Published in final edited form as:

J Stat Plan Inference. 2012 May ; 142(5): 1081–1091. doi:10.1016/j.jspi.2011.11.013.

Testing for Spatial Isotropy Under General Designs

Arnab Maity and

Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695, U.S.A.
amaity@ncsu.edu

Michael Sherman

Department of Statistics, Texas A&M University, College Station, Texas 77843-3143, U.S.A.
sherman@stat.tamu.edu

Abstract

Spatial modeling is typically composed of a specification of a mean function and a model for the correlation structure. A common assumption on the spatial correlation is that it is isotropic. This means that the correlation between any two observations depends only on the distance between those sites and not on their relative orientation. The assumption of isotropy is often made due to a simpler interpretation of correlation behavior and to an easier estimation problem under an assumed isotropy. The assumption of isotropy, however, can have serious deleterious effects when not appropriate. In this paper we formulate a test of isotropy for spatial observations located according to a general class of stochastic designs. Distribution theory of our test statistic is derived and we carry out extensive simulations which verify the efficacy of our approach. We apply our methodology to a data set on longleaf pine trees from an oldgrowth forest in the southern United States.

Keywords

Anisotropy; Covariogram; Isotropy; Spatial bootstrap; Spatial statistics

1 Introduction

Spatial modeling is typically composed of a specification of a mean function and a model for the covariance or correlation structure. Correlation is typical in spatial applications. For example, crop yields on adjacent plots tend to be similar due to common soil composition and watering patterns; wind speeds from nearby sites are often similar due to similar geographical conditions. One of the common key assumptions on the spatial correlation is that it is isotropic. This means that the correlation between any two observations depends only on the distance between those sites and not on their relative orientation. If the correlation structure is not isotropic it is said to be anisotropic. The assumption of isotropy is often made due to a simpler interpretation of correlation behavior and to an easier estimation problem under an assumed isotropy. The assumption of isotropy, however, can have serious deleterious effects when not warranted.

To see concretely how an assumption of isotropy impacts an analysis, consider the following covariance model

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

$$C(x, y) = \exp(-x^2 - \gamma y^2)$$

where $C(x, y)$ denotes the covariance between two observations separated by spatial lag (x, y) . Data coming from this spatial process (with $\gamma \neq 1.0$) yields an anisotropic correlation whose spatial dependence in the x direction is stronger than that in the y direction when $\gamma > 1$. Note that, for example, when $\gamma = 4$ that points $(0, 1)$ and $(2, 0)$ have the same correlation with a point at the origin $(0, 0)$. Suppose four observations have been made of a stationary random field following this covariance structure at locations $(2/3, 0)$, $(0, 2/3)$, $(-2/3, 0)$ and $(0, -2/3)$, and the goal is prediction of the value at the point $(0, 0)$ which is equally distant from the observed four points. Under an assumed isotropic model ($\gamma = 1$) the optimal prediction weights are equal to 0.25 for each of the four observations. This is true for *any* isotropic covariance model.

Ignoring the correct anisotropic structure has three potentially undesirable effects:

1. A researcher who assumes isotropy a priori would not detect this dependence of correlation on direction is present and thus miss the true underlying spatial structure. This could lead to improper understanding of the physical phenomenon under study. In many cases, the lack of isotropy leads to a search for reasons for the directional effects.
2. Kriging weights and variances are greatly affected by anisotropy. Specifically, we can calculate the squared prediction error for any prediction weights. Further, we can calculate the optimal kriging weights for various γ . These are given in the 2nd and 3rd columns of Table 1. Finally, we can calculate the squared prediction error using this (assumed to be) correct anisotropic covariance using the optimal kriging weights as given in the 4th column. To see the effect of ignoring the true anisotropic structure, we also compute the correct prediction variance obtained using the incorrect equal weights. These variances are given in the 5th column.

There are two main effects due to anisotropy in this case. We observe the following:

- a. The kriging weights become more unequal as the degree of anisotropy increases. Note, in particular, that a large degree of anisotropy ($\gamma = 6$) actually leads to negative weights on observations and locations 2 and 4.
 - b. As the degree of anisotropy increases, the true kriging variances become relatively worse under an incorrectly assumed isotropy as can be seen in the 6th column.
3. In a spatial regression model

$$y = \mathbf{X}\beta + \epsilon, \text{Var}(\epsilon) = \Sigma.$$

The OLS estimator, $\widehat{\beta}_{OLS} := (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, has

$$\text{Var}(\widehat{\beta}_{OLS}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Sigma\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

while the GLS estimator, $\widehat{\beta}_{GLS} := (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$, has

$$\text{Var}(\widehat{\beta}_{GLS}) = (\mathbf{X}' \Sigma^{-1} \mathbf{X})^{-1}.$$

The OLS estimator does not require knowledge of Σ while the GLS does. We see, however, that for either the GLS or the OLS estimator, in order to draw inferences using $\widehat{\beta}$ concerning the regression parameter vector β , the matrix Σ needs to be estimated. The accuracy of inferences depends on the accuracy in estimation of the matrix Σ . Thus, we see the appropriateness of inferences drawn to regression parameters depends on correctly specifying the correlation structure. Specifically, if an anisotropic covariance function is estimated parametrically under an incorrectly assumed isotropic model then the resulting standard errors are generally unreliable. This leads to inappropriate inferences. For these three reasons it is important to detect anisotropy when present and properly account for anisotropy in our analyses.

