DSSR

Discussion Paper No. 136

Gaussian semiparametric estimation of two-dimensional intrinsically stationary random fields

Yoshihiro Yajima and Yasumasa Matsuda

July, 2023

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

Gaussian semiparametric estimation of two-dimensional intrinsically stationary random fields

YOSHIHIRO YAJIMA[1,a] and YASUMASA MATSUDA[2,b]

[1]Department of Economics, University of Tokyo, Tokyo, Japan, [a]yajima@e.u-tokyo.ac.jp
[2]Graduate School of Economics and Management, Tohoku University, Sendai, Japan, [b]yasumasa.matsuda.a4@tohoku.ac.jp

We consider Gaussian semiparametric estimation (GSE) for two-dimensional intrinsically stationary random fields (ISRFs) observed on a regular grid and derive its asymptotic properties. Originally GSE was proposed to estimate long memory time series models in a semiparametric way either for stationary or nonstationary cases. We try an extension of GSE for time series to anisotropic ISRFs observed on two dimensional lattice that include isotropic fractional Brownian fields (FBF) as special cases, which have been employed to describe many physical spatial behaviours. The GSE extended to ISRFs is consistent and has a limiting normal distribution with variance independent of any unknown parameters as sample size goes to infinity, under conditions we specify in this paper. We conduct a computational simulation to compare the performances of it with those of an alternative estimator on the spatial domain.

Keywords: Anisotropic Random Fields; Discrete Fourier Transform; Fractional Brownian Fields; Intrinsically Stationary Random Fields; Spectral Density Function

1. Introduction

An intrinsically stationary random field (ISRF) has been applied for modelling of statistical dependence in spatial data and discussed extensively both in theory and practice. See e.g. Chilès and Delfiner (2012), Cressie (1993), Gikhman and Skorokhod (2004), Huang et al. (2011), Itô (1953), Lee et al. (2016), Matheron (1973), Solo (1992), Stein (1999), Yaglom (1957) and the references therein. Let $\{X(\boldsymbol{s}) : \boldsymbol{s} \in \boldsymbol{R}^d\}$ be a d-dimensional random field. Though an ISRF can be difined for any integer d, we specialize to the two-dimensional ISRF, $d = 2$ throughout this paper. Because the case of $d = 2$ is much important in practice and the theoretical results developed in the subsequent sections can be generalized for a larger d but their derivation is prohibitively lengthy.

If the increment $Z_{\boldsymbol{h}}(\boldsymbol{s}) = X(\boldsymbol{s} + \boldsymbol{h}) - X(\boldsymbol{s})$ for any fixed $\boldsymbol{h}(\in \boldsymbol{R}^2) = (h_1, h_2)'$ where $'$ means the transpose is a stationary random field, $\{X(\boldsymbol{s})\}$ is called an ISRF . Then $\{X(\boldsymbol{s})\}$ is characterized by

$$E(X(\boldsymbol{s} + \boldsymbol{h}) - X(\boldsymbol{s})) = 0,$$

$$\mathrm{Var}(X(\boldsymbol{s} + \boldsymbol{h}) - X(\boldsymbol{s})) = 2\gamma(\boldsymbol{h}),$$

where $2\gamma(\boldsymbol{h})$ is the variogram function (Chilès and Delfiner (2012), Cressie (1993)). If $X(\boldsymbol{0}) = 0$ where $\boldsymbol{0} = (0, 0)'$, we have

$$\mathrm{Cov}(X(\boldsymbol{s}), X(\boldsymbol{t})) = \gamma(\boldsymbol{t}) + \gamma(\boldsymbol{s}) - \gamma(\boldsymbol{t} - \boldsymbol{s}). \tag{1}$$

Let $(\boldsymbol{\lambda}, \boldsymbol{h})$ be the inner product, $\lambda_1 h_1 + \lambda_2 h_2$, and $\| \boldsymbol{\lambda} \|$ be the norm, $\sqrt{(\lambda_1^2 + \lambda_2^2}$for $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)'$ and $\boldsymbol{h} = (h_1, h_2)'(\in \boldsymbol{R}^2)$. Then if $2\gamma(\boldsymbol{h})$ is a continuous function on $\boldsymbol{R}^2$ satisfying $\gamma(\boldsymbol{0}) = 0$, it has the spectral representation

$$2\gamma(\boldsymbol{h}) = \int_{\boldsymbol{R}^2} \frac{1-\cos((\boldsymbol{\lambda},\boldsymbol{h}))}{(2\pi)^2} G(d\boldsymbol{\lambda}) + Q(\boldsymbol{h}), \quad (2)$$

where $Q(\boldsymbol{h})(\geq 0)$ is a quadratic form and $G(\boldsymbol{\lambda})$ is a positive, symmetric measure such that $\| \boldsymbol{\lambda} \|^2 G(\boldsymbol{\lambda})$ is continuous at the origin and

$$\int_{\boldsymbol{R}^2} \frac{\| \boldsymbol{\lambda} \|^2}{1 + \| \boldsymbol{\lambda} \|^2} G(d\boldsymbol{\lambda}) < \infty. \quad (3)$$

See Chilès and Delfiner (2012), Cressie (1993), Solo (1992), Yaglom (1957).

Hereafter we assume that $Q(\boldsymbol{h}) \equiv 0$ and $G(\boldsymbol{\lambda})$ is absolutely continuous with density $g(\boldsymbol{\lambda})$. Then (2) and (3) reduce to

$$2\gamma(\boldsymbol{h}) = \int_{\boldsymbol{R}^2} \frac{1-\cos((\boldsymbol{\lambda},\boldsymbol{h}))}{(2\pi)^2} g(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \quad (4)$$

and

$$\int_{\boldsymbol{R}^2} \frac{\| \boldsymbol{\lambda} \|^2}{1 + \| \boldsymbol{\lambda} \|^2} g(\boldsymbol{\lambda}) d\boldsymbol{\lambda} < \infty, \quad (5)$$

respectively.

An ISRF is said to be isotropic if $\gamma(\boldsymbol{h})$ depends only on $\| \boldsymbol{h} \|$, or equivalently when $g(\boldsymbol{\lambda})$ depends only on $\| \boldsymbol{\lambda} \|$. Otherwise it is said to be anisotropic.

An interesting class of ISRFs that has been applied to spatial data analysis is a fractional Brownian field (FBF). See e.g. Adler (1981), Mandelbrot and Van Ness (1968), Samorodnitsky and Taqqu (1994), Zhu and Stein (2002) and the references therein for empirical or theoretical details. FBF is a Gaussian isotropic ISRF with $2\gamma(\boldsymbol{h}) = C \| \boldsymbol{h} \|^{2H}$, which is shown in Yaglom (1957) to correspond to

$$g(\boldsymbol{\lambda}) = CHK_H \|\boldsymbol{\lambda}\|^{-2-2H}, \quad (6)$$

where

$$K_H = \pi 2^{2H+2} \Gamma(H+1)/\Gamma(1-H),$$

C is a scale parameter and H is called the Hurst effect, which is a spatial memory parameter with a larger value corresponding to a stronger correlations. See Chilès and Delfiner (2012), Cressie (1993), Huang et al. (2011), Lee et al. (2016), Matheron (1973), Solo (1992) and the references therein. H must be in $(0, 1)$ to satisfy (5).

The aim of this paper is to establish asymptotic properties of Gaussian semiparametric estimation (GSE) when it is applied to lattice samples of a class of ISRFs which includes the FBF as a special case. Specifically we consider an ISRF whose spectral density is defined by

$$g(\boldsymbol{\lambda}) = g_o(\boldsymbol{\lambda}) \|\boldsymbol{\lambda}\|^{-2-2H}.$$

If $g_o(\boldsymbol{\lambda})$ is a constant, it reduces to FBF. If it is not a constant, especially depends on the direction of $\boldsymbol{\lambda}$, it is an anisotropic ISRF.

Originally GSE was proposed to estimate semiparametric long memory time series models(Künsch (1987), Robinson (1995), Velasco (1999b)). We show that GSE is still consistent and has the limiting normal distribution as the sample size goes to infinity for the ISRF mentioned above.

The rest of paper is organized as follows. In Section 2, we specify a class of ISRFs in a semiparametric way and adjust the original GSE for time series to that for ISRFs in the two kinds of ways. The first one is a straightforward extension, while the second one is modified to achieve a better performance in finite samples to reduce bias caused by aliasing effects. Section 3 shows that the estimators are consistent and have the limiting normal distribution with mean 0 and variance independent of unknown parameters. In Section 4, we conduct some computational experiments to compare the performance of our estimators to those of an alternative one on the spatial domain by Zhu and Stein (2002). Finally concluding remarks are shown in Section 5. The technical lemmas and the propositions are shown in the Supplementary Material(Yajima and Matsuda (2023)). Though some of them have already been given by Yajima and Matsuda (2020b), they are also included in Yajima and Matsuda (2023) to make the paper self-contained.

2. Models and estimators

Hereafter we assume that $\{X(\boldsymbol{s}), \boldsymbol{s} \in R^2\}$ is a Gaussian ISRF on R^2 and the sampling sites are square lattices denoted by $s_{qr} = (q, r), q, r = 1, \dots, n$ and, hence, the sample size is n^2. Then $X(\boldsymbol{s})$ is denoted by $X(s_{qr})$.

