Bernoulli **26**(4), 2020, 2463–2487
https://doi.org/10.3150/19-BEJ1160

A general frequency domain method for assessing spatial covariance structures

MATTHEW VAN HALA[1,*], SOUTIR BANDYOPADHYAY[2], SOUMENDRA N. LAHIRI[3] and DANIEL J. NORDMAN[1,**]

[1]*Department of Statistics, Iowa State University, Ames, IA 50011, USA.*
E-mail: [*]*matthew.vanhala@gmail.com;* [**]*dnordman@iastate.edu*
[2]*Department of Applied Mathematics and Statistics, Colorado School of Mines, Golden, CO 80401, USA.*
E-mail: sbandyopadhyay@mines.edu
[3]*Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA.*
E-mail: snlahiri@ncsu.edu

When examining dependence in spatial data, it can be helpful to formally assess spatial covariance structures that may not be parametrically specified or fully model-based. That is, one may wish to test for general features regarding spatial covariance without presupposing any particular, or potentially restrictive, assumptions about the joint data distribution. Current methods for testing spatial covariance are often intended for specialized inference scenarios, usually with spatial lattice data. We propose instead a general method for estimation and testing of spatial covariance structure, which is valid for a variety of inference problems (including nonparametric hypotheses) *and* applies to a large class of spatial sampling designs with irregular data locations. In this setting, spatial statistics have limiting distributions with complex standard errors depending on the intensity of spatial sampling, the distribution of sampling locations, and the process dependence. The proposed method has the advantage of providing valid inference in the frequency domain without estimation of such standard errors, which are often intractable, and without particular distributional assumptions about the data (e.g., Gaussianity). To illustrate, we develop the method for formally testing isotropy and separability in spatial covariance and consider confidence regions for spatial parameters in variogram model fitting. A broad result is also presented to justify the method for application to other potential problems and general scenarios with testing spatial covariance. The approach uses spatial test statistics, based on an extended version of empirical likelihood, having simple chi-square limits for calibrating tests. We demonstrate the proposed method through several numerical studies.

Keywords: confidence sets; spatial periodogram; spatial testing; spectral moment conditions; stochastic sampling

1. Introduction

Frequency domain analysis is a well-known approach for examining covariance structures of random fields [5,12,37,38]. Recently, a variety of frequency domain methods have been proposed for problems in assessing spatial covariance, particularly for spatial lattice data. For example, [8,9] developed a spectral method for testing for nonstationarity, [10] considered an ANOVA-like test for covariance separability of spatio-temporal processes, and [7] proposed goodness-of-fit tests based on the spectral density for spatial lattice data. In contrast to regular time series and space lattice data [14,16,30,36], the spectral analysis of irregularly located spatial data has received less

1350-7265

attention but is a topic of increasing interest (see, [17] and [24] for spectral density estimation with Gaussian processes, and [1,4,32] for spectral analysis under more general setup). However, testing procedures for assessing spatial covariance structure are lacking in the frequency domain, particularly with irregularly located spatial observations. To address this gap, we propose a *general* spatial methodology for frequency domain estimation and testing of spatial covariance structures which applies to a large class of spatial processes allowing arbitrary stochastic patterns in spatial locations and various rates of infill sampling (i.e., spatial sampling of differing intensity levels).

To describe the methodology, we highlight and explain some of its advantages for such spatial data. Namely, the methodology

(a) is valid without stringent distributional assumptions on the underlying spatial process, such as Gaussianity (e.g., [24]), or strong conditions on the generation of spatial sampling locations, such as uniformity (e.g., isotropy tests of [13] and stationarity tests of [1,4]).
(b) does not require any explicit variance estimation steps for test statistics, where the latter is extremely difficult given the considerations to be described below.
(c) applies in a unified manner to different spatial sampling regimes which can be hard to disambiguate in practice and between which the large sample properties of spatial statistics often differ (i.e., pure increasing domain vs mixed increasing domain structures in the following).
(d) has a flexible formulation that allows for several types of spatial estimation and testing problems to be treated in the frequency domain, including assessments about spatial covariance that may not be fully parametric in nature.

The first three points relate to challenges in the spectral analysis of spatial processes observed at irregularly spaced locations, which do not exist for the more standard cases of equi-spaced time series or spatial lattice data. For example, the latter data cases are associated with a simple compact frequency region (e.g., $[0, 2\pi]$ with time series), which is not true for spatial data with irregular locations in $\mathbb{R}^d$ (i.e., the frequency regions then becomes $\mathbb{R}^d$). A more serious complication is the diversity of sampling schemes for spatial data with irregular locations and its effect on the large sample properties of spatial statistics. In contrast to spatial lattice data, different asymptotic structures can arise with irregularly located spatial data, namely, pure increasing domain (PID) and mixed increasing domain (MID) asymptotics (cf. [6]), depending on the relative size of a spatial sampling region to the number n of spatial observations (see Section 2). The limiting distributions of statistics, and in particular the form of their standard errors, often change dramatically depending on the PID vs MID frameworks (see [20]) and further depend intricately on additional factors, such as the (unknown) distribution of sampling locations and the underlying process dependence structure. Spectral analysis may be simplified by assuming a Gaussian process or a uniform distribution to spatial locations (see [24]), but the standard errors involved remain complicated and difficult to estimate directly without restrictive assumptions. Hence, a motivation for the proposed spatial method with irregular spatial data is its application without direct variance estimation steps and its robustness to the aforementioned spatial factors.

The main idea of the spatial method, related to the last point (d) above, is to prescribe estimation and testing problems in the frequency domain by formulating spectral estimating equations. In particular, if $\theta \in \mathbb{R}^p$ denotes a spatial "parameter" (which may not be model-based), then a practitioner specifies a set of $r \geq p$ estimating functions $G_\theta(\boldsymbol{\omega})$ to link frequencies $\boldsymbol{\omega} \in \mathbb{R}^d$ and values θ under a spectral moment condition $\int_{\mathbb{R}^d} G_\theta(\boldsymbol{\omega})\phi(\boldsymbol{\omega}) = \mathbf{0}_r$, where $\phi(\cdot)$ represents the process spectral density and d is the dimension of spatial sampling. The choice of functions $G_\theta(\cdot)$ reflects the inference intended in the frequency domain. The method's generality stems from being able to formulate estimating functions which satisfy a moment condition under an assumed spatial covariance form that need not be parametric model-based, even if unknown spatial quantities (as θ above) are involved. For stating such functions, the quantities θ can often be based on the normalized spectral distribution

$$\Phi^0(\boldsymbol{t}) = \int_{\mathbb{R}^d} \mathbb{I}_{(-\infty,\boldsymbol{t}]}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\,d\boldsymbol{\omega} \Big/ \int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\,d\boldsymbol{\omega} \quad \text{for } \boldsymbol{t} = (t_1, \dots, t_d)' \in \mathbb{R}^d, \tag{1.1}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function and $(-\infty, \boldsymbol{t}] = (-\infty, t_1] \times \cdots \times (-\infty, t_d]$, or on process autocorrelations $\rho(\boldsymbol{h}) = \sigma(\boldsymbol{h})/\sigma(\mathbf{0})$ where

$$\sigma(\boldsymbol{h}) = \mathrm{Cov}\big[Z(\mathbf{0}), Z(\boldsymbol{h})\big] = \int_{\mathbb{R}^d} \cos\big(\boldsymbol{h}'\boldsymbol{\omega}\big)\phi(\boldsymbol{\omega})\,d\boldsymbol{\omega}, \quad \text{for } \boldsymbol{h} \in \mathbb{R}^d. \tag{1.2}$$

Hence, tests of spatial covariance are found by assessing whether generic spatial quantities θ, related to spectral distributions or correlations, satisfy constraints $\int_{\mathbb{R}^d} G_\theta(\boldsymbol{\omega})\phi(\boldsymbol{\omega}) = \mathbf{0}_r$ imposed by $G_\theta(\cdot)$ under a hypothesized covariance structure. For concreteness, we specifically treat spatial problems in testing isotropy, testing separability, and variogram model fitting in sections to follow; extensions to other problems are also given. Once estimating functions are prescribed, log-ratio-type test statistics are computed in a simple and unified manner using an extended version of spatial empirical likelihood. These statistics are shown to have chi-square limits for calibrating spatial tests, without specification of a joint data distribution or explicit estimation of the confounding factors associated with irregularly located spatial data.

The rest of the paper is organized as follows. Section 2 explains the spatial sampling framework as well as an empirical likelihood (EL) scheme for processing estimating functions. Sections 3–5 develop and numerically illustrate the frequency domain methodology, respectively, for tests of isotropy, separability, and variogram parameters. In Section 6, a broad and unified result is further provided to validate the general methodology for testing spatial covariance structure over a wide range of potential problems. Section 7 highlights some important aspects of the proofs involved, and Section 8 provides concluding remarks. Proof details and extended numerical results can be found in Supplement A and Supplement B [34].

We end this section with background on EL, as a device here for assessing spatial estimating functions in the frequency domain. EL is a resampling-type method that formulates a likelihood function nonparametrically by probability profiling data ([27,28], for iid data). However, the application of EL for dependent data, particularly spatial data, is challenging because EL formulations for independent data typically fail with correlated data [19]. Data blocking is one approach known for extending EL to spatial lattice data (see [18,25]), but this does not extend readily to irregularly located spatial data (outside of inference about the marginal distribution of spatial observations, [35]). See [26] for a review of EL with dependent data. Recently, [3] (henceforth [BLN]) proposed a frequency domain EL using a spatial periodogram. Our spatial testing methodology is based on their initial EL approach, but differs by being grounded on point estimation (i.e., EL maximization) as well as broader (i.e., potentially over-identified) estimating functions. These aspects are crucial here for EL assessments of spatial structure, which represent formal tests of spatial "moment conditions" and require normalizing EL-ratios by their maximized values. The spatial EL framework of [BLN], in contrast, is restricted to tests of parameter values; this approach lacks point estimation steps and is invalid for performing moment tests of spatial structure. However, the proposed spatial method additionally applies to both moment and parameter tests, where the point estimation steps can also induce better performance for parameter testing compared to the EL version of [BLN] (e.g., better power properties in Section 6). See Section 2.2 for some further comparative details and Section 7 for an outline of technical arguments needed.

2. Preliminaries

2.1. Spatial sampling scheme

We adopt a general spatial sampling framework as described in [15] and [BLN]. Suppose that $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ represents a mean zero, real-valued, second-order stationary process, which is observed at n irregularly located locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ over a sampling region $\mathcal{D}_n \subset \mathbb{R}^d$. The spatial region $\mathcal{D}_n = \lambda_n \mathcal{D}_0$ is prescribed by inflating a fixed "template" subset $\mathcal{D}_0 \subset (1/2, 1/2]^d$ (containing the origin) by sequence $\{\lambda_n\}$ of scaling factors ($\lambda_n \to \infty$ as $n \to \infty$). Note that, this formulation allows a variety of sampling region shapes, such as polygonal, ellipsoidal, and star-shaped regions that can be non-convex. In practice, λ_n can be determined by the diameter of a sampling region for use here (cf. [11,15,23,24]). Let $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. To avoid pathological cases, we require that for any sequence of real numbers $\{b_n\}_{n \geq 1}$ such that $b_n \to 0+$ as $n \to \infty$, the number of cubes of the form $b_n(\boldsymbol{j} + [0, 1)^d)$, $\boldsymbol{j} \in \mathbb{Z}^d$ that intersect both $\mathcal{D}_0$ and $\mathcal{D}_0^c$ is of the order $O([b_n]^{-(d-1)})$ as $n \to \infty$. This boundary condition holds for most regions of practical interest.

