 1879–1897. MR2673978

SCHLATHER, M. (2001). Simulation and analysis of random fields. *R News* **1** 18–20.

SCHMITTBUHL, J., VILOTTE, J. P. and ROUX, S. (1995). Reliability of self-affine measurements. *Phys. Rev. E* **51** 131–147.

ŠEVČÍKOVÁ, H., GNEITING, T. and PERCIVAL, D. B. (2011). fractaldim: Estimation of fractal dimensions, R package version 0.8-0.

SMITH, L. (2007). *Chaos: A Very Short Introduction*. Cambridge Univ. Press, New York.

STEIN, M. L. (1999). *Interpolation of Spatial Data*: *Some Theory for Kriging Springer Series in Statistics*. Springer, New York. MR1697409

STEIN, M. L. (2002). Fast and exact simulation of fractional Brownian surfaces. *J. Comput. Graph. Statist.* **11** 587–599. MR1938447

STRANG, G. (1999). The discrete cosine transform. *SIAM Rev.* **41** 135–147 (electronic). MR1669796

TAQQU, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302. MR0400329

TAYLOR, C. C. and TAYLOR, S. J. (1991). Estimating the dimension of a fractal. *J. Roy. Statist. Soc. Ser. B* **53** 353–364. MR1108332

THEILER, J. (1990). Estimating fractal dimension. *J. Opt. Soc. Amer. A* **7** 1055–1073. MR1054127

TURCOTTE, D. L. (1997). *Fractals and Chaos in Geology and Geophysics*, 2nd ed. Cambridge Univ. Press, Cambridge. MR1458893

WEISSEL, J., PRATSON, L. F. and MALINVERNO, A. (1994). The length-scaling properties of topography. *Journal of Geophysical Research* **99** 13997–14012.

WOOD, A. T. A. and CHAN, G. (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Statist.* **3** 409–432. MR1323050

XUE, Y. and XIAO, Y. (2011). Fractal and smoothness properties of space–time Gaussian models. *Front. Math. China* **6** 1217–1248.

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions. Vol. I: Basic Results*. Springer, New York. MR0893393

ZAISER, M., GRASSET, F. M., KOUTSOS, V. and AIFANTIS, E. C. (2004). Self-affine surface morphology of plastically deformed metals. *Phys. Rev. Lett.* **93** 195507.

ZHU, Z. and STEIN, M. L. (2002). Parameter estimation for fractional Brownian surfaces. *Statist. Sinica* **12** 863–883. MR1929968