We have to remark the two points in these assumptions. First it is difficult but important to construct a non Gaussian ISRF. The difficulty is in that a nonlinear transformation of an ISRF is no longer an ISRF. If $\{X(\boldsymbol{s})\}$ is a stationary Gaussian random field and $G(x)$ is a nonlinear function, $\{G(X(\boldsymbol{s}))\}$ is a non Gaussian stationary random field. However it does not hold if $\{X(\boldsymbol{s})\}$ is a Gaussian ISRF. Next it is also interesting to derive asymptotic properties of the GSE under a diffrent sampling scheme like irregulary spaced data or infill asymptotics(Cressie (1993)). These issues are to be considered in future.

We denote $g(\boldsymbol{\lambda})$ by $g(\lambda_1, \lambda_2)$. We consider the following class of the density functions.

Assumption 2.1. *$X(\boldsymbol{s})$ is a Gaussian ISRF with the spectral density $g(\lambda_1, \lambda_2)$ expressed by*

$$g(\lambda_1, \lambda_2) = g_o(\lambda_1, \lambda_2) \parallel \boldsymbol{\lambda} \parallel^{-2H-2}, \quad 0 < H < 1,$$

where $g_o(\lambda_1, \lambda_2)$ is a nonnegative with $g_o(0, 0) > 0$, symmetric, $g_o(\lambda_1, \lambda_2) = g_o(-\lambda_1, -\lambda_2)$, twice continuously differentiable function for $-\infty < \lambda_1, \lambda_2 < \infty$ and is bounded with bounded first and second order partial derivatives.

If $g_o(\lambda_1, \lambda_2)$ is a constant, it reduces to an FBF. Otherwise, it allows aninsotropic models that include a moving average random field of FBFs or an additive model being composed of an FBF and a stationary random field (Yajima and Matsuda (2020a)).

Now we introduce Gaussian semiparametric estimation (GSE) of H. First we construct discrete Fourier transforms (DFTs) of the tapered observations. Tapering observations has been helpful for both time series and spatial data analysis (Dahlhaus (1983), Dahlhaus and Künsch (1983), Guyon (1995), Priestley (1981)) to avoid a leakage effect or reduce bias of estimators. Following Velasco (1999a), we define a sequence of data tapers $\{h_t : t = 1, \dots, n\}$.

Definition 2.1. Let p be a positive integer. $\{h_t : t = 1, \dots, n\}$ is called a sequence of data tapers of order p if it satisfies the following conditions.

(1) h_t is positive and symmetric around $t = n/2$ with $\max_{1 \le t \le n} h_t = 1$.
(2) For any $n > 0$, there exists a constant $b, 0 < b < \infty$, which may depend on n so that $\sum_{t=1}^{n} h_t^2 = bn$ holds.

(3) For $N = n/p$, which we assume as an integer, the kernel $D_p = \sum_{t=1}^{n} h_t \exp(i\lambda t)$ satisfies

$$D_p(\lambda) = \frac{a_n(\lambda)}{n^{p-1}}[\sin(N\lambda/2)/\sin(\lambda/2)]^p, \tag{7}$$

where $a_n(\lambda)$ is a complex function, whose modulus is bounded and bounded away from zero, with $p-1$ derivatives, all bounded in modulus as n increases for $\lambda \in [-\pi, \pi]$.

A few examples of data tapers are given in the Supplementary Material included in Yajima and Matsuda (2023). Hereafter for $p = 1$, we assume $h_t = 1$ for any t. Then $D_1(\lambda)$ reduces to the Dirichlet kernel, given by

$$D_1(\lambda) = \exp(i(n+1)/2)\sin(n\lambda/2)/\sin(\lambda/2).$$

Then we define the tapered DFT $w_p(\omega_{j_1,j_2})$ and the periodogram $I_p(\omega_{j_1,j_2})$ by

$$w_p(\omega_{j_1,j_2}) = \frac{1}{2\pi \sum_{t=1}^{n} h_t^2} \sum_{q,r=1}^{n} h_q h_r X(s_{qr}) \exp(i(\omega_{j_1,j_2}, s_{qr})),$$
$$I_p(\omega_{j_1,j_2}) = |w_p(\omega_{j_1,j_2})|^2,$$

respectively, where ω_{j_1,j_2} is the bivariate Fourier frequency,

$$\omega_{j_1,j_2} = (\omega_{j_1}, \omega_{j_2})' = \left(\frac{2\pi j_1}{n}, \frac{2\pi j_2}{n}\right)', \ -\left[\frac{(n-1)}{2}\right] \le j_1,\ j_2 \le \left[\frac{n}{2}\right],$$

and $[x]$ is the integer part of x. Hence $(\omega_{j_1,j_2}, s_{qr}) = qw_{j_1} + rw_{j_2}$. Next we define the normalized tapered DFT by

$$v_p(\omega_{j_1,j_2}) = \frac{w_p(\omega_{j_1,j_2})}{\left\{G(\omega_{j_1}^2 + \omega_{j_2}^2)^{-(H+1)}\right\}^{1/2}},$$
$$v_{pR}(\omega_{j_1,j_2}) = \mathrm{Re}(v_p(\omega_{j_1,j_2})),$$
$$v_{pI}(\omega_{j_1,j_2}) = \mathrm{Im}(v_p(\omega_{j_1,j_2})),$$

where $G = g_o(0,0)/(8\pi^2)$.

Now we construct the GSE based on the periodogram. We denote by G_0 and H_0 the true parameters, and by G and H any admissible values. Then let $\tilde{v}_{pR}(j_1, j_2)$ and $\tilde{v}_{pI}(j_1, j_2)$ be $v_{pR}(\omega_{j_1 p\xi, j_2 p\xi})$ and $v_{pI}(\omega_{j_1 p\xi, j_2 p\xi})$ evaluated at $H = H_0$ and $G = G_0$ respectively. Define the closed interval of admissible estimators of H_0, $\mathcal{H} = [\Delta_1, \Delta_2]$, where Δ_1 and Δ_2 are numbers chosen such that $0 < \Delta_1 < \Delta_2 < 1$. We can choose Δ_1 and Δ_2 arbitrarily close to 0 and 1 respectively. Consider the objective function

$$Q(G,H) = \frac{1}{m} \sum_{(j_1,j_2)\in S_n} \left\{ \log\left(G(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{-H-1}\right) + \frac{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{H+1}}{G} I_p(\omega_{j_1 p\xi, j_2 p\xi}) \right\}, \tag{8}$$

where S_n is the set of frequencies used for estimation defined by

$$S_n = \left\{ (j_1, j_2) | 0 < \left(\frac{j_1 p\xi}{n}\right)^2 + \left(\frac{j_2 p\xi}{n}\right)^2 \le r_{U,n}^2,\ 0 < j_1,\ j_2, b_L \le \frac{j_2}{j_1} \le b_U \right\},$$

with $0 < b_L < 1 < b_U$ and m is the cardinality of S_n. ξ is a fixed integer or diverges to ∞ as $n \to \infty$. We note that by approximating an integral by its Riemann sum, m satisfies

$$m = \tilde{m} + O(\tilde{r}_{U,n}) = \frac{\theta_U - \theta_L}{2} \tilde{r}_{U,n}^2 + O(\tilde{r}_{U,n}), \tag{9}$$

where

$$\begin{aligned} \tilde{m} &= \int \int_{\tilde{D}_n} dx dy, \\ \tilde{D}_n &= \left\{ (x, y) | 0 \le x^2 + y^2 \le \tilde{r}_{U,n}^2,\ 0 < x, y,\ b_L \le \frac{y}{x} \le b_U \right\}, \\ \theta_U &= \arctan b_U, \left(> \frac{\pi}{4} \right), \quad \theta_L = \arctan b_L, \left(< \frac{\pi}{4} \right), \\ \tilde{r}_{U,n} &= n r_{U,n} / (p\xi), \end{aligned}$$

respectively.

Then the first estimator is defined by

$$(\hat{G}_n, \hat{H}_n) = \arg \min_{0<G<\infty, H \in \mathcal{H}} Q(G, H).$$

An essential difference from time series is that the components of the bivariate Fourier frequency, $\omega_{j_1 p\xi}$ and $\omega_{j_2 p\xi}$ are not able to behave independently since the ratio j_2/j_1 is bounded and bounded away from 0. Furthermore $r_{U,n}$ and ξ have to be chosen appropriately in a way to ensure that

$$\frac{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{H_0+1}}{G_0} I_p(\omega_{j_1 p\xi, j_2 p\xi}) = \tilde{v}_{pR}^2(\omega_{j_1 p\xi, j_2 p\xi}) + \tilde{v}_{pI}^2(\omega_{j_1 p\xi, j_2 p\xi})$$

are independently and identically distributed asymptotically so that $\hat{G}_n$ and $\hat{H}_n$ are consistent and asymptotically normally distributed.

