A test for stationarity for irregularly spaced spatial data

Soutir Bandyopadhyay and Suhasini Subba Rao
Department of Mathematics, Lehigh University
Bethlehem, PA, U.S.A.
Department of Statistics, Texas A&M University
College Station, TX, U.S.A.
sob210@lehigh.edu and suhasini@stat.tamu.edu

June 29, 2015

Abstract

The analysis of spatial data is based on a set of assumptions, which in practice need to be checked. A commonly used assumption is that the spatial random field is second order stationary. In this paper, a test for spatial stationarity for irregularly sampled data is proposed. The test is based on a transformation of the data (a type of Fourier transform), where the correlations between the transformed data is close to zero if the random field is second order stationary. On the other hand, if the random field were second order nonstationary, this property does not hold. Using this property a test for second order stationarity is constructed. The test statistic is based on measuring the degree of correlation in the transformed data. The asymptotic sampling properties of the test statistic is derived under both stationarity and nonstationarity of the random field. These results motivate a graphical tool which allows a visual representation of the nonstationary features. The method is illustrated with simulations and a real data example.

Key words and phrases: Fourier Transforms, Irregular sampling, Nonstationary, Stationary random fields.

1 Introduction

An important problem in spatial statistics is to estimate the underlying covariance function of the spatial random field (see, for example, Cressie [1993], Stein [1999] and Sherman

[2011]). A common assumption that is used is that the spatial random field $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is second order stationary, in the sense that $\mathrm{E}[Z(\boldsymbol{s})] = \mu$ and $\mathrm{cov}[Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)] = c(\boldsymbol{s}_1 - \boldsymbol{s}_2)$ for all $\boldsymbol{s}, \boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathbb{R}^d$. An advantage of imposing second order stationarity is that it greatly simplifies the statistical analysis. However, if this assumption does not hold, then either the covariance function or the mean function is misspecified. This can, for example, lead to inaccurate predictions. Therefore, there is a need to test for second order stationarity of the spatial random field.

To address this issue, Fuentes [2005] extends the test for stationarity of a time series, proposed in Priestley and Subba Rao [1969], to the spatial random field. However, this test is based on testing for spatial random fields which are defined on the regular grid $\mathbb{Z}^d$, and it is unclear how to extend the method to irregularly spaced data. In a recent paper, Jun and Genton [2012] proposed a test for second order stationarity for irregularly spaced spatial data. Their method is based on partitioning the spatial domain in two non-intersecting domains, estimating the autocovariance function at various lags in both domains and then comparing the two autocovariance functions. In common with many tests for stationarity of a time series, the results of the Jun-Genton test depend on the partition of the spatial domain that is used.

In this paper, we propose an alternative approach to test for stationarity based on a characterization of second order stationarity. More precisely, suppose $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is a spatial random process, observed only at a finite number of locations, denoted as $\{\boldsymbol{s}_j\}$, in the region $[-\lambda/2, \lambda/2]^d$, and we observe $\{(\boldsymbol{s}_j, Z(\boldsymbol{s}_j)); j = 1, \ldots, n\}$. We assume that the locations are independent, uniformly distributed random variables. To motivate our approach we recall that in the discrete time series setting if the time series is second order stationary, then the discrete Fourier transforms (DFT) of the time series, defined as $\frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X_t \exp(it\omega)$ (where $\{X_t\}$ is the observed time series) are asymptotically uncorrelated at the fundamental frequencies $2\pi k/T$. Dwivedi and Subba Rao [2011] and Jentsch and Subba Rao [2015] exploit this property to test for second order stationarity of a time series by testing for uncorrelatedness of the DFTs using a Portmanteau type test statistic. It is clear that this type of test can be adapted to the spatial setting on the grid $\mathbb{Z}^d$. However, as far as we are aware, there does not exist an analogous result for irregularly spaced data. In Section 2 we bridge this gap. To define the Fourier transform we follow the approach of Matsuda and Yajima [2009] and Bandyopadhyay and Lahiri [2009], and replace t in the classical DFT with $\boldsymbol{s}_j$ and $T^{-1/2}$ with $\lambda^{d/2}/n$ to give

$$J_n(\boldsymbol{\omega}) = \frac{\lambda^{d/2}}{n} \sum_{j=1}^{n} Z(\boldsymbol{s}_j) \exp(i\boldsymbol{s}_j'\boldsymbol{\omega}), \qquad \boldsymbol{\omega} \in \mathbb{R}^d.$$

We show that if the spatial random field is second order stationary, then the sequence

$\{J_n(\boldsymbol{\omega_k})\}$ is 'near uncorrelated' at the frequencies $\boldsymbol{\omega_k} = (\frac{2\pi k_1}{\lambda}, \ldots, \frac{2\pi k_d}{\lambda})'$, where $\boldsymbol{k} \in \mathbb{Z}^d$. We call $J_n(\boldsymbol{\omega})$, defined at these frequencies, the DFT. In contrast, we show that if the random field is covariance nonstationary, the 'near uncorrelatedness' property between the DFTs does not hold. We exploit these differing properties to define a test for second order stationarity. In Section 3 we define the weighted DFT covariance function which measures the second order dependence between the DFTs. We derive the asymptotic sampling properties of the weighted DFT under the null hypothesis of stationarity and we use this to define the test statistic. In Section 4 the asymptotic sampling properties of the test statistic, under the alternative of nonstationarity is considered. In order to study the power of the test statistic under the alternative of covariance nonstationarity, we define a form of rescaled asymptotic for spatial random fields defined on $\mathbb{R}^d$ and use this construction to obtain an asymptotic expression for the weighted DFT covariance. We use these results to determine which tuning parameters may give power to the test. In Section 6 we discuss the selection of tuning parameters. In Section 7.1 we present an extensive simulation study. Finally, in Section 7.2 we apply the proposed test to test for spatial stationarity of ground ozone from April 1st - September 30th, 2014 in the Ohio Central Valley and South East of the USA.

The R software for implementing the test can be found on the authors' website. Some additional materials and all of the proofs can be found in the supplementary material.

2 Motivation

In this section we define the frequencies where the Fourier transforms are near uncorrelated under covariance stationarity. In the discrete time series setting if the time series is stationary, near uncorrelatedness occurs at the Fourier frequencies $\omega_k = \frac{2\pi k}{T}$ (where T denotes the length of the time series). In the theorem below, we show that if the spatial random field is irregularly sampled, the frequencies where 'near uncorrelation' occurs depends on the size of the spatial domain.

Theorem 2.1 *Let us suppose that $\{\boldsymbol{s}_j : j = 1, \cdots, n\}$ are independent, uniformly distributed random variables. In addition, define the fundamental frequencies $\boldsymbol{\omega_k} = (\frac{2\pi k_1}{\lambda}, \ldots, \frac{2\pi k_d}{\lambda})'$. We define the function $\beta_{1+\delta}(\cdot)$, for some $\delta > 0$, such that $\beta_{1+\delta}(s) \leq C$ for $|s| \in [-1, 1]$, and $|\beta_{1+\delta}(s)| \leq C|s|^{-(2+\delta)}$ for $|s| > 1$, and constant $C > 0$. Define $J_n^{\mu}(\boldsymbol{\omega})$ in the same way as $J_n(\boldsymbol{\omega})$, but with $Z(\boldsymbol{s}) - \mu$ replacing $Z(\boldsymbol{s})$.*

(i) Suppose that $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is second order stationary with $\mathrm{E}[Z(\boldsymbol{s})|\boldsymbol{s}] = \mu$ and $\mathrm{cov}[Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)|\boldsymbol{s}_1, \boldsymbol{s}_2] = c(\boldsymbol{s}_1 - \boldsymbol{s}_2)$ for all $\boldsymbol{s}, \boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathbb{R}^d$. We assume that $|c(\boldsymbol{s})| \leq \prod_{j=1}^d \beta_{1+\delta}(s_j)$ for all $\boldsymbol{s} \in \mathbb{R}^d$. Further suppose that the spectral density of the process $Z(\cdot)$ is given by $f(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} c(\boldsymbol{s}) \exp(i\boldsymbol{s}'\boldsymbol{\omega}) d\boldsymbol{\omega}$. Then,

(a) *For* $\boldsymbol{k} \neq 0$, *we have* $\mathrm{E}[J_n(\boldsymbol{\omega}_{\boldsymbol{k}})] = 0$.

(b) *If* $\boldsymbol{k}_1 = \boldsymbol{k}_2 = \boldsymbol{0}$, *then* $\mathrm{var}\big[J_n(\boldsymbol{\omega}_0)\big] = \mathrm{var}\big[J_n^{\mu}(\boldsymbol{\omega}_0)\big]$ *and* $\mathrm{E}\big[J_n(\boldsymbol{\omega}_0)^2\big] = \mathrm{var}\big[J_n^{\mu}(\boldsymbol{\omega}_0)\big] + \lambda^d\mu^2$.

(c) *If* $\boldsymbol{k}_1 \neq \boldsymbol{0}$ *or* $\boldsymbol{k}_2 \neq \boldsymbol{0}$, *then* $\mathrm{cov}[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})] = \mathrm{E}[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1})\overline{J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})}]$. *In addition, if* $\boldsymbol{k}_1 \neq \boldsymbol{k}_2$, *then* $\mathrm{cov}\big[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})\big] = \mathrm{cov}\big[J_n^{\mu}(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n^{\mu}(\boldsymbol{\omega}_{\boldsymbol{k}_2})\big]$ *or if* $\boldsymbol{k}_1 = \boldsymbol{k}_2$, *then* $\mathrm{var}\big[J_n(\boldsymbol{\omega}_{\boldsymbol{k}})\big] = \mathrm{var}\big[J_n^{\mu}(\boldsymbol{\omega}_{\boldsymbol{k}})\big] + \frac{\mu^2\lambda^d}{n}$.

Using the above we have,

$$\mathrm{cov}\big[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})\big] = \begin{cases} f(\boldsymbol{\omega}_k) + O(\frac{1}{\lambda} + \frac{\lambda^d}{n}) & \boldsymbol{k}_1 = \boldsymbol{k}_2 (= \boldsymbol{k}), \\ O(\frac{1}{\lambda^{d-b}}) & b \text{ elements of the vectors } \boldsymbol{k}_1 \text{ and } \boldsymbol{k}_2 \text{ are the same.} \end{cases}$$

(ii) *Suppose that* $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ *is second order nonstationary with a constant mean and* $\mathrm{cov}\,[Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)|\boldsymbol{s}_1, \boldsymbol{s}_2] = c(\boldsymbol{s}_1, \boldsymbol{s}_2) = \kappa_{\boldsymbol{s}_2}(\boldsymbol{s}_1 - \boldsymbol{s}_2)$. *Let* $f(\boldsymbol{\omega}; \boldsymbol{s}) = \int_{\mathbb{R}^d} \kappa_{\boldsymbol{s}}(\boldsymbol{u}) \exp(i\boldsymbol{u}'\boldsymbol{\omega}) d\boldsymbol{u}$. *We assume that* $\sup_{\boldsymbol{s}} |\kappa_{\boldsymbol{s}}(\boldsymbol{u})| \leq \prod_{j=1}^d \beta_{1+\delta}(u_j)$. *Then we have,*

$$\begin{aligned} \mathrm{cov}\big[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})\big] &= \frac{1}{\lambda^d} \int_{[-\lambda/2,\lambda/2]^d} f(\boldsymbol{\omega}_{\boldsymbol{k}_1}; \boldsymbol{s}) \exp\left(i\frac{2\pi}{\lambda}\boldsymbol{s}'(\boldsymbol{k}_1 - \boldsymbol{k}_2)\right) d\boldsymbol{s} \\ &\quad + \frac{1}{n} \int_{[-\lambda/2,\lambda/2]^d} \kappa_{\boldsymbol{s}}(\boldsymbol{0}) \exp\left(i\frac{2\pi}{\lambda}\boldsymbol{s}'(\boldsymbol{k}_1 - \boldsymbol{k}_2)\right) d\boldsymbol{s} + O(\frac{1}{\lambda}). \end{aligned}$$

The theorem above raises some issues, that we now discuss.

Stationary vs. nonstationary

We see that depending on whether the spatial random field is second order stationary or not, the covariance between the Fourier transforms could behave in very different ways. In particular, if the spatial random field is second order stationary, the Fourier transform $J_n(\boldsymbol{\omega})$ at the frequencies $\boldsymbol{\omega}_{\boldsymbol{k}} = 2\pi\boldsymbol{k}/\lambda$ are near uncorrelated. The situation is different in the nonstationary set-up. To understand why, we observe that $\mathrm{cov}\big[J_n(\boldsymbol{\omega}_{\boldsymbol{k}_1}), J_n(\boldsymbol{\omega}_{\boldsymbol{k}_2})\big] \approx h(\boldsymbol{\omega}_{k_1}, \boldsymbol{k}_1 - \boldsymbol{k}_2)$ where

$$h(\boldsymbol{\omega}; \boldsymbol{r}) = \frac{1}{\lambda^d} \int_{[-\lambda/2,\lambda/2]^d} f(\boldsymbol{\omega}; \boldsymbol{s}) \exp(i\frac{2\pi}{\lambda}\boldsymbol{s}'\boldsymbol{r}) d\boldsymbol{s}.$$

Using basic results in Fourier analysis, it is clear that $h(\boldsymbol{\omega}; \boldsymbol{r}) = 0$, for all $\boldsymbol{r} \neq \boldsymbol{0}$ and all $\boldsymbol{\omega}$ if and only if $f(\boldsymbol{\omega}; \boldsymbol{s}) = f(\boldsymbol{\omega})$, in other words the spatial process is spatially second order stationary. Furthermore, since $f(\boldsymbol{\omega}; \cdot) \in L_2[-\lambda/2, \lambda/2]^d$, the magnitude of the Fourier

transform $\frac{1}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^d} f(\boldsymbol{\omega}_{\boldsymbol{k}_1};\boldsymbol{s})\exp[i\frac{2\pi}{\lambda}\boldsymbol{s}'(\boldsymbol{k}_1-\boldsymbol{k}_2)]d\boldsymbol{s}$ will decrease to zero as $\|\boldsymbol{k}_1-\boldsymbol{k}_2\|_1\to\infty$ (where $\|\cdot\|_1$ denotes the ℓ_1 norm of a vector). In other words, the correlations between the DFTs decrease the further apart the frequencies.