We next consider specification of the locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ within $\mathcal{D}_n$. Independently of $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, let $\{\boldsymbol{X}_k\}_{k \geq 1} \subset \mathcal{D}_0$ be a sequence of independently and identically distributed (i.i.d.) $\mathbb{R}^d$-valued random vectors, with probability density function $f(\boldsymbol{x})$ with support on the closure of $\mathcal{D}_0$. The sampling locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ are then generated as $\boldsymbol{s}_i = \lambda_n \boldsymbol{X}_i$, $i = 1, \ldots, n$. This stochastic sampling design allows sampling sites to have an arbitrary and potentially *non-uniform* density over the sampling region, improving upon the common approach of modeling irregularly spaced sites with a homogeneous Poisson point process. Further, this formulation allows the number n of sampling sites to grow at a different rate than the volume $O(\lambda_n^d)$ of the sampling region $\mathcal{D}_n = \lambda_n \mathcal{D}_0$, leading to different asymptotic structures (see [6,20]). Suppose $c_* = \lim_{n \to \infty} n/\lambda_n^d \in (0, \infty]$. The case of $c_* \in (0, \infty)$, in which the number of spatial observations is proportional to the volume of the region, corresponds to pure increasing domain (PID) asymptotics. On the other hand, the case $c_* = \infty$ corresponds to sampling with a heavy infill component, whereby the number of spatial observations grows at a faster rate than the size of the sampling region; we refer to this as mixed increasing domain (MID). As a complication, limit laws of even simple statistics, such as sample means, typically change with the type of spatial asymptotic structure [20]; see also [6,21,24,29] and the references therein for further details. Note that neither the type of spatial sampling structure (PID/MID) nor the density $f(\boldsymbol{x})$ of locations need to be known or estimated in our testing approach to follow.

The spatial testing method uses a periodogram for irregularly located spatial data, which we define next. Denote the discrete Fourier transform (DFT) of the spatial data $\{Z(\boldsymbol{s}_1),\dots,Z(\boldsymbol{s}_n)\}$ at a frequency $\boldsymbol{\omega}\in\mathbb{R}^d$ as

$$d_n(\boldsymbol{\omega}) = \lambda_n^{d/2} n^{-1} \sum_{j=1}^{n} Z(\boldsymbol{s}_j)\exp\big(\imath\boldsymbol{\omega}'\boldsymbol{s}_j\big), \quad \imath\equiv\sqrt{-1}$$

and define the (raw) periodogram as $I_n(\boldsymbol{\omega}) = |d_n(\boldsymbol{\omega})|^2$. Unlike the (equi-spaced) time series setting, the spatial periodogram here can have a nontrivial bias depending on the spatial asymptotic structure and the spatial sampling density f (cf. [2]). Indeed, [24] showed that

$$\lim_{n\to\infty} \mathrm{E} I_n(\boldsymbol{\omega}) = c_*^{-1}\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega}) \quad \text{for } \boldsymbol{\omega}\in\mathbb{R}^d,$$

where $\phi(\cdot)$ is the spectral density of the process $Z(\cdot)$, $\sigma(\cdot)$ denotes the process covariance function, $K=(2\pi)^d\int_{\mathbb{R}^d} f^2$ and $\lim_{n\to\infty} n/\lambda_n^d = c_*$. Under PID ($c_*\in(0,\infty)$), there exists a nontrivial bias component, which vanishes asymptotically in the MID case ($c_*=\infty$). To address this, we use a bias-corrected periodogram as

$$\widetilde{I}_n(\boldsymbol{\omega}) = I_n(\boldsymbol{\omega}) - n^{-1}\lambda_n^d\widehat{\sigma}_n(\mathbf{0}), \quad \boldsymbol{\omega}\in\mathbb{R}^d$$

where $\widehat{\sigma}_n(\mathbf{0}) = n^{-1}\sum_{j=1}^n (Z(\boldsymbol{s}_j)-\bar{Z}_n)^2$ is the sample variance with $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z(\boldsymbol{s}_i)$.

2.2. Spectral estimating functions

We next describe a framework for setting spectral estimating functions and the device for assessing these with EL for later spatial testing. Suppose a spatial parameter $\theta\in\Theta\subset\mathbb{R}^p$ is connected to the spectral density $\phi(\cdot)$ of the process $Z(\cdot)$ through a system of estimating equations. Specifically, let $G:\mathbb{R}^d\times\Theta\to\mathbb{R}^r$ be a vector of $r\geq p$ estimating functions (i.e., functions of frequencies and parameters) such that $G_\theta(\boldsymbol{\omega})\equiv G(\boldsymbol{\omega};\theta)$ satisfies a spectral moment condition

$$\int_{\mathbb{R}^d} G_\theta(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\,d\boldsymbol{\omega} = \mathbf{0}_r, \tag{2.1}$$

at a true parameter $\theta_0\in\Theta$, where $\mathbf{0}_r\in\mathbb{R}^r$ denotes the zero vector. Because the spectral density $\phi(\cdot)$ is symmetric, we assume that $G_\theta(\cdot)$ is also symmetric, that is, one may replace $G_\theta(\boldsymbol{\omega})$ with $[G_\theta(\boldsymbol{\omega})+G_\theta(-\boldsymbol{\omega})]/2$, $\boldsymbol{\omega}\in\mathbb{R}^d$. The case of "over-identified" estimating functions, with r functions for $p<r$ parameters, will be used to assess whether the spectral moment (2.1) holds for some parameter $\theta_0\in\Theta$, thereby providing tests of spatial covariance structures. Sections 3–5 illustrate such testing problems.

To assess estimating functions satisfying (2.1), we compute the (bias-corrected) spatial periodogram along a discretized set of frequencies in order to mimic the spectral mean (2.1). Such frequencies are selected on a grid in a way that ensures the corresponding periodogram variants are approximately uncorrelated. That is, by evaluating the spatial periodogram on a frequency grid with appropriate frequency spacing, we may obtain an important spatial analog of the independence (or "whitening") property associated with the typical periodogram for regular time series. At the same time, the frequency grid needs to be large enough to adequately approximate the spatial integral in the spectral moment (2.1). To formulate the frequency grid, for $\kappa \in (0,1)$, $\eta \in (\kappa, \infty)$, and $\mathcal{C} \in (0,\infty)$, define a set of Fourier frequencies as

$$\mathcal{N} = \mathcal{N}_n = \left\{ \boldsymbol{j}\lambda_n^{-\kappa} : \boldsymbol{j} \in \mathbb{Z}^d, \boldsymbol{j} \in \left[-\mathcal{C}\lambda_n^{\eta}, \mathcal{C}\lambda_n^{\eta}\right]^d \right\}.$$

Let $N = |\mathcal{N}|$ be the cardinality of $\mathcal{N}$, and let $\boldsymbol{\omega}_{kn}$, $k = 1, \dots, N$ (with arbitrary ordering) denote the elements of $\mathcal{N}$. The set or grid $\mathcal{N}$ has two relevant properties. First, the frequencies $\{\boldsymbol{\omega}_{kn}\}_{k=1}^{N}$ form a regular lattice over the set $[-\mathcal{C}\lambda_n^{\eta-\kappa}, \mathcal{C}\lambda_n^{\eta-\kappa}] \uparrow \mathbb{R}^d$ as $n \to \infty$, which expands to cover the entire frequency domain $\mathbb{R}^d$ asymptotically. Additionally, any pair of frequencies $\boldsymbol{\omega}_{kn}, \boldsymbol{\omega}_{jn} \in \mathcal{N}$ in the set is asymptotically distant (i.e., $\lambda_n \|\boldsymbol{\omega}_{kn} - \boldsymbol{\omega}_{jn}\| \geq \lambda_n^{1-\kappa} \to \infty$), implying their associated periodogram values are approximately independent by results in [2]. It is also worth mentioning that from the above formulation, it is evident that the frequency grid depends directly on the size of the spatial region λ_n but not the spatial sample size n. That is, in large samples and with an increasing sampling region ($\lambda_n \to \infty$ as $n \to \infty$), a grid specified on the basis of spatial region size λ_n along with tuning parameters (e.g., $\mathcal{C}$, κ, and η), should provide valid inference, regardless of the ratio of n and λ_n^d. Hence, our simulation studies presented in Sections 3–5 and in Supplement B are designed to reflect the theory where we always consider basing the grid choice on the spatial region size (not given the sample size n).

To assess the plausibility of a parameter θ, using an estimating function $G_\theta(\cdot)$ fulfilling (2.1), the corresponding (normalized) EL function for θ is defined as

$$\mathcal{R}_n(\theta) = \sup\left\{ \prod_{k=1}^{N} Np_k : \sum_{k=1}^{N} p_k = 1,\, p_k \geq 0, \sum_{k=1}^{N} p_k G_\theta(\boldsymbol{\omega}_{kn}) \widetilde{I}_n(\boldsymbol{\omega}_{kn}) = \mathbf{0}_r \right\},$$

based on spatial periodogram along the frequency grid. The EL function is a multinomial likelihood found by probability profiling the (approximately independent) periodogram variants under a constraint that imitates the moment condition (2.1). This provides a nonparametric way of measuring the strength of evidence in support of θ, analogously to parametric likelihood, and the numeric computation of $\mathcal{R}_n(\theta)$ follows well-known recipes (see [27,28]). Maximizing the function $\mathcal{R}_n(\theta)$ over the parameter space $\Theta \subset \mathbb{R}^p$ produces a point estimator $\widehat{\theta}_n \in \Theta$ (or the maximum EL estimator (MELE) of θ). In what follows, our frequency domain tests are based on modified log-ratio statistics, $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$ and $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$, involving a simple scaling factor

$$\widehat{a}_n = \frac{\sum_{j=1}^{N} \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 \widetilde{I}_n^2(\boldsymbol{\omega}_{jn})}{\sum_{j=1}^{N} \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 I_n^2(\boldsymbol{\omega}_{jn})}. \tag{2.2}$$

We mention some relevant EL context to compare our testing results to follow to the spatial EL version of [BLN]. The latter method involves statistics based on the form $-2a(\theta) \log \mathcal{R}_n(\theta)$ for testing claims about parameter values θ, where $a(\theta)$ is defined by substituting θ for $\widehat{\theta}_n$ in (2.2). These statistics are not valid, however, for testing moment conditions (i.e., testing if the spectral moment (2.1) holds for some θ_0), for which $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$ is valid in contrast. This is because their EL work does not consider steps of point estimation/maximization $\widehat{\theta}_n$, which are essential to general test statistics of spatial covariance (e.g., $\mathcal{R}_n(\widehat{\theta}_n)$-based). The EL methodology changes dramatically by involving maximizers $\widehat{\theta}_n$ and Section 7 briefly outlines some issues in establishing distributional theory for our frequency domain test statistics, where complications arise in that correct "scaling" of test statistics depends intricately on the underlying type of spatial sampling asymptotics (e.g., PID/MID). By using a single data-based factor $\widehat{a}_n$ based on the MELE $\widehat{\theta}_n$ in (2.2), test statistics self-adjust to any effect of the spatial asymptotics and retain simple chi-square limits in both PID/MID sampling regimes. In addition to tests of spatial structure, statistics based on $\widehat{\theta}_n$ can also be formulated here to test hypotheses about parameter values θ, where the EL approach of [BLN] is also applicable. Compared to the latter method, however, test statistics based on $\widehat{\theta}_n$, given by the log-ratio $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$, can exhibit better performance (cf. Section 5.2) and have increased power properties (cf. Section 6) in parameter testing cases.

3. Assessing spatial isotropy

Consider the problem of assessing whether the underlying process exhibits isotropy. Recall the spatial process $Z(\cdot)$ is isotropic if its covariance is only a function of distance, that is, $\sigma(\boldsymbol{h}_1) = \sigma(\boldsymbol{h}_2)$ if $\|\boldsymbol{h}_1\| = \|\boldsymbol{h}_2\|$ or, alternatively its spectral density function $\phi(\boldsymbol{\omega}) \equiv \phi(\|\boldsymbol{\omega}\|)$ is a function of frequency through $\|\boldsymbol{\omega}\|$. Section 3.1 explains the methodology and Section 3.2 numerically examines its performance.