The covariance function or covariogram between observations at locations, s and t , is $C(s, t) = E[(Z(s) - \mu(s))(Z(t) - \mu(t))]$. We assume that the random field is strictly stationary. In particular, this means that any mean function has been accounted for. Specifically, second order stationarity implies that $C(s, t) = C(0, t - s)$ and thus we can write the covariance function as a function of one argument. Formally, our setup is then the following.

Given the observed sample from a random field $\{Z(s) : s \in R^2\}$, we want to test

$$H_0: C(t) = C(\|t\|) \text{ for all } t,$$

i.e., the covariance depends only on the length of the lag vector $\|t\|$ between the two locations and not on the orientation.

2 Definitions and Large Sample Results

2.1 Stochastic Sampling Design

Let $\{Z(s) : s \in R^2\}$ be a random field. Irregularly spaced observations often occur in spatial applications. For this reason, we allow the locations to be irregularly spaced. Formally, we assume that the locations at which $Z(s)$ is generated are random and follows a stochastic sampling design. Specifically, we consider the class of stochastic designs presented by Lahiri and Zhu (2006). Let U_0 be a open connected subset of $(0, 1]^d$ and R_0 be a Borel set satisfying $U_0 \subset R_0 \subset \bar{U}_0$ where $\bar{U}_0$ denotes the closure of U_0 . Lahiri and Zhu (2006) considered R_0 to be a prototype for the sampling region R_n , that is,

$$R_n = \lambda_n R_0,$$

where $\{\lambda_n\}$ is a sequence of positive numbers such that $\lambda_n = o(n^{-\epsilon})$ for some $\epsilon > 0$. In other words λ_n plays the role of a scaling factor while keeping the shape of R_n the same for different values of n . Let $f(x)$ be a probability density function on R_0 and let $X_1, \dots, X_n$ be a sequence of independent and identically distributed random vectors with density $f(x)$. We assume that the sampling sites $s_1, \dots, s_n$ are obtained from a realization $x_1, \dots, x_n$ of the random vectors $X_1, \dots, X_n$ following the relation, for $i = 1, \dots, n$

$$s_i = \lambda_n x_i.$$

We observe the sample $\{Z(s_1), \dots, Z(s_n)\}$ along with the sampling sites. This is the so-called stochastic design as opposed to the common equally spaced assumption on the sampling sites. Note that this formulation allows for a large variety of sampling locations. For example, if $f(x)$ is the uniform density then the sampling locations follow a homogeneous Poisson process. For nonuniform densities, however, we have locations according to an inhomogeneous Poisson process with intensities given by $f(x)$. Further, we allow for data to be observed on a large class of domain shapes given by R_0 .

2.2 Covariogram Estimation and Asymptotic Properties

Our testing procedure, which we describe in Section 3, requires us to estimate the covariogram $C(t)$ for a given set of lag-vectors $t_1, \dots, t_K \in R^2$. For a given lag-vector t , we propose to estimate $C(t)$ using a nonparametric kernel method similar to Hall and Patil

(1994). Let $K(t)$ denote a kernel function and let h be a bandwidth. Let $\bar{Z} = n^{-1} \sum Z(s_i)$, $Z_{ij} = \{Z(s_i) - \bar{Z}\} \{Z(s_j) - \bar{Z}\}$ and $s_{ij} = s_i - s_j$. We estimate $C(t)$ by

$$\widehat{C}(t) = \left[\sum_{ij} K\{(t - s_{ij})/h\} \right]^{-1} \left[\sum_{ij} K\{(t - s_{ij})/h\} Z_{ij} \right].$$

Under regularity conditions, Hall and Patil (1994) showed that for any given $c > 0$,

$$\lambda_n^2 \int_{|t| \leq c} \{\widehat{C}(t) - C(t)\}^2 dt \rightarrow \int_{|t| \leq c} Y^2(t) dt$$

in distribution, where $Y(t)$ is a nonstationary Gaussian random field. However, we require the joint distribution of $\widehat{C}(t_1), \dots, \widehat{C}(t_K)$

For a given set of lag-vectors $t_1, \dots, t_K \in R^2$, let $G = [C(t_1), \dots, C(t_K)]$ denote the vector of true values of the covariogram and define $\hat{G}_n = [\widehat{C}(t_1), \dots, \widehat{C}(t_K)]$ to be the nonparametric estimates. We show the following result.