Now we show a more explicit computational form of $\hat{G}_n$ and $\hat{H}_n$, which we call the first estimator ignoring aliasing effects. Hereafter for notational simplicity $\sum$ implies $\sum_{(j_1, j_2) \in S_n}$ unless otherwise specified. If we solve the following equation on G given a fixed H

$$\frac{\partial Q}{\partial G} = \frac{1}{m} \sum \left(\frac{1}{G} - \frac{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{H+1} I_p(\omega_{j_1 p\xi, j_2 p\xi})}{G^2} \right) = 0,$$

the solution is

$$\hat{G}(H) = \frac{1}{m} \sum (\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{H+1} I_p(\omega_{j_1 p\xi, j_2 p\xi}).$$

By substituting $\hat{G}(H)$ to G in (8),

$$Q(\hat{G}(H), H) = \log \hat{G}(H) - \frac{H+1}{m} \sum \log(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2) + 1.$$

Consequently

$$\hat{H}_n = \arg \min_{H \in \mathcal{H}} R(H),$$

where

$$R(H) = \log \hat{G}(H) - \frac{H+1}{m} \sum \log(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2).$$

Let us introduce the second estimator accounting for aliasing effects. The first estimator ignores aliasing effects caused by lattice sampling of continuous data. Because the spectral density function of the ISRF on square lattices is given by

$$\tilde{g}(\lambda) = g(\lambda) + \sum_{(p,q)\neq(0,0)} g(\lambda + 2\pi(p,q)), -\pi < \lambda_1, \lambda_2 \leq \pi,$$

where the second term on the right hand side is caused by aliasing, which is not negligible for processes when high frequency components are significant. Hence we incorporate it in the objective function by the integral approximation,

$$(2\pi)^{-2} \int_{||\lambda||>2\pi} G||\lambda||^{-2H-2} d\lambda = G\left\{2H(2\pi)^{2H+1}\right\}^{-1} = Gf(H),$$

which is expected to ease the bias caused from aliasing. Consequently the objective function $Q(G,H)$ is defined by

$$\begin{aligned} &Q^*(G,H) \\ &= \frac{1}{m} \sum_{(j_1,j_2)\in S_n} \log\left[G\left\{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{-H-1} + f(H)\right\}\right] + \frac{I_p(\omega_{j_1 p\xi, j_2 p\xi})}{G\left\{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{-H-1} + f(H)\right\}} \end{aligned} \quad (10)$$

Then the estimator is defined by

$$(G_n^*, H_n^*) = \arg \min_{0<G<\infty, H\in\mathcal{H}} Q^*(G,H).$$

Similar to $\hat{H}_n$, H_n^* is given by

$$H_n^* = \arg\min_H R^*(H),$$

where

$$\begin{aligned} R^*(H) &= \log G^*(H) + \frac{1}{m} \sum \log\{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{-H-1} + f(H)\}, \\ G^*(H) &= \frac{1}{m} \sum \frac{I_p(\omega_{j_1 p\xi, j_2 p\xi})}{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{-H-1} + f(H)}. \end{aligned}$$

3. Theoretical results

We assumed that $X(\mathbf{0}) = 0$. If $X(\mathbf{0}) \neq 0$, we replace $X(\boldsymbol{s})$ by $\tilde{X}(\boldsymbol{s}) = X(\boldsymbol{s}) - X(\mathbf{0})$ because $\tilde{X}(\boldsymbol{s})$ has the same variogram $2\gamma(\boldsymbol{h})$ as $X(\boldsymbol{s})$ and $\tilde{X}(\mathbf{0}) = 0$. Practically $X(\mathbf{0})$ is likely to be unknown. However it does not lose any generality because our estimator is based on the DFT's of the Fourier frequencies and hence

$$\begin{aligned}
&\sum_{q,r} h_q h_r \exp\left\{i(\omega_{j_1 p, j_2 p}, s_{qr})\right\} \\
&= \left\{\sum_q h_q \exp(i\omega_{j_1 p} q)\right\}\left\{\sum_r h_r \exp(i\omega_{j_2 p} r)\right\} \\
&= 0,
\end{aligned}$$

for $j_1 \neq 0$ or $j_2 \neq 0$. Consequently the DFTs of $X(\boldsymbol{s})$ are identical to those of $\tilde{X}(\boldsymbol{s})$, which is an advantage of our estimators.

The following assumptions on $r_{U,n}$, $\tilde{r}_{U,n}$ and ξ are introduced. Then the assumption on m is determined by (9).

Assumption 3.1. (i) *As n tends to* ∞*, for* $p = 1$*,* $r_{U,n} \to 0$*,* $\tilde{r}_{U,n} \to \infty$ *and* $\log n = O(\log \tilde{r}_{U,n})$*,*
(ii) *As n tends to* ∞*, for* $p \geq 2$*,* $r_{U,n} \to 0$*,*$\tilde{r}_{U,n} \to \infty$ *and* $\xi \to \infty$*.*

Assumption 3.2. (i) *As n tends to* ∞*, for* $p = 1$*,* $r_{U,n} = o(\tilde{r}_{U,n}^{-1/2})$ *,*$\tilde{r}_{U,n} \to \infty$ *and* $\log n = O(\log \tilde{r}_{U,n})$*.*
Furthermore $\xi^{-1} = o\tilde{r}_{U,n}^{-1/2-\epsilon})$ *for some* $\epsilon > 0$ *if* $H \leq 1/2$ *and* $\xi^{-1} = o(\tilde{r}_{U,n}^{-(H/(2(1-H))-\epsilon})$ *for some* $\epsilon > 0$ *if* $H > 1/2$*,*
(ii) *As n tends to* ∞*, for* $p \geq 2$*,* $r_{U,n} = o(\tilde{r}_{U,n}^{-1/2})$*,* $\tilde{r}_{U,n} \to \infty$ *and* $\xi^{-p} = o(\tilde{r}_{U,n}^{-1})$*.*

Then we have the following asymptotic properties of the estimators.

Theorem 3.1. *Under Assumption 3.1,* $\hat{H}_n$ *and* H_n^* *converge to* H_0 *in probability as* $n \to \infty$*.*

Theorem 3.2. *Under Assumptions 3.2,* $m^{1/2}(\hat{H}_n - H_0)$ *and* $m^{1/2}(H_n^* - H_0)$ *converge to* $N(0,1)$ *in distribution as* $n \to \infty$*.*

Remark 3.1. Here we give some remarks on the assumptions.

(1) It follows from Assumption 3.1.(i) and (ii) that for the consistency of the estimator ξ can be fixed for $p = 1$ contrary to $p \geq 2$. Because Proposition 3.2.(2) of the Supplementary Material in Yajima and Matsuda (2023) shows that the covariance between $v^*_{j_1 p\xi, j_2 p\xi}$ and $v^*_{k_1 p\xi, k_2 p\xi}$ is small if $\max(j_i, k_i), (i = 1, 2)$ is large for $p = 1$ while for $p \geq 2$, it can be large even if $\max(j_i, k_i), (i = 1, 2)$ is large with (j_1, j_2) and (k_1, k_2) being closer to each other and ξ being fixed. However to ensure the asymptotic normality of the estimator for $p = 1$, ξ has to diverge to ∞ and moreover the speed of its divergence depends on the unknown H if $H > 1/2$.
(2) We give examples of $r_{U,n}$ and ξ that satisfy Assumptions 3.1 and 3.2. Put $r_{U,n} = c_1 n^{-\tau_U}$ and $\xi = c_2 n^{\tau_\xi}$ with positive constants, $c_i (i = 1, 2)$. First consider Assumption 3.1. For $p = 1$, the assumptin is satisfied if $0 < \tau_U < 1 - \tau_\xi$. τ_ξ can be 0, which implies that ξ is fixed. While $\tau_\xi > 0$ is necessary for $p \geq 2$.
(3) Assumption 3.2 holds if τ_U and τ_ξ satisfy

$$\begin{aligned}
&\max((1-\tau_\xi)/3, 1-\tau_\xi(1+2/(1+2\epsilon)) < \tau_U < 1-\tau_\xi, \quad p = 1, H \leq 1/2, \\
&\max((1-\tau_\xi)/3, 1-\tau_\xi(1+2(1-H)/(H+2(1-H)\epsilon))) < \tau_U < 1-\tau_\xi, \\
&\quad p = 1, H > 1/2,
\end{aligned}$$

$$\max((1-\tau_\xi)/3, 1-(p+1)\tau_\xi)) < \tau_U < 1-\tau_\xi, \quad p \geq 2.$$

(4) The spatial domain estimator proposed by Zhu and Stein (2002), which is used as the benchmark in the simulation studies of the next section, has the consistent order of $1/n$ if the true underlying ISRF is FBF. While for $p \geq 2$, the consistrncy order of ou estimator is $m^{1/2} = O(1/n^{1-\tau_U-\tau_\xi})$ with $1-\tau_U-\tau_\xi < 2/3$, which implies that our estimator is less efficient than that of Zhu and Stein (2002). However if the true underlying ISRF is not FBF, it does hold no longer. As the simulation studies of the next section reveals that the estimator of Zhu and Stein (2002) is much biased but our estimator is still consistent and has the same consistecy order. It shows that our estimator is more robust than their estimator.

Now we prove these theorems. Hereafter C is a generic constant which can change depending on each context. We only consider $(\hat{G}_n, \hat{H}_n)$ because the assertion for (G_n^*, H_n^*) is shown in a similar way.