3.1. Methodology

For testing purposes, we prescribe over-identified estimating functions which satisfy the spectral moment condition (2.1) under assumptions of isotropy. We base these functions on the normalized spectral distribution $\Phi^0(\cdot)$ from (1.1). To illustrate, for a vector $\boldsymbol{t}_1 \in \mathbb{R}^d$, select r_1 orthogonal $d \times d$ matrices, say $\boldsymbol{Q}_{1,1}, \dots, \boldsymbol{Q}_{1,r_1}$ (i.e., $\boldsymbol{Q}'_{1,i}\boldsymbol{Q}_{1,i} = I_{d\times d}$). Under isotropy, it holds that $\theta_1 = \Phi^0(\boldsymbol{Q}_{1,i}\boldsymbol{t}_1)$ for each $i = 1, \dots, r_1$, and the r_1 functions $G^*_{\theta_1}(\boldsymbol{\omega}) = [\widetilde{G}_{\theta_1}(\boldsymbol{\omega}) + \widetilde{G}_{\theta_1}(-\boldsymbol{\omega})]/2$ for $\widetilde{G}_{\theta_1}(\boldsymbol{\omega}) = [\mathbb{I}_{(-\infty,\boldsymbol{Q}_{11}\boldsymbol{t}_1]}(\boldsymbol{\omega}) - \theta_1, \dots, \mathbb{I}_{(-\infty,\boldsymbol{Q}_{1,r_1}\boldsymbol{t}_1]}(\boldsymbol{\omega}) - \theta_1]$ fulfill the moment condition (2.1) with $p = 1$ parameter θ_1. In general, one may choose p vectors $\boldsymbol{t}_1, \dots, \boldsymbol{t}_p \in \mathbb{R}^d$ and r_i orthogonal matrices $\boldsymbol{Q}_{i,1}, \dots, \boldsymbol{Q}_{i,r_i}$ for each $\boldsymbol{t}_i$ to formulate p parameters $\theta = (\theta_1, \dots, \theta_p)'$ as

$$\theta_i = \Phi^0(\boldsymbol{Q}_{i,j}\boldsymbol{t}_i), \quad j = 1, \dots, r_i, i = 1, \dots, p$$

and analogously develop $r = \sum_{i=1}^p r_i > p$ estimating functions

$$G^{\text{iso}}_\theta(\boldsymbol{\omega}) = \big[G^*_{\theta_1}(\boldsymbol{\omega}), \dots, G^*_{\theta_p}(\boldsymbol{\omega})\big]'. \tag{3.1}$$

Typically, one may choose small p and r (e.g., $p = 1$, $r = 2$ in Section 3.2). With the estimating functions (3.1), a test statistic $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$ is formulated from EL (Section 2.2) to test for isotropy (i.e., if $\int G_\theta^{\text{iso}}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \mathbf{0}_r$ holds at some θ_0).

To state a formal result on the test's validity, we require some mild assumptions on the dependence of the mean zero, second-order stationary process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$, expressed in terms of mixing/moment conditions. For brevity, these regularity conditions (denoted as Conditions (R.1)–(R.5)) are described in Supplement A. For the result, note that the distribution of the test statistic $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$ depends on two types of randomness, due to the spatial process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ and the random sequence $\boldsymbol{X} \equiv \{\boldsymbol{X}_i\}_{i\geq 1} \subset \mathbb{R}^d$ determining the spatial locations (Section 2.1). In the following, $P(\cdot) \equiv P(\cdot|\boldsymbol{X})$ denotes probability conditional on the collection of random vectors $\boldsymbol{X}$, while $P_{\boldsymbol{X}}$ denotes the joint distribution of $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$.

Theorem 1 (PID or MID cases). *Suppose Conditions* (R.1)–(R.5) *hold*, *the estimating functions* (3.1) *are used*, *and* $\lim_{n\to\infty} n/\lambda_n^d \in (0, \infty]$. *Then*, *under* H_0: *"*$\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ *has an isotropic covariance structure*,*"*

$$-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n) \xrightarrow{d} \chi^2_{r-p} \quad \text{as } n \to \infty, \text{a.s. } (P_{\boldsymbol{X}}).$$

That is, regardless of the spatial sampling locations $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, the distribution of the test statistic based on the MELE $\widehat{\theta}_n$ is guaranteed to have a chi-square limit under the null hypothesis. One can additionally establish that the test statistic diverges to $+\infty$ if the estimating functions $G_\theta^{\text{iso}}(\boldsymbol{\omega})$ in (3.1) no longer satisfy the moment condition $\int G_\theta^{\text{iso}}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \mathbf{0}_r$ (i.e., isotropy is violated). Importantly, this distributional result for the test statistic holds equally for *both* PID and MID spatial sampling structures (i.e., $c_* \in (0, \infty)$ or $c_* = \infty$ for $c_* = \lim_{n\to\infty} n/\lambda_n^d$), which is not often possible in the frequency domain with irregularly located spatial data (cf. [24]), and again requires no variance estimation steps. We next consider a numerical study of our test for isotropy.

Remark 1. Estimating functions could be based on correlations (1.2) to prescribe tests of isotropy and the result in Theorem 1 would still hold. For example, by considering a single correlation $\theta_1 = \rho(\boldsymbol{h}_1)$ (i.e., $p = 1$) and $r > 1$ lags $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_r \in \mathbb{R}^d$ satisfying $\|\boldsymbol{h}_i\| = \|\boldsymbol{h}_1\|$, $i = 2, \ldots, r$, the resulting estimating functions $G_{\theta_1}(\boldsymbol{\omega}) = [\cos(\boldsymbol{h}_1'\boldsymbol{\omega}) - \theta_1, \ldots, \cos(\boldsymbol{h}_r'\boldsymbol{\omega}) - \theta_1]'$ would satisfy the moment condition (2.1). More generally, one may formulate estimating functions $G_\theta(\boldsymbol{\omega})$ based on p correlations $\theta = (\theta_1, \ldots, \theta_p)'$ with lag sets for each θ_i. However, numerical studies (not shown here) indicate that the estimating functions (3.1) have greater power for assessing isotropy than those based on correlations. Intuitively, for our tests involving the spatial periodogram, the spectral distribution function is more natural, without a translation step of the frequency domain (spectral distribution) into the spatial domain (covariances).

Remark 2. In formulating results, we assume the process $Z(\cdot)$ has mean zero for simplicity. In practice, observations $\{Z(\boldsymbol{s}_i)\}_{i=1}^n$ can be replaced with sample centered versions $\{Z(\boldsymbol{s}_i) - \bar{Z}_n\}_{i=1}^n$, $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z(\boldsymbol{s}_i)$, in computing the periodogram $I_n(\cdot)$. The results still hold and we implement this centering in simulation studies to follow.

3.2. Illustration

In this section, we present a simulation study of the testing method for spatial isotropy. Considering sampling regions $\mathcal{D}_n = \lambda_n[-1/2, 1/2)^2$, $\lambda_n = 24, 36$ and sample sizes $n = 1200, 1800, 2400, 3600$, we generated i.i.d. locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n \in \mathcal{D}_n$ for $\boldsymbol{s}_i = \lambda_n \boldsymbol{X}_i$ with $\boldsymbol{X}_i$'s drawn from a uniform distribution on $[-0.5, 0.5]^2$ and a Gaussian process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^2\}$. We evaluated the (sample mean-centered) periodogram on a frequency grid $\mathcal{N}_n = \{\lambda_n^{-\kappa} \boldsymbol{j} : \boldsymbol{j} \in \mathbb{Z}^2 \cap [-\mathcal{C}\lambda_n, \mathcal{C}\lambda_n]^2\}$, with varying values $\mathcal{C} = 1, 1.5, 2$ and $\kappa = 0.05, 0.1$, where $\mathcal{C}$ and κ respectively control the number and spacing of frequencies for the EL device (Section 2.2). These $\mathcal{C}$, κ values induce a set of frequencies that is large but with sufficient spacing to ensure periodogram variants are approximately independent (e.g., choices of κ roughly create spacings between frequencies of 0.5 in horizontal/vertical directions).

For the simulation study presented below, we considered an isotropic exponential correlation function $\sigma(\boldsymbol{h}) = \exp(-\|\boldsymbol{h}\|)$, $\boldsymbol{h} = (h_1, h_2)' \in \mathbb{R}^2$ or a directional (anisotropic) correlation $\sigma(\boldsymbol{h}) = \exp(-2h_2)$; the results are invariant to the mean and variance of the process. We applied the spectral distribution-based estimating functions $G_\theta^{\text{iso}}(\cdot)$ from (3.1) with vector sets of the form $\{\pm \boldsymbol{t}_1\}$ for $\boldsymbol{t}_1 = (1, -1)'$ (vector set B1) or $\boldsymbol{t}_1 = (2, -2)'$ (vector set B2); these provide $r = 2$ functions with $p = 1$ parameter under isotropy (corresponding to two orthogonal matrices $\boldsymbol{Q}_{1,1}$, $\boldsymbol{Q}_{1,2}$ in (3.1) as the identity matrix and a 180° rotation). We also considered a further vector set (B3) by combining B1/B2 sets. Empirical size and power results for the tests of isotropy with vector sets B1 and B2 appear in Tables 1–2. It should be noted again that the EL testing

Table 1. Empirical size for tests of isotropy based on the spectral distribution function (vector sets B1 and B2); based on 1000 runs with nominal size 0.1

			B1				B2			
			n				n			
λ_n	$\mathcal{C}$	κ	1200	1800	2400	3600	1200	1800	2400	3600
24	1	0.05	0.193	0.203	0.194	0.173	0.224	0.259	0.295	0.256
		0.1	0.185	0.179	0.156	0.125	0.266	0.245	0.207	0.164
	1.5	0.05	0.168	0.186	0.198	0.181	0.190	0.230	0.266	0.254
		0.1	0.169	0.192	0.161	0.133	0.206	0.238	0.223	0.182
	2	0.05	0.179	0.150	0.187	0.187	0.204	0.187	0.231	0.243
		0.1	0.172	0.155	0.193	0.125	0.190	0.208	0.222	0.198
36	1	0.05	0.163	0.142	0.173	0.196	0.169	0.173	0.211	0.227
		0.1	0.169	0.156	0.171	0.175	0.172	0.202	0.216	0.189
	1.5	0.05	0.150	0.130	0.160	0.169	0.155	0.148	0.177	0.181
		0.1	0.157	0.133	0.157	0.142	0.166	0.145	0.180	0.182
	2	0.05	0.129	0.142	0.132	0.144	0.139	0.143	0.148	0.147
		0.1	0.145	0.141	0.142	0.139	0.161	0.154	0.163	0.162

Table 2. Empirical power for tests of isotropy based on the spectral distribution function (vector sets B1 and B2); based on 1000 runs with nominal size 0.1

			B1				B2			
			n				n			
λ_n	$\mathcal{C}$	κ	1200	1800	2400	3600	1200	1800	2400	3600
24	1	0.05	0.973	0.972	0.959	0.896	0.920	0.954	0.960	0.885
		0.1	0.978	0.972	0.958	0.898	0.963	0.966	0.957	0.892
	1.5	0.05	0.967	0.968	0.962	0.891	0.890	0.949	0.964	0.885
		0.1	0.979	0.967	0.956	0.891	0.947	0.962	0.954	0.895
	2	0.05	0.963	0.970	0.953	0.892	0.851	0.940	0.950	0.888
		0.1	0.981	0.977	0.955	0.887	0.931	0.968	0.952	0.893
36	1	0.05	0.901	0.958	0.949	0.903	0.745	0.889	0.927	0.898
		0.1	0.971	0.971	0.959	0.909	0.868	0.954	0.951	0.903
	1.5	0.05	0.846	0.949	0.944	0.894	0.675	0.834	0.898	0.894
		0.1	0.946	0.965	0.958	0.898	0.793	0.934	0.942	0.904
	2	0.05	0.807	0.913	0.933	0.893	0.593	0.779	0.869	0.893
		0.1	0.904	0.957	0.960	0.893	0.729	0.894	0.937	0.893

methodology is fully non-parametric (i.e., no assumptions required about joint distributions or even the spatial pattern of locations), which can add distortion to test calibrations under spatial dependence for moderate sample sizes. However, better sizes and powers resulted from using small vector sets (e.g., B1 and B2) compared to larger, more complicated sets (e.g., B3). In general, regardless of the vector set, size and power improved for larger sampling regions and the power often improved as the sample size n increased. The testing method's sensitivity to choices of $\kappa, \mathcal{C}$ seemingly decreased as n increased.