Theorem 1—Under regularity conditions stated in the appendix, we have

$$\lambda_n (\hat{G}_n - G) \rightarrow \text{Normal}(0, \mathbf{V})$$

in distribution as $n \rightarrow \infty$, where $\mathbf{V}$ is the (standardized) $K \times K$ covariance matrix of $\hat{G}_n$ with (k, k') -th element given by

$$V_{kk'} = c(f) \int E [\{Z(0)Z(t_k) - C(t_k)\} \{Z(u)Z(u+t_{k'}) - C(t_{k'})\}] du, \quad (1)$$

where $c(f) = (\int f^2)^{-2} \int f^4$

See the Appendix for a proof.

Remark 1—We have stated Theorem 1 specifically assuming that the lag vectors $t_1, \dots, t_K \in R^2$. However, a more general result can be stated for $t_1, \dots, t_K \in R^d$, $d \geq 2$ as well. In fact, for general dimension d , the result will stay the same except the scaling factor of $(\hat{G}_n - G)$ changes to $\lambda_n^{d/2}$

3 Testing for Spatial Isotropy

Recall that, given the observed sample from a random field $\{Z(s) : s \in R^2\}$, we want to test

$$H_0: C(t) = C(\|t\|) \text{ for all } t,$$

i.e., the covariance depends only on the length of the lag vector $\|t\|$ between the two locations and not on the orientation.

There are several graphical methods to assess directional effects in covariances. The idea is to examine the direction specific sample variograms or empirical covariance functions informally, see e.g., Diggle (1981) and Isaaks and Srivastava (1989). While often informative, these graphical methods are difficult to assess and open to interpretation.

There is a growing literature on testing for isotropy in various spatial settings. A formal testing procedure for investigating isotropy was proposed by Baczowski and Mardia (1990) under an assumed “doubly geometric” model. Cabana (1987) developed an F test for isotropy for a class of affine processes. Lu and Zimmerman (2001) developed tests based on the asymptotic distribution of sample semi-variograms. Guan et al. (2004) proposed a formal test for isotropy for a class of random fields where point locations were assumed to be equally spaced or generated from a homogeneous random (i.e., Poisson) process. The primary idea in Lu and Zimmerman and Guan et al. is to compare the sample variogram for a given set of lags and utilize the asymptotic normality of the variogram estimates.

To test H_0 , the null hypothesis of isotropy, we will utilize Lu and Zimmerman’s formulation, but in a general setting where sample locations are generated from a density $f(\cdot)$. We propose to compare values of $C(t)$ for $t \in \Lambda$, where $\Lambda = \{t_1, \dots, t_K\}$ is set of user chosen spatial lags. In general, we test for

$$AG=0,$$

where $G = \{C(t) : t \in \Lambda\}$ and A is a known matrix. For example, we may take $\Lambda = \{(0, 1), (1, 0), (1, 2), (2, 1)\}$ and test for

$$C\{(0, 1)\} - C\{(1, 0)\} = 0 \text{ and } C\{(1, 2)\} - C\{(2, 1)\} = 0.$$

In this case,

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Using Theorem 1, we have that

$$\lambda_n A (\widehat{G}_n - G) \rightarrow \text{Normal}(0, A V A^T),$$

in distribution. We propose to use the test statistic

$$T_n = (\widehat{A G}_n)^T (\widehat{A V}_0 A^T)^{-1} (\widehat{A G}_n),$$

where $\widehat{\mathbf{V}}_0$ is an estimate of $\lambda_n^{-2}\mathbf{V}$.

Utilizing the asymptotic formula for $\mathbf{V}$ to construct $\widehat{\mathbf{V}}_0$ is not a simple matter since the density $f(\cdot)$ and the inflation factor λ_n are generally unknown. For this reason we consider bootstrap methods.

There is a large literature in model free bootstrapping and subsampling for temporally and spatially correlations observations. Kunsch (1989) and Liu and Singh proposed the block bootstrap for stationary time series while Carlstein (1986) justified subsampling for variance estimation in this setting. Subsampling methodology for equally spaced spatial data has been developed in Possolo (1991), Politis and Romano (1994), Sherman (1996) and Lahiri, Kaiser, Cressie and Hsu (1999). Politis and Sherman (2001), and Politis, Paparoditis and Romano (1998) considered unequally spaced data with locations generated by stationary processes. In our setting, with locations generated by a general $f(\cdot)$ these methods may not be reliable.

Instead, we use a particular spatial bootstrap, developed by Lahiri and Zhu (2006) to construct $\widehat{\mathbf{V}}_0$. They consider two resampling procedures, namely the data site shifted block bootstrap (DSSBB) and grid based block bootstrap (GBBB) methods. We use the GBBB to estimate $\mathbf{V}$ as it allows for nonuniformity in the spatial sampling design. Specifically, we carry out the GBBB algorithm on the data set B times. For $b = 1, \dots, B$, let $\hat{G}_{n,b}$ denote the estimate of G based on the b -th bootstrapped data set. Then we set $\widehat{\mathbf{V}}_0$ equal to the sample covariance of the bootstrap replicates $\{\hat{G}_{n,1}, \dots, \hat{G}_{n,B}\}$. The effectiveness of this procedure does depend on the choice of block size, we examine the effects of this choice in our simulation study in the following section.