Proof of Theorem 3.1. We shall show the result by following the proof of Theorem 1 of Robinson (1995). However we have to prove some assertions in different ways from those for time series models. For example the summation by parts formula used in the proof of Theorem 1 of Robinson (1995) is not applicable to ISRFs. Define $S(H)$ by

$$\begin{aligned} S(H) &= R(H) - R(H_0) \\ &= \log \hat{G}(H) - \log \hat{G}(H_0) + \frac{1}{m}\sum \log\{(\omega_{j_1 p\xi}^2 + \omega_{j_2 p\xi}^2)^{H_0 - H}\}. \end{aligned}$$

For $\delta > 0$, let $N_\delta = \{H : |H - H_0| < \delta\}$ and $\overline{N}_\delta = (-\infty, \infty) - N_\delta$. Then

$$\begin{aligned} P(|\hat{H}_n - H_0| \geq \delta) &= P(\hat{H}_n \in \overline{N}_\delta \cap \mathcal{H}) \\ &= P(\inf_{\overline{N}_\delta \cap \mathcal{H}} R(H) \leq \inf_{N_\delta \cap \mathcal{H}} R(H)) \\ &\leq P(\inf_{\overline{N}_\delta \cap \mathcal{H}} S(H) \leq 0). \end{aligned}$$

Now define $\mathcal{H}_1 = \{H : \Delta \leq H \leq \Delta_2\}$ where $\Delta = \Delta_1$ when $H_0 < \frac{1}{2} + \Delta_1$ and $H_0 \geq \Delta > H_0 - \frac{1}{2}$ otherwise. When $H_0 \geq \frac{1}{2} + \Delta_1$, define $\mathcal{H}_2 = \{H : \Delta_1 \leq H < \Delta\}$, and otherwise take $\mathcal{H}_2$ to be empty. Then

$$P(|\hat{H}_n - H_0| \geq \delta) \leq P(\inf_{\mathcal{H}_1 \cap \overline{N}_\delta} S(H) \leq 0) + P(\inf_{\mathcal{H}_2} S(H) \leq 0). \tag{11}$$

First we shall prove that the first probability on the right hand side of (11) converges to 0. We apply the method developed by Walker (1964). Because the formula on the summation by parts, which is used by (3.2) of Robinson (1995) for a time series model, is not applicable to ISRFs. We shall show that $S(H_1)(= R(H_1) - R(H_0))$ converges to a positive value in probability for any $H_1 (\neq H_0)$. Then if we choose any constant, $K(H_1, H_0)$(say) that is less than it, we have

$$\lim_{n\to\infty} P(S(H_1) > K(H_1, H_0)) = 1. \tag{12}$$

Next we show that there exists $L_{\delta,n}(H_1)$ such that

$$|R(H_2) - R(H_1)| \leq L_{\delta,n}(H_1) \tag{13}$$

for all $H_1 \in \mathcal{H}_1$, $H_2 \in \mathcal{H}_1$ with $|H_1 - H_2| < \delta$ (δ possibly depending on H_1) and

$$\lim_{n\to\infty} P(L_{\delta,n}(H_1) < K(H_1, H_0)) = 1, \tag{14}$$

for a sufficiently small δ. Then applying (12), (13) and (14) instead of (15), (18) and (19) in Lemma 2 of Walker (1964), we have the assertion.

$S(H)$ can be expressed as

$$S(H) = U(H) - T(H),$$

where

$$\begin{aligned}
T(H) &= T_1(H) + T_2(H) + T_3(H),\\
T_1(H) &= \log\{\frac{\hat{G}(H_0)}{G_0}\} - \log\{\frac{\hat{G}(H)}{G(H)}\},\\
T_2(H) &= \log\{\frac{n^2}{\tilde{m}} \int\int_{D_n} \{(2\pi p\xi)^2(x^2+y^2)\}^{H-H_0} dxdy\}\\
&\quad -\log\{\frac{1}{m}\sum(\omega^2_{j_1 p\xi} + \omega^2_{j_2 p\xi})^{H-H_0}\},\\
T_3(H) &= (H-H_0)\frac{1}{m}\sum \log\{\omega^2_{j_1 p\xi} + \omega^2_{j_2 p\xi}\}\\
&\quad -(H-H_0)\frac{n^2}{\tilde{m}}\int\int_{D_n} \log\{(2\pi p\xi)^2(x^2+y^2)\}dxdy,\\
D_n &= \{(x,y)|0 \le x^2+y^2 \le (\frac{r_{U,n}}{p\xi})^2,\ 0 < x, y,\ b_L \le \frac{y}{x} \le b_U\},\\
U(H) &= \log\{\frac{n^2}{\tilde{m}} \int\int_{D_n} \{(2\pi p\xi)^2(x^2+y^2)\}^{H-H_0} dxdy\}\\
&\quad -(H-H_0)\frac{n^2}{\tilde{m}}\int\int_{D_n} \log\{(2\pi p\xi)^2(x^2+y^2)\}dxdy,\\
G(H) &= G_0 \frac{1}{m}\sum(\omega^2_{j_1 p\xi} + \omega^2_{j_2 p\xi})^{H-H_0}.
\end{aligned}$$

Then

$$\begin{aligned}
U(H) &= \log\{\frac{n^2}{\tilde{m}}\int\int_{D_n}(x^2+y^2)^{H-H_0}dxdy\}\\
&\quad -(H-H_0)\frac{n^2}{\tilde{m}}\int\int_{D_n}\log(x^2+y^2)dxdy\\
&= \log\{(\theta_U-\theta_L)\frac{n^2}{\tilde{m}}\frac{1}{2(H-H_0+1)}[r^{2(H-H_0+1)}]_0^{r_{U,n}/(p\xi)}\}\\
&\quad -(H-H_0)(\theta_U-\theta_L)\frac{n^2}{\tilde{m}}[r^2\log r - \frac{r^2}{2}]_0^{r_{U,n}/(p\xi)}\\
&= (H-H_0) - \log\{(H-H_0)+1\}.
\end{aligned} \tag{15}$$

The first equality follows from $nD_n = \tilde{D}_n$ and hence

$$\int\int_{D_n} dxdy = \frac{\tilde{m}}{n^2}.$$

Then it follows from (15) and Propositions 3.3-3.5 of the Supplementary Material in Yajima and Matsuda (2023) that for any $H_1(\neq H_0)$, $S(H_1)$ converges in probability to $(H_1 - H_0) - \log\{(H_1 - H_0) + 1\}$, a positive constant as $n \to \infty$.

While we have

$$R(H_2) - R(H_1) = U(H_2) - U(H_1) - \sum_{i=1}^{3}(T_i(H_2) - T_i(H_1)).$$

Next set $K(H_1, H_0)$ so that $0 < K(H_1, H_0) < (H_1 - H_0) - \log\{(H_1 - H_0) + 1$ and choose τ and δ of Proposition 3.6 of the Supplementary Material in Yajima and Matsuda (2023) so that $(C_U + C_2 + C_3)\delta + \tau < K(H_1, H_0)$. Then we define $L_{\delta,n}(H_1) = (C_U + C_2 + C_3)\delta + L_{1,\delta,n}(H_1)$. Consequently $L_{\delta,n}(H_1)$ satisfies (13) and (14).

Now we shall show that the second probability on the right hand side of (11) converges to 0 as $n \to \infty$. Define

$$\hat{Y}(H) = \frac{1}{m}\sum(\frac{j_1^2 + j_2^2}{q})^{H-H_0}(j_1^2 + j_2^2)^{H_0+1} I_p(\omega_{j_1 p\xi, j_2 p\xi}),$$

where $q = \exp(\frac{1}{m}\sum\log(j_1^2 + j_2^2))$. Note $q \leq \tilde{r}_{U,n}^2$.