For comparison, we also evaluated the test of isotropy from [23] (hereafter [MS]), which involves a kernel estimator $\widehat{\sigma}(\boldsymbol{h})$ of the process covariances $\sigma(\boldsymbol{h})$ (cf. [15]). The [MS] method computes standard errors for $\widehat{\sigma}(\boldsymbol{h})$ using a spatial block bootstrap [22], which is computationally quite demanding. Because of this, we considered this method only for a sample size $n = 600$; Supplement A provides implementation details. Based on a lag set $\{(0,1),(1,0),(0,5),(5,0)\}$ (i.e., lags for evaluating the kernel estimator $\widehat{\sigma}(\cdot)$ which differ from the vectors $\boldsymbol{t}_i$ in (3.1)), the [MS] method had empirical size and power of 0.128 and 1, respectively, at a nominal level $\alpha = 0.1$. However, the size and power of our tests based on the spectral distribution could often perform comparably, for example, size 0.147 and power 0.927 with vector set B1 and $(\mathcal{C}, \kappa) = (1.5, 0.1)$ when $n = 600$. Further, our method has far greater computational advantages with a better capacity to scale to large data; using a computer with a 3.00 GHz processor, our tests were approximately 40 and 100 times faster those of for sample sizes $n = 600$ and $n = 1200$. Note the [MS] method is specifically designed for tests of isotropy, while our frequency domain approach is not tailored to a particular spatial testing problem.

4. Assessing spatial covariance separability

To illustrate a different problem in spatial testing, consider the common issue of assessing whether the process covariance function $\sigma(\cdot)$ from (1.2) is separable, that is, whether there exist valid functions $\sigma_i(\cdot)$, $i = 1, 2$, such that $\sigma(\boldsymbol{h}) = \prod_{i=1}^d \sigma_i(h_i)$ for $\boldsymbol{h} = (h_1, \ldots, h_d)' \in \mathbb{R}^d$, or alternatively whether the process spectral density $\phi(\boldsymbol{\omega}) = \prod_{i=1}^d \phi_i(\omega_i)$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)' \in \mathbb{R}^d$ factors component-wise for appropriate functions $\phi_i(\cdot)$. Section 4.1 presents our test, with numerical illustration in Section 4.2.

4.1. Methodology

We base test assessments on the feature that, under separability, the normalized spectral distribution, for $\boldsymbol{t} = (t_1, \ldots, t_d)' \in \mathbb{R}^d$, factors into d univariate integrals

$$\Phi^0(\boldsymbol{t}) = \prod_{i=1}^d \Phi_i^0(t_i), \qquad \Phi_i^0(t_i) = \int_{-\infty}^{t_i} \phi_i(\omega)\, d\omega \Big/ \int_{-\infty}^{\infty} \phi_i(\omega)\, d\omega, \quad i = 1, \ldots, d,$$

satisfying $\Phi_i^0(-t_i) = 1 - \Phi_i^0(t_i)$ by the evenness of $\phi_i(\cdot) \geq 0$. For a given $\boldsymbol{t} \in \mathbb{R}^d$, one can define $p = d$ parameters as $\theta_i = \Phi_i^0(t_i)$, $i = 1, \ldots, d$, for which the component-wise indicator $\mathbb{I}_{i,\boldsymbol{t}}(\boldsymbol{\omega}) \equiv I(\omega_i \leq t_i)$ fulfills $\int [\mathbb{I}_{i,\boldsymbol{t}}(\boldsymbol{\omega}) - \theta_i]\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \int [\mathbb{I}_{i,-\boldsymbol{t}}(\boldsymbol{\omega}) - (1 - \theta_i)]\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = 0$ for each $i = 1, \ldots, d$ under separability. This leads to a general class of estimating functions given as $G_\theta^{\text{sep}}(\boldsymbol{\omega}) = [\widetilde{G}_\theta(\boldsymbol{\omega}) + \widetilde{G}_\theta(-\boldsymbol{\omega})]/2$ for

$$\widetilde{G}_\theta(\boldsymbol{\omega}) = \left[\prod_{i=1}^d g_{1,i}^*(\boldsymbol{\omega}) - \prod_{i=1}^d \theta_{1,i}^*, \ldots, \prod_{i=1}^d g_{r,i}^*(\boldsymbol{\omega}) - \prod_{i=1}^d \theta_{r,i}^*\right]', \quad \boldsymbol{\omega} \in \mathbb{R}^d, \tag{4.1}$$

defined by $r > p$ components of the form $\prod_{i=1}^d g_{j,i}^*(\boldsymbol{\omega}) - \prod_{i=1}^d \theta_{j,i}^*$; here each $g_{j,i}^*(\boldsymbol{\omega})$ is chosen as either $\mathbb{I}_{i,\boldsymbol{t}}(\boldsymbol{\omega})$, $\mathbb{I}_{i,-\boldsymbol{t}}(\boldsymbol{\omega})$, or 1 and, correspondingly, each $\theta_{j,i}^*$ is set to θ_i, $1-\theta_i$ or 1, respectively. This construction (4.1) ensures that the spectral moment condition (2.1) holds under separability, that is, $\int G_\theta^{\text{sep}}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \boldsymbol{0}_r$ at some θ_0. For illustration, with spatial sampling in $\mathbb{R}^2$ ($d = 2$), we follow (4.1) for a given $\boldsymbol{t} = (t_1, t_2)' \in \mathbb{R}^2$ and define two different forms of $r = 3 > p = 2$ estimating functions, given in (4.2)–(4.3), as

$$\begin{aligned} \widetilde{G}_\theta(\boldsymbol{\omega}) = \big[&\mathbb{I}_{(-\infty,(t_1,t_2)]}(\boldsymbol{\omega}) - \theta_1\theta_2, \mathbb{I}_{(-\infty,(-t_1,t_2)]}(\boldsymbol{\omega}) \\ &- (1 - \theta_1)\theta_2, \mathbb{I}_{(-\infty,(t_1,-t_2)]}(\boldsymbol{\omega}) - \theta_1(1 - \theta_2)\big]', \end{aligned} \tag{4.2}$$

$$\begin{aligned} \widetilde{G}_\theta(\boldsymbol{\omega}) = \big[&\mathbb{I}(\omega_1 \leq t_1) - \theta_1, \mathbb{I}_{(-\infty,(-t_1,t_2)]}(\boldsymbol{\omega}) \\ &- (1 - \theta_1)\theta_2, \mathbb{I}_{(-\infty,(t_1,-t_2)]}(\boldsymbol{\omega}) - \theta_1(1 - \theta_2)\big]', \end{aligned} \tag{4.3}$$

where $\theta_i = \Phi^0(t_i)$, $i = 1, 2$. Estimating functions may also be extended to incorporate further vector sets $\boldsymbol{t}$.

To test for separability, we create a test statistic $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$, as in Section 2.2, using estimating functions $G_\theta^{\text{sep}}(\cdot)$. The following result establishes the test's validity.

Theorem 2 (PID or MID cases). *Suppose Conditions* (R.1)–(R.5) *hold, the estimating functions based on* (4.1) *are used, and* $\lim_{n\to\infty} n/\lambda_n^d \in (0, \infty]$. *Then, under* H_0: "$\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ *has a separable covariance structure,*"

$$-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n) \xrightarrow{d} \chi^2_{r-p} \quad \text{as } n \to \infty, \text{a.s. } (P_{\boldsymbol{X}}).$$

Again, for any outcome of the spatial sampling locations $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$, the test statistic based on the MELE $\widehat{\theta}_n$ has a simple chi-square limit for assessing separability, which holds in a unified manner for both PID and MID spatial sampling structures. One could alternatively formulate test statistics for separability based on estimating functions defined by process correlations $\rho(\cdot)$. For example, for some $\boldsymbol{h} = (h_1, \ldots, h_d)' \in \mathbb{R}^d$, define $p = d$ parameters $\theta = (\theta_1, \ldots, \theta_d)' \equiv [\rho_1(h_1), \ldots, \rho_d(h_d)]'$ where $\rho(\boldsymbol{h}) = \prod_{i=1}^d \rho_i(h_i)$ under covariance separability and set $r = p+1$ estimating functions as $G_\theta(\boldsymbol{\omega}) = [\cos(h_1\omega_1) - \theta_1, \ldots, \cos(h_r\omega_r) - \theta_r, \cos(\boldsymbol{h}'\boldsymbol{\omega}) - \prod_{i=1}^d \theta_i]'$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)' \in \mathbb{R}^d$. Such functions satisfy the moment condition (2.1) under separability, and variations are possible by adding further correlation parameters for other lags $\boldsymbol{h} \in \mathbb{R}^d$. However, similarly to Remark 1, test statistics based on the normalized spectral distribution (4.1) typically exhibited better power than those based on correlations in our numerical studies. The next section demonstrates the tests for separability using the estimating functions $G_\theta^{\text{sep}}(\cdot)$ from Theorem 2.

4.2. Illustration

We provide a numerical study of our tests for separability using the basic simulation design from Section 3.2 ($d = 2$). For this study, we considered a separable exponential correlation function $\sigma(\boldsymbol{h}) = \exp(-2h_1)\exp(-h_2)$ for size calculations and a non-separable exponential correlation $\sigma(\boldsymbol{h}) = \exp(-0.01\|\boldsymbol{h}\|)$ for power determination. The test statistics were based on the estimating functions from (4.2)–(4.3) (i.e., $r = 3$ functions for $p = 2$ parameters) using the normalized spectral distribution. We refer to these as Type 1 and Type 2 functions, respectively, and consider two different vector sets for implementing each: $\boldsymbol{t} = (t_1, t_2)'$ as (2, 1) or (1, 2). Tables 3–4 show the sizes and power, respectively, for tests of separability based on Type 2 functions.

In assessing separability, Type 2 functions emerged better than Type 1 in both size and power, though performances became similar for larger sampling regions. To explain this performance difference, note the Types 1 and 2 functions in (4.2)–(4.3) differ in their first components, where Type 2 functions estimate a parameter (i.e., θ_1, a marginal spectral distribution value) that exists even when separability assumptions are violated. This well-defined parameter is then used to set the remaining Type 2 functions in (4.3) for checking separability, which produced better outcomes in our simulations. From Tables 3–4, the separability tests with Type 2 functions were fairly insensitive to the frequency grid used in the EL device (e.g., $(\mathcal{C}, \kappa)$) and to the vector set used for $\boldsymbol{t}$. Power generally improved for increasing sample sizes n, regardless of the size ($\lambda_n = 24, 36$) of the sampling region.