Given the consistency of $\lambda_n^2 \widehat{\mathbf{V}}_0$ for $\mathbf{V}$ we have, under the null hypothesis, that T_n has a large sample χ_v^2 distribution, where v denotes the row rank of $\mathbf{A}$. Hence we can construct p-values using χ_v^2 critical values.

4 Simulation Study

In this section, we conduct a simulation study to assess the performance of our proposed test. We generate realizations from a zero mean second order stationary Gaussian random field using a specified covariance function. We consider two different covariance functions:

$$\begin{aligned} C_1(t) &= \exp(-r); \\ C_2(t) &= \exp\left(-r/8 - r^2/8\right), \end{aligned}$$

with $r = (\mathbf{t}^T \mathbf{M} \mathbf{t})^{1/2}$, where $\mathbf{M}$ is a 2×2 positive definite rotation matrix inducing geometric anisotropy in the random field. The graphs of $C_1(r)$ and $C_2(r)$ for different values of r in $(0, 8)$ are given in Figure 1.

The matrix $\mathbf{M}$ is determined by alignment of the main anisotropic axes compared to the x-y axes (counter clock-wise rotation angle) and the anisotropic ratio (the ratio of lengths of the main axes). Let $\mathbf{I}_k$ denote the $k \times k$ identity matrix and let $\mathbf{J}_k$ denote the $k \times k$ matrix with all entries equal to one. We consider seven different choices of $\mathbf{M}$:

C1: Isotropic (rotation = 0, anisotropic ratio = 1). $\mathbf{M} = \mathbf{I}_2$.

C2: Rotation = 0, anisotropic ratio = $\sqrt{2}$ $\mathbf{M} = \text{diag}(1, 2)$.

C3: Rotation = $3\pi/4$, anisotropic ratio = $\sqrt{2}$. $\mathbf{M} = 2\mathbf{I}_2 - 0.5\mathbf{J}_2$.

C4: Rotation = 0, anisotropic ratio = 2. $\mathbf{M} = \text{diag}(1, 4)$.

C5: Rotation = $3\pi/4$, anisotropic ratio = 2. $\mathbf{M} = 4\mathbf{I}_2 - 1.5\mathbf{J}_2$.

C6: Rotation = 0, anisotropic ratio = 3. $\mathbf{M} = \text{diag}(1, 9)$.

C7: Rotation = $3\pi/4$, anisotropic ratio = 3. $\mathbf{M} = 9\mathbf{I}_2 - 4\mathbf{J}_2$.

The sample locations, $s = (x, y)$, are generated according to the density $f_{xy} = f(x)f(y)$. We consider two different choices of design density $f(\cdot)$:

$$\begin{aligned} \text{Uniform: } f(\cdot) &= U(0, 1) \quad ; \\ \text{Truncated normal mixture: } f(\cdot) &= 0.5TN(\mu=1/2, \sigma=1/2, \ell=0, u=1) \\ &\quad + 0.5TN(\mu=3/4, \sigma=\sqrt{2}, \ell=0, u=1), \end{aligned}$$

where $TN(\mu, \sigma, \ell, u)$ denotes a normal density with mean μ and standard deviation σ truncated in $[\ell, u]$. We use two combinations of sample sizes and scaling factors: (1) $n = 350$, $\lambda = 6$ and (2) $n = 600$, $\lambda = 8$. These two situations allow for approximately the same number of observations per unit area.

For each of the above combinations, we construct our test statistic based on three sets of lags

$$\begin{aligned} \Lambda_1 &= \{(1, 0), (0, 1)\} ; \\ \Lambda_2 &= \{(1, 0), (0, 1), (1, 1)/\sqrt{2}, (1, -1)/\sqrt{2}\} ; \\ \Lambda_3 &= \{(1, 0), (0, 1), (1, 1)/\sqrt{2}, (1, -1)/\sqrt{2}, (2, 1)/\sqrt{5}, (1, 2)/\sqrt{5}, (-2, 1)/\sqrt{5}, (-1, 2)/\sqrt{5}\}, \end{aligned} \quad (2)$$

and we compare $(1, 0)$ with $(0, 1)$, $(1, 1)/\sqrt{2}$ with $(1, -1)/\sqrt{2}$, $(2, 1)/\sqrt{5}$ with $(-1, 2)/\sqrt{5}$ and $(1, 2)/\sqrt{5}$ with $(-2, 1)/\sqrt{5}$.

