

ESTIMATION OF SMOOTHNESS OF A STATIONARY GAUSSIAN RANDOM FIELD

Wei-Ying Wu and Chae Young Lim

National Dong Hwa University and Seoul National University

Abstract: For a stationary Gaussian random field, the decay rate of the spectral density as the frequency becomes large determines the smoothness of the random field. The decay rate of the spectral density is also related to the fractal dimension, which is used to measure the surface smoothness of a random field. We propose an estimator of the decay rate using the periodogram when the observations are on a grid and investigate the asymptotic properties under the fixed domain asymptotic setting. A bias-reduced estimate is proposed based on the theoretical property of the estimator found in this work. A simulation study and a data example are presented.

Key words and phrases: Fractal dimension, fractal index, Gaussian random fields, infill asymptotics, periodogram, smoothness, spectral density.

1. Introduction

Analyzing surface smoothness (or roughness) is of interest in many applications. Examples are coastlines of islands, surfaces of manufactured products, moon craters and so on. Fractal dimension, or Fractal index, is related to smoothness of sample paths of a random field and there is an extensive literature on Fractal-based analysis to investigate surface smoothness (e.g., Butler, Lane, and Chandler (2001); Deems, Fassnacht, and Elder (2006); Klinkenberg (1994); Mandelbrot (1982); Milne (1992); Vázquez, Miranda, and Conzález (2005)).

For a stationary random field, the spectral density $f(\boldsymbol{\lambda})$ provides the smoothness information of a random field - its decay rate of the spectral density as $|\boldsymbol{\lambda}|$ increases is related to the smoothness of its corresponding random field. For example, if the covariance function $K(\boldsymbol{x})$ of a stationary Gaussian random field, $Z(\boldsymbol{s})$, on \mathbb{R}^d satisfies

$$K(\boldsymbol{x}) \sim K(\mathbf{0}) - k|\boldsymbol{x}|^\alpha \quad \text{as } |\boldsymbol{x}| \rightarrow 0 \quad (1.1)$$

for some k and $0 < \alpha \leq 2$, α is the fractal index that measures the roughness of sample paths of Z . For such covariance function, the corresponding spectral density satisfies

$$f(\boldsymbol{\lambda}) \sim c|\boldsymbol{\lambda}|^{-\theta} \quad \text{as } |\boldsymbol{\lambda}| \rightarrow \infty, \quad (1.2)$$

for some $c > 0$ and $\theta = \alpha + d$, by an Abelian-type theorem. Further discussion of these connections can be found in Adler and Taylor (2007) and Xue and Xiao (2011).

We are interested in estimating θ in (1.2) to obtain the smoothness information of a realization of a random field. There are several approaches to estimate a fractal index. One of them is based on box counting (Block, Von Bloh, and Schellnhuber (1990); Hall and Wood (1993); Liebovitch and Toth (1989)). For example, Hall and Wood (1993) observed that the limit of log ratio of the total area of the boxes at scale ϵ , $A(\epsilon)$, to its scale, ϵ , is a linear function of the fractal index so that one can estimate it by regressing $\log A(\epsilon)$ on $\log \epsilon$. Another approach is based on observing that logarithm of the variogram is a linear function of the fractal index at small lags, so that one can get an estimator from a regression fit (Constantine and Hall (1994); Kent and Wood (1997); Davies and Hall (1999); Chan and Wood (2000, 2004); Zhu and Stein (2002)). As a generalized version of a variogram estimator, a variation estimator is introduced that makes use of the variogram of order p , $\gamma_p(h) = (1/2)E|Z(s) - Z(s+h)|^p$. When $p = 2$, this is the usual variogram. The logarithm of the variogram of order p is a linear function of the fractal index for a Gaussian process and a regression fit can be used to find an estimator (Emery (2005)).

In the spectral domain, Chan, Hall, and Poskitt (1995) introduced a periodogram-based estimator using the relationship between the spectral density function for a stationary process and the fractal index. A recent review paper, Gneiting, Ševčíková, and Percival (2012), gives a detailed comparison of the above mentioned methods for estimating a fractal index.

On the other hand, the estimation of the differentiability of a stationary Gaussian random field is not always equivalent to the estimation of a fractal index, since the smoothness parameter θ is related to a fractal index only when θ is in $(d, 2 + d]$. Thus, when θ is outside this range, the estimation methods for a fractal index may not work.

The estimation of the differentiability of a stationary Gaussian random field was investigated in Wu, Lim, and Xiao (2013). They proposed a local Whittle likelihood-*type* estimator to estimate the smoothness parameter θ in (1.2); it is related to the differentiability of the random field. Their objective function can be viewed as a weighted asymptotic local Whittle-likelihood. In Wu, Lim, and Xiao (2013), the consistency and asymptotic normality of the estimators of c and θ were shown when the other parameter is assumed to be known. They also obtained the consistency of the estimator of θ when c is any fixed positive quantity, but simulation results show significant bias depending on the choice of c .

Estimation on the spectral domain typically requires data on a regular lattice. This can be a limitation given that there are many applications with irregularly

spaced data, but data on a regular lattice is popular, for example, in the areas of remote sensing, computer vision, biomedical imaging, surface meteorology, etc. There are also several studies dealing with irregularly spaced data on the spectral domain (Fuentes (2007); Matsuda and Yajima (2009); Bandyopadhyay and Lahiri (2009)). These references investigate Whittle likelihood or the discrete Fourier transform under stochastic sampling design, but do not lead estimates of the smoothness parameter. Here we consider data on a regular lattice.

The proposed estimation of a smoothness parameter is a modification of the estimation procedure of Wu, Lim, and Xiao (2013), yet it maintains the same asymptotic properties. We are able to verify additional asymptotic features of the modified estimator, which leads us to develop a bias-reduced estimate of θ when c is any given positive value.

Our asymptotic results are based on a fixed domain asymptotic setting (or infill asymptotic setting) (Cressie (1993); Stein (1999)), where the observation domain is fixed but the distance of neighboring observations tends to zero. A main reason to work on the fixed domain asymptotic setting is to bring out the information of smoothness (tail behavior of the spectral density) from the estimate of the spectral density.

The remainder of the paper is as follows. Section 2 introduces the proposed estimation method and its asymptotic properties. Also, a bias-reduced estimate of θ is proposed. Section 3 verifies the proposed method through simulations. A data example is given as well. Section 4 gives the summary of our findings.