Table 3. Empirical size for tests of separability based on normalized spectral distribution (Type 2 functions with $\boldsymbol{t} = (1, 2)$ or $(2, 1)$); from 1000 runs with nominal size 0.1

			(1, 2)				(2, 1)			
			n				n			
λ_n	$\mathcal{C}$	κ	1200	1800	2400	3600	1200	1800	2400	3600
24	1	0.05	0.114	0.123	0.150	0.149	0.122	0.140	0.134	0.133
		0.1	0.109	0.150	0.133	0.165	0.116	0.129	0.125	0.179
	1.5	0.05	0.111	0.116	0.115	0.124	0.120	0.113	0.101	0.127
		0.1	0.113	0.125	0.121	0.144	0.111	0.120	0.112	0.144
	2	0.05	0.111	0.109	0.121	0.133	0.110	0.119	0.111	0.133
		0.1	0.118	0.131	0.111	0.152	0.120	0.134	0.110	0.141
36	1	0.05	0.068	0.092	0.105	0.088	0.061	0.090	0.107	0.096
		0.1	0.075	0.109	0.075	0.104	0.074	0.108	0.080	0.102
	1.5	0.05	0.073	0.090	0.094	0.101	0.075	0.099	0.106	0.106
		0.1	0.076	0.099	0.085	0.101	0.072	0.102	0.091	0.105
	2	0.05	0.106	0.085	0.093	0.102	0.113	0.098	0.094	0.099
		0.1	0.084	0.093	0.093	0.088	0.083	0.099	0.088	0.097

Table 4. Empirical power for tests of separability based on normalized spectral distribution (Type 2 functions with $\boldsymbol{t} = (1, 2)$ or $(2, 1)$); from 1000 runs with nominal size 0.1

			(1, 2)				(2, 1)			
			n				n			
λ_n	$\mathcal{C}$	κ	1200	1800	2400	3600	1200	1800	2400	3600
24	1	0.05	0.772	0.866	0.881	0.902	0.766	0.868	0.874	0.902
		0.1	0.792	0.881	0.909	0.908	0.783	0.879	0.901	0.908
	1.5	0.05	0.669	0.797	0.856	0.892	0.663	0.798	0.856	0.890
		0.1	0.697	0.820	0.876	0.906	0.681	0.819	0.872	0.903
	2	0.05	0.584	0.718	0.792	0.871	0.590	0.715	0.798	0.873
		0.1	0.612	0.745	0.829	0.884	0.603	0.748	0.834	0.877
36	1	0.05	0.599	0.729	0.781	0.879	0.601	0.738	0.793	0.873
		0.1	0.598	0.732	0.798	0.884	0.598	0.727	0.808	0.886
	1.5	0.05	0.467	0.622	0.711	0.830	0.462	0.617	0.706	0.831
		0.1	0.470	0.644	0.689	0.825	0.474	0.643	0.692	0.827
	2	0.05	0.429	0.544	0.634	0.768	0.436	0.545	0.635	0.776
		0.1	0.399	0.530	0.584	0.769	0.406	0.520	0.572	0.761

5. Variogram model fitting

In contrast to previous sections (e.g., Sections 3–4), here we illustrate our method for testing spatial parameter values rather than for assessing a spatial covariance form. Section 5.1 describes the methodology applied to variogram model fitting and Section 5.2 numerically demonstrates its performance.

5.1. Methodology

Variogram estimation plays an important role in spatial prediction. Suppose $\{2\gamma^*(\cdot;\theta) : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^p$ denotes a class of variogram models for the true variogram $2\gamma^*(\boldsymbol{h}) \equiv \mathrm{Var}(Z(\boldsymbol{h}) - Z(\mathbf{0}))$, $\boldsymbol{h} \in \mathbb{R}^d$ of the process $Z(\cdot)$. Let $2\gamma(\cdot;\theta) \equiv 2\gamma^*(\cdot;\theta)/\sigma(\mathbf{0})$ and $2\gamma(\cdot) \equiv 2\gamma^*(\cdot)/\sigma(\mathbf{0})$ denote their scale-invariant versions. Least squares estimation (cf. [6]) is a common approach for fitting variogram models, but the resulting point estimators can have complex limiting distributions. As an alternative, our frequency domain testing method can be applied to least squared-type fitting, without requiring such limit laws. Motivated by the consideration that the population criterion $\sum_{i=1}^m \{2\gamma(\boldsymbol{h}_i) - 2\gamma(\boldsymbol{h}_i;\theta)\}^2$ (based on some fixed lags $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_m \in \mathbb{R}^d$) is minimized at the true parameter $\theta = \theta_0$, we can define $r = p$ estimating functions

$$G_\theta^{\mathrm{var}}(\boldsymbol{\omega}) = \sum_{i=1}^m \bigl\{1 - \cos(\boldsymbol{h}_i'\boldsymbol{\omega}) - \gamma(\boldsymbol{h}_i;\theta)\bigr\} \nabla\bigl[2\gamma(\boldsymbol{h}_i;\theta)\bigr], \quad \boldsymbol{\omega} \in \mathbb{R}^d, \tag{5.1}$$

where $\nabla[2\gamma(\boldsymbol{h};\theta)]$ denotes the $p \times 1$ vector of first order partial derivatives of $2\gamma(\boldsymbol{h};\theta)$ with respect to θ. Under mild conditions on the variogram model, these estimating functions fulfill the moment condition (2.1) at $\theta = \theta_0$, that is, $\int G_{\theta_0}^{\mathrm{var}}(\boldsymbol{\omega})\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} = \mathbf{0}_p$.

Using the EL scheme (Section 2.2) with functions (5.1), we formulate a test statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$ for the parameter hypothesis $H_0 : \theta = \theta_0 \in \mathbb{R}^p$. The form of this test statistic resembles a log-ratio similar to parameter assessments with parametric likelihood and also differs slightly from previous test statistics for evaluating moment conditions (e.g., $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$ in Theorems 1–2).

Theorem 3 (PID or MID cases). *Suppose Conditions* (R.1)–(R.5) *hold, the estimating functions based on* (5.1) *are used, and the variogram* $2\gamma(\cdot;\theta)$ *is twice continuously differentiable in a neighborhood of* θ_0. *Then, under* $H_0 : \theta = \theta_0 \in \Theta \subset \mathbb{R}^p$,

$$-2\widehat{a}_n \log\bigl[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)\bigr] \xrightarrow{d} \chi_p^2 \quad \textit{as } n \to \infty, \textit{a.s. } (P_{\boldsymbol{X}}).$$

Hence, the frequency domain test for variogram model parameters involves a simple chi-square calibration, without difficult variance estimation steps as common to alternative approaches for fitting variograms with irregular spatial observations in PID/MID sampling schemes (cf. [11], kernel variogram estimators). By inverting the test in Theorem 3, an approximate $100(1-\alpha)\%$ confidence region for variogram model parameters is given as

$$\bigl\{\theta \in \Theta : -2\widehat{a}_n \log\bigl[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)\bigr] \le \chi^2_{p,1-\alpha}\bigr\} \tag{5.2}$$

using a $(1-\alpha)$ chi-square percentile $\chi^2_{p,1-\alpha}$. We next demonstrate such confidence regions through numerical studies.

Remark 3. For evaluating parameter claims $H_0 : \theta = \theta_0$, an alternative EL test statistic of [BLN] can also be applied. This approach does not involve EL maximizers $\widehat{\theta}_n$ and, rather than (5.2), a confidence region is set as $\{\theta \in \Theta : -2a_n(\theta)\log \mathcal{R}_n(\theta) \le \chi^2_{p,1-\alpha}\}$ with a factor $a_n(\theta)$ defined by substituting θ for $\widehat{\theta}_n$ in (2.2). However, by using an extended version of EL with point estimation, the proposed confidence region (5.2) is computationally less involved (i.e., $\widehat{a}_n \equiv a_n(\widehat{\theta}_n)$ is evaluated only once at $\widehat{\theta}_n$) and can improve performance, as considered next in simulation studies for the variogram. Theoretical results in Section 6 also show that test statistics based on $\widehat{\theta}_n$ generally have better large-sample power compared to the counterpart statistics from [BLN] for parameter testing.

5.2. Illustration

We next study the test statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$ and the confidence region (5.2) for variogram model fitting and we also include the EL approach from [BLN] (cf. Remark 3). Using the simulation design of Section 3.2 ($d = 2$), we considered processes having variograms defined by either (a) an exponential correlation model with range parameter $r = 1$ or (b) a Matérn correlation model (cf. [31]) with smoothness parameter $\nu = 1$ and range parameter $r = 1.5$. We computed coverage probabilities of 90% confidence regions for these model parameters using the proposed approach (5.2) and the method of [BLN]. For variogram fitting, we applied estimating functions $G^{\mathrm{var}}_\theta(\cdot)$ from (5.1) based on a lag set $\{\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3, \boldsymbol{h}_4\} = \{(1, 1), (2, 2), (3, 3), (4, 4)\}$ to span a sequence of distances. The resulting empirical coverages appear in Tables 5 and 6. The proposed testing approach produced confidence sets with generally better accuracy than those without point estimation, irrespective of the frequency grid $(\mathcal{C}, \kappa)$ and sampling configuration. This provides evidence supporting the new frequency domain approach for parameter testing considered here.

6. Extensions to general frequency domain tests

While previous sections have treated several spatial testing scenarios, the proposed frequency domain method has the potential to be extended to other testing problems about spatial dependence forms or parameters. For the testing issue of interest, the central idea is to stipulate spatial estimating functions $G_\theta(\cdot)$ that fulfill a spatial expectation condition (2.1) in the frequency domain. Then, the EL scheme of Section 2.2 provides two types of test statistics. The test statistic $-2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n)$ can be applied to assess the null hypothesis H_0: "functions $G_\theta(\cdot)$, at some parameter θ_0, satisfy the moment condition (2.1)." This provides a general basis for assessing spatial covariance structures, using $r > p$ estimating functions for p spatial parameters. To test a hypothesis about specific parameter values $H_0 : \theta = \theta_0$, the test statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)]$ can be used and subsequently inverted to set confidence regions $\{\theta \in \Theta : -2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)] \le \chi^2_{p,1-\alpha}\}$.

Table 5. Coverages of 90% confidence regions for variogram parameters with (5.2) or without [BLN] point estimation; using exponential correlation ($r = 1$) and 1000 simulation runs

			n							
			1200		1800		2400		3600	
λ_n	$\mathcal{C}$	κ	with	w/o	with	w/o	with	w/o	with	w/o
24	1	0.05	0.905	0.944	0.892	0.923	0.897	0.914	0.897	0.897
		0.10	0.889	0.923	0.892	0.915	0.890	0.910	0.869	0.873
	1.5	0.05	0.910	0.952	0.904	0.941	0.895	0.919	0.911	0.926
		0.10	0.892	0.938	0.898	0.936	0.893	0.917	0.887	0.896
	2	0.05	0.910	0.946	0.916	0.957	0.916	0.953	0.903	0.939
		0.10	0.886	0.940	0.920	0.954	0.901	0.921	0.904	0.919
36	1	0.05	0.910	0.970	0.912	0.950	0.901	0.951	0.892	0.932
		0.10	0.888	0.945	0.902	0.940	0.907	0.939	0.901	0.937
	1.5	0.05	0.901	0.950	0.923	0.972	0.908	0.955	0.888	0.941
		0.10	0.888	0.947	0.911	0.958	0.924	0.959	0.903	0.940
	2	0.05	0.887	0.952	0.910	0.953	0.906	0.960	0.884	0.941
		0.10	0.891	0.953	0.895	0.964	0.909	0.962	0.906	0.960

Table 6. Coverages of 90% confidence regions for variogram parameters with (5.2) or without [BLN] point estimation; using Matérn correlation ($\nu = 1$, $r = 1.5$) and 1000 simulation runs

			n							
			1200		1800		2400		3600	
λ_n	$\mathcal{C}$	κ	with	w/o	with	w/o	with	w/o	with	w/o
24	1	0.05	0.807	0.860	0.792	0.854	0.848	0.880	0.815	0.869
		0.10	0.806	0.855	0.801	0.861	0.840	0.879	0.824	0.871
	1.5	0.05	0.801	0.854	0.793	0.837	0.867	0.911	0.850	0.903
		0.10	0.788	0.832	0.782	0.827	0.875	0.902	0.840	0.877
	2	0.05	0.870	0.908	0.894	0.931	0.861	0.901	0.843	0.887
		0.10	0.879	0.918	0.859	0.913	0.873	0.909	0.838	0.883
36	1	0.05	0.869	0.926	0.862	0.910	0.894	0.940	0.893	0.931
		0.10	0.877	0.914	0.860	0.922	0.905	0.945	0.867	0.905
	1.5	0.05	0.848	0.897	0.852	0.909	0.898	0.930	0.915	0.957
		0.10	0.855	0.905	0.851	0.902	0.905	0.936	0.886	0.921
	2	0.05	0.867	0.929	0.882	0.939	0.895	0.940	0.896	0.947
		0.10	0.871	0.934	0.897	0.947	0.899	0.939	0.885	0.930

The following theorem establishes the validity of these frequency domain tests. Due to its generality, the result requires additional conditions on the estimating functions (denoted as Conditions (R.6)–(R.8) in Supplement A) which hold for functions in Theorem 1–3 based on the spectral distribution or autocorrelations.