We use the product kernel

$$K\{(s_1, s_2) - (x, y); (h_x, h_y)\} = K_0\{(s_1 - x)/h_x\} K_0\{(s_2 - y)/h_y\}$$

to estimate the covariance function, where $K_0(\cdot)$ denotes the Epanechnikov kernel and use the band-widths $h_x = \kappa \text{std}(D_x)N^{-1/5}$ and $h_y = \kappa \text{std}(D_y)N^{-1/5}$ where D_x and D_y denote the differences between x - and y -coordinates, respectively, N denotes the total number of such differences and κ is a scaling constant. The results displayed here correspond to $\kappa = 1$. We use block length, $b_n = 2$ for performing the spatial resampling and generated $B = 200$ resamples for each generated random field. Although these choices of bandwidth and block length are not theoretically optimal they give reasonable results as will be seen. The problem of finding the optimal bandwidth in covariance function estimation and finding the optimal block length in the GBBB spatial bootstrap, respectively, are open problems to the best of our knowledge and are beyond the scope of this paper.

The proportions of rejections at a nominal level of 0.05 for our test based on 500 simulation runs for each setting are given in Table 2 for the cases C1-C7 for different combinations of sample sizes ($n = 350$, $\lambda = 6$ and $n = 600$, $\lambda = 8$), covariance functions ($C_1(\cdot)$ and $C_2(\cdot)$), design densities (uniform and truncated normal).

Several key observations can be made from the results. To start with, the level of the test, regardless of the lag set used, remains very similar and is shown to be close to the nominal level. The power of the test using only two lags (Λ_1) is relatively good when there is no rotation (C2, C4, C6), but suffers greatly when there is rotation of the main anisotropic axis (C3, C5 and C7). This problem is not observed in the performance of the tests using four (Λ_2) or eight (Λ_3) spatial lags. Their powers increase as the anisotropic ratio increases. In fact, the powers for tests using four (Λ_2) and eight (Λ_3) lags are not much worse than the test based on Λ_1 even when there is no rotation, although Λ_1 is superior in the no rotation case. In addition, the power of all tests increases significantly as the sample size increases, except for the lag set Λ_1 in cases C3, C5 and C7, when this choice has almost no power beyond the size of the test. These conclusions are true for both the cases of the uniform and the truncated normal densities. This suggests that the performance of the test is not overly sensitive to the specific spatial design.

Remark

We have compared results from our test to that proposed in Guan et al. (2004). We have used their test statistic, $\widetilde{TS}_{IR,n}$, in the irregularly spaced location case, as defined in their Section 4.2.2. To construct the test statistic, we employed a product kernel with separate bandwidths, as in our test, with subsampling block length equal to 2. In the Uniform design with $n = 350$ and $\lambda = 6$, in observing the type I error (case C1) with covariance function $C_1(\cdot)$, we have seen inflated type I error for all three lags considered (0.122 for Λ_1 , 0.134 for Λ_2 and 0.125 for Λ_3) with a nominal level of 0.05. We observed similar results in other settings, as well. Moreover, for the truncated normal design, a case where Guan et al. approach is not theoretically supported, the inflation was seen to be even more severe. Compared to these results, our proposed test produces type I error rates closer to the nominal level.

4.1 Effect of Bandwidth and Block Size Choice

To assess the impact of bandwidth choice numerically, we consider the truncated normal design with $n = 600$, $\lambda_n = 8$ and covariance function $C_1(\cdot)$. We used $h_x = \kappa \text{std}(D_x)N^{-1/5}$, and $h_y = \kappa \text{std}(D_y)N^{-1/5}$ where D_x and D_y denote the differences between the x -coordinates and the y -coordinates, respectively, N denotes the total number of such differences and we consider four scaling constants, namely $\kappa = 0.5, 1.0, 1.5, 2.0$. The block size is fixed at $b_n = 2$. For each value of κ , we compute the power of the three test statistics for the cases C1 (isotropic) and C5. The results are displayed in Table 3. It is evident that the level of the test remains close to the nominal level over the range of bandwidths. For the anisotropic case C5, the power of the tests remains roughly similar for $\kappa = 1, 1.5$ and 2 but decreases appreciably for $\kappa = 0.5$. This indicates that excessive undersmoothing may not be a favorable option in this case.

To assess the effect of block length choice, we fix $\kappa = 1$ and consider different the values of $b_n = 1, 2$ and 4. We compute the power of the three resulting test statistics in the cases C1, C5 and C7 for each value of b_n . The results are displayed in Table 4. We see that for $b_n = 1$ and $b_n = 2$, the level of the test remains close to the nominal level for the case C1 and that the power of the test remains similar for C5 and C7. However, for the block length $b_n = 4$, the level of the test increases significantly. This suggests that choosing an overly large block length is detrimental to the level of the test.

5 Data Example

We test for isotropy in a data set on longleaf pine trees in an old-growth forest in the southern United States. This data set contains the diameters of 584 longleaf pine trees at

breast height along with their locations in a 200×200 grid. A bubble plot of the data set is displayed in Figure 2. Standard methods for the analysis of point patterns, namely using the K-function or nearest neighbor distances, strongly demonstrate that the locations of the trees do not follow a Poisson process. Strong clustering is clearly indicated. Further, inspection suggests that the density generating the locations appears to be quite non uniform, with a higher intensity of trees on the top (in the east) of the grid than on the bottom (in the west).