2. Estimation and Asymptotic Properties

For a stationary Gaussian random field, $Z(\mathbf{s})$ on \mathbb{R}^d , we assume that the spectral density, $f(\boldsymbol{\lambda})$, satisfies (1.2). We also assume that $Z(\mathbf{s})$ is observed on the lattice $\phi\mathbf{J}$ with $\mathbf{J} \in T_m = \{1, \dots, m\}^d$ so that ϕ is the distance between two grid points in each direction. Let \mathbb{Z}^d be the set of d -dimensional integer-valued vectors. Then, we can define the lattice process $Y_\phi(\mathbf{J}) \equiv Z(\phi\mathbf{J})$ for $\mathbf{J} \in \mathbb{Z}^d$, with spectral density $\bar{f}_\phi(\boldsymbol{\lambda}) = \phi^{-d} \sum_{\mathbf{Q} \in \mathbb{Z}^d} f((\boldsymbol{\lambda} + 2\pi\mathbf{Q})/\phi)$ for $\boldsymbol{\lambda} \in (-\pi, \pi]^d$.

As $\phi \rightarrow 0$, $\bar{f}_\phi(\boldsymbol{\lambda})$ increases near $\boldsymbol{\lambda} = 0$ which makes the spectral density unbounded. We apply a differencing operator to avoid non-integrability of the spectral density. For $d = 1$, we consider the simple differencing operator and for $d > 1$, we use a discrete Laplacian operator defined as $\Delta_\phi(Z(\mathbf{s})) = \sum_{j=1}^d Z(\mathbf{s} + \mathbf{e}_j) - 2Z(\mathbf{s}) + Z(\mathbf{s} - \mathbf{e}_j)$, where \mathbf{e}_j is the unit vector whose j th entry is 1.

We introduce a lattice process $Y_\phi^\tau(\mathbf{J})$ obtained by applying the Laplacian operator τ times. That is, $Y_\phi^\tau(\mathbf{J}) \equiv (\Delta_\phi)^\tau(Z(\mathbf{J}))$. The spectral density of Y_ϕ^τ is then given by

$$\bar{f}_\phi^\tau(\boldsymbol{\lambda}) = \left\{ \sum_{j=1}^d 4 \sin^2\left(\frac{\lambda_j}{2}\right) \right\}^{2\tau} \bar{f}_\phi(\boldsymbol{\lambda}). \quad (2.1)$$

Here we have

$$\lim_{\phi \rightarrow 0} \phi^{d-\theta} \bar{f}_\phi^\tau(\boldsymbol{\lambda}) = c \left\{ \sum_{j=1}^d 4 \sin^2 \left(\frac{\lambda_j}{2} \right) \right\}^{2\tau} \sum_{\mathbf{Q} \in \mathbb{Z}^d} |(\boldsymbol{\lambda} + 2\pi\mathbf{Q})|^{-\theta} \mathbb{I}_{\{\boldsymbol{\lambda} \neq \mathbf{0}\}}, \quad (2.2)$$

for $\boldsymbol{\lambda} \in (-\pi, \pi]^d$. We let $g_{c,\theta}(\boldsymbol{\lambda})$ be the limit in (2.2) that is obtained from the condition (1.2). Here \mathbb{I}_A is the indicator function of the set A . $g_{c,\theta}(\boldsymbol{\lambda})$ is integrable by choosing τ such that $4\tau - \theta > -d$ for $d > 1$, and $2\tau - \theta > -d$ for $d = 1$. For simplicity, we assume that we have $Y_\phi^\tau(\mathbf{J})$ at $\mathbf{J} \in T_m = \{1, \dots, m\}^d$ after differencing $Z(\mathbf{s})$ τ times, and ϕ is set as $\phi = m^{-1}$ so that we are considering a fixed domain asymptotic setting.

The periodogram, a nonparametric estimate of the spectral density of $Y_\phi^\tau(\mathbf{J})$ is defined as $I_m^\tau(\boldsymbol{\lambda}) = (2\pi m)^{-d} \left| \sum_{\mathbf{J} \in T_m} Y_\phi^\tau(\mathbf{J}) \exp\{-i\boldsymbol{\lambda}^T \mathbf{J}\} \right|^2$. A smoothed periodogram at Fourier frequencies is defined as $\hat{I}_m^\tau(2\pi\mathbf{J}/m) = \sum_{\mathbf{K} \in \mathcal{T}_m} W_h(\mathbf{K}) I_m^\tau(2\pi(\mathbf{J} + \mathbf{K})/m)$, where $\mathcal{T}_m \equiv \{-\lfloor (m-1)/2 \rfloor, \dots, m - \lfloor m/2 \rfloor\}^d$ is the set of Fourier frequencies and $\lfloor x \rfloor$ is the largest integer not greater than x . $W_h(\mathbf{K})$ is a weight defined as $W_h(\mathbf{K}) = \Lambda_h(2\pi\mathbf{K}/m) / \sum_{\mathbf{L} \in \mathcal{T}_m} \Lambda_h(2\pi\mathbf{L}/m)$, where $\Lambda_h(\mathbf{s}) = (1/h)\Lambda(\mathbf{s}/h) \mathbb{I}_{\{\|\mathbf{s}\| \leq h\}}$ for a nonnegative symmetric continuous function Λ on \mathbb{R}^d with $\Lambda(\mathbf{0}) > 0$, and $\|\cdot\|$ is the max norm $\|\mathbf{s}\| = \max\{|s_1|, |s_2|, \dots, |s_d|\}$. We refer the reader to Wu, Lim, and Xiao (2013) for more discussion on the spectral density and its limit, as well as the periodogram.

We need some assumptions on the spectral density $f(\boldsymbol{\lambda})$.

Assumption 1. Let $f(\boldsymbol{\lambda})$ be the spectral density of a stationary Gaussian random field $Z(\mathbf{s})$ on \mathbb{R}^d .

- (A) $f(\boldsymbol{\lambda})$ satisfies $f(\boldsymbol{\lambda}) \sim c |\boldsymbol{\lambda}|^{-\theta}$ as $|\boldsymbol{\lambda}| \rightarrow \infty$, for some $c > 0$, $\theta > d$.
- (B) $f(\boldsymbol{\lambda})$ is twice differentiable and there exists a positive constant C such that, for $|\boldsymbol{\lambda}| > C$,

$$\left| \frac{\partial}{\lambda_j} f(\boldsymbol{\lambda}) \right| (1 + |\boldsymbol{\lambda}|)^{\theta+1} \quad \text{and} \quad \left| \frac{\partial^2}{\lambda_j \lambda_k} f(\boldsymbol{\lambda}) \right| (1 + |\boldsymbol{\lambda}|)^{\theta+2} \quad (2.3)$$

are uniformly bounded for $j, k = 1, \dots, d$.