Then

$$S(H) = \log(\hat{Y}(H)/\hat{Y}(H_0)).$$

We define a_{j_1,j_2} by

$$a_{j_1,j_2} = \begin{cases} (\frac{j_1^2+j_2^2}{q})^{\Delta - H_0}, & 0 < j_1^2 + j_2^2 \leq q \\ (\frac{j_1^2+j_2^2}{q})^{\Delta_1 - H_0}, & q < j_1^2 + j_2^2 \leq \tilde{r}_{U,n}^2 \end{cases}$$

Since

$$\inf_{\mathcal{H}_2} \hat{Y}(H) \geq \frac{1}{m}\sum a_{j_1,j_2}(j_1^2 + j_2^2)^{H_0+1} I_p(\omega_{j_1 p\xi, j_2 p\xi}),$$

$$\begin{aligned} & P(\inf_{\mathcal{H}_2} S(H) \leq 0) \\ & = P(\inf_{\mathcal{H}_2} \hat{Y}(H) \leq \hat{Y}(H_0)) \\ & \leq P(\frac{1}{m}\sum(a_{j_1,j_2} - 1)(j_1^2 + j_2^2)^{H_0+1} I_p(\omega_{j_1 p\xi, j_2 p\xi}) \leq 0). \end{aligned} \tag{16}$$

We shall show the probability of (16) converges to 0 as $n \to \infty$. First we evaluate $\frac{1}{m}\sum(a_{j_1,j_2} - 1)$. We have

$$\log q = \frac{\int\int_{\tilde{D}_n} \log(x^2 + y^2)dxdy + O(\tilde{r}_{U,n}\log\tilde{r}_{U,n})}{m} \tag{17}$$

$$= \frac{2(\theta_U - \theta_L)[r^2 \log r - r^2/2]_0^{\tilde{r}_{U,n}} + o(\tilde{r}_{U,n}^2)}{(\theta_U - \theta_L)\tilde{r}_{U,n}^2 + o(\tilde{r}_{U,n}^2)}$$

$$= \log \tilde{r}_{U,n}^2 - 1 + o(1).$$

While

$$\sum_{0<j_1^2+j_2^2\le q, b_L \le j_2/j_1 \le b_U} a_{j_1,j_2}$$

$$= q^{H_0-\Delta} \int_{\theta_L}^{\theta_U} \int_0^{\sqrt{q}} r^{2(\Delta-H_0)+1} d\theta dr + o(\tilde{r}_{U,n}^2)$$

$$= \frac{(\theta_U - \theta_L)q}{2(\Delta - H_0 + 1)} + o(\tilde{r}_{U,n}^2). \tag{18}$$

From (17),

$$\sum_{q\le j_1^2+j_2^2\le \tilde{r}_{U,n}^2, b_L \le j_2/j_1 \le b_U} a_{j_1,j_2}$$

$$= q^{H_0-\Delta_1} \int_{\theta_L}^{\theta_U} \int_{\sqrt{q}}^{\tilde{r}_{U,n}} r^{2(\Delta_1-H_0)+1} d\theta dr + o(\tilde{r}_{U,n}^2)$$

$$= \frac{(\theta_U - \theta_L)q^{H_0-\Delta_1}(\tilde{r}_{U,n}^{2(\Delta_1-H_0+1)} - q^{\Delta_1-H_0+1})}{2(\Delta_1 - H_0 + 1)} + o(\tilde{r}_{U,n}^2)$$

$$= \frac{(\theta_U - \theta_L)(\exp(\Delta_1 - H_0 + 1) - 1)q}{2(\Delta_1 - H_0 + 1)} + o(\tilde{r}_{U,n}^2). \tag{19}$$

By noting (9), (17), (18) and (19),

$$\frac{1}{m}\sum a_{j_1,j_2} = \frac{1}{e(\Delta - H_0 + 1)} + \frac{\exp(\Delta_1 - H_0 + 1) - 1}{e(\Delta_1 - H_0 + 1)} + o(1).$$

Hence there exists Δ and η so that $\Delta - H_0 + 1 > 1/2$ and

$$\frac{1}{m}\sum(a_{j_1,j_2} - 1)$$

$$> \eta + o(1).$$

Then the probability of (16) satisfies the following relations,

$$P(\frac{1}{m}\sum(a_{j_1,j_2} - 1)(\tilde{v}_{pR}(j_1,j_2)^2 + \tilde{v}_{pI}(j_1,j_2)^2) \le 0) \tag{20}$$

$$= P(\frac{1}{m}\sum(a_{j_1,j_2} - 1)(\tilde{v}_{pR}(j_1,j_2)^2 + \tilde{v}_{pI}(j_1,j_2)^2 - 1)$$

$$\le -\frac{1}{m}\sum(a_{j_1,j_2} - 1))$$

$$\le P(\frac{1}{m}|\sum(a_{j_1,j_2} - 1)(\tilde{v}_{pR}(j_1,j_2)^2 + \tilde{v}_{pI}(j_1,j_2)^2 - 1)|$$

$$\geq \frac{1}{m}\sum(a_{j_1,j_2}-1))$$

$$\leq P(\frac{1}{m}|\sum(a_{j_1,j_2}-1)(\tilde{v}_{pR}(j_1,j_2)^2+\tilde{v}_{pI}(j_1,j_2)^2-1)|\geq \eta+o(1))$$

$$\leq \frac{E[\{\frac{1}{m}|\sum(a_{j_1,j_2}-1)(\tilde{v}_{pR}(j_1,j_2)^2+\tilde{v}_{pI}(j_1,j_2)^2-1)|\}^2]}{(\eta+o(1))^2}$$

Now we consider the second order moment of the numerator of the last inequality. In the same way for $A(H)$ of Proposition 3.3 of the Supplementary Material in Yajima and Matsuda (2023), we express it as

$$E[\{\frac{1}{m}|\sum(a_{j_1,j_2}-1)(\tilde{v}_{pR}(j_1,j_2)^2+\tilde{v}_{pI}(j_1,j_2)^2-1)|\}^2] \tag{21}$$

$$= E[\{\frac{1}{m}|\sum_{(j_1,j_2)\in S'_n}+\sum_{(j_1,j_2)\in S_n\backslash S'_n}|\}^2]$$

$$= E[\{\frac{1}{m}|V_1+V_2|\}^2],\ \text{(say)},$$

where

$$S'_n=\{(j_1,j_2)|0<(\frac{j_1p\xi}{n})^2+(\frac{j_2p\xi}{n})^2\leq r^2_{L,n},\ 0<j_1,\ j_2, b_L\leq \frac{j_2}{j_1}\leq b_U\}.$$

Put $\tilde{r}_{L,n}=nr_{L,n}/(p\xi)$ and assume that $r_{L,n}/r_{U,n}\longrightarrow 0$ and $\tilde{r}_{L,n}\longrightarrow\infty$ as $n\longrightarrow\infty$.

Then

$$E\left(\frac{V_1^2}{m^2}\right) \tag{22}$$

$$=\frac{1}{m^2}[\sum_{(j_1,j_2)\in S'_n}(a_{j_1,j_2}-1)E\{\tilde{v}^2_{pR}(j_1,j_2)+\tilde{v}^2_{pI}(j_1,j_2)-1\}]^2$$

$$+\frac{1}{m^2}\sum_{(j_1,j_2),(l_1,l_2)\in S'_n}(a_{j_1,j_2}-1)(a_{l_1,l_2}-1)$$

$$\times 2[\{E(\tilde{v}_{pR}(j_1,j_2)\tilde{v}_{pR}(l_1,l_2))\}^2+\{E(\tilde{v}_{pR}(j_1,j_2)\tilde{v}_{pI}(l_1,l_2))\}^2$$

$$+\{E(\tilde{v}_{pI}(j_1,j_2)\tilde{v}_{pR}(l_1,l_2))\}^2+\{E(\tilde{v}_{pI}(j_1,j_2)\tilde{v}_{pI}(l_1,l_2))\}^2].$$

It follows from Proposition 3.2 of the Supplementary Material in Yajima and Matsuda (2023) that every moment is bounded unifomly in $j_i, l_i (i=1,2)$. While by noting (17)and $\Delta-H_0+1>1/2$, similar to (18).

$$\frac{1}{m}\sum_{(j_1,j_2)\in S'_n}|a_{j_1,j_2}-1|$$

$$\leq \frac{1}{m}\sum_{(j_1,j_2)\in S'_n}(a_{j_1,j_2}+1)$$

$$= O\left(\frac{q^{H_0-\Delta}\tilde{r}_{L,n}^{2(\Delta-H_0+1)}}{\tilde{r}_{U,n}^2}\right) + O\left((\frac{\tilde{r}_{L,n}}{\tilde{r}_{U,n}})^2\right)$$

$$= O\left((\frac{\tilde{r}_{L,n}}{\tilde{r}_{U,n}})^{2(\Delta-H_0+1)}\right) + O\left((\frac{\tilde{r}_{L,n}}{\tilde{r}_{U,n}})^2\right)$$

$$= o(1).$$

Consequently $E(V_1^2/m^2) = o(1)$.

Now we consider $E(V_2^2/m^2)$. It can be expanded in the same way as $E(V_1^2/m^2)$ by replacing $\sum_{(j_1,j_2)\in S_n'}$ by $\sum_{(j_1,j_2)\in S_n\backslash S_n'}$. It follows from (18) and (19) that $\sum_{(j_1,j_2)\in S_n\backslash S_n'} a_{j_1,j_2} = O(\tilde{r}_{U,n}^2) = O(m)$. Similarly $\sum_{(j_1,j_2)\in S_n\backslash S_n'} a_{j_1,j_2}^2 = O(\tilde{r}_{U,n}^2) = O(m)$. Then noting that $\tilde{r}_{L.n} \to \infty$ as $n \to \infty$, it is shown in the same way as for $E\{A_2^2(H)\}$ of Proposition 3.3 of the Supplementary Material in Yajima and Matsuda (2023) that $E(V_2^2/m^2) = o(1)$. Hence the probability of (20) converges to 0 as $n \to \infty$, which assures that $\hat{H}_n$ is a consistent estimator of H_0.