Asymptotic set-up

From the above theorem we observe that in the stationary case the DFTs become more uncorrelated as $\lambda\to\infty$ (as the spatial domain grows). In addition, we observe that the variance of the DFT converges to the spectral density as $\lambda\to\infty$ and $\lambda^d/n\to 0$ (i.e., as the spatial domain grows the number of observations should become denser on the spatial domain). This is the mixed (a combination of increasing domain and infill) spatial asymptotics set-up described in Hall and Patil [1994], Lahiri [2003], Matsuda and Yajima [2009], Bandyopadhyay and Lahiri [2009], and Bandyopadhyay et al. [2015]. In this paper, the results will be derived under this asymptotic set-up.

3 The test statistic

In this section we will develop the test statistic that will allow us to test the hypothesis H_0 : The spatial random field is second order stationary, against the alternative H_A : The spatial random field is second order nonstationary.

Summarizing Theorem 2.1, we observe that under second order stationarity, the Fourier transforms $\{J_n(\boldsymbol{\omega}_{\boldsymbol{k}})\}_{\boldsymbol{k}}$ are close to uncorrelated. Therefore, in some respects $\{J_n(\boldsymbol{\omega}_{\boldsymbol{k}})\}_{\boldsymbol{k}}$ can be considered as a near uncorrelated frequency series. To measure the degree of correlation between the DFTs at lag $\boldsymbol{r}$ we define the weighted DFT covariance as

$$\begin{aligned} a_\lambda(g;\boldsymbol{r}) &= \frac{1}{\lambda^d}\sum_{k_1,\ldots,k_d=-a}^{a} g(\boldsymbol{\omega}_{\boldsymbol{k}})J_n(\boldsymbol{\omega}_{\boldsymbol{k}})\overline{J_n(\boldsymbol{\omega}_{\boldsymbol{k}+\boldsymbol{r}})} \\ &\quad -\left[\frac{1}{n}\sum_{k_1,\cdots,k_d=-a}^{a} g(\boldsymbol{\omega}_{\boldsymbol{k}})\times\frac{1}{n}\sum_{j=1}^{n} Z(\boldsymbol{s}_j)^2\exp(-i\boldsymbol{s}_j'\boldsymbol{\omega}_{\boldsymbol{r}})\right], \end{aligned} \tag{1}$$

where the limits of the sum satisfy $a^d=O(n)$, $\boldsymbol{k}=(k_1,\ldots,k_d)'$ and g is a given Lipschitz continuous function with $\sup_{\boldsymbol{\omega}\in\mathbb{R}^d}|g(\boldsymbol{\omega})|<\infty$. In the definition of (1) we removed the nugget term to reduce the influence of the variance (which may be problematic in the case that the observations have been corrupted by independent, but possibly heterogeneous noise). We note that Masry [1978] and Matsuda and Yajima [2009] removed a similar term in their definition of the spectral density estimator. Examples of weight functions are discussed in Sections 4 and 6. In what follows we assume that the random field has a zero mean. However,

the analysis below also holds for spatial random fields which have been de-trended with a parametric mean.

The asymptotic sampling properties of $a_\lambda(g;\boldsymbol{r})$ are derived using the following assumptions.

Assumption 3.1 *(i)* $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ *is a second order stationary random field.*

(ii) $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ *is a Gaussian random field.*

(iii) $f(\cdot)$ *satisfies* $\int_{\mathbb{R}^d} f(\boldsymbol{\omega})d\boldsymbol{\omega} < \infty$ *and* $\int_{\mathbb{R}^d} f^2(\boldsymbol{\omega})d\boldsymbol{\omega} < \infty$.

(iv) For some $\delta > 0$, *define the monotonic function* $\beta_\delta : \mathbb{R} \to \mathbb{R}$ *where for some constant* C, $\beta_\delta(s) = C$ *if* $s \in [-1, 1]$ *and* $|\beta_\delta(s)| \leq C|s|^{-(1+\delta)}$ *if* $|s| > 1$. *Let* $\beta_\delta(\boldsymbol{\omega}) = \prod_{j=1}^d \beta_\delta(\omega_j)$. *Then*

(a) $f(\boldsymbol{\omega}) \leq \beta_\delta(\boldsymbol{\omega})$.

(b) For all $1 \leq j \leq d$, *the partial derivatives satisfy* $|\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_j}| \leq \beta_\delta(\boldsymbol{\omega})$.

(c) $|\frac{\partial^d f(\boldsymbol{\omega})}{\partial \omega_1 \ldots \partial \omega_d}| \leq \beta_\delta(\boldsymbol{\omega})$.

Remark 3.1 *Assumptions 3.1(iii,iv) appear quite technical, but it is satisfied by a wide range of spatial covariance functions. For example, we now show that the exponential covariance (which belongs to the Matérn class), defined by* $c(\|\boldsymbol{s}\|_2) = \phi \exp(-\|\boldsymbol{s}\|_2)$ *(where* $\|\cdot\|_2$ *denotes the Euclidean distance) satisfies these assumptions. To see why, we consider the case* $d = 1$ *and* $d = 2$. *For* $d = 1$, *the spectral density function of the exponential covariance is*

$$f(\omega) = \frac{\phi}{1 + \omega^2},$$

whereas for $d = 2$, *the exponential covariance has the spectral density*

$$f(\omega_1, \omega_2) = \frac{2\pi\phi}{(1 + \omega_1^2 + \omega_2^2)^{3/2}}.$$

It is straightforward to show that these spectral density functions satisfy Assumptions 3.1(iii,iv).

Using the above set of assumptions we obtain the following two theorems.

Theorem 3.1 *Suppose that Assumption 3.1(i,iii,iv(a,b,c)) holds. Then for* $\boldsymbol{r} \in \mathbb{Z}^d/\{\boldsymbol{0}\}$, *where* $\{m_1, \ldots, m_{d-b}\}$ *is the subset of non-zero values in* $\boldsymbol{r} = (r_1, \ldots, r_d)'$, *we have*

$$\mathrm{E}\left[a_\lambda(g;\boldsymbol{r})\right] = \begin{cases} O\left(\frac{1}{\lambda^{d-b}} \prod_{j=1}^{d-b} \left(\log \lambda + \log |m_j|\right)\right), & \boldsymbol{r} \in \mathbf{Z}^d/\{\boldsymbol{0}\} \\ \frac{1}{(2\pi)^d} \int_{\boldsymbol{\omega} \in \mathbb{R}^d} f(\boldsymbol{\omega}) g(\boldsymbol{\omega}) d\boldsymbol{\omega} + O\left(\frac{\log \lambda + \log \|\boldsymbol{r}\|_1}{\lambda} + \frac{1}{n}\right), & \boldsymbol{r} = \boldsymbol{0}. \end{cases} \tag{2}$$

where $a/\lambda \to \infty$ *as* $a \to \infty$ *and* $\lambda \to \infty$.

Theorem 3.1 shows that if the process is second order stationary, then $a_\lambda(g;\boldsymbol{r})$ is an asymptotically unbiased estimator of zero.

In the following theorem we obtain an expression for the asymptotic variance of $a_\lambda(g;\boldsymbol{r})$, which together with Theorem 3.1, show that under stationarity, $a_\lambda(g;\boldsymbol{r})$ converges in probability to zero.

Theorem 3.2 *Suppose Assumption 3.1(i,iii,iv(a,b)) holds and* $\boldsymbol{r}_1, \boldsymbol{r}_2 \in \mathbb{Z}^d/\{\boldsymbol{0}\}$. *Then we have*

$$\lambda^d \mathrm{cov}\left[\Re a_\lambda(g;\boldsymbol{r}_1), \Re a_\lambda(g;\boldsymbol{r}_2)\right] = \begin{cases} c_{\lambda,1} + O\left(\ell_{\lambda,a,n} + \frac{\|\boldsymbol{r}\|_1}{\lambda}\right) & \boldsymbol{r}_1 = \boldsymbol{r}_2 (= \boldsymbol{r}) \\ \Re c_{\lambda,2} + O\left(\ell_{\lambda,a,n} + \frac{\|\boldsymbol{r}\|_1}{\lambda}\right) & \boldsymbol{r}_1 = -\boldsymbol{r}_2 (= \boldsymbol{r}), \\ O(\ell_{\lambda,a,n}) & \textit{otherwise} \end{cases}$$

$$\lambda^d \mathrm{cov}\left[\Im a_\lambda(g;\boldsymbol{r}_1), \Im a_\lambda(g;\boldsymbol{r}_2)\right] = \begin{cases} c_{\lambda,1} + O\left(\ell_{\lambda,a,n} + \frac{\|\boldsymbol{r}\|_1}{\lambda}\right) & \boldsymbol{r}_1 = \boldsymbol{r}_2 (= \boldsymbol{r}) \\ -\Re c_{\lambda,2}(\boldsymbol{r}) + O\left(\ell_{\lambda,a,n} + \frac{\|\boldsymbol{r}\|_1}{\lambda}\right) & \boldsymbol{r}_1 = -\boldsymbol{r}_2 (= \boldsymbol{r}), \\ O(\ell_{\lambda,a,n}) & \textit{otherwise} \end{cases}$$

$$\lambda^d \mathrm{cov}\left[\Re a_\lambda(g;\boldsymbol{r}_1), \Im a_\lambda(g;\boldsymbol{r}_2)\right] = \begin{cases} \Im c_{\lambda,2} + O\left(\ell_{\lambda,a,n} + \frac{\|\boldsymbol{r}\|_1}{\lambda}\right) & \boldsymbol{r}_1 = -\boldsymbol{r}_2 (= \boldsymbol{r}) \\ O(\ell_{\lambda,a,n}) & \boldsymbol{r}_1 \neq -\boldsymbol{r}_2 \end{cases},$$

where

$$c_{\lambda,j} = \frac{1}{2(2\pi)^d} \int_{2\pi[-a/\lambda, a/\lambda]^d} f^2(\boldsymbol{\omega}) h_j(\boldsymbol{\omega}) d\boldsymbol{\omega} \text{ and } h_j(\boldsymbol{\omega}) = \begin{cases} |g(\boldsymbol{\omega})|^2 + g(\boldsymbol{\omega})\overline{g(-\boldsymbol{\omega})} & j = 1 \\ g(\boldsymbol{\omega})g(-\boldsymbol{\omega}) + g(\boldsymbol{\omega})^2 & j = 2 \end{cases}, \tag{3}$$

with

$$\ell_{\lambda,a,n} = \log^2 a \left(\frac{\log a + \log \lambda}{\lambda}\right) + \frac{\lambda^d}{n}.$$

We now show the asymptotic normality of $a_\lambda(g;\boldsymbol{r})$.

Theorem 3.3 *Suppose Assumption 3.1(i,ii,iii,iv(a,b)) holds. In addition, suppose that* m *is finite,* $\boldsymbol{r}_j \in \mathbb{Z}^d/\{\boldsymbol{0}\}$ *are fixed and are such that* $\boldsymbol{r}_i \neq \boldsymbol{r}_j$ *or* $-\boldsymbol{r}_j$ *for all* $1 \leq i, j \leq m$. *Then we have*

$$c_{\lambda,1}^{-1/2} \lambda^{d/2} \left(\Re a_\lambda(g;\boldsymbol{r}_1), \Im a_\lambda(g;\boldsymbol{r}_1), \ldots, \Re a_\lambda(g;\boldsymbol{r}_m), \Im a_\lambda(g;\boldsymbol{r}_m) \right)' \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{0}, I_{2m})$$

with $\ell_{\lambda,a,n} \to 0$, $\lambda^d/[n\log^{2d} a] \to 0$ *as* $\lambda, a, n \to \infty$. I_{2m} *denotes the* $2m$*-dimensional identity matrix.*

Theorem 3.3 allows us to check for correlations between the DFTs. We start by selecting a finite set $\mathcal{S} \in \mathbb{Z}^d/\{\mathbf{0}\}$, which surrounds the origin (see Figure 2a, for an example). To measure the correlations between the DFTs we define the test statistic

$$\mathcal{T}_{\mathcal{S}} = \frac{\lambda^d \max_{\boldsymbol{r}\in\mathcal{S}} |a_\lambda(g;\boldsymbol{r})|^2}{c_{\lambda,1}}.$$

Using Theorem 3.3, asymptotically, under the null, the distribution of $\mathcal{T}_{\mathcal{S}}$ converges to $F(x) = [1-\exp(-x/2)]^{|\mathcal{S}|}$, where $|\mathcal{S}|$ denotes the cardinality of the set $\mathcal{S}$. However, $c_{\lambda,1}$ is unknown, therefore we need to define an estimator of it. Using Theorem 3.2 we see that when $\boldsymbol{r}$ is is in a neighbourhood of zero, the coefficients $\{a_\lambda(g;\boldsymbol{r})\}$ are almost uncorrelated and with similar variance. Therefore we define a set $\mathcal{S}' \in \mathbb{Z}^d/\{\mathbf{0}\}$ which is close to zero and such that $\mathcal{S} \cap \mathcal{S}' = \varnothing$ (see Figure 2a, for an example). We estimate $c_{\lambda,1}$ with

$$\widehat{c}_\lambda(\mathcal{S}') = \frac{\lambda^d}{(2|\mathcal{S}'|-1)} \sum_{\boldsymbol{r}\in\mathcal{S}'} \left([\Re a_\lambda(g;\boldsymbol{r}) - \bar{a}]^2 + [\Im a_\lambda(g;\boldsymbol{r}) - \bar{a}]^2 \right), \tag{4}$$

where $\bar{a} = \frac{1}{2|\mathcal{S}'|}\sum_{\boldsymbol{r}\in\mathcal{S}'}[\Re a_\lambda(g;\boldsymbol{r}) + \Im a_\lambda(g;\boldsymbol{r})]$. Replacing $c_{\lambda,1}$ in $\mathcal{T}_{\mathcal{S}}$ with $\widehat{c}_\lambda(\mathcal{S}')$ and using Theorem 3.3 gives

$$\mathcal{T}_{\mathcal{S},\mathcal{S}'} = \frac{\lambda^d \max_{\boldsymbol{r}\in\mathcal{S}} |a_\lambda(g;\boldsymbol{r})|^2}{\widehat{c}_\lambda(\mathcal{S}')} \xrightarrow{\mathcal{D}} \frac{\max_{1\le i\le|\mathcal{S}|}(Z_{2i-1}^2 + Z_{2i}^2)}{\frac{1}{2|\mathcal{S}'|-1}\sum_{j=2|\mathcal{S}|+1}^{2(|\mathcal{S}|+|\mathcal{S}'|)}(Z_j - \bar{Z})^2}, \tag{5}$$

where $\{Z_j; 1 \le j \le 2(|\mathcal{S}| + |\mathcal{S}'|)\}$ are iid Gaussian random variables and $A_\lambda \xrightarrow{\mathcal{D}} B$ means that asymptotically the distribution of A_λ converges to the distribution of B.