Theorem 4 (PID or MID cases). *Suppose Conditions* (R.1)–(R.8) *hold*; *the* $r \geq p$ *estimating functions satisfy the moment* (2.1) *at a true* $\theta_0 \in \Theta \subset \mathbb{R}^p$; $D_{\theta_0} \equiv \int_{\mathbb{R}^d} [\partial G_{\theta_0}(\boldsymbol{\omega})/\partial\theta]\phi(\boldsymbol{\omega})\, d\boldsymbol{\omega}$ *has full column rank* p; *and* $\lim_{n\to\infty} n/\lambda_n^d \in (0, \infty]$. *Then, as* $n \to \infty$,

$$-2\widehat{a}_n \log\bigl[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)\bigr] \xrightarrow{d} \chi_p^2 \quad \textit{and} \quad -2\widehat{a}_n \log \mathcal{R}_n(\widehat{\theta}_n) \xrightarrow{d} \chi_{r-p}^2 \quad \textit{a.s. } (P_{\boldsymbol{X}}).$$

Without stringent assumptions on the underlying process or spatial sampling design, the proposed frequency domain approach allows for a wide range of moment and parameter assessments with irregularly located spatial data, as tools for diagnosing dependence structures with such data.

We mention two further properties related to parameter tests based on the log-ratio statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$ in Theorem 4. The first is that profile log-ratio statistics can be specified for testing parameter subsets of θ in an analogous way to parametric likelihood. Decomposing the spatial parameter $\theta = (\theta_1, \theta_2) \in \Theta \in \mathbb{R}^p$ into a $q \times 1$ subvector θ_1 and a $(p-q) \times 1$ subvector θ_2, the profile EL statistic for a specified value of $\theta_1 \in \mathbb{R}^q$ is given as

$$-2\widehat{a}_n \log\bigl[\mathcal{R}_n\bigl(\theta_1, \widehat{\theta}_{2,n}^{\theta_1}\bigr)/\mathcal{R}_n(\widehat{\theta}_n)\bigr]$$

where, given θ_1, the estimator $\widehat{\theta}_{2,n}^{\theta_1}$ maximizes $\mathcal{R}_n(\theta_1, \theta_2)$ with respect to $\theta_2 \in \mathbb{R}^{p-q}$. Corollary 1 establishes that this profile statistic is valid with an intuitive limit.

Corollary 1. *Suppose the conditions of Theorem* 4 *hold*. *Then*, *under* $H_0 : \theta_1 = \theta_{01} \in \mathbb{R}^q$,

$$-2\widehat{a}_n \log\bigl[\mathcal{R}_n\bigl(\theta_{01}, \widehat{\theta}_{2,n}^{\theta_{01}}\bigr)/\mathcal{R}_n(\widehat{\theta}_n)\bigr] \xrightarrow{d} \chi_q^2 \quad \textit{as } n \to \infty \textit{ a.s. } (P_{\boldsymbol{X}}).$$

As a second property, the log-ratio statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta)/\mathcal{R}_n(\widehat{\theta}_n)]$ from Theorem 4 can generally be shown to have better local power than an alternative EL statistic $-2a(\theta)\log \mathcal{R}_n(\theta)$ from [BLN] for testing hypotheses about a spatial parameter $\theta \in \Theta \subset \mathbb{R}^p$. As described in Section 2.2, both test statistics are based on $r \geq p$ estimating functions and a frequency grid for the spatial periodogram, though the statistic of [BLN] lacks point estimation and the adjustment $a_n(\theta)$ is defined by θ in place of $\widehat{\theta}_n$ in (2.2); see also Remark 3. To describe local power properties, we require a factor $b_n^2 \equiv (\lambda_n^d/n)^2 N_n + \lambda_n^{\kappa d}$ that depends on the number n of sampling sites, the volume λ_n^d of spatial sampling region, the number $N \equiv N_n$ of frequency grid spacings and the volume $\lambda_n^{-\kappa d}$ between such spacings. While the size of b_n changes with the spatial asymptotic scheme (e.g., PID vs. MID), it holds that $b_n \to \infty$ with $b_n \lambda_n^{-\kappa d} \to 0$ as $n \to \infty$. If $\theta_0 \in \mathbb{R}^p$ denotes the true parameter value, Corollary 2 next determines the asymptotic power of both test statistics along a sequence of alternative parameters as shifts of θ_0 by $O(b_n \lambda_n^{-\kappa d})$.

Corollary 2. *Suppose Theorem* 4 *conditions hold and define* $\theta_n \equiv \theta_0 + b_n \lambda_n^{-\kappa d} v$, $n \geq 1$, *in terms of the true parameter* $\theta_0 \in \mathbb{R}^p$ *and an arbitrary vector* $v \in \mathbb{R}^p$. *Then*, *given any sub-sequence*

$\{n_j\} \subset \{n\}$, *there exists a further sub-sequence* $\{n_k\} \subset \{n_j\}$ *along with a positive definite* $r \times r$ *matrix* U_2 *and a constant* $a \in \{1, 2\}$ *such that, as* $n_k \to \infty$,

$$-2\widehat{a}_{n_k} \log\big[\mathcal{R}_{n_k}(\theta_{n_k})/\mathcal{R}_{n_k}(\widehat{\theta}_{n_k})\big] \xrightarrow{d} \chi_p^2(\tau) \quad \textit{and} \quad -2a(\theta_{n_k}) \log \mathcal{R}_{n_k}(\theta_{n_k}) \xrightarrow{d} \chi_r^2(\tau)$$

hold a.s. $(P_{\mathbf{X}})$, *where* $\tau \equiv a^{-1} v' U_2^{-1} v$ *is a non-centrality parameter. As described in Proposition 7.1, values of* U_2, a *vary by the spatial asymptotic structure* (*PID-MID/slow infill vs. MID/fast infill*) *and* U_2 *may further depend on the sub-sequence* $\{n_k\}$.

The statement of Corollary 2 is complicated by the issue that the estimating functions determine if the non-centrality parameter τ (through U_2) can potentially vary with the sub-sequence $\{n_k\}$; see also Proposition 7.1 and its proof. For many estimating functions though, including those of Sections 3–5, the value τ does not change with the sub-sequence; in which case, Corollary 2 implies convergence to non-central chi-squared distributions. However, regardless of the exact subsequence, Corollary 2 gives that p-values from the test statistic $-2\widehat{a}_n \log[\mathcal{R}_n(\theta_n)/\mathcal{R}_n(\widehat{\theta}_n)]$ will be smaller than those from the counterpart statistic $-2a(\theta_n) \log \mathcal{R}_n(\theta_n)$ along local alternatives $\theta_n = \theta_0 + b_n \lambda_n^{-\kappa d} v$: for any constant $c > 0$, the probability that the latter test statistic exceeds a threshold c is asymptotically given by $P(\chi_r^2(\tau) > c)$, which exceeds the asymptotic p-value $P(\chi_p^2(\tau) > c)$ of the first statistic (e.g., $\chi_r^2(\tau) \stackrel{d}{=} \chi_p^2(\tau) + \chi_{r-p}^2(\tau)$ for an independent sum). Hence, the spatial EL version with point estimation $\widehat{\theta}_n$ can have power advantages for parameter tests over the EL analog without estimation $\widehat{\theta}_n$. Additionally, as shown in Corollary 1, such point estimation also leads to natural profile test statistics for spatial parameters, which are not possible in the EL framework lacking maximization steps.

7. Overview of theoretical development and proofs

This section aims to briefly outline technical details in showing chi-square limit laws for the spatial EL test statistics based on point maximization $\widehat{\theta}_n$. Section 7.1 sets up ingredients for establishing these asymptotic distributions. At issue, the testing method is again intended to be unified and valid across differing asymptotic forms for spatial sampling, and Proposition 7.1 indicates the rather complex large-sample behavior of periodogram-based quantities important to the method. This result culminates with the distributional feature that appropriate scaling for maximized EL log-ratio statistics (in Proposition 7.1(ii)) varies with the underlying spatial asymptotics. Section 7.2 then proves the general testing result in Theorem 4 for assessing spatial structure, showing that estimated scaling factors automatically adjust to the asymptotic type of spatial sampling and also previewing some intricacies of the spatial periodogram. A complete description of regularity conditions and technical supporting lemmas appear in Supplement A along with proofs of Corollaries 1–2.

7.1. A key result for the spatial test statistics

Proposition 7.1 collects some basic components for establishing the distributional limits of spatial test statistics based on EL. To describe the result, recall that the differing PID or MID spatial sampling forms correspond respectively to $\lim_{n\to\infty} n/\lambda_n^d \in (0,\infty)$ or $\lim_{n\to\infty} n/\lambda_n^d = \infty$, where n is the number of sampling sites and λ_n^d is the sampling region volume. The asymptotic type of spatial sampling turns out to impact the properties of test statistics, as can the volume $N\lambda_n^{-\kappa d}$ of the frequency grid used in the testing method (cf. Section 2.2), where $N \equiv N_n$ and $\lambda_n^{-\kappa d}$ denote the number of, and volume between, frequency grid spacings. In particular, additional complexities arise in establishing the spatial statistics under MID sampling where two distributional subcases merge depending how fast n grows relative to $\lambda_n^d(N_n\lambda_n^{-\kappa d})^{1/2}$. Note that, under MID (i.e., $\lim_{n\to\infty} n/\lambda_n^d = \infty$), the sample size n already grows faster than the volume λ_n^d of the spatial sampling region but the subcases require further considering whether or not n grows faster than λ_n^d expanded by a factor involving the frequency grid volume $N\lambda_n^{-\kappa d} \uparrow \infty$ as $n \to \infty$. When n increases faster than $\lambda_n^d(N_n\lambda_n^{-\kappa d})^{1/2}$, this corresponds a MID subcase with a "fast rate" of infill sampling, where n that grows substantially faster than the sampling region volume λ_n^d (increased by a factor $(N\lambda_n^{-\kappa d})^{1/2}$); a MID subcase with a slower infill rate results when $\lambda_n^d(N_n\lambda_n^{-\kappa d})^{1/2}$ dominates n. These differences impact the testing method through the spatial periodogram, which has a bias that decreases at a rate λ_n^d/n. To summarize the effects of different spatial asymptotics on the EL-based testing method, Proposition 7.1(i) first describes a stochastic expansion of the point estimator $\widehat{\theta}_n$, which is used for establishing the main outcome of maximized EL test statistics in Proposition 7.1(ii). Appropriate "scaling" is an issue in both Proposition 7.1(i)–(ii), where scaling for the estimator $\widehat{\theta}_n$ involves a factor $b_n^2 \equiv (\lambda_n^d/n)^2 N_n + \lambda_n^{\kappa d}$ that varies in size by the case of spatial asymptotics. Details on the proof appear in Supplement A.