□

Proof of Theorem 3.2. We shall show the result by following the proof of Theorem 2 of Robinson (1995) with different evaluations for some intermediate assertions. Theorem 3.1 holds under the present conditions and implies that with probability approaching 1 as $n \to \infty$, $\hat{H}_n$ satisfies

$$0 = \frac{dR(\hat{H}_n)}{dH} = \frac{dR(H_0)}{dH} + \frac{d^2R(\tilde{H}_n)}{dH^2}(\hat{H}_n - H_0), \tag{23}$$

where $|\tilde{H}_n - H_0| \le |\hat{H}_n - H_0|$. For the assertion, it suffices to show that

$$\begin{aligned}\frac{d^2R(\tilde{H}_n)}{dH^2} &= \frac{d^2R(H_0)}{dH^2} + o_p(1)\\ &= 1 + o_p(1),\end{aligned} \tag{24}$$

and $m^{1/2}\frac{dR(H_0)}{dH}$ converges to $N(0,1)$ in distribution as $n \to \infty$. (24) are proved by Proposition 3.7 of the Supplementary Material in Yajima and Matsuda (2023).

Now we consider $m^{1/2}\frac{dR(H_0)}{dH}$. Define

$$\begin{aligned}\nu(j_1,j_2) &= \log(\omega_{j_1p\xi}^2 + \omega_{j_2p\xi}^2) - \frac{1}{m}\sum_{l_1,l_2}\log(\omega_{l_1p\xi}^2 + \omega_{l_2p\xi}^2)\\ &= \log(j_1^2 + j_2^2) - \frac{1}{m}\sum_{l_1,l_2}\log(l_1^2 + l_2^2).\end{aligned}$$

By noting $\sum \nu(j_1,j_2) = 0$ and $\hat{G}(H_0) = G(H_0) + o_p(1)$, which follows from the proof of Proposition 3.3 of the Supplementary Material in Yajima and Matsuda (2023), similar to (4.11) of Robinson (1995),

$$\begin{aligned}m^{1/2}\frac{dR(H_0)}{dH} &= \frac{\frac{1}{m^{1/2}}\sum \nu(j_1,j_2)(\omega_{j_1p\xi}^2 + \omega_{j_2p\xi}^2)^{H_0+1}I_p(\omega_{j_1p\xi,j_2p\xi})}{\frac{1}{m}\sum(\omega_{j_1p\xi}^2 + \omega_{j_2p\xi}^2)^{H_0+1}I_p(\omega_{j_1p\xi,j_2p\xi})}\\ &= m^{-1/2}\sum \nu(j_1,j_2)(\tilde{v}_R^2(j_1,j_2) + \tilde{v}_I^2(j_1,j_2) - 1)(1 + o_p(1)).\end{aligned}$$

Similar to (21), we express

$$\begin{aligned}
&m^{-1/2}\sum \nu(j_1,j_2)(\tilde{v}_R^2(j_1,j_2)+\tilde{v}_I^2(j_1,j_2)-1)\\
&=m^{-1/2}\sum_{(j_1,j_2)\in S_n'} + m^{-1/2}\sum_{(j_1,j_2)\in S_n\backslash S_n'}\\
&=T_1+T_2, \text{(say)},
\end{aligned}$$

where

$$S_n' = \{(j_1,j_2)|0<(\frac{j_1 p\xi}{n})^2+(\frac{j_2 p\xi}{n})^2 \le r_{L,n}^2,\ 0<j_1,\ j_2, b_L \le \frac{j_2}{j_1} \le b_U\}.$$

Hereafter we set $r_{L,n}=(p\xi/n)^{1-\tau} r_{U,n}^{\tau}$ with $\max(0,1/2-1/p)<\tau<1/2$. Then note that $\tilde{r}_{L,n}=\tilde{r}_{U,n}^{\tau}$.

Now we show that $T_1=o_p(1)$ and T_2 converges to $N(0,1)$ as $m\to\infty$.

First for any $\epsilon>0$,

$$\begin{aligned}
&P(|T_1|>\epsilon)\\
&\le E(T_1^2)/\epsilon^2.
\end{aligned}$$

Then similar to (22),

$$\begin{aligned}
&E(T_1^2)\\
&=\frac{1}{m}[\sum_{(j_1,j_2)\in S_n'} \nu(j_1,j_2)E\{\tilde{v}_{pR}^2(j_1,j_2)+\tilde{v}_{pI}^2(j_1,j_2)-1\}]^2\\
&+\frac{1}{m}\sum_{(j_1,j_2),(l_1,l_2)\in S_n'} \nu(j_1,j_2)\nu(l_1,l_2)\\
&\times 2[\{E(\tilde{v}_{pR}(j_1,j_2)\tilde{v}_{pR}(l_1,l_2))\}^2+\{E(\tilde{v}_{pR}(j_1,j_2)\tilde{v}_{pI}(l_1,l_2))\}^2\\
&+\{E(\tilde{v}_{pI}(j_1,j_2)\tilde{v}_{pR}(l_1,l_2))\}^2+\{E(\tilde{v}_{pI}(j_1,j_2)\tilde{v}_{pI}(l_1,l_2))\}^2].
\end{aligned}$$

It follows from Proposition 3.2 of the Supplementary Material in Yajima and Matsuda (2023) that every moment is bounded uniformly in $j_i, l_i (i=1,2)$. Hence

$$E(T_1^2)=O\left(\frac{\{(\log \tilde{r}_{L,n}^2)\tilde{r}_{L,n}^2\}^2}{m}\right).$$

Since $\tilde{r}_{L,n}=\tilde{r}_{U,n}^{\tau}=O(m^{\tau/2})$ with $\tau<1/2$, we have $E(T_1^2)=o(1)$, which implies $T_1=o_p(1)$.

Next consider T_2. The cardinality of $S_n\backslash S_n'$ is $m(1+o(1))$ and hence, hereafter it can be assumed to be m. Now we shall show that m, $\nu(j_1,j_2)$, $\tilde{v}_{pR}(j_1,j_2)$ and $\tilde{v}_{pI}(j_1,j_2)$ satisfy the conditions imposed on m^*, $\alpha(j_1,j_2)$,$v_1(j_1,j_2)$ and $v_2(j_1,j_2)$ of Proposition 3.8 of the Supplementary Material in Yajima and Matsuda (2023) respectively.

First evaluate $\nu(j_1,j_2)$. We have $\max|\nu(j_1,j_2)|=O(\log m)=o(m)$.

Next

$$\sum_{S_n\backslash S_n'} \nu(j_1,j_2)^2$$

$$
\begin{aligned}
&= \sum_{S_n} - \sum_{S'_n} \\
&= \sum_{S_n} - O(m^\tau \log m).
\end{aligned}
$$

Using

$$
\begin{aligned}
x \log x &= \frac{1}{2}\frac{d}{dx}[x^2(\log x - \frac{1}{2})], \\
x(\log x)^2 &= \frac{d}{dx}[\frac{x^2}{2}((\log x)^2 - \log x + \frac{1}{2})],
\end{aligned}
$$

we have

$$
\begin{aligned}
&\sum_{S_n} \nu^2_{j_1,j_2} \\
&= \sum (\log(j_1^2 + j_2^2))^2 \\
&\quad - \frac{(\sum_{(k_1,k_2)\in S_n} \log(k_1^2+k_2^2))^2}{m} \\
&= \int_0^{\tilde{r}_{U,n}} \int_{\theta_L}^{\theta_U} r(\log(r^2))^2 dr d\theta + O(\tilde{r}_{U,n}(\log \tilde{r}_{U,n})^2) \\
&\quad - \frac{[\int_0^{\tilde{r}_{U,n}} \int_{\theta_L}^{\theta_U} r \log(r^2) dr d\theta + O(\tilde{r}_{U,n} \log \tilde{r}_{U,n})]^2}{\int_0^{\tilde{r}_{U,n}} \int_{\theta_L}^{\theta_U} dr d\theta + O(\tilde{r}_{U,n})} \\
&= \frac{\theta_U - \theta_L}{2} \tilde{r}^2_{U,n} + O(\tilde{r}_{U,n}(\log \tilde{r}_{U,n})^2) \\
&= \frac{\theta_U - \theta_L}{2} \tilde{r}^2_{U,n} + o(\tilde{r}^2_{U,n}) \\
&= m + o(\tilde{r}^2_{U,n}).
\end{aligned}
$$

Next

$$
\begin{aligned}
&\sum_{(j_1,j_2)\in S_n \setminus S'_n} |\nu_{j_1,j_2}|^l \\
&\le \sum_{(j_1,j_2)\in S_n} |\nu_{j_1,j_2}|^l \\
&= \sum_{(j_1,j_2)\in S_n} |\log((j_1^2+j_2^2)/\tilde{r}^2_{U,n}) - \frac{\sum_{(k_1,k_2)\in S_n} \log((k_1^2+k_2^2)/\tilde{r}^2_{U,n}))}{m}|^l \\
&\le 2^l \sum_{(j_1,j_2)\in S_n} |\log((j_1^2+j_2^2)/\tilde{r}^2_{U,n})|^l \\
&= O(\tilde{r}^2_{U,n} \int_{x^2+y^2\le 1} |\log(x^2+y^2)|^l dx dy)
\end{aligned}
$$

$$= O(m).$$

Next define the 2-dimensional vector $\tilde{v}_p(j_1, j_2)$ by $\tilde{v}_p(j_1, j_2) = (\tilde{v}_{pR}(j_1, j_2), \tilde{v}_{pI}(j_1, j_2))'$ and the 2×2 covariance matrix $\Sigma_{jp,kp}$ by $E[\tilde{v}_p(j_1, j_2)\tilde{v}_p(k_1, k_2)']$.