In the following lemma we show that for a set $\mathcal{S}'$ that grows at a sufficient rate, $\widehat{c}_\lambda(\mathcal{S}') \xrightarrow{\mathcal{P}} c_{\lambda,1}$. One implication of this is that when $\mathcal{S}'$ is sufficiently large, the distributions of $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ and $\mathcal{T}_{\mathcal{S}}$ are asymptotically equivalent.

Lemma 3.1 *Suppose the set* $\mathcal{S}' \in \mathbb{Z}^d/\{0\}$ *is such that if* $\boldsymbol{r} \in \mathcal{S}'$ *then* $-\boldsymbol{r} \notin \mathcal{S}'$ *and the vector* $\boldsymbol{r}$ *contains at most* $\lfloor d/2 \rfloor$*-zeros (*$\mathcal{S}_0 \subset \mathcal{S}'$ *consisting of vectors which contain zeros). Then we have*

$$\mathrm{E}\left[\widehat{c}_\lambda(\mathcal{S}') - c_{\lambda,1}\right]^2 = O\left(\left[\left[\frac{1}{|\mathcal{S}'|} \sum_{\boldsymbol{r}\in\mathcal{S}'} \frac{\|\boldsymbol{r}\|_1}{\lambda} \right]^2 + \frac{|\mathcal{S}_0|}{|\mathcal{S}'|} + \frac{1}{\lambda^d} \right]^2 + \frac{1}{|\mathcal{S}'|} + \ell_{\lambda,a,n} \right).$$

Using the above lemma we can determine how the set $\mathcal{S}'$ should grow such that $\widehat{c}_\lambda(\mathcal{S}')$ is

mean squared consistent as $\lambda \to \infty$. Define the semi-norm $\|\boldsymbol{r}\|_{\max} = \max(|r_i|; 1 \leq i \leq d)$ and let $|\mathcal{S}'|_{\max} = \max_{\boldsymbol{r}\in\mathcal{S}'}(\|\boldsymbol{r}\|_{\max})$. Using these definitions we construct a set $\mathcal{S}'$ such that $|\mathcal{S}'|_{\max} = \lambda^{1-\delta}$, where $0 < \delta < 1$ and suppose $\mathcal{S}'$ expands at the rate $|\mathcal{S}'| = O(\lambda^{d(1-\delta)})$. Then by using the above lemma we see that

$$\mathrm{E}\left[\widehat{c}_\lambda(\mathcal{S}') - c_{\lambda,1}\right]^2 = O\left(\frac{1}{\lambda^{2\delta}} + \frac{1}{\lambda^{d(1-\delta)}} + \ell_{\lambda,a,n}\right),$$

which leads to a mean squared consistent estimator of $c_{\lambda,1}$.

Remark 3.2 *An important assumption that we have made is that the spatial locations are uniformly distributed. This assumption is relatively standard for nonparametric methods in spatial statistics. For example, Guan et al. [2004], Li et al. [2008] and Jun and Genton [2012] develop tests for spatial data and model the irregularly spaced locations using a homogeneous Poisson point process (which is equivalent to uniform sampling).*

Non-uniform locations can induce correlations between the DFTs (see, Bandyopadhyay and Lahiri [2009] and Subba Rao [2014]), which may lead us to false rejections of the null. In the case the spatial locations are highly non-uniform it is possible to induce the 'near uncorrelated' property (for spatially stationary random fields) by dividing the spatial observations by the density of locations and use $X(\boldsymbol{s}) = Z(\boldsymbol{s})/h(\boldsymbol{s})$ (where $h(\boldsymbol{s})$ is the density of locations) as the 'observations'. Of course, in practice the density $h(\cdot)$ is unlikely to be known and will have to be estimated from the data. The simulations in Section 7.1 suggest that our method is relatively robust to deviations from uniformity, therefore we have not pursued the above method in this paper. Nonetheless, we suggest checking the influence of locations on the outcome of the test by using the method outlined at the end of Section 7.2.

In order to simplify notations we have assumed the observations are uniformly distributed on a d-dimensional cube. The same results apply for d-dimensional hyper-rectangles of the type $\prod_{j=1}^d [-\lambda_j/2, \lambda_j/2]$, with the obvious adjustment of the fundamental frequencies to $\boldsymbol{\omega_k} = (\frac{2\pi k_1}{\lambda_1}, \ldots, \frac{2\pi k_d}{\lambda_d})'$ (see Section 7.2).

Remark 3.3 *In the case that the locations $\{\boldsymbol{s}_j; j = 1, \ldots, n\}$ are not that dense on $[-\lambda/2, \lambda/2]^d$, $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ is not particularly sensitive to changes in the covariance. In such cases the method can be modified to simply detect changes in the spatial variance. We do this by using Theorem 2.1 and adapting the DFT $J_n(\boldsymbol{\omega}_k)$ to detect changes in the spatial variation. Specifically, let*

$$v_\lambda(\boldsymbol{r}) \;=\; \frac{1}{n}\sum_{j=1}^n |Z(\boldsymbol{s}_j)| \exp(-i\boldsymbol{s}_j'\boldsymbol{\omega}_{\boldsymbol{r}}).$$

The results in Theorem 2.1 can be used to obtain the sampling properties of $v_\lambda(\boldsymbol{r})$ under the null and alternative hypotheses. Details can be found in the supplementary material.

The test for non-Gaussian stationary random fields

We now show that similar results hold in the case that the spatial random field is stationary but not necessarily Gaussian. The main difference is that the asymptotic variance of $a_\lambda(g;\boldsymbol{r})$ contains an additional term. Here we give a summary of the result, which requires additional assumptions and definitions. Suppose Assumption 3.1(i,iii,iv) are satisfied. By relaxing the Gaussianity assumption, additional terms are involved and we define the tri-spectral density as

$$f_4(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\boldsymbol{\omega}_3)=\int_{\mathbb{R}^{3d}}\kappa_4(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3)\exp(-i\sum_{j=1}^{3}\boldsymbol{s}_j'\boldsymbol{\omega}_j)d\boldsymbol{s}_1d\boldsymbol{s}_2d\boldsymbol{s}_3,$$

where $\kappa_4(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3)=\mathrm{cum}[Z(\boldsymbol{0}),Z(\boldsymbol{s}_1),Z(\boldsymbol{s}_2),Z(\boldsymbol{s}_3)]$. We assume that the spatial tri-spectral density function is such that $|f_4(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\boldsymbol{\omega}_3)|<\prod_{j=1}^{3d}\beta_\delta(\omega_j)$ and for $1\leq i\leq 3d$, $|\frac{\partial f_4(\omega_1,\ldots,\omega_{3d})}{\partial\omega_i}|\leq\prod_{j=1}^{3d}\beta_\delta(\omega_j)$.

Under these conditions, we use Subba Rao [2014], Theorem 3.6, to show an analogous version of Theorem 3.2, that is

$$\lambda^d\mathrm{cov}\left[\Re a_\lambda(g;\boldsymbol{r}_1),\Re a_\lambda(g;\boldsymbol{r}_2)\right]=\begin{cases}c_{\lambda,1}+d_{\lambda,1}+O(\ell^{(2)}_{\lambda,a,n}+\frac{\|\boldsymbol{r}\|_1}{\lambda}) & \boldsymbol{r}_1=\boldsymbol{r}_2(=\boldsymbol{r})\\ \Re[c_{\lambda,2}+d_{\lambda,2}]+O(\ell^{(2)}_{\lambda,a,n}+\frac{\|\boldsymbol{r}\|_1}{\lambda}) & \boldsymbol{r}_1=-\boldsymbol{r}_2(=\boldsymbol{r}),\\ O(\ell^{(2)}_{\lambda,a,n}) & \text{otherwise}\end{cases}$$

where $c_{\lambda,1}$ and $c_{\lambda,2}$ are defined in Theorem 3.2,

$$d_{\lambda,j}\;=\;\frac{1}{2(2\pi)^{2d}}\int_{2\pi[-a/\lambda,a/\lambda]^{2d}}h_{2,j}(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)f_4(-\boldsymbol{\omega}_1,-\boldsymbol{\omega}_2,\boldsymbol{\omega}_2)d\boldsymbol{\omega}_1d\boldsymbol{\omega}_2,$$

$$h_{2,j}(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)=\begin{cases}g(\boldsymbol{\omega}_1)\overline{g(\boldsymbol{\omega}_2)} & j=1\\ g(\boldsymbol{\omega}_1)g(\boldsymbol{\omega}_2) & j=2\end{cases}\quad\text{and}\quad\ell^{(2)}_{\lambda,a,n}=\ell_{\lambda,a,n}+\frac{a^d\lambda^d}{n^2}+\frac{\log^3\lambda}{\lambda}.$$

Analogous results hold for $\lambda^d\mathrm{cov}\left[\Im a_\lambda(g;\boldsymbol{r}_1),\Im a_\lambda(g;\boldsymbol{r}_2)\right]$ and $\lambda^d\mathrm{cov}\left[\Re a_\lambda(g;\boldsymbol{r}_1),\Im a_\lambda(g;\boldsymbol{r}_2)\right]$.

In addition, with sufficient mixing conditions on the spatial random field Theorem 2.1 holds. The only difference is that $c_{\lambda,1}+d_{\lambda,1}$ replaces $c_{\lambda,1}$. These results imply that the same test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ as defined in (5), can be used to test for stationarity. Furthermore, a version of Lemma 3.1 can also be shown. More precisely, for $3\leq m\leq 8$, we define the m-th order spatial spectral density $f_m(\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_{m-1})$, in the same way as the tri-spectral density f_4. If the conditions $|f_m(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\ldots,\boldsymbol{\omega}_{m-1})|<\prod_{j=1}^{(m-1)d}\beta_\delta(\omega_j)$ and for $1\leq i\leq(m-1)d$, $|\frac{\partial f_m(\omega_1,\ldots,\omega_{(m-1)d})}{\partial\omega_i}|\leq\prod_{j=1}^{(m-1)d}\beta_\delta(\omega_j)$ hold, then $\widehat{c}_\lambda(\mathcal{S}')$ is a mean squared consistent estimator of $c_{\lambda,1}+d_{\lambda,1}$.

4 The behaviour of the test statistic under the alternative hypothesis

In this section we consider the behaviour of the test statistic when the spatial random process is second order nonstationary. More precisely, $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is a constant mean random process whose covariance $\mathrm{cov}[Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)|\boldsymbol{s}_1, \boldsymbol{s}_2]$ is not a function of the difference $(\boldsymbol{s}_1 - \boldsymbol{s}_2)$. To fix notation we let $\kappa_{\boldsymbol{s}_2}(\boldsymbol{s}_1 - \boldsymbol{s}_2) = \mathrm{cov}[Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)|\boldsymbol{s}_1, \boldsymbol{s}_2]$.

Our objective is to understand what $a_\lambda(g; \boldsymbol{r})$ is estimating when the spatial covariance is nonstationary. Unfortunately, in general if a process is nonstationary, it is usually not possible to obtain asymptotically consistent estimators, since the global character of the spatial random field changes as we increase the spatial domain, and a limit cannot be easily defined. A possible solution to this problem can be found in the regular sampled time series framework. In this setting, Dahlhaus (1997, 2012) address this issue by conducting the asymptotics in so called 'rescaled time' not 'real time' and show that by using rescaled asymptotics, one can evaluate the asymptotic limit of an estimator constructed from a 'locally stationary' time series. This allows us to understand the sampling properties of an estimator, even though asymptotic consistency in real time cannot be achieved. Fuentes (2002, 2005) use a similar rescaling devise to study the nonstationary behaviour of estimators on the spatial grid $\mathbb{Z}^d$. Motivated by these approaches we propose a form of rescaled asymptotics for spatial processes defined on $\mathbb{R}^d$, which we use to understand how $a_\lambda(g; \boldsymbol{r})$ behaves under the nonstationary set-up.