We apply our method to assess the assumption of isotropy for these data. We consider three sets of spatial lags

$$\begin{aligned}\Lambda_1 &= \{(1, 0), (0, 1)\}; \\ \Lambda_2 &= \{(1, 0), (0, 1), (1, 1), (1, -1)\}; \\ \Lambda_3 &= \{(1, 0), (0, 1), (1, 1), (1, -1), (2, 1), (1, 2), (-2, 1), (-1, 2)\},\end{aligned}$$

and we compare (1, 0) with (0, 1), (1, 1) with (1, -1), (2, 1) with (-1, 2) and (1, 2) with (-2, 1). We employ the product kernel

$$K\{(s_1, s_2) - (x, y); (h_x, h_y)\} = K_0\{(s_1 - x)/h_x\} K_0\{(s_2 - y)/h_y\}$$

to estimate the covariance function, where $K_0(\cdot)$ denotes the Epanechnikov kernel and use the band-widths $h_x = \text{std}(D_x)N^{-1/5}$, and $h_y = \text{std}(D_y)N^{-1/5}$ with D_x and D_y denoting the differences between x -coordinates and y -coordinates, respectively, and N denoting total number of such differences. We consider different values of block length, namely $b_n = 10, 20$ and 40 for performing the spatial resampling and for each block length, we generate 200 resamples. The p-values of our test are given in Table 5. It is evident that we can not reject the null hypothesis of isotropy for this data set.

To assess the effect of bandwidth on the test results, we used different bandwidths of the form $h_x = k_x \cdot \text{std}(D_x)N^{-1/5}$, and $h_y = k_y \cdot \text{std}(D_y)N^{-1/5}$ where we use different values of k_x and k_y ranging from 0.5 to 2.0. We compute p-values using each of these bandwidth combinations. The results are displayed in Figure 3 - Figure 5 for Λ_1 , Λ_2 and Λ_3 , respectively. We use a block size of 20 in this analysis. In each of these figures, we use black to plot p-values less than 0.05 and use white otherwise. It is evident that the results depend on the bandwidth combination used in all three cases. We see, however, that the dependence reduces appreciably when a larger number of lags are used (as in Λ_3 , Figure 5) compared to the other two lag sets.

Acknowledgments

This research was supported in part by NIH grant R00 ES017744. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Appendix

:

Assumptions and Conditions

We assume the following conditions are satisfied:

1. The stationary random field $Z(\cdot)$ is weakly dependent in the sense of Hall and Patil (1994). Specifically, see their conditions (3.2) and (3.3). Note that, these conditions

are satisfied when $Z(\cdot)$ is a Gaussian random field with $\text{cov}\{Z(t), Z(0)\} = 0$ for all sufficiently large $\|t\|$.

2. The kernel $K(\cdot)$ is bounded with compact support and is Hölder continuous with the property

$$\int t_i^j K(t) dt = \delta_{j0} \quad i=1, 2; j=1, \dots, r-1,$$

where r is the order of the kernel $K(\cdot)$, $t = (t_1, t_2)$ and $\delta_{j1,j2}$ is the Kronecker delta.

3. The sampling design density $f(\cdot)$ has a compact support and has r bounded derivatives.
4. For some $\eta > 0$, $h + \lambda_n^{-1} + h^{2r} \lambda_n^2 + (n^{1-\eta} h \lambda_n^{-2}) \rightarrow 0$

Sketch of Theorem 1

We will use several results from Hall and Patil (1994) throughout the proof. Let $\Lambda = (t_1, \dots, t_K)$ denote a set of lags. Recall that

$$\widehat{G}_n - G = [\widehat{C}(t_1) - C(t_1), \dots, \widehat{C}(t_K) - C(t_K)].$$

We start by noting that for any given t ,

$$\widehat{C}(t) - C(t) = \frac{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\} \{Z_{ij} - C(t)\}}{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\}} = R_1(t) + R_2(t),$$

where

$$R_1(t) = \frac{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\} \{Z_{ij} - C(s_{ij})\}}{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\}},$$

$$R_2(t) = \frac{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\} \{C(s_{ij}) - C(t)\}}{\sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\}}$$

We start by noting that we can write

$$R_2(t) = \{R_{21}(t)\}^{-1} \{R_{22}(t) - R_{21}(t) C(t)\}, \quad (3)$$

where

$$R_{21}(t) = \sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\};$$

$$R_{22}(t) = \sum_{ij} K\left\{\frac{(t - s_{ij})}{h}\right\} C(s_{ij}).$$

Using the expansions given in equations (4.7) and (4.8) of Hall and Patil (1994), we have that for each $c > 0$, uniformly in $\|t\| < c$,

$$R_{21}(t) = n(n-1)(\lambda^{-1}h)^2 \left\{ \int \int K(w) f(u) f\{u - \lambda^{-1}(t - hw)\} du dw + o(\lambda^{-1}) \right\}; \quad (4)$$

$$R_{22}(t) = n(n-1)(\lambda^{-1}h)^2 \left\{ C(t) \int \int K(w) f(u) f\{u - \lambda^{-1}(t - hw)\} du dw + o(\lambda^{-1}) \right\}. \quad (5)$$