Now we evaluate $\Sigma_{jp,kp}$. First assume that $p \geq 2$. From$1/2 - 1/p < \tau$ and Proposition 3.2 of the Supplementary Material in Yajima and Matsuda (2023),

$$\Sigma_{jp,jp} = \frac{1}{2}I_2 + O(m^{-\tau}\xi^{-2}) + O(r_{U,n}^2) = \frac{1}{2}I_2 + o(m^{-1/2}),$$
$$\Sigma_{jp,kp} = O(\xi^{-p}) = o(m^{-1/2}).$$

Hence the limiting distribution of T_2 is $N(0, 1)$.

Next consider the case of $p = 1$. For $H \leq 1/2$, if we set τ sufficiently close to $1/2$ so that $m^{1/2} = o(m^{\tau/2}\xi(\log n)^{-2})$, then Proposition 3.2 of the Supplementary Material in Yajima and Matsuda (2023) assures that

$$\Sigma_{j,j} = \frac{1}{2}I_2 + O(m^{-(\tau/2)}\xi^{-1}) + O(r_{U,n}^2) = \frac{1}{2}I_2 + o((m^{-1/2}),$$
$$\Sigma_{j,k} = O((\log n)^2 m^{-(\tau/2)}\xi^{-1}) = o(m^{-1/2}).$$

Similarly for $H > 1/2$, if we set τ sufficiently close to $1/2$ so that $m^{1/2} = o((m^{\tau/2}\xi)^{2-2H}(\log n)^{-2})$, then Proposition 3.2 of the Supplementary Material in Yajima and Matsuda (2023) assures that

$$\Sigma_{j,j} = \frac{1}{2}I_2 + O((m^{\tau/2}\xi)^{2H-2}) + O(r_{U,n}^2) = \frac{1}{2}I_2 + o(m^{-1/2}),$$
$$\Sigma_{j,k} = O((\log n)^2 (m^{\tau/2}\xi)^{2H-2}) = o(m^{-1/2}).$$

Consequently we have the assertion. □

4. Simulation studies

This section examines empirical properties of Gaussian semiparametric estimation (GSE) in (8) and (10) for a spatial memory parameter H in Assumption 2.1 in comparisons with alternative estimation by spatial domains. Here we focused on the spatial domain estimator by Zhu and Stein (2002) as the benchmark, who assumed that g_0 in Assumption 2.1 is a constant. Our main interests are in the comparisons between them especially when g_0 is not a constant.

To conduct the comparisons, we simulate spatial data on lattice points. We examine the following two cases for $g_0(x)$ in Assumption 2.1. In Case 1, the constant function

$$g_0(\omega_1, \omega_2) = 1,$$

is employed, while in Case 2, the function that depends on frequencies as

$$g_0(\omega_1, \omega_2) = |1 + e^{-i\omega_1} + e^{-i\omega_2}|^2$$

is designed. We simulate $X_H(s, t)$, a pure fractional Brownian field of Case 1, over 200×200 lattice points via the frequency domain method by Stein (2002) for $0 < H \leq 0.75$, while we obtain $Y(s, t)$ of Case 2 by the moving average of $X_H(s, t)$ given as

$$Y(s, t) = X_H(s, t) + X_H(s - 1, t) + X_H(s, t - 1),$$

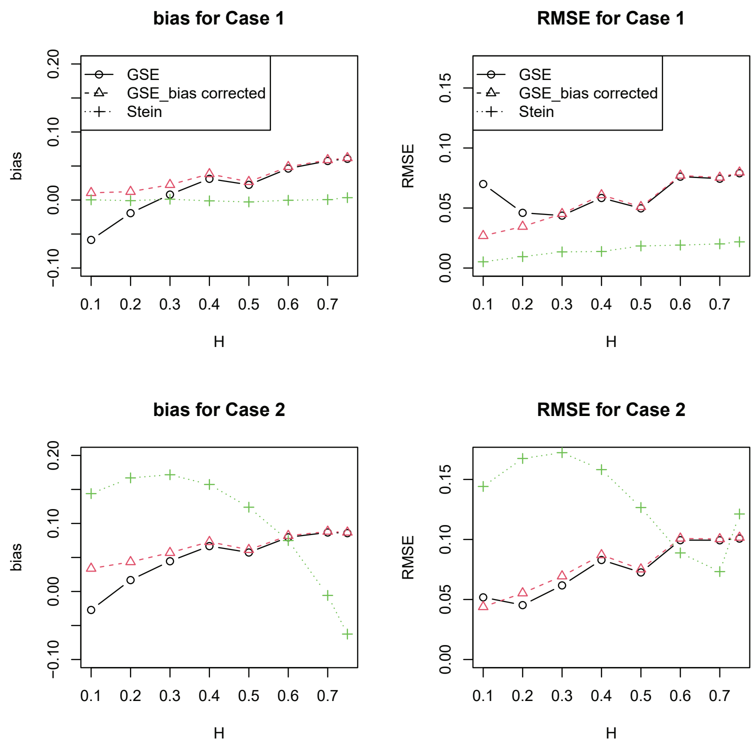

Figure 1. Empirical bias and root mean squared error (RMSE) as functions of H for our Gaussian semiparametric estimator in (8) and the bias corrected version in (10), and those of Zhu and Stein (2002), evaluated by 100 simulations for Cases 1 and 2.

which has the function g_0 indicated in Case 2. We simulated 100 data sets of Cases 1 and 2 with $H = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ and 0.75, for which we examined the empirical comparisons.

We calculated GSE with and without the bias correction term together with Zhu and Stein (2002) for the simulated data, where we designed the bandwidth as $\{||\omega|| < \pi/4\}$ for GSE and chose the filter 1 with $M = 2$ for $k = 1, 2, 3, 4$ for Zhu and Stein (2002). In Figure 1, we show empirical bias and root mean squared error (RMSE) as functions of H for the three estimators.

We find from Figure 1 that GSE without the bias correction term is negatively biased for a small H less than 0.2, which is corrected by the bias correction term, and that GSE has similar performances in terms of bias and RMSE for both Cases 1 and 2. On the other hand, Stein estimator is seriously biased with larger RMSE for Case 2. Constancy for g_0 is critical for Zhu and Stein to work efficiently, which needs care when it is not guaranteed. We claim that GSE is a good alternative to the estimator of Zhu and Stein (2002) when the function g_0 in Assumption 2.1 depends on frequencies.

5. Concluding remarks

This paper proposed Gaussian semiparametric estimation (GSE) for anisotropic ISRFs which include an FBF as special cases. GSE is consistent and asymptotically normally distributed with known variance independent of unknown parameters. The simulation study shows that though estimators based on empirical variograms show better performance than GSE if the true underlying model is a pure FBF, GSE is more robust against anisotropic deviations from a pure FBF in the sense that it is much less biased and has a smaller mean squared error.

Throughout this paper we considered an FBF and its anisotropic extensions. Our GSE is applicable to more general random fields beyond them. In the following we consider the two examples and outline the estimation procedure for them. The first one is the model for an ISRF proposed by Istats (2007). In this model, $g(\lambda_1, \lambda_2)$ is defined by

$$g(\lambda_1, \lambda_2) = \| \boldsymbol{\lambda} \|^{-2H-2} g_0(\lambda_1, \lambda_2),\ 0 < H < 1,$$

where $g_0(\lambda_1, \lambda_2) = h(\lambda_1/\| \boldsymbol{\lambda} \|, \lambda_2/\| \boldsymbol{\lambda} \|)$ with a positive function $h(x_1, x_2)$ and $g_0(0, 0)$ can take any nonnegative value.

Then the anisotropicity depends only on the ratio $\lambda_2/\lambda_1 (= \beta)$, say, since

$$h(\lambda_1/\| \boldsymbol{\lambda} \|, \lambda_2/\| \boldsymbol{\lambda} \|) = h(1/\sqrt{1+\beta^2}, \beta/\sqrt{1+\beta^2}).$$

Then the objective function is defined by

$$Q(G_\beta, H) = \frac{1}{m} \sum_{j_1=1}^{m} \left\{ \log \left(G_\beta (\omega_{j_1 p \xi}^2 + \omega_{j_1 p \beta \xi}^2)^{-H-1} \right) + \frac{(\omega_{j_1 p \xi}^2 + \omega_{j_1 p \beta \xi})^{H+1}}{G_\beta} I_p(\omega_{j_1 p \xi}, \omega_{j_1 p \beta \xi}) \right\},$$

where $G_\beta = h(1/\sqrt{1+\beta^2}, \beta/\sqrt{1+\beta^2})/(8\pi^2)$. Then we define $\hat{G}_{\beta,n},\ \hat{H}_{\beta,n}$ by

$$(\hat{G}_{\beta,n}, \hat{H}_{\beta,n}) = \arg \min_{0 < G_\beta < \infty, H \in [\Delta_1, \Delta_2]} Q(G_\beta, H),$$

with $0 < \Delta_1 < \Delta_2 < 1$.