To motivate our approach, we return to the location dependent spectral density $f(\boldsymbol{\omega}; \boldsymbol{s})$, defined in Theorem 2.1. Taking the Fourier transform of $f(\boldsymbol{\omega}; \boldsymbol{s})$ over location $\boldsymbol{s}$, we observe that $f(\boldsymbol{\omega}; \boldsymbol{s})$ can be written as

$$f(\boldsymbol{\omega}; \boldsymbol{s}) = \frac{1}{(2\pi)^d} \sum_{\boldsymbol{j} \in \mathbb{Z}^d} \zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega}) \exp\left(i2\pi \frac{\boldsymbol{s}'\boldsymbol{j}}{\lambda}\right), \tag{6}$$

where $\zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega}) = \frac{1}{\lambda^d} \int_{[-\lambda/2,\lambda/2]^d} f(\boldsymbol{\omega}; \boldsymbol{s}) \exp(-i2\pi \boldsymbol{s}'\boldsymbol{j}/\lambda) d\boldsymbol{s}$. Recall $\zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega}) = 0$ for all $\boldsymbol{j} \neq \boldsymbol{0}$ and all $\boldsymbol{\omega} \in \mathbb{R}^d$ if and only if the process is second order stationary. Since $\zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega})$ depends on λ, using mixed asymptotics (described in Section 2) $\zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega})$ cannot be consistently estimated. Therefore, we place some structure on the coefficients $\zeta_{\lambda,\boldsymbol{j}}(\boldsymbol{\omega})$ as $\lambda \to \infty$.

For $\boldsymbol{s} \in [-1/2, 1/2]^d$, define the function

$$\widetilde{f}(\boldsymbol{\omega}; \boldsymbol{s}) = \sum_{j_1,\ldots,j_d=-\infty}^{\infty} \zeta_{\boldsymbol{j}}(\boldsymbol{\omega}) \exp\left(i2\pi \boldsymbol{s}'\boldsymbol{j}\right) \text{ where } \zeta_{\boldsymbol{r}}(\boldsymbol{\omega}) = \int_{[-1/2,1/2]^d} \widetilde{f}(\boldsymbol{\omega}; \boldsymbol{s}) \exp\left(-i2\pi \boldsymbol{s}'\boldsymbol{r}\right) d\boldsymbol{s} \tag{7}$$

and let

$$\widetilde{\kappa}_{\boldsymbol{s}}(\boldsymbol{u}) = \int_{\mathbb{R}^d} \widetilde{f}(\boldsymbol{\omega}; \boldsymbol{s}) \exp(-i2\pi \boldsymbol{u}'\boldsymbol{\omega}) d\boldsymbol{\omega}.$$

Given the above definitions we define a sequence of nonstationary spatial processes $\{Z_\lambda(\boldsymbol{s})\}$ (we use the term 'sequence' loosely, since λ is defined on $\mathbb{R}^+$ not $\mathbb{Z}^+$), where for each $\lambda > 0$ and $\boldsymbol{s} \in [-\lambda/2, \lambda/2]^d$ the covariance is

$$\mathrm{cov}[Z_\lambda(\boldsymbol{s}), Z_\lambda(\boldsymbol{s}+\boldsymbol{u})] = \kappa_{\lambda,\boldsymbol{s}}(\boldsymbol{u}) = \widetilde{\kappa}_{\boldsymbol{s}/\lambda}(\boldsymbol{u}),$$

and the location dependent spectral density is $f_\lambda(\boldsymbol{\omega}; \boldsymbol{s}) = \widetilde{f}(\boldsymbol{\omega}; \frac{\boldsymbol{s}}{\lambda})$. Notice that the Fourier coefficients $\{\zeta_{\boldsymbol{j}}(\cdot); \boldsymbol{j} \in \mathbb{Z}^d\}$ remain the same for all λ, but the covariance changes with λ. We will use this set-up to obtain the asymptotic limit of $a_\lambda(g; \boldsymbol{r})$. Below we state the assumptions that we require.

Assumption 4.1 *Suppose $\{Z_\lambda(\boldsymbol{s})\}$ is a series of constant mean ($\mathrm{E}[Z_\lambda(\boldsymbol{s})|\boldsymbol{s}] = \mu$) Gaussian random fields whose location dependent spectral density $f_\lambda(\boldsymbol{\omega}; \boldsymbol{s})$ satisfies (7). The coefficients $\{\zeta_{\boldsymbol{j}}(\cdot)\}$ defined in (7) satisfy*

(i) $\sup_{\boldsymbol{\omega}\in\mathbb{R}^d} |\zeta_{\boldsymbol{j}}(\boldsymbol{\omega})| \leq \ell(\boldsymbol{j})$ *and* $\int_{\mathbb{R}^d} |\zeta_{\boldsymbol{j}}(\boldsymbol{\omega})| d\boldsymbol{\omega} \leq \ell(\boldsymbol{j})$,

(ii) For $1 \leq i \leq d$, *we have* $\sup_{\boldsymbol{\omega}\in\mathbb{R}^d} |\frac{\partial \zeta_{\boldsymbol{j}}(\boldsymbol{\omega})}{\partial \omega_i}| \leq \ell(\boldsymbol{j})$ *and* $\int_{\mathbb{R}^d} |\frac{\partial \zeta_{\boldsymbol{j}}(\boldsymbol{\omega})}{\partial \omega_i}| d\boldsymbol{\omega} \leq \ell(\boldsymbol{j})$,

where $\ell(\boldsymbol{j}) = \prod_{i=1}^d \ell(j_i)$ *with* $\ell(j) = C|j|^{-1}$ *if* $j \neq 0$ *and* $\ell(0) = C$ ($|C| < \infty$).

We now study the sampling properties of $a_\lambda(g; \boldsymbol{r})$ under this asymptotic set-up.

Theorem 4.1 *Let us suppose that Assumptions 4.1 is satisfied, and let $a_\lambda(g; \boldsymbol{r})$ be defined as in (1), respectively. Then we have*

$$\mathrm{E}[a_\lambda(g; \boldsymbol{r})] \;=\; a(g; \boldsymbol{r}) + O\left(\frac{\log^2(\lambda + \|\boldsymbol{r}\|_1)}{\lambda}\right), \tag{8}$$

and

$$\lambda^d \mathrm{cov}\left[a_\lambda(g; \boldsymbol{r}_1), a_\lambda(g; \boldsymbol{r}_2)\right] = O\left(\ell_5(\boldsymbol{r}_1 - \boldsymbol{r}_2) + \frac{\lambda^d}{n}\right), \tag{9}$$

as $\ell_{\lambda,a,n} \to 0$, where $\ell_5(\boldsymbol{r}) = \prod_{i=1}^d \ell_5(r_i)$, with $\ell_5(r) = C$ if $|r| \leq e$ and $\ell_5(r) = C\log^5(|r|)/r$ if $|r| > e$, and

$$a(g; \boldsymbol{r}) = \frac{1}{(2\pi)^d} \int_{[-a/\lambda, a/\lambda]^d} \zeta_{\boldsymbol{r}}(\boldsymbol{\omega}) g(\boldsymbol{\omega}) d\boldsymbol{\omega}. \tag{10}$$

We use Theorem 4.1 to assess the power of the test. We rewrite $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ as

$$\begin{aligned}\mathcal{T}_{\mathcal{S},\mathcal{S}'} &= \frac{\lambda^d}{\widehat{c}_\lambda(\mathcal{S}')}\max_{\boldsymbol{r}\in\mathcal{S}}\Big[|a_\lambda(g;\boldsymbol{r})-a(g;\boldsymbol{r})|^2+|a(g;\boldsymbol{r})|^2 \\ &\quad +2\Re[a_\lambda(g;\boldsymbol{r})-a(g;\boldsymbol{r})]\Re a(g;\boldsymbol{r})+2\Im[a_\lambda(g;\boldsymbol{r})-a(g;\boldsymbol{r})]\Im a(g;\boldsymbol{r})\Big]. \quad (11)\end{aligned}$$

We see that $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ depends on two components: (a) $a_\lambda(g;\boldsymbol{r})$ and (b) the variance estimator $\widehat{c}_\lambda(\mathcal{S}')$. Theorem 4.1 explains how $a_\lambda(g;\boldsymbol{r})$ behaves under the alternative, and we now use this result to understand the limit of $\widehat{c}_\lambda(\mathcal{S}')$ under the alternative. Let $\|\boldsymbol{r}\|_{\min}=\min(|r_i|;1\le i\le d)$ and $|\mathcal{S}'|_{\min}=\min_{\boldsymbol{r}\in\mathcal{S}'}(\|\boldsymbol{r}\|_{\min})$.

Lemma 4.1 *Suppose Assumption 4.1 holds. Then we have*

$$\mathrm{E}[\widehat{c}_\lambda(\mathcal{S}')]=O\left(1+\left[\frac{\lambda^d\log^2|\mathcal{S}'|_{\max}}{|\mathcal{S}'||\mathcal{S}'|_{\min}}+\frac{\lambda^d}{|\mathcal{S}'|^2}\left(\log|\mathcal{S}'|_{\max}\log\frac{|\mathcal{S}'|_{\max}}{|\mathcal{S}'|_{\min}}\right)^{2d}\right]\right)$$

and

$$\mathrm{var}\,[\widehat{c}_\lambda(\mathcal{S}')]=O\left(\frac{1}{|\mathcal{S}'|}+\frac{\lambda^d}{n}+\frac{\log^{6d}a}{\lambda}+\frac{C\log^{4d}a}{|\mathcal{S}'|}\left[\log|\mathcal{S}'|_{\max}\left(\log\frac{|\mathcal{S}'|_{\max}}{|\mathcal{S}'|_{\min}}\right)\right]^d\right).$$

We use this result to understand how $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ will behave for different choices of $\mathcal{S}'$ under the null and alternative. In Section 3 we showed that for $\widehat{c}_\lambda(\mathcal{S}')$ to be a consistent estimator of $c_{\lambda,1}$, $\mathcal{S}'$ should be such that $|\mathcal{S}'|_{\max}=O(\lambda^{1-\delta})$ and $|\mathcal{S}'|=O(\lambda^{d(1-\delta)})$ for some $0<\delta<1$. In the discussion below we will assume that $\mathcal{S}'$ satisfies these specifications. Further, we will assume that the set $\mathcal{S}$ is such that for at least one $\boldsymbol{r}\in\mathcal{S}$, $a(g;\boldsymbol{r})\neq 0$. From Lemma 4.1 we observe that if we keep the lower part of the set $\mathcal{S}'$ fixed, i.e., $|\mathcal{S}'|_{\min}=C$, where $C\ge 0$ is a fixed constant (hence the smallest values in $\mathcal{S}$ remain fixed), then $\widehat{c}_\lambda(\mathcal{S}')=O_p(\lambda^{d\delta}\log^2\lambda)$. This together with (11) implies that $\mathcal{T}_{\mathcal{S},\mathcal{S}'}=O_p(\lambda^d\sup_{\boldsymbol{r}\in\mathcal{S}}|a(g;\boldsymbol{r})|^2/\lambda^{d\delta}\log^2\lambda)$. On the other hand, if we define $\mathcal{S}'$ in such a way that $|\mathcal{S}'|_{\min}=O(\lambda^{d\delta}\log^2\lambda)$, then $\widehat{c}_\lambda(\mathcal{S}')=O_p(1)$ and $\mathcal{T}_{\mathcal{S},\mathcal{S}'}=O(\lambda^d\sup_{\boldsymbol{r}\in\mathcal{S}}|a(g;\boldsymbol{r})|^2)$. In other words, the power is greatest when we allow the set $\mathcal{S}'$ to grow in such a way that its inner values are moving away from zero. This is reasonable, since $a_\lambda(g;\boldsymbol{r})$ are estimating the Fourier coefficients, $a(g;\boldsymbol{r})$ and $|a(g;\boldsymbol{r})|\to 0$ as $\|\boldsymbol{r}\|_1\to\infty$. Hence, for maximum power we require that the set $\mathcal{S}$ to neighbour zero so that $a(g;\boldsymbol{r})$ is largest. Whereas, $\mathcal{S}'$ should be further from zero so the corresponding $a_\lambda(g;\boldsymbol{r})$ and $\widehat{c}_\lambda(\mathcal{S}')$ are small. Of course, in order that $\widehat{c}_\lambda(\mathcal{S}')$ is a consistent estimator of $c_{\lambda,1}$ (under the null), $\mathcal{S}'$ should not be too far from zero. In Section 6, we give some guidelines on selecting $\mathcal{S}$ and $\mathcal{S}'$.

To understand what the test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ is able to detect, we use (7) to rewrite $a(g;\boldsymbol{r})$

(defined in (10)) as

$$a(g;\boldsymbol{r}) = \int_{[-\lambda/2,\lambda/2]^d} \exp\left(\frac{-i2\pi\boldsymbol{s}'\boldsymbol{r}}{\lambda}\right) \frac{1}{(2\pi)^d} \int_{[-a/\lambda,a/\lambda]^d} f_\lambda(\boldsymbol{\omega};\boldsymbol{s}) g(\boldsymbol{\omega}) d\boldsymbol{\omega} d\boldsymbol{s}, \tag{12}$$

which is the Fourier transform of $(2\pi)^{-d}\int_{[-a/\lambda,a/\lambda]^d} f_\lambda(\boldsymbol{\omega};\boldsymbol{s})g(\boldsymbol{\omega})d\boldsymbol{\omega}$ over $\boldsymbol{s}$ at frequency $\boldsymbol{r}$. We use this transformation to suggest examples of weight functions.