- (C) $f(\boldsymbol{\lambda})$ satisfies $f(\boldsymbol{\lambda}) \asymp (1 + |\boldsymbol{\lambda}|)^{-\theta}$ for all possible $\boldsymbol{\lambda}$.

The Assumption 1 (A) means $f(\boldsymbol{\lambda})/c |\boldsymbol{\lambda}|^{-\theta} \rightarrow 1$ as $|\boldsymbol{\lambda}| \rightarrow \infty$, and 1 (C) means there exist C_1 and C_2 such that $0 < C_1 \leq f(\boldsymbol{\lambda})/(1 + |\boldsymbol{\lambda}|)^{-\theta} \leq C_2 < \infty$ for all possible $\boldsymbol{\lambda}$. There are several models that satisfy the Assumption 1. Two examples are considered in the simulation study section: a damped oscillation model (Yaglom (1987)) and a Matérn model (Stein (1999)). A Matérn model and its variants are commonly used in modeling spatial data.

Wu, Lim, and Xiao (2013) introduced a local Whittle likelihood-type estimator for θ that utilizes Fourier frequency information around some fixed non-zero frequency (e.g. $(\pi/2)\mathbf{1}_d$). The difference from the local Whittle likelihood is that Wu, Lim, and Xiao (2013) considered a weighted local Whittle likelihood using a kernel function. Fourier frequencies used to obtain the estimator in Wu, Lim, and Xiao (2013) are getting close to that fixed non-zero frequency as the sample size increases. This motivates us to consider a modified objective function to estimate the parameters (c, θ) :

$$R_m(c, \theta) = \log \left(m^{d-\theta} g_{c,\theta} \left(\frac{2\pi\mathbf{J}}{m} \right) \right) + \frac{1}{m^{d-\theta}} \frac{\hat{I}_m^\tau(2\pi\mathbf{J}/m)}{g_{c,\theta}(2\pi\mathbf{J}/m)}, \quad (2.4)$$

where $g_{c,\theta}(2\pi\mathbf{J}/m)$ is introduced as (2.2). This objective function considers Fourier frequencies near the fixed non-zero frequency, say, $2\pi\mathbf{J}/m$, through the smoothed periodogram \hat{I}_m^τ . The fixed non-zero frequency can be chosen such that $2\pi\mathbf{J}/m$ is the closest to $(\pi/2)\mathbf{1}_d$ as in Wu, Lim, and Xiao (2013). Although it looks like it is using information at only one non-zero Fourier frequency, it is not. The estimators of c and θ obtained by minimizing (2.4) have similar asymptotic properties to the ones introduced in Wu, Lim, and Xiao (2013). To avoid confusion in notation, we use c_0 and θ_0 for the true values of c and θ , respectively. Then, the estimators of c and θ are given by

$$c_m = \arg \min_{c \in \mathcal{C}} R_m(c, \theta_0), \quad (2.5)$$

$$\theta_m = \arg \min_{\theta \in \Theta} R_m(c^*, \theta), \quad (2.6)$$

where \mathcal{C} and Θ are the parameter spaces of c and θ , respectively. Here $c^* > 0$ is a fixed value that may not be equal to the true value, c_0 . Theorem 1 gives the consistency and asymptotic distribution of c_m . The consistency and convergence rate of θ_m are given in Theorem 2. The asymptotic distribution for θ_m when $c^* = c_0$, is given in Theorem 3.

We write \xrightarrow{P} for the convergence in probability, and \xrightarrow{d} for the convergence in distribution.

Theorem 1. *For a stationary Gaussian random field $Z(\mathbf{s})$ on \mathbb{R}^d satisfying Assumption 1, suppose $4\tau > \theta_0 - 1$ and $h = \kappa m^{-\gamma}$ for some $\kappa > 0$, where γ satisfies $\max\{0, (d-2)/d\} < \gamma < 1$. Assume that \mathbf{J} satisfies $\mathbf{J} = \lfloor (m/4)\mathbf{1}_d \rfloor$, and that c_0 is in the interior of the parameter space \mathcal{C} , a closed interval. Then,*

$$c_m \xrightarrow{P} c_0, \quad \text{and} \quad (2.7)$$

$$m^\eta(c_m - c_0) \xrightarrow{d} \mathcal{N} \left(0, c_0^2 \frac{\Lambda_2}{\Lambda_1^2} \left(\frac{2\pi}{\kappa} \right)^d \right), \quad (2.8)$$

where $\Lambda_r = \int_{[-1,1]^d} \Lambda^r(\mathbf{s}) d\mathbf{s}$ and $\eta = d(1 - \gamma)/2$.

Theorem 2. *If the conditions of Theorem 1 hold, θ_0 is in the interior of the parameter space Θ , a closed interval, and c^* is chosen from a closed interval \mathcal{C} , then*

$$\theta_m \xrightarrow{p} \theta_0. \quad (2.9)$$

Moreover, there exists M_ϵ for $0 < \epsilon < 1, \delta > 0$ such that for all $m > M_\epsilon$,

$$P \left\{ (\log(m))^{-1} \log \frac{c^*}{c_0(1+\epsilon)} \leq \theta_m - \theta_0 \leq (\log(m))^{-1} \log \frac{c^*}{c_0(1-\epsilon)} \right\} \geq 1 - \delta. \quad (2.10)$$

Remark 1. Theorem 2 implies that $m^{-\theta_m + \theta_0} c^* / c_0 - 1 = o_p(1)$ and, equivalently, $\theta_0 - \theta_m = (\log(m))^{-1} \log(c_0/c^*) + o_p((\log(m))^{-1})$. When $c^* = c_0$, we get $\theta_m - \theta_0 = o_p((\log(m))^{-1})$, a result in Wu, Lim, and Xiao (2013).

Different from the estimator in Wu, Lim, and Xiao (2013), we have strong consistency results when $d \geq 2$.

Corollary 1. *For $\max\{0, (d-2)/d\} < \gamma < (d-1)/d$ with $d \geq 2$, $\theta_m \rightarrow \theta_0$ a.e..*

When the true value of c is known, the asymptotic normality of the proposed estimator of θ can be shown.