Let $\beta_k (k = 1, \ldots, K)$ be integers with $\beta_k \neq \beta_{k'} (k \neq k')$. Then under some conditions on $p, r_{U,n}, \xi, m$ and $h(x_1, x_2)$, $\hat{H}_{\beta_k,n}$ converges to H_0 in probability as $n \to \infty$, and

$$m^{1/2}(\hat{H}_{\beta_k,n} - H_0), k = 1, \ldots, K,$$

are asymptotically independent and their limiting distribution is $N(0, 1)$ as $n \to \infty$. Hence the estimator

$$\hat{H}_n = \sum_{k=1}^{K} \hat{H}_{\beta_k,n}/K$$

has the limiting distribution of $m^{1/2}(\hat{H}_n - H)$ is $N(0, 1/K)$.

The second one is a fractional Brownian sheet (FBS), which is another popular class of random fields. A FBS $\{X(\boldsymbol{t}); \boldsymbol{t} \in \boldsymbol{R}^d\}$ is defined by Adler (1981), Herbin (2006) as a Gaussian random field whose covariance function is given by

$$\mathrm{Cov}(X(\boldsymbol{s}), X(\boldsymbol{t})) = C \prod_{i=1}^{d} (s_i^{2H_i} + t_i^{2H_i} - |s_i - t_i|^{2H_i}), \quad 0 < H_i < 1,$$

where $\boldsymbol{s} = (s_1, \cdots, s_d)'$ and $\boldsymbol{t} = (t_1, \cdots, t_d)'$. For $d = 2$, define the objective function by

$$Q(G, H_1, H_2) = \frac{1}{m^2} \sum_{(j_1,j_2)\in S_{1n}\times S_{2n}} \left\{ \log\left(G(\omega_{j_1 p\xi}^{-2H_1-1} \omega_{j_2 p\xi}^{-2H_2-1}) \right) + \frac{(\omega_{j_1 p\xi}^{2H_1+1} \omega_{j_2 p\xi}^{2H_2+1})}{G} I_p(\omega_{j_1 p\xi, j_2 p\xi}) \right\}$$

where $G = C/(16\pi^2)$, and m is the cardinality of S_{in} for

$$S_{in} = \left\{ j_i : 0 < \frac{j_i p\xi}{n} \le r_{U,n} \right\}, i = 1, 2.$$

Then we define $\hat{G}_n$, $\hat{H}_{in} (i = 1, 2)$ by

$$(\hat{G}_n, \hat{H}_{1n}, \hat{H}_{2n}) = \arg \min_{0<G<\infty, H_i\in[\Delta_1,\Delta_2] (i=1,2)} Q(G, H_1, H_2),$$

with $0 < \Delta_1 < \Delta_2 < 1$. Under some condtions on p, $r_{U,n}$, ξ and m, $\hat{H}_{in}$ converges to H_{i0} in probability as $n \to \infty$ for $i = 1, 2$, and

$$m(\hat{H}_{1n} - H_{10}, \hat{H}_{2n} - H_{20})'$$

converges in distribution to $N(\mathbf{0}, I_2)$ as $n \to \infty$.

Recently Shen and Hsing (2020) considered estimation of a mutifractional Browninan motion (MFBM), a generalization of the FBF in which the Hurst effect H depends on $\boldsymbol{t}$. Let

$$D(H) = \left(\int_{\boldsymbol{R}^d} \frac{1 - \cos x_1}{\| \boldsymbol{x} \|^{2H+d}} d\boldsymbol{x} \right)^{1/2}, 0 < H < 1,$$

where x_1 is the first component of $\boldsymbol{x}$. Then an MFBM $\{X(\boldsymbol{t}) : \boldsymbol{t} \in (0, 1)^d\}$ is defined as a Gaussian random field whose covariance function is given by

$$\mathrm{Cov}\,(X(\boldsymbol{s}), X(\boldsymbol{t})) = \sigma^2 \mathcal{D}\,(H(\boldsymbol{s}), H(\boldsymbol{t})) \times \left(\| \boldsymbol{s} \|^{H(\boldsymbol{s})+H(\boldsymbol{t})} + \| \boldsymbol{t} \|^{H(\boldsymbol{s})+H(\boldsymbol{t})} - \| \boldsymbol{s} - \boldsymbol{t} \|^{H(\boldsymbol{s})+H(\boldsymbol{t})} \right),$$

where $\sigma^2 \in (0, \infty)$, $H(\boldsymbol{t})$ is a Hölder continuous function with the range in $(0, 1)$ and

$$\mathcal{D}(H(\boldsymbol{s}), H(\boldsymbol{t})) = [2D(H(\boldsymbol{s})D(H(\boldsymbol{t}))]^{-1} D^2((H(\boldsymbol{s}) + H(\boldsymbol{t}))/2).$$

They proposed an estimation procedure of $H(\boldsymbol{t})$ based on the differenced data of $X(\boldsymbol{t})$ and investigated their asymptotic properties in the framework of infill asymptotics. It is an important and interesting issue in future to consider whether our GSE is applicable to estimate $H(\boldsymbol{t})$ and to compare its asymptotic properties with those of their estimator.

Supplementary Material

Supplement to Gaussian semiparametric estimation of two-dimensional intrinsically stationary fields

The lemmas and propositions that are necessary to prove Theorems 3.1 and 3.2 are given.

References

Adler,R.J. (1981). *The Geometry of Random Fields*. New York:Wiley **23** 1684-1711.

Chilès, J.-P. and Delfiner,P. (2012). *Geostatistics :Modeling Spatial Uncertainty*. 2nd ed. New York: Wiley.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. Revised ed. New York: Wiley.

Dahlhaus, R. (1983). Spectral analysis with tapered data. *J. Time Ser. Anal.* **4** 163-175.

Dahlhaus, R. and Künsch, H. R. (1983). Edge effect and efficient parameter estimation for stationary random fields. *Biometrika* **74** 877-882.

Gikhman, I.I. and Skorokhod,A.V. (2004). *The Theory of Stochastic Processes*. Translated from the Russian by S.Kotz. Corrected Printings of 1st ed. Berlin: Springer.

Guyon, X. (1995). *Random Fields on a Network Modeling, Statistics, and Applications*. translated by C.Ludeña. New York: Springer.

Herbin, E. (2006). From N parameter fractional Brownian motions to N parameter multifracional Brownian motions. *Rocky Mountain J. of Math.* **36** 1249-1284,

Huang, C., Hsing, T. and Cressie, N. (2011). Nonparametric estimation of the variogram and its spectrum. *Biometrika* **98** 775-789.

Istas, J (2007). Identifying the anisotropical function of a d-dimensional Gaussian self-similar processes with stationary increments. *Stat. Infer. Stoch. Proc.* **10** 97-106.

Itô, K (1953). Stationary random distributions. *Mem. College Science, Univ. of Kyoto. Ser A* **28** 209-223.

Künsch, H. R. (1987). Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society* Yu. Prohorov and V.V.Sazanov, eds. **1**, 67-74. Utrecht:VNU Science Press.

Lee, M., Genton, M.G. and Jun, M. (2016). Testing self-similarity through Lamperti transformations. *J. Agr. Biol. and Envior. Statist.* **21** 426-447.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractal noises and applications. *SIAM. Rev.* **10** 422-437.

Matheron, G. (1973). The intrinsic random functions and their applications. *Adv. Appl. Probab.* **5** 439-468.

Priestley, M.B. (1981). *Spectral Analysis and Time Series*. London:Academic Press.

Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23** 1630-1661.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes:Stochastic Models with Infinite Variance*. London:Chapman& Hall/CRC.

Shen , J. and Hsing, T. (2020). Hurst function estimation. *Ann. Statist.* **48** 838-862.

Solo,V. (1992). Intrinsic stationary random functions and the paradox of $1/f$ noise. *SIAM J. Appl. Math.* **52** 270-291.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer.

Stein, M. L. (2002). Fast and exact simulation of fractional Brownian surfaces. *J. Comp. Graphical Stat.* **11**, 587-599.

Velasco, C. (1999a). Non-stationary log-periodogram regression. *J. Econometrics* **91** 325-371.

Velasco, C. (1999b). Gaussian semiparametric estimation of non-stationary time series. *J. Time Ser. Anal.* **20** 87-127.

Walker, A. M. (1964). Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time-series. *J. Aust. Math. Soc.* **4** 363-384.

Yaglom, A. M. (1957). Some classes of random fields in n-dimensional space, related to stationary random processes. translated by R.A.Silverman *Theory Probab. Appl.* **2** 273-320.

Yajima, Y. and Matsuda, Y. (2020a). Log-periodogram regression of two-dimensional intrinsically stationary random fields. *J.Statist. and Data Science* **3** 333-347.

Yajima, Y. and Matsuda, Y. (2020b). Supplementary Material to Log-priodogram regression of two-dimensional intrinsically stationary random fields". https://doi.org/10.1007/s42081-020-00078-9

Yajima, Y. and Matsuda, Y. (2023). Supplement to "Gaussian semiparametric estimation of two-dimensional intrinsically stationary random fields."

Zhu, Z. and Stein, M. L. (2002). Parametric estimation for fractional Brownian surfaces. *Statist. Sinica* **12** 863-883.