Hence we derive

$$R_{22}(t) - R_{21}(t)C(t) = o\left\{n^2(\lambda^{-1}h)^2\lambda^{-1}\right\}$$

Next, in view of (4), it is easy to see that for $\|t\| < c$,

$$R_{21}(t) \sim n^2(\lambda^{-1}h)^2 \int f^2(x) dx. \quad (6)$$

Hence using (3) we derive

$$R_2(t) = o(\lambda^{-1}), \quad (7)$$

Next, by results in Hall and Patil (1994) (see their step (f) in the proof of Theorem 1), we have that

$$\left\{n^2(\lambda^{-1}h)^2 \left(\int f^2(x) dx\right)\right\}^{-1} \lambda R_{21}[R_1(t_1), \dots, R_1(t_k)] \rightarrow \text{Normal}(0, \mathbf{V})$$

in distribution as $n \rightarrow \infty$, where $\mathbf{V}$ is a $k \times k$ covariance matrix as given in (1). This, combined with (6), implies that

$$\lambda[R_1(t_1), \dots, R_1(t_k)] \rightarrow \text{Normal}(0, \mathbf{V}) \quad (8)$$

in distribution. Note that (7) implies that $\lambda R_2(t) = o(1)$. Hence, combined with (8) we have that

$$\lambda(\widehat{G}_n - G) \rightarrow \text{Normal}(0, \mathbf{V})$$

in distribution as $n \rightarrow \infty$. Hence the result.

References

- Baczkowski AJ, Mardia KV. A test of spatial symmetry with general application. *Communications in Statistics-Theory and Methods*. 1990; 19:555–572.
- Cabana EM. Affine processes: A test of isotropy based on level sets. *SIAM Journal of Applied Mathematics*. 1987; 47:886–891.
- Carlstein E. The use of subseries values for estimating the variance of a general statistic from a Stationary Sequence. *Annals of Statistics*. 1986; 14:1171–1179.
- Diggle PJ. Binary mosaics and the spatial pattern of heather. *Biometrics*. 1981; 37:531–539.

- Guan Y, Sherman M, Calvin J. A Nonparametric test for spatial isotropy using subsampling. *Journal of the American Statistical Association*. 2004; 99:810–821.
- Hall P, Patil P. Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields*. 1994; 99:399–423.
- Isaaks, EH.; Srivastava, RM. *An Introduction to Applied Geostatistics*. Oxford University Press; Oxford: 1989.
- Kunsch HR. The Jackknife and the bootstrap for general stationary observations. *Annals of Statistics*. 1989; 17:1217–1241.
- Lahiri SN, Kaiser MS, Cressie N, Hsu NJ. Prediction of spatial cumulative distribution functions using subsampling (with discussion). *Journal of the American Statistical Association*. 1999; 94:86–110.
- Lahiri SN, Zhu J. Resampling methods for spatial regression models under a class of stochastic designs. *Annals of Statistics*. 2006; 34:1774–1813.
- Liu, RY.; Singh, K. Moving blocks jackknife and bootstrap capture weak dependence. In: Lepage, R.; Billard, L., editors. *Exploring the Limits of Bootstrap*. Wiley; New York: 1992. p. 225-248.
- Lu, H.; Zimmerman, DL. Testing for isotropy and other directional symmetry properties of spatial correlation. 2001. Preprint
- Politis DN, Paparoditis E, Romano JP. Large sample inference for irregularly spaced dependent observations based on subsampling. *Sankhya*. 1998; 60:274–292.
- Politis DN, Romano JP. Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*. 1994; 22:2031–2050.
- Politis DN, Sherman M. Moment estimation for statistics from marked point processes. *Journal of the Royal Statistical Society, Ser. B*. 2001; 63:261–275.
- Possolo, A. Subsampling a random field. In: Possolo, A., editor. *Spatial Statistics and Imaging*. IMS; Hayward, CA: 1991. p. 286-294.
- Sherman M. Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society, Ser. B*. 1996; 58:509–523.

Figure 1. Graph of the covariance functions used for the simulation study described in Section 4. The solid line corresponds to $C_1(\cdot)$ and the dashed line corresponds to $C_2(\cdot)$.

Figure 2. Results for the analysis of the longleaf pine tree data. Displayed is the bubble plot of the longleaf pine tree data (each circle represents a tree with location given by the corresponding x and y coordinates and diameter proportional to the diameter of the corresponding tree).

Figure 3.

Results for analysis of longleaf pine tree data. Displayed are the p-values of our isotropy test for different choices of bandwidth when we use the lags given in Λ_1 . We use bandwidths of the form $h_x = k_x \cdot \text{std}(D_x) N^{-1/5}$ and $h_y = k_y \cdot \text{std}(D_y) N^{-1/5}$. The x-axis and y-axis represent the different values of the scaling factors (k_x and k_y) used in the bandwidths and the p-values less than 0.05 are plotted as black and white otherwise.