Theorem 3. *Under the conditions of Theorem 2, if $c^* = c_0$, we have*

$$\log(m) m^\eta (\theta_m - \theta) \xrightarrow{d} \mathcal{N} \left(0, \frac{\Lambda_2}{\Lambda_1^2} \left(\frac{2\pi}{\kappa} \right)^d \right),$$

where $\eta = d(1 - \gamma)/2$.

Remark 2. The condition on γ in Theorems 1–3, $\max\{0, (d-2)/d\} < \gamma < 1$, is weaker than the condition on γ in Wu, Lim, and Xiao (2013) since it allows a better convergence rate, η , for the estimator. For example, when $d = 1$, we need $0 < \gamma < 1$ for the asymptotic results of the proposed estimator so that $\eta = (1 - \gamma)/2$ satisfies $0 < \eta < 1/2$, where $1/3 < \gamma < 1$ is needed for the asymptotic results in Wu, Lim, and Xiao (2013).

The proofs of Theorems 1 and 3 are similar to those in Wu, Lim, and Xiao (2013) so their proofs are omitted. Proofs of Theorem 2 and Corollary 1 are presented in the online supplementary material along with the proof of the next result.

Theorem 4. *Under the conditions of Theorem 2, (i) $c^* > c_0$, there exists M such that $P(\theta_0 \leq \theta_m) = 1$ for $m > M$; (ii) $c^* < c_0$, there exists M such that $P(\theta_0 \geq \theta_m) = 1$ for $m > M$.*

Remark 3. By Theorems 2 and 4, θ_m will over-estimate or under-estimate θ_0 when $c^* \neq c_0$ for a large sample size m . As the sample size increases, θ_m is getting close to θ_0 from the one side of θ_0 . On the other hand, when $c^* = c_0$, θ_m can be on the both sides of the true value, θ_0 . This property is also empirically verified through the simulation study in section 3.

2.1. Plug-in approach to estimate c and θ

Asymptotic properties of estimators of c given θ_0 , and θ given c^* , are investigated in Theorems 1–4. As we have a consistent estimator of θ even if we do not know the true c , we need a consistent estimator of c without assuming θ is known. As c_m has an explicit expression, $c_m = (1/m^{d-\theta_0}) \hat{I}_m^\tau(2\pi\mathbf{J}/m) / g_{1,\theta_0}(2\pi\mathbf{J}/m)$, we can replace θ_0 with θ_m to get an estimator of c when θ is unknown. It is, however, challenging to show its consistency. One could profile out c to estimate θ and estimate c after that, but this does not work since we get a constant objective function after c is profiled out. We consider a modified plug-in approach to produce a consistent estimator of c when θ is not known.

Recall that we have assumed the number of grid points where data are observed is m^d but, for simplicity, we use the term ‘sample size’ to refer to m . Suppose m can be written as $m \equiv m_k = \lfloor a^k \rfloor$ for some positive integers a and k . Then, consider $m'_k = \lfloor a^{k\rho} \rfloor$ for some $0 < \rho < 1$ so that we have a sub-sample of size m'_k by selecting every $b = \lfloor a^{k-k\rho} \rfloor$ -th grid point for each direction. Then, we consider the estimator for c .

$$c_{m_k} = \frac{1}{m_k^{d-\theta_{m_k}}} \frac{\hat{I}_{m'_k}^\tau(2\pi\mathbf{J}/m'_k)}{g_{1,\theta_{m_k}}(2\pi\mathbf{J}/m'_k)}, \quad (2.11)$$

where $\theta_{m_k} = \arg \min_{\theta \in \Theta} R_{m_k}(c^*, \theta)$ is the estimator of θ obtained by using m_k samples and $\hat{I}_{m'_k}^\tau(2\pi\mathbf{J}/m'_k)$ is a smoothed periodogram using m'_k samples. The consistency of θ_{m_k} as $k \rightarrow \infty$ follows from Theorem 2. The consistency of c_{m_k} follows from

$$\frac{1}{m_k^{d-\theta_0}} \frac{\hat{I}_{m'_k}^\tau(2\pi\mathbf{J}/m'_k)}{g_{1,\theta_{m'_k}}(2\pi\mathbf{J}/m'_k)} \frac{g_{1,\theta_{m'_k}}(2\pi\mathbf{J}/m'_k)}{g_{1,\theta_{m_k}}(2\pi\mathbf{J}/m'_k)} \xrightarrow{p} c_0,$$

$$\log(m_k^{\theta_0-\theta_{m_k}}) = (\theta_0 - \theta_{m_k}) \log(m'_k) = O_p\left(\frac{\log(m'_k)}{\log(m_k)}\right) = o_p(1).$$

Although this approach provides a consistent estimator of (c, θ) when both parameters are unknown, a more efficient approach is needed to improve the performance of the estimator in finite samples.

Another issue is the bias of θ_m . From simulations, we found that the further the bias grows the further c^* is from c_0 . This affects the estimation of c using the above approach. To reduce such impact, we propose a one-step bias reduction approach to estimating θ .

2.2. Bias reduction in the estimation of θ

We have $\theta_0 - \theta_m = (\log(m))^{-1} \log(c_0/c^*) + o_p((\log(m))^{-1})$ from Remark 1. Consider a sub-sample of size $m_1 < m$ and an estimate constructed using m_1 samples. Then, we have the alternative relationship $\theta_0 - \theta_{m_1} = (\log(m_1))^{-1} \log(c_0/c^*) + o_p((\log(m_1))^{-1})$. Combining the two, we have

$$\log\left(\frac{c_0}{c^*}\right) = \frac{\theta_m - \theta_{m_1}}{(\log(m_1))^{-1} - (\log(m))^{-1}} + o_p(1),$$

and this leads us to propose a bias-reduced estimate of θ ,

$$\theta_m^* = \theta_m + (\log(m))^{-1} \left(\frac{\theta_{m_1} - \theta_m}{(\log(m))^{-1} - (\log(m_1))^{-1}} \right).$$