Proposition 7.1. *Assume Theorem 4 conditions. Further, define a constant $a \in \{1,2\}$ where $a = 2$ if either* $\lim_{n\to\infty} n/\lambda_n^d \in (0,\infty)$ (PID) *or if* $\lim_{n\to\infty} n/\lambda_n^d = \infty$ *with $n \ll \lambda_n^d(N_n\lambda_n^{-\kappa d})^{1/2}$* (MID/slow infilling); *and where $a = 1$ if* $\lim_{n\to\infty} n/\lambda_n^d = \infty$ *with $n \gg \lambda_n^d(N_n\lambda_n^{-\kappa d})^{1/2}$* (MID/fast infilling).

(i) *Then, given any sub-sequence $\{n_j\} \subset \{n\}$, there exists a further sub-sequence $\{n_k\} \subset \{n_j\}$ and a positive definite $r \times r$ matrix V (possibly depending on $\{n_k\}$ and $a \in \{1,2\}$) such that*

$$b_{n_k} J_{n_k,\theta_0} \equiv \frac{1}{b_{n_k}^2} \sum_{j=1}^{N_{n_k}} G_{\theta_0}(\boldsymbol{\omega}_{jn_k}) \widetilde{I}_{n_k}(\boldsymbol{\omega}_{jn_k}) \xrightarrow{d} N(\mathbf{0}_r, aV) \quad a.s.\ (P_{\boldsymbol{X}}),$$

holds with $b_{n_k}^2 \equiv (\lambda_{n_k}^d/n_k)^2 N_{n_k} + \lambda_{n_k}^{\kappa d} \to \infty$ and $\lambda_{n_k}^{\kappa d}/b_{n_k} \to \infty$ as $n_k \to \infty$, and the maximizer $\widehat{\theta}_{n_k} \in \mathbb{R}^p$ along the sub-sequence satisfies

$$\begin{pmatrix} b_{n_k} \boldsymbol{t}_{\widehat{\theta}_{n_k}} \\ (\widehat{\theta}_{n_k} - \theta_0)\dfrac{\lambda_{n_k}^{\kappa d}}{b_{n_k}} \end{pmatrix} = \begin{pmatrix} U_1 \\ -U_2 D_{\theta_0}' V^{-1} \end{pmatrix} b_{n_k} J_{n_k,\theta_0} + o_p(1) \quad a.s.\ (P_{\boldsymbol{X}})$$

for a random vector $\boldsymbol{t}_{\widehat{\theta}_{n_k}} \in \mathbb{R}^r$ *(an EL Langrange multiplier for* $\widehat{\theta}_{n_k}$*) and non-singular matrices* $U_1 = V^{-1} - V^{-1}D_{\theta_0}U_2 D'_{\theta_0}V^{-1}$ *and* $U_2 = (D'_{\theta_0}V^{-1}D_{\theta_0})^{-1}$.

(ii) *As* $n \to \infty$, *it holds that*

$$-\frac{2}{a}\log\big[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)\big] \xrightarrow{d} \chi^2_p \quad \textit{and}$$

$$-\frac{2}{a}\log\mathcal{R}_n(\widehat{\theta}_n) \xrightarrow{d} \chi^2_{r-p} \quad \textit{a.s. } (P_{\boldsymbol{X}}).$$

In Proposition 7.1(ii), note that the appropriate scaling (i.e., -1 or -2) for test statistics changes with the asymptotic form of spatial sampling. Similarly to EL theory in [BLN], a dichotomy in scaling exists according to whether n grows substantially faster than λ_n^d ("fast infill" MID) or not (i.e., PID or "slow infill" MID cases). The estimator $\widehat{\theta}_n$ in the spatial test statistics here creates challenges compared to the EL form of [BLN] that lacks point estimation. The spatial peridogram $\widetilde{I}_n(\cdot)$ and estimating functions $G_\theta(\cdot)$ induce the large-sample properties of the point estimator $\widehat{\theta}_n$ in Proposition 7.1(i), which in turn play an essential role in the main distributional limits of Proposition 7.1(ii). However, in contrast to other EL applications with estimating functions (cf. [19,28]), the spatial point estimator $\widehat{\theta}_n$ is not asymptotically guaranteed to have a normal limit. That is, the behavior of $\widehat{\theta}_n$ (particularly its limit variance) can vary across different sample sub-sequences so that a corresponding limit distribution may not even exist for $\widehat{\theta}_n$, though the estimator remains consistent. This aspect owes in part to the unbounded frequency domain in the analysis of irregularly located spatial data, as described in [33]. However, while the large-sample behavior of point estimators in Proposition 7.1(i) depends on sub-sequences, the limit distribution of test statistics in Proposition 7.1(ii) does *not*, which is of main interest here. However, correct scaling for such test statistics still depends on the underlying spatial asymptotics.

Ultimately, the general test statistics $-2\widehat{a}_n\log[\mathcal{R}_n(\theta_0)/\mathcal{R}_n(\widehat{\theta}_n)]$ or $-2\widehat{a}_n\log\mathcal{R}_n(\widehat{\theta}_n)$ in Theorem 4 need to automatically adjust to the spatial sampling structure with estimated scaling $\widehat{a}_n$ from (2.2). We show this to be true and provide a proof of Theorem 4 from Proposition 7.1(ii) next.

7.2. Proof of Theorem 4

In the following, all probabilistic convergence (e.g., o_p, O_p, $\xrightarrow{p}$) refers to probability conditional $P(\cdot) \equiv P(\cdot|\boldsymbol{X})$ for the spatial process $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{R}^d\}$ given a collection of random vectors $\boldsymbol{X} \equiv \{\boldsymbol{X}_i\}_{i\geq 1} \subset \mathbb{R}^d$ determining spatial locations (Section 2.1); we generally suppress that such convergence holds a.s. $(P_{\boldsymbol{X}})$, where again $P_{\boldsymbol{X}}$ denotes the joint distribution of $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ (cf. Section 3.1). From Proposition 7.1(ii) and the constant $a \in \{1, 2\}$ there, Theorem 4 will follow by establishing $2\widehat{a}_n \xrightarrow{p} 2/a$ for $\widehat{a}_n$ from (2.2) or equivalently that

$$\widehat{a}_n \xrightarrow{p} \begin{cases} \frac{1}{2} & \text{for PID or MID with } n \ll \lambda_n^d (N_n\lambda_n^{-\kappa d})^{1/2} \\ 1 & \text{MID with } n \gg \lambda_n^d (N_n\lambda_n^{-\kappa d})^{1/2} \end{cases}$$

as $n \to \infty$ (a.s. (P_X)). Set $c_n = n/\lambda_n^d$ and $b_n^2 = c_n^{-2}N + \lambda_n^{\kappa d}$ (with $N \equiv N_n$), where $b_n^{-1} + \lambda_n^{-\kappa d} b_n \to 0$ holds, and recall $I_n(\cdot)$ and $\widetilde{I}_n(\cdot) = I_n(\cdot) - c_n^{-1}\widehat{\sigma}(\mathbf{0})$ denote the ordinary and bias corrected periodograms (Section 2.1). We shall refer to some conditions (e.g., (R.3)) and lemmas appearing in Supplement A, and write $s_n \sim t_n$ to denote $\lim_{n\to\infty}(s_n/t_n) = 1$ for generic positive sequences $\{s_n\}$, $\{t_n\}$.

As $\|\widehat{\theta}_n - \theta_0\| = O_p(\lambda_n^{-\kappa d} b_n) = o_p(1)$ by Proposition 7.1(i), we have

$$\begin{aligned}
&\left| b_n^{-2} \sum_{j=1}^{N} \left(\left\| G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn}) \right\|^2 - \left\| G_{\theta_0}(\boldsymbol{\omega}_{jn}) \right\|^2 \right) I_n^2(\boldsymbol{\omega}_{jn}) \right| \\
&\quad \le C \|\widehat{\theta}_n - \theta_0\| b_n^{-2} \sum_{j=1}^{N} \left(\widetilde{I}_n^2(\boldsymbol{\omega}_{jn}) + \widehat{\sigma}(\mathbf{0})^2 c_n^{-2} \right) \\
&\quad = o_p(1) O_p\left(1 + b_n^{-2} N c_n^{-2}\right) = o_p(1)
\end{aligned}$$

using the differentiability of $G_\theta(\cdot)$ in a neighborhood of θ_0 under Condition (R.7) along with $b_n^{-2} \sum_{j=1}^{N} \widetilde{I}_n^2(\boldsymbol{\omega}_{jn}) = O_p(1)$ (cf. Lemma 3), $b_n^{-2} N c_n^{-2} = O(1)$, and the fact that $\widehat{\sigma}_n(\mathbf{0}) \equiv n^{-1} \sum_{j=1}^{n} (Z(\boldsymbol{s}_j) - \bar{Z}_n)^2 = O_p(1)$; it also follows that

$$\left| b_n^{-2} \sum_{j=1}^{N} \left(\left\| G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn}) \right\|^2 - \left\| G_{\theta_0}(\boldsymbol{\omega}_{jn}) \right\|^2 \right) \widetilde{I}_n^2 \right| = o_p(1).$$

From these, we have

$$\begin{aligned}
b_n^{-2} \sum_{j=1}^{N} \left\| G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn}) \right\|^2 \widetilde{I}_n^2(\boldsymbol{\omega}_{jn}) &= b_n^{-2} \sum_{j=1}^{N} \left\| G_{\theta_0}(\boldsymbol{\omega}_{jn}) \right\|^2 \left(A_n(\boldsymbol{\omega}_{jn})^2 + K^2 \phi(\boldsymbol{\omega}_{jn})^2 \right) + o_p(1), \\
b_n^{-2} \sum_{j=1}^{N} \left\| G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn}) \right\|^2 I_n^2(\boldsymbol{\omega}_{jn}) &= b_n^{-2} \sum_{j=1}^{N} \left\| G_{\theta_0}(\boldsymbol{\omega}_{jn}) \right\|^2 2 A_n(\boldsymbol{\omega}_{jn})^2 + o_p(1)
\end{aligned} \tag{7.1}$$

using periodogram expansions (cf. Lemma 2) given by

$$\begin{aligned}
\sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jn})^2 \left[\widetilde{I}_n^2(\boldsymbol{\omega}_{jn}) - \left(A_n(\boldsymbol{\omega}_{jn})^2 + K^2 \phi(\boldsymbol{\omega}_{jn})^2 \right) \right] &= o_p\left(b_n^2\right), \\
\sum_{j=1}^{N} G_{\theta_0}(\boldsymbol{\omega}_{jn})^2 \left[I_n^2(\boldsymbol{\omega}_{jn}) - 2 A_n(\boldsymbol{\omega}_{jn})^2 \right] &= o_p\left(b_n^2\right),
\end{aligned}$$

where $A_n(\boldsymbol{\omega}) = c_n^{-1}\sigma(\mathbf{0}) + K\phi(\boldsymbol{\omega})$, $\boldsymbol{\omega} \in \mathbb{R}^d$ for $K = (2\pi)^d \int f^2$. Because $\lambda_n^{-\kappa d} \sum_{j=1}^N \phi(\boldsymbol{\omega}_{jn}) \to \int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\, d\boldsymbol{\omega}$ from $\int_{\mathbb{R}^d} \phi(\boldsymbol{\omega})\, d\boldsymbol{\omega} < \infty$ and Condition (R.3), we also have bounds

$$\sum_{j=1}^{N} \phi(\boldsymbol{\omega}_{jn}) = O\big(\lambda_n^{\kappa d}\big) = O\big(b_n^2\big), \qquad \sum_{j=1}^{N} \phi^2(\boldsymbol{\omega}_{jn}) = O\big(\lambda_n^{\kappa d}\big) = O\big(b_n^2\big), \tag{7.2}$$

by $\sup_{1\le j \le N} |\phi(\boldsymbol{\omega}_{jn})| \le C$ from (R.3).