Example 4.1 (Exponential weight functions) *From the above arguments it can be shown that by using the weight function $g(\boldsymbol{\omega}) = \exp(i\boldsymbol{v}'\boldsymbol{\omega})$, we have $\int_{\mathbb{R}^d} f_\lambda(\boldsymbol{\omega};\boldsymbol{s})\exp(i\boldsymbol{v}'\boldsymbol{\omega})d\boldsymbol{\omega} = \kappa_{\lambda,\boldsymbol{s}}(\boldsymbol{v})$. Therefore under Assumption 4.1 we have,*

$$a(e^{i\boldsymbol{v}\cdot};\boldsymbol{r}) = \int_{[-1/2,1/2]^d} \widetilde{\kappa}_{\boldsymbol{s}}(\boldsymbol{v})\exp(-i\boldsymbol{r}'\boldsymbol{s})d\boldsymbol{s} + o_p(1). \tag{13}$$

Thus, $a_\lambda(e^{i\boldsymbol{v}\cdot};\boldsymbol{r})$ estimates $a(e^{i\boldsymbol{v}\cdot};\boldsymbol{r})$, the Fourier transform of the location dependent spatial covariance at covariance lag $\boldsymbol{v}$. Eqn.(13) gives guidelines on how to select $\boldsymbol{v}$ for weight functions which are the sum of exponentials, $g(\boldsymbol{\omega}) = \exp(-i\boldsymbol{v}'\boldsymbol{\omega})$. Since $\kappa_{\lambda,\boldsymbol{s}}(\boldsymbol{v})$ is the nonstationary covariance function at covariance lag $\boldsymbol{v}$ and the magnitude of the covariance will decrease as $\|\boldsymbol{v}\|_1 \to \infty$, we need to select relatively small values of $\boldsymbol{v}$ in order to ensure that $a_\lambda(g;\boldsymbol{r})$ is large (suggestions are given in Sections 6 and 7.2).

Remark 4.1 (Spatially dependent mean) *We now consider the case that the spatial random field has not been detrended and $\mathrm{E}[Z(\boldsymbol{s})] = \mu(\boldsymbol{s})$, where $\mu(\cdot)$ is a function which depends on location. We now show that in the case that the mean spatial function varies over space, then regardless of whether the covariance function is stationary or not, $\mathrm{E}[a_\lambda(g;\boldsymbol{r})]$ will be 'large'. We define $a^\mu_\lambda(g;\boldsymbol{r})$ just as $a_\lambda(g;\boldsymbol{r})$ but with $Z(\boldsymbol{s}) - \mu$ replacing $Z(\boldsymbol{s})$ in the definition of the $J_n(\boldsymbol{\omega})$. It can be shown that*

$$\mathrm{E}\big[a_\lambda(g;\boldsymbol{r})\big] \approx \sum_{k_1,\ldots,k_d=-a}^{a} g(\boldsymbol{\omega}_k)\mu_F(\boldsymbol{k})\mu_F(-\boldsymbol{k}-\boldsymbol{r}) + \mathrm{E}\big[a^\mu_\lambda(g;\boldsymbol{r})\big],$$

where

$$\mu_F(\boldsymbol{r}) = \frac{1}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^d} \mu(\boldsymbol{s})\exp(i2\pi\frac{\boldsymbol{s}'\boldsymbol{r}}{\lambda})d\boldsymbol{s}.$$

Therefore, we see that $\mathrm{E}\big[a_\lambda(g;\boldsymbol{r})\big]$ is unlikely to be zero as $\lambda \to \infty$. Also note that the performance of the test also depends on the magnitude of $\widehat{c}_\lambda(\mathcal{S}')$.

5 Diagnostic plots

Besides motivating a formal test, Theorem 3.3 allows us to visualize the nonstationarities in the random field. We observe that if the null is true then asymptotically the standardized coefficients,

$$t_R(\boldsymbol{r}) = \lambda^{d/2}\frac{\Re a_\lambda(g;\boldsymbol{r})}{\sqrt{\widehat{c}_\lambda(\mathcal{S})}} \quad \text{and} \quad t_I(\boldsymbol{r}) = \lambda^{d/2}\frac{\Im a_\lambda(g;\boldsymbol{r})}{\sqrt{\widehat{c}_\lambda(\mathcal{S})}},$$

are uncorrelated and follow a t-distribution with $(2|\mathcal{S}'|-1)$ degrees of freedom. We use this result to make a plot of $t_R(\boldsymbol{r})$ and $t_I(\boldsymbol{r})$ which will allow us to see which $a_\lambda(g;\boldsymbol{r})$ coefficients are significant and identify nonstationary features in the data. In Figure 1a we plot $t_R(\boldsymbol{r})$ and $t_I(\boldsymbol{r})$ for $\boldsymbol{r} \in \mathrm{R} = \{\boldsymbol{r} = (r_1, r_2)' : r_j \in \{0, 1, \cdots, 5\}, j = 1, 2; \boldsymbol{r} \neq \boldsymbol{0}, \boldsymbol{r}_\ell + \boldsymbol{r}_k \neq \boldsymbol{0}, \forall \ell, k\}$ from one realisation of a stationary random field, where $n = 1000$ and $\lambda = 5$ (a Gaussian process with exponential covariance and range parameter $\rho = 1$). We observe that most of the standardized real and imaginary coefficients are blue or green (meaning that their t-transforms lie between -1.75 and 1.75, where 1.75 is the 95-th percentile of a t distribution with df = 15) and are statistically insignificant.

In Figure 1b we plot a $t_R(\boldsymbol{r})$ and $t_I(\boldsymbol{r})$ for $\boldsymbol{r} \in \mathrm{R}$, from one realisation of a nonstationary random field where $n = 1000$ and $\lambda = 5$ (a Gaussian process with covariance model NS2, as defined in Section 7.1.3). In contrast to Figure 1a we observe that some of the standardized coefficients lie outside the interval $[-1.75, 1.75]$, some are even lying outside the interval $[-2.95, 2.95]$, where 2.95 is the 99.5-th percentile of a t distribution with df = 15, which suggests nonstationarity in the random field.

To summarize, it is often difficult to visually discriminate between stationary and nonstationary random fields. However, the plots of $t_R(\boldsymbol{r})$ and $t_I(\boldsymbol{r})$ do allow us to visually discriminate between them. Moreover, $a_\lambda(g;\boldsymbol{r})$ can convey information about the nonstationarity. Using Theorem 4.1, heuristically we see that

$$\mathrm{E}\left(\sum_{\boldsymbol{r}\in\mathbb{Z}^d} a_\lambda(g;\boldsymbol{r})e^{i\boldsymbol{s}'\boldsymbol{r}}\right) \approx \int_{[-a/\lambda, a/\lambda]^d} g(\boldsymbol{\omega}) f_\lambda(\boldsymbol{\omega};\boldsymbol{s}) d\boldsymbol{\omega}.$$

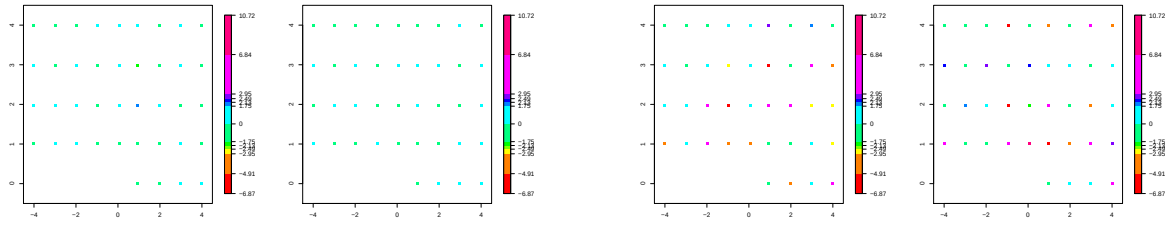
Using this result certain configurations of significant $a_\lambda(g;\boldsymbol{r})$ can suggest where the nonstationarities lie. Examples include:

- Significant coefficients of $a_\lambda(g;\boldsymbol{r})$ ($\boldsymbol{r} = (r_1, r_2)$) which only lie on the r_1 and r_2 axis suggest that the location dependent spectral density has an additive structure of the form $f_\lambda(\boldsymbol{\omega};\boldsymbol{s}) = f_{\lambda,1}(\boldsymbol{\omega}; s_x) + f_{\lambda,2}(\boldsymbol{\omega}; s_y)$ ($\boldsymbol{s} = (s_x, s_y)'$).

- Significant coefficients which only lie on the $r_1 = r_2$ diagonal suggest the location

dependent spectral density has the form $f_\lambda(\boldsymbol{\omega}; \boldsymbol{s}) = f_\lambda(\boldsymbol{\omega}; s_x + s_y)$.

It is worth noting that in the case the null is rejected and a diagnostic plot is made it is instructive to change the grid $\mathcal{S}'$ in the definition of $\widehat{c}_\lambda(\mathcal{S}')$. Moving the grid $\mathcal{S}'$ further from the origin will highlight statistically significant coefficients.



(a) Stationary: $t_R(\boldsymbol{r})$ (left) and $t_I(\boldsymbol{r})$ (right). (b) Nonstationary: $t_R(\boldsymbol{r})$ (left) and $t_I(\boldsymbol{r})$ (right).

Figure 1: Representative plots of standardized coefficients for stationary and nonstationary processes, where $\widehat{c}_\lambda(\mathcal{S}')$ is calculated using the grid given in Section 6 (in total there are 16 real and imaginary terms in $\mathcal{S}'$). Scale is according to a t-distribution with df = 15.

6 Implementation Issues

In this section, we briefly discuss implementation of the test. The test depends on the selection of the tuning parameters, $\mathcal{S}$ and $\mathcal{S}'$ in the definition of the test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ and the choice of weight function $g(\cdot)$ in the definition of $a_\lambda(g; \boldsymbol{r})$. We now briefly discuss how these user-chosen parameters can be selected. We recall that in the definition of $a_\lambda(g; \boldsymbol{r})$ we need to choose the size of the frequency grid, a, over which we test for correlation. In all the simulations and data analysis we used $a = \sqrt{n}/2$ in the definition of $a_\lambda(g; \boldsymbol{r})$. Further, in all the simulations we removed the sample mean from the observations.

1. Selection of $\mathcal{S}$ and $\mathcal{S}'$: The analysis of the test under the alternative shows that the test tends to have more power when $\boldsymbol{r}$ is small. Therefore it makes sense only to choose $\boldsymbol{r}$ values which are close to the origin. For practical purpose one might only want to consider the 4 or at the most 12 points nearest to the origin on the r grid to check for the nonstationarity. In our simulation study we have considered the nearest 4 points to $(0, 0)$, i.e., $\mathcal{S} = \{(-1, 1), (0, 1), (1, 1), (1, 0)\}$, noting that by Theorem 3.2 we cannot use both $\boldsymbol{r}$ and $-\boldsymbol{r}$. The set $\mathcal{S}$ is illustrated in Figure 2a, and this choice of $\mathcal{S}$ appears to work well in detecting nonstationarity.

Simulations show that using the grid configuration $\mathcal{S}'$ given in Figure 2a to define $\widehat{c}_\lambda(\mathcal{S}')$, estimates the variance well. Therefore in all the simulations we have used this grid. However,

(a) r grid (b) v grid

Figure 2: (a) The r grid gives possible $\boldsymbol{r}$ values that can be used in the test. The real and imaginary '$\bullet$' values are used in the testing procedure (this is set $\mathcal{S}$), whereas the real and imaginary 'Δ' values are used to calculate the standard error (this is set $\mathcal{S}'$). (b) The v grid used to construct $g(\cdot)$ for the test.

it is worth bearing in mind that if the alternative were true and the significant coefficients lie only in certain regions of $\mathbb{Z}^2$ (for example, only on the r_1 and r_2 axis, which arises when the nonstationarity has an additive structure), then the variance $\widehat{c}_\lambda(\mathcal{S}')$ may be too large compared with $\lambda^d \max_{\boldsymbol{r}\in\mathcal{S}} |a_\lambda(g;\boldsymbol{r})|^2$, thus reducing the power of the test. Therefore we suggest comparing the p-values of the test with respect to other configurations of $\mathcal{S}'$.

2. Selecting the weight function $g(\cdot)$: We recall from Example 4.1 if $g(\boldsymbol{\omega}) = \exp(-i\boldsymbol{v}'\boldsymbol{\omega})$, then $a_\lambda(g;\boldsymbol{r})$ estimates changes in the covariance at lag $\boldsymbol{v}$. In practice, we are unlikely to know which $\boldsymbol{v}$ to use, therefore we measure the aggregate change of several covariance lags together. In this case we use the weight function $g(\boldsymbol{\omega}) = \sum_{j=1}^{p} e^{-i\boldsymbol{v}_j'\boldsymbol{\omega}}$.

The choice of $\boldsymbol{v}_j$ should depend on the density of the sampling region. More precisely, since there are n observations in $[-\lambda/2, \lambda/2]^d$, the 'average spacing' between the observations on each axis is $s = \lambda/n^{1/d}$. Therefore, we do not have enough data to reliably estimate the nonstationary covariance at covariance lags far smaller than s. In all the simulations, we define the v grid as $\mathbb{V} = \{\boldsymbol{v}_j = (v_{j1}, v_{j2})' \in \mathbb{R}^d : v_{jk} = -s, -s/2, 0, s/2, s, \text{ for } k = 1, 2\}$ such that $\boldsymbol{v}_j + \boldsymbol{v}_{j'} \neq \boldsymbol{0}$ for $\boldsymbol{v}_j, \boldsymbol{v}_{j'} \in \mathbb{V}$. In Figure 2b we give the grid for $n = 1000$, $\lambda = 5$ ($s \approx 0.16$). We should mention that if the support of the empirical covariance of the data appears far greater than $s = \lambda/n^{1/d}$, then using a wider v grid is appropriate.

If changes in the covariance function happen mainly at lags much smaller than $s = \lambda/n^{1/d}$, then data is not available to detect changes in the covariance structure.

7 Numerical Study

7.1 Simulation study

In this section we illustrate the performance of the test $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ and discuss its applicability to various situations. We will use the specifications given in Section 6. All the tests are done at the 5% level and all results are based on 500 replications.