For a fixed c^* , we have $\theta_m - \theta_0 = O_p((\log(m))^{-1})$. With the bias-reduced estimate, θ_m^* , we have

$$\begin{aligned} \theta_m^* - \theta_0 &= \theta_m + (\log(m))^{-1} \left(\frac{\theta_{m_1} - \theta_m}{(\log(m))^{-1} - (\log(m_1))^{-1}} \right) - \theta_0 \\ &= -(\log(m))^{-1} \log\left(\frac{c_0}{c^*}\right) + o_p((\log(m))^{-1}) \\ &\quad + (\log(m))^{-1} \left(\log\left(\frac{c_0}{c^*}\right) + o_p(1) \right) \\ &= o_p((\log(m))^{-1}). \end{aligned}$$

With this bias-reduced estimate, we can obtain a new estimate for c , say, c_m^* by plugging in θ_m^* ,

$$c_m^* = \frac{1}{m^{d-\theta_m^*}} \frac{\hat{I}_m^\tau(2\pi\mathbf{J}/m)}{g_{\theta_m^*}(2\pi\mathbf{J}/m)}.$$

This estimate is consistent since $(1/m^{d-\theta_0}) \hat{I}_m^\tau(2\pi\mathbf{J}/m)/g_{\theta_0}(2\pi\mathbf{J}/m) \rightarrow^p c_0$ and $\theta_m^* - \theta_0 = o_p((\log(m))^{-1})$. As with earlier results, we can show that

$$m^\eta (c_m^* - a_m c_0) \xrightarrow{d} \mathcal{N}\left(0, c_0^2 \frac{\Lambda_2}{\Lambda_1^2} \left(\frac{2\pi}{\kappa}\right)^d\right),$$

where $a_m = ((g_{\theta_0}/g_{\theta_m^*})m^{\theta_m^* - \theta_0}) \rightarrow^p 1$ under similar conditions as the those of Theorem 1. However, we have been unable to show the asymptotic normality of c_m^* since it is not easy to obtain the convergence rate of a_m .

3. Numerical Results

Some simulations were conducted to evaluate performance of the proposed estimation method and the bias-reduced estimate of θ . A data example was also carried out. The expression of $g_{c,\theta}(\boldsymbol{\lambda})$ in (2.2) involves an infinite sum so that

we need to truncate it in practice. We used $\mathbf{Q} \in [-n, n]^d \cap \mathbb{Z}^d$ with $n = 60$ for truncation. The results with $n = 30$ were also looked at and not much difference in the results was found.

3.1. Simulation studies on \mathbb{R}

We investigated performance of the bias-reduced estimate, and compared it with approaches lacking bias correction, using two covariance models on \mathbb{R} . The first was a damped oscillation covariance function on \mathbb{R} (Yaglom (1987)), with covariance function

$$K(x) = \sigma^2 \exp(-\beta|x|) \cos(\omega_0 x), \quad (3.1)$$

for $x \in \mathbb{R}$, where β and ω_0 are positive. The spectral density for (3.1) satisfies $f(\lambda) \sim c|\lambda|^{-\theta}$ as $|\lambda| \rightarrow \infty$, where $c = \sigma^2\beta/\pi$ and $\theta = 2$. We considered $(\sigma^2, \beta, \omega_0) = (\pi^2, 1, 1)$ so that $c_0 = \pi$ and $\theta_0 = 2$.

We generated 100 data sets from the Gaussian process with covariance K . The observation domain was $[0, 10]$ with $m = 2,000$. For bias-reduced estimates, we used $m_1 = 500$ and 1,000. For the number of differencing steps, we took $\tau = 1, 2$. To obtain the smoothed periodogram, we used $\kappa = 5$ and $\gamma = 1/3$. We compared the bias-reduced estimates with the estimator without bias-reduced update introduced here, as well as the estimator from Wu, Lim, and Xiao (2013).

Bias was calculated as the average of difference between the estimates and the true value. STD is the standard deviation of the estimates. Table 1 shows the results for the comparison of estimates.

The second choice was a Matérn covariance model (Stein (1999)). We considered the isotropic Matérn covariance

$$K(|\mathbf{x}|) = \frac{\phi (\beta|\mathbf{x}|)^\nu}{\Gamma(\nu + d/2)2^{\nu-1}} \mathcal{K}_\nu(\beta|\mathbf{x}|), \quad (3.2)$$

where $\mathbf{x} \in \mathbb{R}^d$ and \mathcal{K}_ν is a modified Bessel function of the second kind. The corresponding spectral density is $f(\boldsymbol{\lambda}) = \phi\beta^{2\nu}\pi^{-d/2}(\beta^2 + |\boldsymbol{\lambda}|^2)^{-(\nu+d/2)}$. We took $(\phi, \beta, \nu) = (2, 1, 1.3)$ with $d = 1$ so that $(c_0, \theta_0) = (1.13, 3.6)$. All other settings were same as those in the first example except that $\tau = 2, 3$, as τ needs to satisfy $2\tau > \theta_0 - 1 = 3$. The results are given in Table 2.

In both tables, results in the SINGLE column are comparable to those in the WLX column. This suggests that the same asymptotic results hold for both approaches. Also, we can see the negative bias for $c^* < c_0$ and the positive bias for $c^* > c_0$, in the SINGLE columns in Tables 1 and 2, as supporting Theorem 4. Although we were not able to prove Theorem 4 for the WLX approach, we conjecture that the same results holds based on the simulation results.

Tables 1 and 2 show that the bias for the estimate of θ gets larger as c^* deviates more from the true value for both SINGLE and WLX approaches. The

Table 1. Bias and standard deviation of the estimates of θ under different c^* s for the covariance model given in (3.1). $m = 2,000$ is considered. The first two columns under ‘Bias-reduced’ category are when $m_1 = 500$ is used. The next two columns under ‘Bias-reduced’ category are when $m_1 = 1,000$ is used. The value of Bias and STD are scaled by 10^2 .