Figure 4.

Results for analysis of longleaf pine tree data. Displayed are the p-values of our isotropy test for different choices of bandwidth when we use the lags given in Λ_2 . We use bandwidths of the form $h_x = k_x \cdot \text{std}(D_x) N^{-1/5}$ and $h_y = k_y \cdot \text{std}(D_y) N^{-1/5}$. The x-axis and y-axis represent the different values of the scaling factors (k_x and k_y) used in the bandwidths and the p-values less than 0.05 are plotted as black and white otherwise.

Figure 5.

Results for analysis of longleaf pine tree data. Displayed are the p-values of our isotropy test for different choices of bandwidth when we use the lags given in Λ_3 . We use bandwidths of the form $h_x = k_x \cdot \text{std}(D_x) N^{-1/5}$ and $h_y = k_y \cdot \text{std}(D_y) N^{-1/5}$. The x-axis and y-axis represent the different values of the scaling factors (k_x and k_y) used in the bandwidths and the p-values less than 0.05 are plotted as black and white otherwise.

Table 1

The effect of isotropy on kriging weights and prediction variance. The degree of anisotropy is given by γ and λ_i 's are the kriging weights. "krvar" denotes the prediction variance using the true anisotropic covariance while "isovar" denotes the true prediction variance using an incorrectly assumed isotropic covariance.

γ	$\lambda_1 = \lambda_3$	$\lambda_2 = \lambda_4$	krvar	isovar	ratio
1.0	0.25	0.25	.2155	.2155	1.00
2.0	0.42	0.08	.2877	.3542	0.81
4.0	0.498	0.002	.3021	.5152	0.59
6.0	0.516	-.016	.3011	.5827	0.52

Table 2

Results of the simulation study described in Section 4. Displayed are the proportion of rejections (multiplied by 1,000) of the tests based on three lag-sets (Λ_1, Λ_2 and Λ_3) for the cases C1-C7 for different combinations of sample sizes ($n = 350, \lambda = 6$ and $n = 600, \lambda = 8$), covariance functions ($C_1(\cdot)$ and $C_2(\cdot)$) and design densities (uniform and truncated normal). All values are computed based on the nominal level of 0.05.

$n = 350, \lambda = 6$													$n = 600, \lambda = 8$					
$C_1(t)$			$C_2(t)$						$C_1(t)$						$C_2(t)$			
Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3				
Uniform design																		
C1	44	66	54	42	42	26	44	68	48	24	32	20						
C2	102	118	104	108	108	64	164	156	132	146	108	86						
C3	52	136	130	76	170	124	80	222	250	42	242	186						
C4	340	244	176	474	388	328	466	372	358	550	472	418						
C5	58	328	264	38	478	400	68	456	424	42	654	552						
C6	518	386	346	804	774	564	714	640	516	942	880	816						
C7	56	438	386	34	834	742	56	706	596	20	956	936						
Truncated normal design																		
C1	53	65	60	60	70	51	36	56	51	48	64	56						
C2	128	110	138	206	230	154	174	154	138	120	140	134						
C3	52	140	134	58	180	196	80	234	204	50	272	222						
C4	294	248	236	442	382	308	444	306	294	574	470	402						
C5	68	326	304	24	534	444	48	470	406	22	678	610						
C6	486	406	428	840	756	656	732	634	524	950	906	878						
C7	74	482	438	38	900	842	76	766	652	22	968	942						

Table 3

Results of the simulation study described in Section 4.1. Displayed are the proportion of rejections (multiplied by 1,000) for different bandwidth choices (corresponding to different values of κ) for the cases C1 and C5 under the truncated normal design with $n = 600$, $\lambda = 8$. All values are computed based on the nominal level of 0.05.

κ	Case: C1			Case: C5		
	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3
0.5	36	64	86	16	282	322
1.0	36	56	51	48	470	406
1.5	44	66	66	60	550	464
2.0	50	66	76	36	574	484

Table 4

Results of the simulation study described in Section 4.1. Displayed are the proportion of rejections (multiplied by 1,000) for different block-length choices (b_n) for the cases C1, C5 and C7 under the truncated normal design with $n = 600$, $\lambda = 8$. All values are computed based on the nominal level of 0.05.

b_n	Case: C1			Case: C5			Case: C7		
	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3	Λ_1	Λ_2	Λ_3
1	26	44	54	44	444	396	36	686	648
2	36	56	51	48	470	406	76	766	652
4	116	180	206	102	574	620	92	746	764

Table 5

Results of the analysis of the longleaf pine data. Displayed are the p-values for our test for different block sizes (b_n) and sets of lags (Λ) based on a nominal level of 0.05.

b_n	Λ_1	Λ_2	Λ_3
10	0.30	0.16	0.41
20	0.25	0.20	0.34
40	0.41	0.26	0.44