7.1.1 Models under the null

We consider different choices of (n, λ) pairs. In particular, we consider $n = 50, 100, 500, 1000$ and 2000 with $\lambda = 5$ to include a wide spectrum of (n, λ) combinations. Further the sampling locations are generated from a uniform distribution as well as a non-uniform distribution on $[-\lambda/2, \lambda/2]^2$ (See Figure 3a and 3b). To check for robustness of our test to deviations from uniformity of the locations, we sample from this non-uniform distribution. Spatial locations are often clustered and we model this using a mixture distribution. More precisely, $\{\boldsymbol{s}_j\}$ are iid random variables with two-clusters modelled with

$$\boldsymbol{s} = \begin{pmatrix} s_x \\ s_y \end{pmatrix} = \frac{\lambda}{3}\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} + \frac{\lambda}{3}\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} + \frac{\lambda}{3}\begin{pmatrix} W_1 \\ W_2 \end{pmatrix}, \tag{14}$$

where $U_1, U_2, \ldots, W_2$ are independent random variables; U_1 and U_2 are uniform random variable defined on $[-1/2, 1/2]$, $V_1 \sim TN(1/4, 1/10, (-1/2, 1/2))$, $V_2 \sim TN(1/4, 1/10, (-1/2, 1/2))$, $W_1 \sim TN(-1/4, 1/10, (-1/2, 1/2))$, $W_2 \sim TN(-1/4, 1/10, (-1/2, 1/2))$ and $TN(\mu, \sigma^2, (-1/2, 1/2))$ denotes a truncated normal distribution with non-truncated mean μ, variance σ^2 and corresponding truncation region $[-1/2, 1/2]$. Note that the two bivariate normal distributions with means $(1/4, 1/4)$ and $(-1/4, -1/4)$ model the centers of the two clusters.

In order to determine the size of our test we focus on random fields with the exponential covariance

$$c(\boldsymbol{h}) = \exp\left(-||\boldsymbol{h}||/\rho\right), \tag{15}$$

with different range parameters ρ. In particular, for empirical size calculations we consider $\rho = 2, 4/3, 1, 2/3, 1/3$; thus the percentage of the range of covariance will be between $7\% - 40\%$ of the spatial domain (the size of the spatial domain in all simulations under the null is $\lambda = 5$). A plot of the correlations is given in Figure 4.

To assess the performance of the test for non-Gaussian spatial random fields, we consider the logarithm of the squares of the Gaussian random field. The resulting process will be a non-Gaussian stationary random field with a different correlation structure to the original

Gaussian spatial random field. We also consider the case that the original observations are perturbed by a small amount of measurement error, ie. $Y_j = Z(\boldsymbol{s}_j) + \epsilon_j$, where $\{\epsilon_j\}$ are iid normal random variables with mean zero and variance 0.1^2 (the measurement error is about 10% of the standard deviation of the random field).

7.1.2 Discussion of simulations under the null

In Table 1 we report the results of $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ under the null hypothesis. Focussing first on the

	Gaussian					Non-Gaussian					Non-uniform				
ρ	n					n					n				
	50	100	500	1000	2000	50	100	500	1000	2000	50	100	500	1000	2000
2	0.09	0.05	0.13	0.14	0.18	0.05	0.03	0.06	0.07	0.08	0.03	0.03	0.01	0.01	0.01
	0.08	0.04	0.12	0.13	0.18	0.05	0.03	0.06	0.07	0.08	0.02	0.03	0.01	0.02	0.01
4/3	0.05	0.05	0.08	0.10	0.09	0.04	0.05	0.06	0.06	0.05	0.02	0.03	0.02	0.01	0.01
	0.05	0.04	0.07	0.10	0.09	0.03	0.05	0.05	0.06	0.06	0.02	0.02	0.02	0.01	0.01
1	0.04	0.04	0.04	0.08	0.08	0.04	0.04	0.06	0.08	0.08	0.02	0.02	0.01	0.01	0.01
	0.05	0.03	0.04	0.08	0.07	0.03	0.05	0.06	0.07	0.08	0.02	0.02	0.01	0.01	0.01
2/3	0.02	0.03	0.05	0.06	0.09	0.03	0.04	0.04	0.06	0.06	0.01	0.02	0.01	0.01	0.01
	0.02	0.04	0.05	0.06	0.09	0.03	0.04	0.04	0.06	0.06	0.01	0.02	0.01	0.01	0.01
1/3	0.03	0.03	0.03	0.05	0.07	0.03	0.04	0.04	0.04	0.04	0.02	0.02	0.02	0.01	0.01
	0.03	0.03	0.03	0.05	0.06	0.03	0.04	0.04	0.04	0.05	0.02	0.02	0.02	0.01	0.01

Table 1: Empirical type I errors (at 5% level) based on $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ with $\lambda = 5$ for Gaussian, Non-Gaussian, and data with non-uniform locations based on Exponential correlation functions. For each ρ the first row corresponds to the original data and the second row corresponds to the data with measurement error.

simulations in the Gaussian case using the test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$, we observe that in the case that the range of the covariance function is large with respect to the size of the spatial domain (the case $\rho = 2$ corresponds to 40% of the spatial domain), the empirical type I errors are inflated (see Figure 4). This is not surprising because (a) as can be seen from the proof of Theorem 2.1(i) the approximation errors in the decorrelation result depend on the rate of decay of the covariance tails; if the covariance is thick-tailed and λ is quite small the DFTs will still have some correlation, and (b) asymptotic normality of $a_\lambda(g; \boldsymbol{r})$ requires the range of dependence being far smaller than then the size of the spatial domain. We also observe that when the spatial domain is densely sampled and the range of the covariance is large with respect to the size of spatial domain then there is also a slight inflation of the type I error (this is due to the quality of the variance estimator $\widehat{c}_\lambda(\mathcal{S}')$ which tends to underestimate the variance when n is extremely large). However, it is reassuring to see in the case that ρ is moderately large with respect to λ, the empirical type I errors are quite close to the 5% level (for example, when $\rho = 1$ which corresponds to 20% of the range).

The non-Gaussian random field (logarithm of the squared data) has empirical type I errors which are close to the 5% level even when the range of the covariance is very large

with respect to the size of the domain. We further observe that when the locations are sampled from a mixture distribution (not uniform), empirical type I errors are below the 5%. Some what surprisingly, this suggests the test is conservative to certain deviations from uniformity of the locations.

7.1.3 Models under the alternative

To assess the power of the test based on $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ we consider two different nonstationary spatial models (see, NS1 and NS2 described below). Both models are constructed such that the variance is constant over the random field and the nonstationarity arises *only* in the spatial correlation.

We generate Gaussian data with constant mean $\mu = 0$ and the following nonstationary correlation structures:

(NS1) The nonstationary correlation function considered in Paciorek and Schervish [2006] and Jun and Genton [2012] which has the following form:

$$c_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2) = |\Sigma\left(\frac{\boldsymbol{s}_1}{\lambda}\right)|^{1/4} |\Sigma\left(\frac{\boldsymbol{s}_2}{\lambda}\right)|^{1/4} \left|\frac{\Sigma(\frac{\boldsymbol{s}_1}{\lambda}) + \Sigma(\frac{\boldsymbol{s}_2}{\lambda})}{2}\right|^{-1/2} \exp[-\sqrt{Q_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2)}],$$

where $|\cdot|$ denotes the determinant of a matrix, $Q_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2) = 2(\boldsymbol{s}_1 - \boldsymbol{s}_2)'[\Sigma(\frac{\boldsymbol{s}_1}{\lambda}) + \Sigma(\frac{\boldsymbol{s}_2}{\lambda})]^{-1}(\boldsymbol{s}_1 - \boldsymbol{s}_2)$ and $\Sigma(\frac{\boldsymbol{s}}{\lambda}) = \Gamma(\frac{\boldsymbol{s}}{\lambda})\Lambda\Gamma(\frac{\boldsymbol{s}}{\lambda})'$, where

$$\Gamma\left(\frac{\boldsymbol{s}_1}{\lambda}\right) = \begin{bmatrix} \gamma_1(\boldsymbol{s}/\lambda) & -\gamma_2(\boldsymbol{s}/\lambda) \\ \gamma_2(\boldsymbol{s}/\lambda) & \gamma_1(\boldsymbol{s}/\lambda) \end{bmatrix}, \ \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix},$$

with $\gamma_1(\boldsymbol{s}/\lambda) = \log(s_x/\lambda + 0.75)$, $\gamma_2(\boldsymbol{s}/\lambda) = (s_x/\lambda)^2 + (s_y/\lambda)^2$, and $\boldsymbol{s} = (s_x, s_y)'$. The functions $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are chosen such that the nonstationary correlation changes smoothly through space (see, Jun and Genton [2012]).

(NS2) Whereas in the previous model we considered smooth changes over space, in the second example we consider a piecewise stationary model. The spatial domain $[-\lambda/2, \lambda/2]^2$ is partitioned into four squares. The spatial random field in each square is independent of the others and stationary with exponential covariance function (defined in (15)) and range parameters $\rho = 1, 2/3, 1/2$, and $1/3$, respectively.

To understand the influence that the sample size n has on the power of the test we consider the above models with the following (n, λ)-combinations, ranging from a sparser set-up to a denser set-up:

(a) Model NS1: $n = 50, 100, 500, 1000, 2000$ and $\lambda = 20, 40$.

(b) Model NS2: $n = 50, 100, 500, 1000, 2000$ and $\lambda = 5$.

7.1.4 Discussion of simulations under the alternative

The empirical powers based on $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ are given Table 2.

		Gaussian					Non-Gaussian				
		n					n				
Model	λ	50	100	500	1000	2000	50	100	500	1000	2000
NS1	20	0.02	0.05	0.19	0.47	0.70	0.04	0.04	0.07	0.12	0.51
		0.03	0.04	0.18	0.46	0.71	0.04	0.04	0.07	0.12	0.51
	40	0.03	0.04	0.10	0.35	0.85	0.04	0.05	0.05	0.05	0.18
		0.04	0.04	0.09	0.36	0.85	0.04	0.05	0.05	0.05	0.17
NS2	5	0.05	0.04	0.11	0.24	0.36	0.05	0.04	0.05	0.11	0.15
		0.05	0.04	0.11	0.24	0.36	0.06	0.04	0.05	0.11	0.15

Table 2: Empirical powers based on $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ for the non-stationary models NS1 and NS2. For each λ the first row corresponds to the original data and the second row corresponds to the data with measurement error.

We observe that for any given model and λ, the power increases as the average spacing $\lambda/\sqrt{n}$ decreases. This is expected, because when the spatial random field is more densely sampled we have more information about the covariance structure. In particular, if the average spacing is within the main support of the covariance then the test is better able to discriminate the changing covariance structures. Comparing powers, we do observe that the power for the non-Gaussian spatial random field is a lot lower than the power for the Gaussian spatial random field. This is due to the way that the non-Gaussian spatial random field is defined. We recall that the non-Gaussian spatial covariance is $\mathrm{cov}[X(\boldsymbol{s}_1), X(\boldsymbol{s}_2)] = 4\mathrm{cov}[\log|Z(\boldsymbol{s}_1)|, \log|Z(\boldsymbol{s}_2)|]$, where $Z(\boldsymbol{s})$ is a non-stationary, constant variance, Gaussian random field with covariance function (NS1). Therefore, when the difference between $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ is small $\mathrm{cov}[X(\boldsymbol{s}_1), X(\boldsymbol{s}_2))]$ is close to constant (since the variance of $\log|Z(\boldsymbol{s})|$ is constant), and when the difference between $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ is relatively large, the log-transform tends to reduce/flattern the correlation over the random field, rendering the differences in the correlation structure over the random field small and difficult to detect. The power tends to be a lot lower for model NS2 than NS1. This is probably because the local covariance structure in model NS2 only ranges from 1/3 to 1 and is confined to a limited spatial domain $\lambda = 5$, again making it difficult to detect the nonstationarity.

Finally, we mention that in all the simulations (under the null and alternative) we have used $a = \sqrt{n}/2$ as the number of a-frequencies in definition of $a_\lambda(g; \boldsymbol{r})$. This seemed to give reasonable results. However, when the support of the empirical covariance is large with respect the size of the spatial domain and n is large, using $a = \sqrt{n}/2$ can sometimes lead to $\widehat{c}_\lambda(\mathcal{S})$ underestimating the variance of $\{a_\lambda(g; \boldsymbol{r}); \boldsymbol{r} = (0,1), (1,1), (1,0), (-1,1)\}$. The

reason for this can also be seen from the theoretical expression of the variance of $a_\lambda(g;\boldsymbol{r})$. Subsequently, this leads to an inflation of the type I error. Therefore, in such situations it is reasonable to do the test using $a = \sqrt{n}$. However, some care needs to be taken, as using large a can lead to a loss in power.

7.2 Ground ozone

Tropospheric ozone (ground ozone) is a major constituent of photochemical smog. It is a powerful oxidant that damages human health and natural ecosystems. It is also an important greenhouse gas. Unlike many other pollutants ozone is not directly emitted - it is a secondary pollutant formed by sunlight-driven chemical reactions involving carbon monoxide, volatile organic compounds and nitrogen oxide. The concentration of ozone in any given location and time depends on several factors, including the existence of heavy industry, density of vehicles and weather factors, in particular high temperatures (which drive the chemical reaction) but also wind (which can spread the ozone); see The Royal Society Special Policy Report 15/08 for further details.

In order to monitor air quality, the United States Environmental Protection Agency (EPA) measures the hourly/daily ground ozone from over 900 locations (both rural and urban) in the United States (see `http://www.epa.gov/airtrends/ozone.html`). In this section, our objective is to use this data to understand some of the spatial features observed in ground ozone. The data we consider is the daily average and maximum surface ozone (the data is measured hourly from 9 AM until 4 PM and the average and maximum evaluated from these eight observations) from April 1st, 2014 - September 30th, 2014 (183 days); see `http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html#Raw`. We focus on the Ohio Central Valley and South East of the US, in particular the rectangle: longitude $[-100, -80]$ and latitude $[30, 43]$. We choose this area because there was a high concentration of locations, which consistently gives measurements throughout this period. There are 489 stations in this region, and on any given day (from April 1st, 2014 - September 30th, 2014) between 472-489 stations are observed. The ground ozone is measured in parts per million (PPM).