$(\sigma^2, \beta, \omega_0) = (\pi^2, 1, 1), (c_0, \theta_0) = (\pi, 2)$								
$\tau = 1$								
	Bias-reduced				SINGLE		WLX	
	Bias	STD	Bias	STD	Bias	STD	Bias	STD
$c^* = 0.5$	6.14	9.48	9.40	17.20	-29.73	1.00	-29.74	1.00
2	0.81	9.91	4.73	17.80	-7.38	1.02	-7.39	1.03
π	-1.51	10.04	2.76	17.99	0.03	1.03	0.02	1.03
5	-4.09	10.16	0.51	18.08	7.72	1.04	7.71	1.04
25	-14.47	10.40	-8.30	18.50	34.66	1.06	34.65	1.06
$\tau = 2$								
	Bias-reduced				SINGLE		WLX	
	Bias	STD	Bias	STD	Bias	STD	Bias	STD
$c^* = 0.5$	6.08	10.84	8.66	19.97	-29.75	1.01	-29.75	1.01
2	0.75	11.35	3.98	20.68	-7.40	1.03	-7.40	1.04
π	-1.57	11.50	1.98	20.90	0.006	1.04	0.01	1.05
5	-4.15	11.63	-0.27	21.01	7.69	1.05	7.71	1.06
25	-14.52	11.92	-9.89	19.16	34.64	1.08	34.63	1.08

bias-reduced estimates introduced in Section 2.2 reduce the bias from that seen in the other approaches. However, the variance gets larger. In looking at this as a bias-variance trade-off, the MSE can be obtained from Tables 1 and 2. The increase in MSE from SINGLE and WLX by moving c^* away from the true value is larger than the increase in MSE from the bias-reduced approach. For example, MSE for the bias-reduced approach when $\tau = 1$ and $c^* = 0.5$ with $m_1 = 500$ (Table 1) is 0.01, while MSE for SINGLE and WLX for the corresponding case is 0.09. Also, MSEs from the bias-reduced approach are more stable for various c^* compared to SINGLE and WLX. When the fixed c^* is the true c_0 , the SINGLE and WLX approaches provide the best results. But in terms of MSE, the bias-reduced approach provides a better result overall.

3.2. Simulation studies on \mathbb{R}^2

This section compares the performance of simulation results for the proposed approach with those of other approaches on \mathbb{R}^2 . Recall that the proposed method estimates the smoothness parameter θ while the other methods were designed to estimate a fractal index. We took $\theta = 2.6$, which corresponds to the fractal index $\alpha = 0.6$, and $\theta = 4.4$, which case does not have a corresponding fractal index.

Table 2. Bias and standard deviation of estimates of θ under different c^* s for the covariance model given in (3.2). $m = 2,000$ is considered. The first two columns under ‘Bias-reduced’ category are when $m_1 = 500$ is used. The next two columns under ‘Bias-reduced’ category are when $m_1 = 1,000$ is used. The value of Bias and STD are scaled by 10^2 .

$(\phi, \beta, \nu) = (2, 1, 1.3), (c_0, \theta_0) = (1.13, 3.6)$								
$\tau = 2$								
	Bias-reduced				SINGLE		WLX	
	Bias	STD	Bias	STD	Bias	STD	Bias	STD
$c^* = 0.5$	5.06	12.71	7.80	19.42	-14.44	1.17	-14.28	1.26
1.13	-0.75	12.73	2.78	19.44	-0.33	1.17	-0.18	1.26
5	-11.51	12.76	-6.50	19.47	25.42	1.17	25.57	1.26
10	-16.57	12.77	-10.86	19.48	37.43	1.17	37.59	1.26
20	-21.64	12.77	-15.24	19.49	49.45	1.17	49.60	1.26
$\tau = 3$								
	Bias-reduced				SINGLE		WLX	
	Bias	STD	Bias	STD	Bias	STD	Bias	STD
$c^* = 0.5$	3.35	12.75	7.15	19.53	-14.19	1.21	-14.07	1.30
1.13	-2.47	12.78	2.13	19.55	-0.09	1.22	0.04	1.30
5	-13.23	12.80	-7.15	19.58	25.66	1.22	25.79	1.30
10	-18.29	12.81	-11.51	19.60	37.68	1.22	37.80	1.30
20	-23.37	12.81	-15.89	19.60	49.70	1.22	49.83	1.30

We considered the approaches of Davies and Hall (1999) and Zhu and Stein (2002) for the comparison. Both use the least square estimate from the linear relation of the logarithm of variogram with a fractal index by replacing the variogram with its estimate. Depending on the variogram estimate, Davies and Hall cover both isotropic and anisotropic Gaussian fields when the data are observed on a square grid. Zhu and Stein’s approach considers different types of filters when calculating the variogram. In our simulation study, we used filter 1 as described in Zhu and Stein (2002).

The Gaussian random field with the covariance function as (3.2) with $(\sigma^2, \beta, \nu) = (2, 2.5, 0.3)$ and $(\sigma^2, \beta, \nu) = (1.5, 2.1, 1.2)$ was generated 100 times for each set of parameters. The corresponding values of (c_0, θ_0) were $(1.10, 2.60)$ and $(2.83, 4.40)$, respectively. The fractal index is well-defined for the first set of parameters while the second set of parameters yields a smooth surface. That is not characterized by a fractal index. The observation domain was $[0, 1]^2$ with the number of grids $n = m^2 = 100^2$. For bias-reduced estimates, we used $m_1 = 60$ for the one-step update.

Our approach requires a choice of the number of differencing steps (τ), and the number of frequencies used to calculate smoothed periodogram that is controlled by κ and the value of c . We chose $\tau = 1, 2$, $\kappa = 1, 2, 3, 4$, and three values

Table 3. Bias and standard deviation of estimates of θ for the covariance model given in (3.2) with $d = 2$ and $m = 100$. ‘Bias-reduced’ is our proposed bias-reduced estimate, ‘ZS’ is the Filter 1 approach of Zhu and Stein (2002) and ‘DH’ is the isotropic approach of Davies and Hall (1999). The value of Bias and STD are scaled by 10^2 .

$(\sigma^2, \beta, \nu) = (2, 2.5, 0.3)$ $(c_0, \theta_0) = (1.10, 2.60)$			$(\sigma^2, \beta, \nu) = (1.5, 2.1, 1.2)$ $(c_0, \theta_0) = (2.83, 4.40)$		
	Bias	STD		Bias	STD
Bias-reduced	5.08	12.06	Bias-reduced	4.35	14.23
ZS	0.17	2.42	ZS	-	-
DH	-0.30	1.78	DH	-52.05	2.24

of c . For each such choice, we obtained a bias-reduced estimate and calculated bias and standard deviation using 100 datasets. We used the `fractaldim` package in **R** to obtain the estimates from the approaches of Davies and Hall (1999) and Zhu and Stein (2002). Table 3 gives the results.

The results for the first set of parameters show that DH and ZS are better than our approach on \mathbb{R}^2 , while our approach is better for the second set of parameters. The `fractaldim` package in **R** did not produce meaningful values for the ZS approach. Simulations suggest that the approaches developed for estimating a fractal dimension work well for estimating θ when the range of θ is restricted to define a fractal index but do not work well otherwise. Our approach can estimate θ reasonably well regardless of the range of θ .