Under PID (i.e., $\lim_{n\to\infty} c_n \in (0,\infty)$) or under MID (i.e., $\lim_{n\to\infty} c_n = \infty$) with $n \ll \lambda_n^d (N_n \lambda_n^{-\kappa d})^{1/2}$, it holds that $b_n^2 \sim N c_n^{-2}$. In these cases, it follows from the definition of scaling $\widehat{a}_n$ from (2.2), upon using (7.1)–(7.2) with Condition (R.3) (i.e., $\liminf_{n\to\infty} N^{-1} \times \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 > 0$), that

$$\begin{aligned}
\widehat{a}_n &= \frac{b_n^{-2} \sum_{j=1}^N \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 \widetilde{I}_n^2(\boldsymbol{\omega}_{jn})}{b_n^{-2} \sum_{j=1}^N \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 I_n^2(\boldsymbol{\omega}_{jn})} = \frac{b_n^{-2} \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 c_n^{-2} \sigma(\mathbf{0})^2 + o_p(1)}{b_n^{-2} \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 2 c_n^{-2} \sigma(\mathbf{0})^2 + o_p(1)} \\
&= \frac{1}{2} + o_p(1).
\end{aligned}$$

Under MID with $n \ll \lambda_n^d (N_n \lambda_n^{-\kappa d})^{1/2}$, it holds instead that $b_n^2 \sim \lambda_n^{\kappa d}$ with $c_n^{-1} \to 0$. Using these rates along with (7.1)–(7.2) and Condition (R.3) (i.e., $\lim_{n\to\infty} \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 \phi(\boldsymbol{\omega}_{jn})^2 > 0$), we have

$$\begin{aligned}
\widehat{a}_n &= \frac{b_n^{-2} \sum_{j=1}^N \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 \widetilde{I}_n^2(\boldsymbol{\omega}_{jn})}{b_n^{-2} \sum_{j=1}^N \|G_{\widehat{\theta}_n}(\boldsymbol{\omega}_{jn})\|^2 I_n^2(\boldsymbol{\omega}_{jn})} = \frac{b_n^{-2} \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 2K^2 \phi(\boldsymbol{\omega}_{jn})^2 + o_p(1)}{b_n^{-2} \sum_{j=1}^N \|G_{\theta_0}(\boldsymbol{\omega}_{jn})\|^2 2K^2 \phi(\boldsymbol{\omega}_{jn})^2 + o_p(1)} \\
&= 1 + o_p(1)
\end{aligned}$$

in this case. This completes the proof.

8. Conclusions

We have developed a general spatial frequency domain method for irregularly spaced data, which is applicable to a broad class of spatial processes and to a variety of inference problems about spatial covariance. Using an extended empirical likelihood device, test statistics were shown to have simple chi-square limits under mild conditions, without explicit assumptions or estimation concerning the process distribution, the concentration of sampling locations, or the exact nature of the spatial asymptotics (e.g., the amount of infill sampling). Depending on the spatial inference problem of interest, the testing method requires specification of appropriate estimating functions. For concreteness, we formally treated examples for testing isotropy or separability as well as fitting variogram models, but general extensions (cf. Section 6) make further applications possible. These include, for example, potentially novel goodness-of-fit assessments for the parametric form of variogoram models or spectral densities (e.g., Whittle estimation). In this sense, the proposed methodology provides a unified platform for inference across many scenarios of spatial covariance assessment.

Open questions remain for the testing method about the best choices of estimating functions for inference problems of interest (e.g., the number and structure of such functions). This issue requires investigation through further applications. We have found that, where possible, the formulation of spectral estimating functions based on the normalized spectral distribution, rather than process correlations directly, often produced tests with better robustness properties in performance and implementation.

Acknowledgements

The authors are grateful to two reviewers and an Associate Editor for thoughtful comments and constructive criticism which led to significant improvements in the manuscript. S. Bandyopadhyay is supported in part by NSF DMS-1811384 and DMS-1854181. S.N. Lahiri is supported in part by NSF DMS-1310068 and DMS-1613192. D.J. Nordman is supported in part by NSF DMS-1406747.

Supplementary Material

Supplement A: "A general frequency domain method for assessing spatial covariance structures" (DOI: 10.3150/19-BEJ1160SUPPA; .pdf). Details of regularity conditions and proofs of distributional results, along with additional numerical results.

Supplement B: "A general frequency domain method for assessing spatial covariance structures" (DOI: 10.3150/19-BEJ1160SUPPB; .pdf). Discussion and extended simulation results under further spatial sampling designs.

References

[1] Bandyopadhyay, S., Jentsch, C. and Subba Rao, S. (2017). A spectral domain test for stationarity of spatio-temporal data. *J. Time Series Anal.* **38** 326–351. MR3611747 https://doi.org/10.1111/jtsa.12222

[2] Bandyopadhyay, S. and Lahiri, S.N. (2009). Asymptotic properties of discrete Fourier transforms for spatial data. *Sankhyā* **71** 221–259. MR2639292

[3] Bandyopadhyay, S., Lahiri, S.N. and Nordman, D.J. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *Ann. Statist.* **43** 519–545. MR3316189 https://doi.org/10.1214/14-AOS1291

[4] Bandyopadhyay, S. and Subba Rao, S. (2017). A test for stationarity for irregularly spaced spatial data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 95–123. MR3597966 https://doi.org/10.1111/rssb.12161

[5] Bochner, S. (1959). *Lectures on Fourier Integrals. With an Author's Supplement on Monotonic Functions, Stieltjes Integrals, and Harmonic Analysis. Annals of Mathematics Studies* **42**. Princeton, NJ: Princeton Univ. Press. Translated by Morris Tenenbaum and Harry Pollard. MR0107124

[6] Cressie, N.A.C. (1993). *Statistics for Spatial Data*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. New York: Wiley. Revised reprint of the 1991 edition. MR1239641 https://doi.org/10.1002/9781119115151

[7] Crujeiras, R.M., Fernández-Casal, R. and González-Manteiga, W. (2010). Goodness-of-fit tests for the spatial spectral density. *Stoch*. *Environ*. *Res*. *Risk Assess*. **24** 67–79.

[8] Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika* **89** 197–210. MR1888368 https://doi.org/10.1093/biomet/89.1.197

[9] Fuentes, M. (2005). A formal test for nonstationarity of spatial stochastic processes. *J*. *Multivariate Anal*. **96** 30–54. MR2202399 https://doi.org/10.1016/j.jmva.2004.09.003

[10] Fuentes, M. (2006). Testing for separability of spatial-temporal covariance functions. *J*. *Statist*. *Plann*. *Inference* **136** 447–466. MR2211349 https://doi.org/10.1016/j.jspi.2004.07.004

[11] García-Soidán, P. (2007). Asymptotic normality of the Nadaraya–Watson semivariogram estimators. *TEST* **16** 479–503. MR2365173 https://doi.org/10.1007/s11749-006-0016-8

[12] Gikhman, I.I. and Skorokhod, A.V. (2015). *The Theory of Stochastic Processes I*. Springer.

[13] Guan, Y., Sherman, M. and Calvin, J.A. (2004). A nonparametric test for spatial isotropy using subsampling. *J*. *Amer*. *Statist*. *Assoc*. **99** 810–821. MR2090914 https://doi.org/10.1198/016214504000001150

[14] Guyon, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika* **69** 95–105. MR0655674 https://doi.org/10.1093/biomet/69.1.95

[15] Hall, P. and Patil, P. (1994). Properties of nonparametric estimators of autocovariance for stationary random fields. *Probab*. *Theory Related Fields* **99** 399–424. MR1283119 https://doi.org/10.1007/BF01199899

[16] Heyde, C.C. and Gay, R. (1993). Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stochastic Process*. *Appl*. **45** 169–182. MR1204868 https://doi.org/10.1016/0304-4149(93)90067-E

[17] Im, H.K., Stein, M.L. and Zhu, Z. (2007). Semiparametric estimation of spectral density with irregular observations. *J*. *Amer*. *Statist*. *Assoc*. **102** 726–735. MR2381049 https://doi.org/10.1198/016214507000000220

[18] Kaiser, M.S. and Nordman, D.J. (2012). Blockwise empirical likelihood for spatial Markov model assessment. *Stat*. *Interface* **5** 303–318. MR2979210 https://doi.org/10.4310/SII.2012.v5.n3.a3

[19] Kitamura, Y. (1997). Empirical likelihood methods with weakly dependent processes. *Ann*. *Statist*. **25** 2084–2102. MR1474084 https://doi.org/10.1214/aos/1069362388

[20] Lahiri, S.N. (2003). Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **65** 356–388. MR2028905

[21] Lahiri, S.N. and Mukherjee, K. (2004). Asymptotic distributions of M-estimators in a spatial regression model under some fixed and stochastic spatial sampling designs. *Ann*. *Inst*. *Statist*. *Math*. **56** 225–250. MR2067154 https://doi.org/10.1007/BF02530543

[22] Lahiri, S.N. and Zhu, J. (2006). Resampling methods for spatial regression models under a class of stochastic designs. *Ann*. *Statist*. **34** 1774–1813. MR2283717 https://doi.org/10.1214/009053606000000551

[23] Maity, A. and Sherman, M. (2012). Testing for spatial isotropy under general designs. *J*. *Statist*. *Plann*. *Inference* **142** 1081–1091. MR2879753 https://doi.org/10.1016/j.jspi.2011.11.013

[24] Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **71** 191–217. MR2655530 https://doi.org/10.1111/j.1467-9868.2008.00685.x

[25] Nordman, D.J. and Caragea, P.C. (2008). Point and interval estimation of variogram models using spatial empirical likelihood. *J*. *Amer*. *Statist*. *Assoc*. **103** 350–361. MR2420238 https://doi.org/10.1198/016214507000001391

[26] Nordman, D.J. and Lahiri, S.N. (2014). A review of empirical likelihood methods for time series. *J. Statist. Plann. Inference* **155** 1–18. MR3264452 https://doi.org/10.1016/j.jspi.2013.10.001

[27] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120. MR1041387 https://doi.org/10.1214/aos/1176347494

[28] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22** 300–325. MR1272085 https://doi.org/10.1214/aos/1176325370

[29] Stein, M. (1989). Asymptotic distributions of minimum norm quadratic estimators of the covariance function of a Gaussian random field. *Ann. Statist.* **17** 980–1000. MR1015134 https://doi.org/10.1214/aos/1176347252

[30] Stein, M.L. (1995). Fixed-domain asymptotics for spatial periodograms. *J. Amer. Statist. Assoc.* **90** 1277–1288. MR1379470

[31] Stein, M.L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging. Springer Series in Statistics*. New York: Springer. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

[32] Subba Rao, S. (2018). Statistical inference for spatial statistics defined in the Fourier domain. *Ann. Statist.* **46** 469–499. MR3782374 https://doi.org/10.1214/17-AOS1556

[33] Van Hala, M., Bandyopadhyay, S., Lahiri, S.N. and Nordman, D.J. (2017). On the non-standard distribution of empirical likelihood estimators with spatial data. *J. Statist. Plann. Inference* **187** 109–114. MR3638046 https://doi.org/10.1016/j.jspi.2017.02.007

[34] Van Hala, M., Bandyopadhyay, S., Lahiri, S.N. and Nordman, D.J. (2019). Supplement to "A general frequency domain method for assessing spatial covariance structures." https://doi.org/10.3150/19-BEJ1160SUPPA, https://doi.org/10.3150/19-BEJ1160SUPPB.

[35] Van Hala, M., Nordman, D.J. and Zhu, Z. (2015). Empirical likelihood for irregularly located spatial data. *Statist. Sinica* **25** 1399–1420. MR3409073

[36] Whittle, P. (1954). On stationary processes in the plane. *Biometrika* **41** 434–449. MR0067450 https://doi.org/10.1093/biomet/41.3-4.434

[37] Yaglom, A.M. (1957). Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory Probab. Appl.* **2** 273–320.

[38] Yaglom, A.M. (1987). *Correlation Theory of Stationary and Related Random Functions. Vol. I: Basic Results. Springer Series in Statistics*. New York: Springer. MR0893393

Received December 2018 and revised June 2019