For each day between April 1st - September 30th we apply the test for stationarity. We use the test statistics $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ defined in (5) with the $\mathcal{S}$ and $\mathcal{S}'$ defined in Section 6. Since $[-100, -80] \times [30, 43]$ is a rectangle, we let $\lambda_1 = 20$ and $\lambda_2 = 13$ and use as the average spacing on each axis $s_1 = \lambda_1/\sqrt{n} \approx 0.59$ and $s_2 = \lambda_2/\sqrt{n} \approx 0.89$ (since n is between $472 - 489$). In the definition of $a_\lambda(g;\boldsymbol{r})$ (which is used to define $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$) we use

$$g(\boldsymbol{\omega}) = \sum_{\boldsymbol{j}=(j_1,j_2)\in\mathbb{V}} \exp\left[\frac{i}{2}(5s_1j_1\omega_1 + 3s_2j_2\omega_2)\right],$$

where $\mathbb{V} = \{\{-2,-1,0,1,2\} \times \{1,2\}, \{0,1,2\} \times \{0\}\}$. Note that compared to Section 6 we used a wider v grid, stretching from $[-3s_1, 3s_1] \approx [-2.7, 2.7]$ and $[-5s_2, 5s_2] \approx [-3, 3]$. This is because, in general the range of the empirical covariance of the ozone data was relatively wide (usually over 2).

In Figure 5 we plot the value of the test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ over time, for both average and maximum ozone. Note that under the null the 5%, 2%, 1%, 0.5% and 0.1% rejection regions correspond to 11.56, 15, 18, 21 and 29, respectively. We observe that there seems to be some correlation between the average and maximum test statistics (this is also seen in Figure 6). Furthermore, there seems to be some clustering over time in the large test statistics. For example, for four days towards the end of September the spatial random fields appear highly nonstationary.

The rejection rates for the tests on maximum and average ozone are given in Table 3, where we observe that there are more rejections than we would expect under the null. We note that if we were to use the Bonferonni correction at the 5% level, then a statistically significant test occurs when the p-value is less than $5/183 = 0.027\%$ (which corresponds to a critical value of over 36). We observe from Figure 5 that none of the test statistics are larger than 36. However, given that a large number of tests were done and the sample autocorrelation of the test statistics over time suggest short term positive correlation (see Figure 6) using a Bonferonni correction is a conservative method for multiple testing.

Rejection level	5%	2%	1%	0.5%	0.1%
$\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ for Average ozone	12.5%	6.0%	2.7%	1.6%	0%
$\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ for Maximum ozone	9.8%	4.4%	2.7%	1.6%	0.54%

Table 3: Rejection rates for the test statistics of average and maximum ozone

Instead, to understand whether there is any evidence of spatial nonstationarity (at least at some time points), in Figure 7 we give a quantile-quantile plot (QQplot) of the empirical quantiles of $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ against the theoretical quantiles under the null. We recall, if the spatial process were stationary on all but a few days, the majority of the tests would lie on the $x = y$ line and only in the larger quantiles would there be a deviation from this line (see Efron [2012], Chapter 3). However, we observe that overall there is not a particularly good fit between the data and the empirical distribution under the null. To check if this is simply because $\widehat{c}_\lambda(\mathcal{S}')$ is underestimating the variance (which results in the points lying on a larger gradient) we redo the test using $a = \sqrt{n}$ in the definition of $a_\lambda(g; \boldsymbol{r})$ (recall from Section 7.1 that using too few a-frequencies in the definition in $a_\lambda(g; \boldsymbol{r})$ sometimes results in $\widehat{c}_\lambda(\mathcal{S}')$ underestimating the variance). However, this does not give a better fitting QQplot.

Therefore, the mismatch between the empirical and theoretical quantiles possibly suggests that the ozone data is spatially nonstationary over the entire time duration. To investigate

this further, we make a QQplot of the data against the theoretical distribution of a simple alternative. To do this, we recall some of the properties of the test statistic and $a_\lambda(g;\boldsymbol{r})$ under the null and alternative. In Section 3 we showed that under the null of spatial stationarity

$$\mathcal{T}_{\mathcal{S},\mathcal{S}'} = \frac{\lambda^d \max_{\boldsymbol{r}\in\mathcal{S}} |a_\lambda(g;\boldsymbol{r})|^2}{\widehat{c}_\lambda(\mathcal{S}')} \xrightarrow{\mathcal{D}} \frac{\max_{1\leq i\leq|\mathcal{S}|}(Z_{2i-1}^2 + Z_{2i}^2)}{\frac{1}{2|\mathcal{S}'|-1}\sum_{j=2|\mathcal{S}|+1}^{2(|\mathcal{S}|+|\mathcal{S}'|)}(Z_j-\bar{Z})^2}$$

where $\{Z_i; 1 \leq i \leq 2(|\mathcal{S}| + |\mathcal{S}'|)\}$ are iid standard normal random variables. However, under the alternative of nonstationarity, in Section 4, we showed that $a_\lambda(g;\boldsymbol{r})$ would have a non-zero mean. Thus to crudely reproduce the test statistic under a specific alternative we obtain the quantiles of the shifted distribution

$$Y = \frac{\max_{1\leq i\leq|\mathcal{S}|}(X_{2i-1}^2 + X_{2i}^2)}{\frac{1}{2|\mathcal{S}'|-1}\sum_{j=2|\mathcal{S}|+1}^{2(|\mathcal{S}|+|\mathcal{S}'|)}(X_j-\bar{X})^2}, \tag{16}$$

where $\{X_i; 1 \leq i \leq 2(|\mathcal{S}| + |\mathcal{S}'|)\}$ are independent normal random variables with variance one, $\mathrm{E}[X_1] = 1$, $\mathrm{E}[X_2] = 1$ and for $3 \leq i \leq 2(|\mathcal{S}|+|\mathcal{S}'|)$, $\mathrm{E}[X_i] = 0$. We use that for $1 \leq i \leq 2$ $\mathrm{E}[X_i] = 1$ to model a mean shift in one of the coefficients $a_\lambda(g;\boldsymbol{r})$. The QQplot of $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ against the distribution of this alternative is given in Figure 8. There is a relatively good fit of the empirical quantiles with the distribution Y (though the empirical quantiles in the upper tail are larger than expected under this alternative). This suggests that the ozone data is spatially nonstationary and could be modelled with the same spatially nonstationary model (note this is a very tentative conjecture). Diagnostic plots of $a_\lambda(g;\boldsymbol{r})$ calculated using the maximum ozone at several time points are given in the supplementary material. They are not all the same, but certain patterns can be seen. For example, the imaginary coefficient corresponding to $\boldsymbol{r} = (0,1)'$ are statistically significant in the majority of cases with high positive values. This adds further support to the conjecture that the ozone data is spatially nonstationary but the nonstationary features are similar over time.

However, even if this conjecture were true, it is useful to understand what causes the test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ to be either large or small. To see whether large or small $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ possibly imply different features in the spatial random field we compare two extreme examples. In particular, the maximum ozone on April 4th (when $\mathcal{T}_{\mathcal{S},\mathcal{S}'} = 2.54$) and the maximum ozone on April 6th (when $\mathcal{T}_{\mathcal{S},\mathcal{S}'} = 18.79$). In Figures 10 and 11 we plot both the ozone for these days and the corresponding sample covariograms (estimated using the function `variog` in R). We observe the two covariograms are very different; the empirical covariogram for the maximum ozone on April 4th plateaus to zero at large lags, whereas the the empirical covariogram for the maximum ozone on April 6th does not appear to plateau. This could be due to nonstationarity in the spatial mean or covariance. Therefore for both these days

we partition the region into 4 equal size quarters and evaluated the covariogram over each quarters (see, Figure 10 and 11). In the case of April 4th, the short region covariograms tend to follow the large region covariogram (with the exception of one region). In the case of April 6th, the short region covariograms are very different to the large region covariogram. This suggests that when $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ is large the 'degree' of nonstationarity is a lot greater than when $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ is small (though we stress that the choice of function $g(\boldsymbol{\omega})$ determines the nonstationary features $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ is able to detect).

To summarize, our analysis suggests that daily ozone in the region $[-100,-80]\times[30,43]$ is spatially nonstationary, with possibly similar nonstationary features over time. However, it appears that the degree of nonstationarity varies over time and in many instances it is difficult to discriminate between the spatial field being stationary or nonstationary. In the section below we check whether the locations of the stations have any influence on the results of the test.

The influence of the measurement station locations on the test

In Figures 10 and 11 we see that the location of the stations do not appear to be uniform. We recall that an underlying mathematical assumption in our procedure is that the locations are uniformly distributed. Therefore it is important to determine the influence these locations may have on the test. To assess the finite sample properties of the test statistics over the given set of observation stations on the rectangle $[-100,-80]\times[30,43]$, we simulate a stationary Gaussian distribution with exponential covariance and range parameter $\rho = 2.8$. For each realisation we apply the test statistic using exactly the same specifications as those used for the ozone data (i.e., same $\mathcal{S}$, $\mathcal{S}'$, $g(\boldsymbol{\omega})$ and $a = \sqrt{n}/2$). The empirical rejection rates over 1000 simulations at the 5%, 2%, 1% and 0.5% level were 6.8%, 1.7%, 0.8% and 0.6% respectively. A QQplot of the empirical quantiles against the theoretical ones under the null is given in Figure 9. We see that the for lower quantiles there is not an exact match between the empirical and theoretical quantiles. However, for the larger quantiles there is reasonable match, noting that the empirical rejections are close to the theoretical rejections. To check whether the distribution of Y, defined in (16), gives a better fit, in Figure 9 we give a QQplot of the empirical distribution again the distribution of Y, and did not see an improvement in fit.

Our analysis suggests that the lack of uniformity of the locations does not have a huge impact on the distribution under the null (at least for this exponential spatial covariance).

Acknowledgements

SB's work has been partially supported by the National Science Foundation, DMS-1406622. SB wishes to thank Dr. Douglas Nychka for several useful suggestions regarding the `fields` package in R. SSR's work has been partially supported by the National Science Foundation, DMS-1106518. SSR gratefully acknowledges the Institute of Mathematical Sciences, National University of Singapore where the initial part of this work was done while she was visiting there in June 2012. The authors are extremely grateful to the Associate Editor and two anonymous referees whose suggestions greatly improved the proposed methodology and presentation of the paper.

References

S. Bandyopadhyay and S. N. Lahiri. Asymptotic properties of discrete fourier transforms for spatial data. *Sankhya, Series A*, 71:221–259, 2009.

S. Bandyopadhyay, S. N. Lahiri, and D. Nordman. A frequency domain empirical likelihood method for irregular spaced spatial data. *Annals of Statistics*, 43(2):519–545, 2015.

N. Cressie. *Statistics for Spatial Data*. Wiley, New York, 1993.

R. Dahlhaus. Fitting time series models to nonstationary processes. *Annals of Statistics*, 16: 1–37, 1997.

R. Dahlhaus. *Handbook of Statistics, volume 30*, chapter Locally Stationary processes, pages 351–413. Elsevier, 2012.

Y. Dwivedi and S. Subba Rao. A test for second order stationarity based on the Discrete Fourier Transform. *Journal of Time Series Analysis*, 32:68–91, 2011.

B. Efron. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction*. Cambridge University Press, Cambridge, 2012.

M. Fuentes. Spectral methods for nonstationary spatial processes. *Biometrika*, 89:197–210, 2002.

M. Fuentes. A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis*, 96:30–54, 2005.

Y. Guan, M.l Sherman, and J. A. Calvin. A nonparametric test for spatial isotropy using subsampling. *Journal of the American Statistical Association*, 99(467):810–821, 2004.

P. Hall and P. Patil. Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields*, 99:399–424, 1994.

C. Jentsch and S. Subba Rao. A test for second order stationarity of a multivariate time series. *Journal of Econometrics*, 185:124–161, 2015.

M. Jun and M. Genton. A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Statistica Sinica*, 22:1737–1764, 2012.

S. N. Lahiri. Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhya*, 65:356–388, 2003.

B. Li, M. G. Genton, and M. Sherman. Testing the covariance structure of multivariate random fields. *Biometrika*, 95:813–829, 2008.

E. Masry. Poisson sampling and spectral estimation of continuous time processes. *IEEE Transactions on Information Theory*, 24:173–183, 1978.

Y. Matsuda and Y. Yajima. Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *Journal of the Royal Statistical Society (B)*, 71:191–217, 2009.

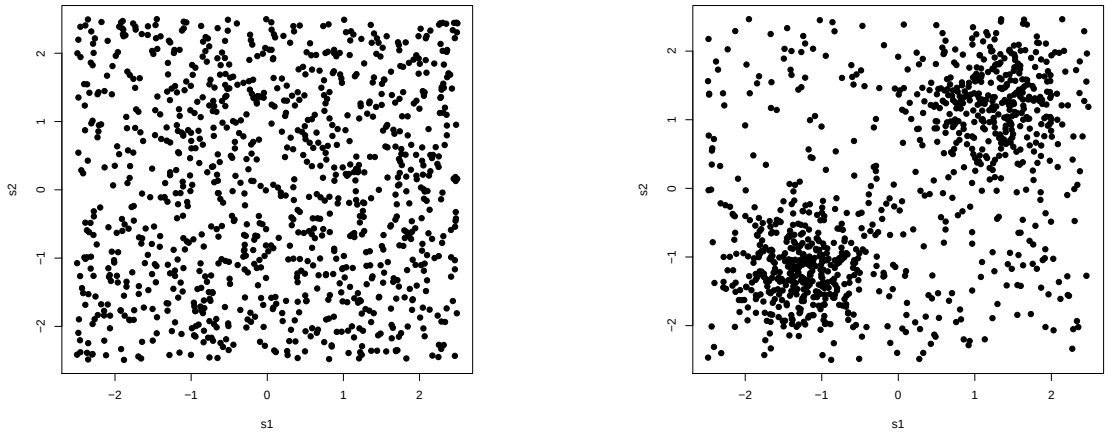
C. J. Paciorek and M. J. Schervish. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17(5):483–506, 2006.