3.3. Simulation studies on non-Gaussian or non-stationary processes

This section compares the results of the proposed approach with DH and ZS approaches on \mathbb{R}^2 for a non-Gaussian random field and a non-stationary random field. We wanted to investigate the robustness of our approach under violations of Gaussianity and stationarity.

For a non-Gaussian random field, we considered a χ^2 random field constructed by taking a square of a Gaussian random field. For the simulation, we generated 100 sets of a simulated χ^2 random field from the Gaussian random field with covariance function as (3.2) using $(\sigma^2, \beta, \nu) = (2, 2.5, 0.3)$ with the same observation domain and the number of grids as in section 3.2. For a non stationary random field, we considered a fractional Brownian surface that has a variogram $\gamma(\mathbf{h}) = C|\mathbf{h}|^\alpha$, where $\alpha \in (0, 2)$. We took $\alpha = 1.4$ in the simulation study.

As described in Section 3.2, we obtained the estimates using our proposed approach, DH and ZS for both cases. Table 4 shows the bias and standard deviation results. They indicate that DH and ZS are better overall than our approach. Here, for a χ^2 random field, we started from a Gaussian random field

Table 4. Bias and standard deviation of estimates from θ for a fractional Brownian surface (non-stationary) and a χ^2 field (non Gaussian). ‘Bias-reduced’ is our proposed bias-reduced estimate, ‘ZS’ is the Filter 1 approach of Zhu and Stein (2002) and ‘DH’ is the isotropic approach of Davies and Hall (1999). The value of Bias and STD are scaled by 10^2 .

	χ^2 field		Fractional Brownian surface		
	Bias	STD	Bias	STD	
Bias-reduced	-7.41	14.82	Bias-reduced	7.54	12.89
ZS	-4.24	3.84	ZS	-0.05	2.55
DH	-5.89	3.36	DH	-1.05	4.58

with a well-defined fractal index. Thus, one expects a similar result as in Table 3. Despite this, the results are reasonable since the simulated data violated our assumption of a stationary Gaussian random field.

3.4. Data examples

For a data example, we considered the arctic sea-ice profile data analyzed by Gneiting, Ševčíková, and Percival (2012). Various estimation methods of the fractal dimension are reviewed and applied to their data.

The data are sonar measurements of the underwater surface of sea-ice in the Arctic Ocean by submarines. The data are collected by the U.S. Navy and are available from the National Snow and Ice Data Center at http://nsidc.org/data/docs/noaa/g01360_upward_looking_sonar. The resolution of the data is one meter. For comparisons, we used the same profiles (sc98drft.002, 003 and 005-008) that were used in Gneiting, Ševčíková, and Percival (2012). Thus, we used the same blocks of six profiles that were shown in their Figure 23. Each block is 1,024 meters long so that the sample size is 1,024. We reproduce Figure 23 in our Figure 1. Although some blocks clearly show non-Gaussianity and/or non-stationarity, we did not do any transformation to alleviate such characteristics to compare the results provided in Gneiting, Ševčíková, and Percival (2012).

To obtain the estimate from the proposed method, we considered $\kappa = 5, 10, 20, 25$ and $\gamma = 1/3$ for bandwidth in computing smoothed periodograms. We set $\tau = 1, 2, 3$ and $c^* = 0.5, 2, 5, 10, 20$. For the bias-reduced estimate, we chose $m_1 = m/2 = 528$. Then, we obtained $4 \times 3 \times 5 = 60$ estimates for each block from six sea-ice profiles. We report the averaged value and standard deviation among these 60 values (in parentheses) in Table 5. Note that this standard deviation is different from the standard deviation we reported for the simulation study since the standard deviation in the simulation study is calculated among estimates from repeated data sets. The values of estimates from other methods are also given in the Table 5; they are taken from Gneiting, Ševčíková, and Percival (2012).

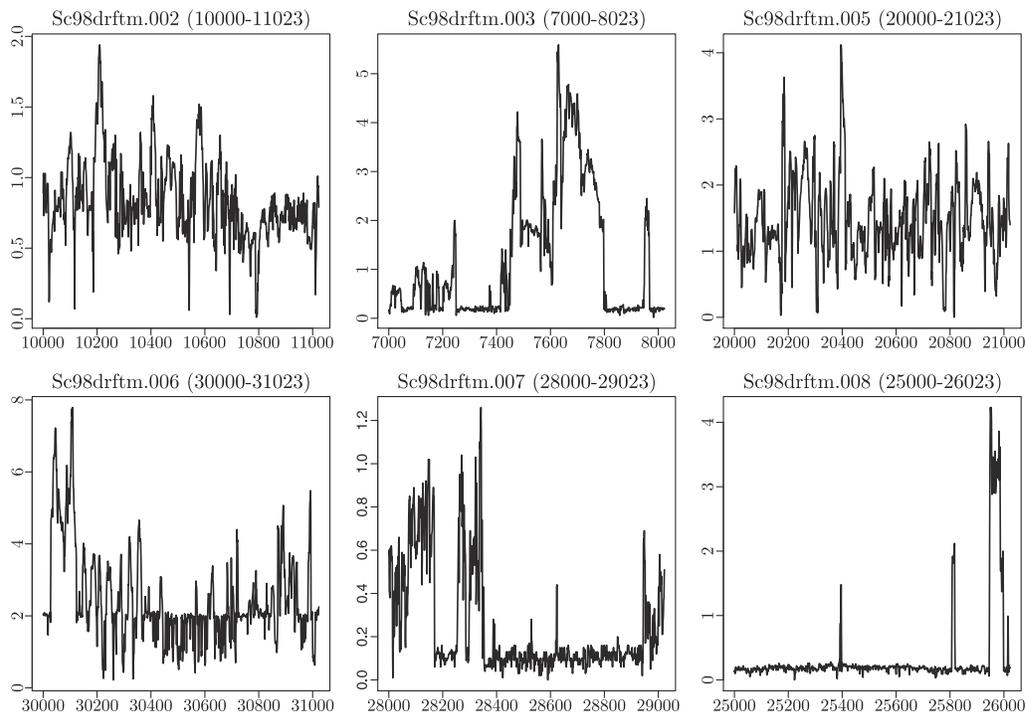


Figure 1. 6 artic sea-ice profiles: Each plot corresponds to a 1,024-length window from each profile.

The proposed bias-reduced estimates are comparable with estimates from other methods. Our estimates are similar to those from other methods for a possibly non-stationary profile (`sc98drftm.007`) and a non-Gaussian profile (`sc98drftm.008`), indicating that our approach is comparable to the other methods for some non-stationary or non-Gaussian data.

4. Discussion

For the bias-reduced estimate, we need to choose m_1 , a sub-sample size, to have another estimate to do the one-step update. Simulations show that the results are comparable for some different choices of m_1 , but this could be investigated further.

The proposed estimation method of estimating the smoothness parameter, θ , can be used to estimate a fractal index, as seen in simulations and a data example.

The objective function in Wu, Lim, and Xiao (2013) is not a local Whittle likelihood but a local Whittle-type objective function given that (asymptotic) Whittle likelihood has not been verified as a valid likelihood for some limiting

Table 5. Fractal dimension estimate of sea-ice profiles. The estimates from other methods were adopted from Gneiting, Ševčíková, and Percival (2012).

Method	sc98drftm.002	sc98drftm.003	sc98drftm.005
Variogram	1.43	1.37	1.38
Variation(1)	1.37	1.38	1.29
Variation(0.5)	1.32	1.37	1.24
Hall-Wood	1.39	1.35	1.30
Proposed approach	1.47(0.19)	1.32(0.17)	1.32(0.06)
Method	sc98drftm.006	sc98drftm.007	sc98drftm.008
Variogram	1.40	1.42	1.30
Variation(1)	1.32	1.42	1.33
Variation(0.5)	1.27	1.38	1.31
Hall-Wood	1.30	1.43	1.32
Proposed approach	1.39(0.06)	1.41(0.28)	1.31(0.18)

process under infill asymptotics. Even without the validity of the Whittle likelihood, one can still consider using more than local frequencies. However, to show the asymptotic properties of estimators, we need asymptotic properties of the sum of periodograms at those frequencies; this is not available and fairly challenging to show. Not many properties are known under infill asymptotics and we expect loss of efficiency in using a local Whittle Likelihood-type approach. As we did not assume any parametric spectral density model, efficiency loss may be unavoidable.

Supplementary Materials

The online supplementary material contains the proofs of some results in Section 2.

Acknowledgements

The authors thank the Editor, an associate editor, and the referees for their valuable suggestions in improving this paper. This work is partially supported by Ministry of Science and Technology of Taiwan under Grant No. 101-2118-M-259-004-MY2 and Michigan State University, USA under Grant No. 08-IRGP-1532.

References

- Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer.
- Bandyopadhyay, S. and Lahiri, S. N. (2009). Asymptotic properties of discrete Fourier transforms for spatial data. *Sankhyā Ser. A* **71**, 221-259.
- Block, A., Von Bloh, W. and Schellnhuber, H. J. (1990). Efficient box-counting determination of generalized fractal dimensions. *Phys. Rev. A* **42**, 1869-1874.

- Butler, J. B., Lane, S. N. and Chandler, J. H. (2001). Characterization of the structure of river-bed gravels using two-dimensional fractal analysis. *Math. Geol.* **33**, 301-330.
- Chan, G. and Wood, A. T. A. (2000). Increment-based estimator of fractal dimension for two-dimensional surface data. *Statist. Sinica* **10**, 343-376.
- 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.
- Chan, G., Hall, P. and Poskitt, D. S. (1995). Periodogram-based estimators of fractal properties. *Ann. Statist.* **23**, 1684-1711.
- 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.
- Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.
- Davies, S. and Hall, P. (1999). Fractal analysis of surface roughness by using spatial data. *J. Roy. Statist. Soc. Ser. B* **61**, 3-37.
- Deems, J. S., Fassnacht, S. R. and Elder, K. J. (2006). Fractal distribution of snow depth from lidar data. *J. Hydrometeor.* **7**, 285-297.
- Emery, X. (2005). Variograms of order ω : A tool to validate a bivariate distribution model. *Math. Geol.* **37**, 163-181.
- Folland, G. B. (1999). *Real Analysis*. Wiley-Interscience Publication, New York.
- Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *J. Amer. Statist. Assoc.* **102**, 321-331.
- Gneiting, T., Ševčíková, H. and Percival, D. B. (2012). Estimators of fractal dimension: Assessing the roughness of time series and spatial data. *Statist. Sci.* **77**, 247-277.
- Hall, P. and Wood, A. (1993). On the performance of boxcounting estimators of fractal dimension, *Biometrika*. **80**, 246-252.
- Kent, J. T. and Wood, 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.
- Klinkenberg, B. (1994). A review of methods used to determine the fractal dimension of linear features. *Math. Geol.* **26**, 23-46.
- Liebovitch, L. S. and Toth, T. (1989). A fast algorithm to determine fractal dimensions by box counting. *Phys. Lett. A*. **141**, 386-390.
- Lim, C. and Stein, M. L. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *J. Multivariate Anal.* **99**, 1962-1984.
- Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman and Co., San Francisco, CA.
- Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on R^d . *J. Roy. Statist. Soc. Ser. B* **71**, 191-217.
- Milne, B. T. (1992). Spatial aggregation and neutral models in fractal landscapes, *Amer. Naturalist*. **139**, 32-57.
- Stein, M. L. (1999). *Interpolation of Spatial Data*. Springer, New York.
- Vázquez, E. V., Miranda, J. G. V. and Conzález, A. P. (2005). Characterizing anisotropy and heterogeneity of soil surface microtopography using fractal models. *Ecological Modeling*. **182**, 337-353.
- Wu, W., Lim, C. and Xiao, Y. (2013). Tail estimation of the spectral density for a stationary gaussian random field. *J. Multivariate Anal.* **116**, 74-91.

- Xue, Y. and Xiao, Y. (2011). Fractal and smoothness properties of space-time Gaussian models. *Frontiers Math.* **6**, 1217-1246.
- Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions I: Basic Results*. Springer, New York.
- Zhu, Z. and Stein, M. (2002). Parameter estimation for fractional Brownian surfaces. *Statist. Sinica* **12**, 863-883.

Department of Applied Mathematics, National Dong Hwa University No. 1, Sec. 2, Da Hsueh Rd., Shoufeng, Hualien 97401, Taiwan.

E-mail: wuweiyong1011@mail.ndhu.edu.tw

Department of Statistics, Seoul National University, Seoul, Korea.

E-mail: limc@stats.snu.ac.kr

(Received November 2014; accepted December 2015)