M. B. Priestley and T. Subba Rao. A test for non-stationarity of a time series. *Journal of the Royal Statistical Society (B)*, 31:140–149, 1969.

M. Sherman. *Spatial Statistics and Spatio-Temporal Data*. Wiley, Chichester, 2011.

M. L. Stein. *Interpolation of spatial data: some theory for kriging*. Springer, 1999.

S. Subba Rao. Fourier based statistics for irregular spaced spatial data. *arXiv:1405.5240*, 2014.



(a) uniform location (b) non-uniform locations

Figure 3: Examples of (a) uniform and (b) locations sampled from the mixture distribution in (14) with $n = 1000$ and $\lambda = 5$.

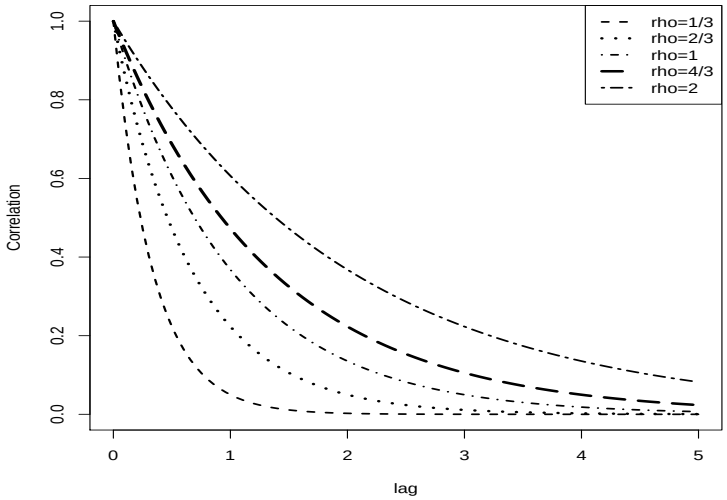


Figure 4: Plot of the Exponential correlation function with the range parameters used in the simulations.

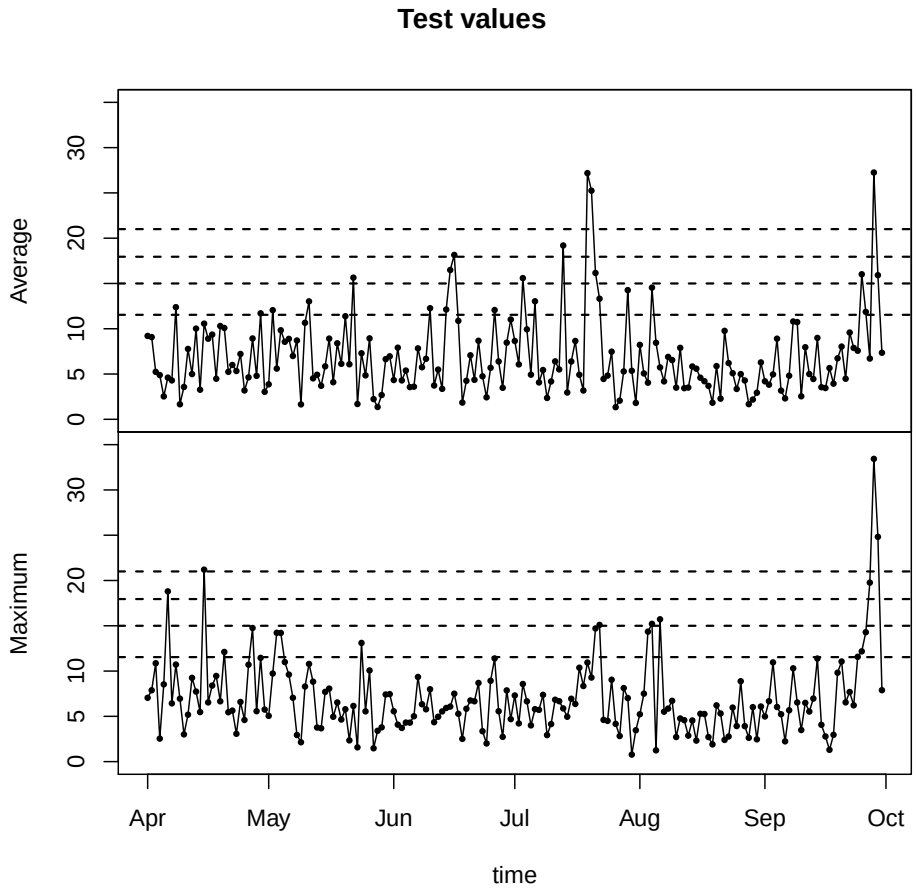


Figure 5: The test statistic $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ plotted over time. The horizontal lines denote the 5%, 2%, 1% and 0.5% rejection regions. Top: Test statistic for average ozone. Bottom: Test statistic for maximum ozone.

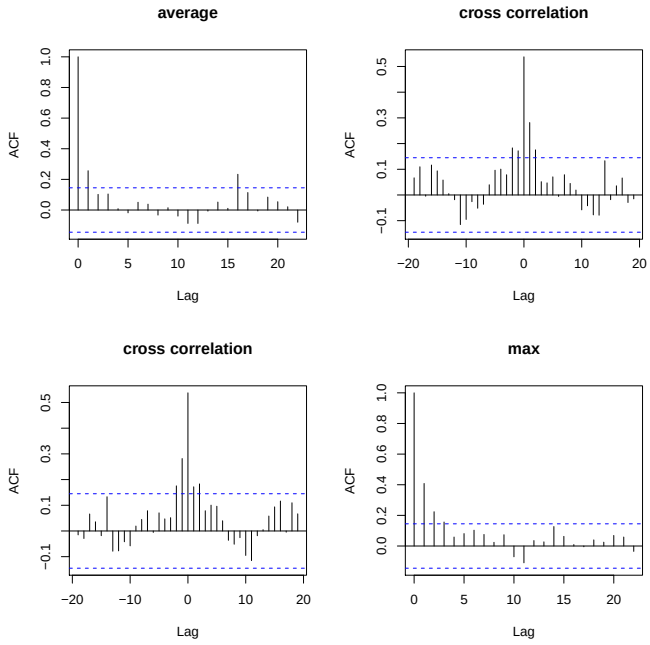


Figure 6: Sample autocorrelation and cross-correlations of the test statistics.

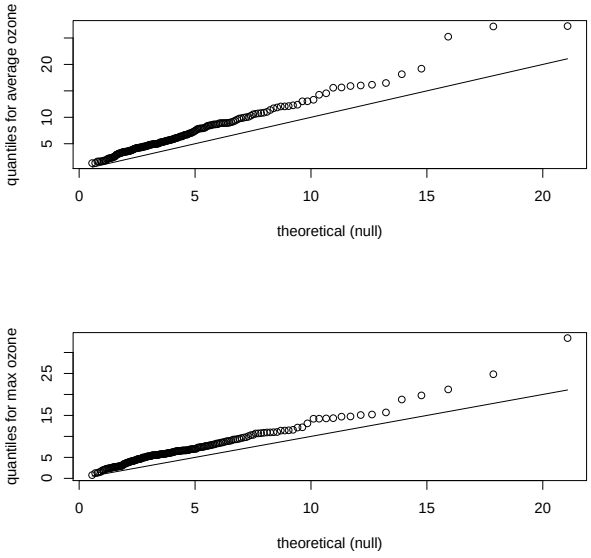


Figure 7: QQplot of $\mathcal{T}_{\mathcal{S},\mathcal{S}'}$ against the theoretical distribution under the null. Top: Average. Bottom: Maximum.

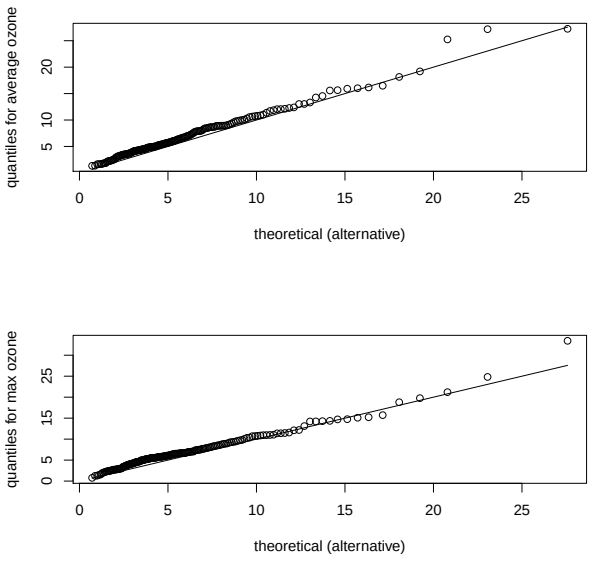


Figure 8: QQplot of data against the theoretical distribution under the alternative in equation (16). Top: Average. Bottom: Maximum.

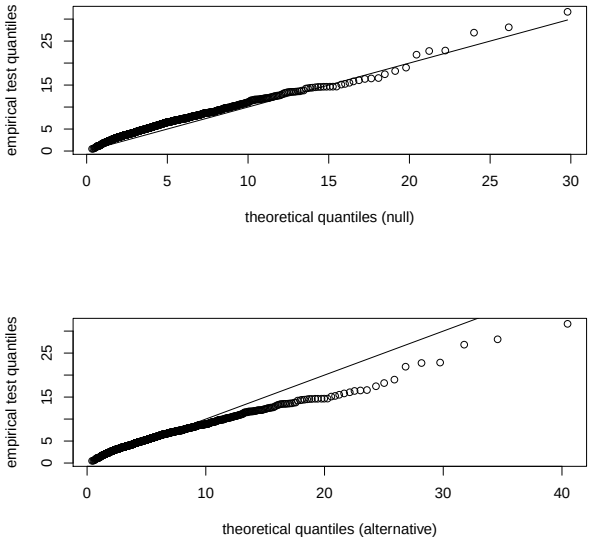


Figure 9: Top: QQplot of test statistic for simulated data against the theoretical distribution under null. Bottom: QQplot of test statistic for simulated data against the alternative in equation (16)

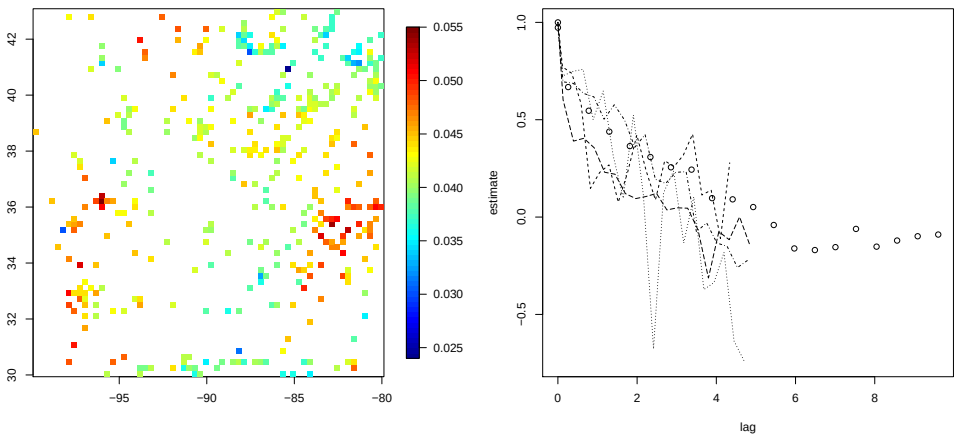


Figure 10: Summary of maximum ozone on 4th April, 2014. Left: Plot of maximum ozone. Right: The estimated covariograms over the entire region and the corresponding quarters.

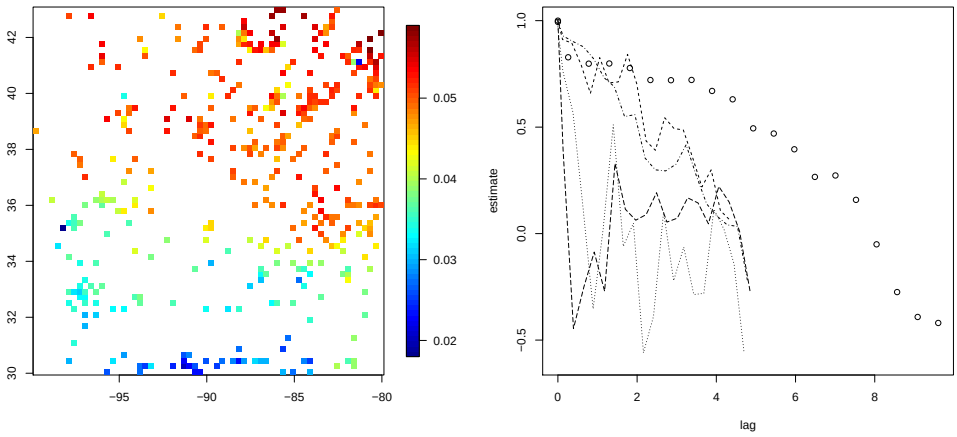


Figure 11: Summary of maximum ozone on 6th April, 2014. Left: Plot of maximum ozone. Right: The estimated covariograms over the entire region and the corresponding quarters.