A Nonparametric Assessment of Properties of Space–Time Covariance Functions

Bo LI, Marc G. GENTON, and Michael SHERMAN

We propose a unified framework for testing various assumptions commonly made for covariance functions of stationary spatio-temporal random fields. The methodology is based on the asymptotic normality of space–time covariance estimators. We focus on tests for full symmetry and separability in this article, but our framework naturally covers testing for isotropy and Taylor's hypothesis. Our test successfully detects the asymmetric and nonseparable features in two sets of wind speed data. We perform simulation experiments to evaluate our test and conclude that our method is reliable and powerful for assessing common assumptions on space–time covariance functions.

KEY WORDS: Asymptotic normality; Covariance; Full symmetry; Random field; Separability; Stationarity.

1. INTRODUCTION

Modeling space–time data often relies on parametric covariance models (see, e.g., Cressie and Huang 1999; Gneiting 2002; Stein 2005) and various assumptions, such as full symmetry and separability. These assumptions are important because they simplify the structure of the model and its inference and ease the possibly extensive computational burden associated with space–time data sets. But they are not appropriate in situations involving a lack of full symmetry due to, for example, prevailing winds, water flows, and atmosphere circulation in geoscience, meteorology, and ecology. For example, Gneiting, Genton, and Guttorp (2007) suggested a lack of full symmetry and separability for the Irish wind data described by Haslett and Raftery (1989) based on graphical evidence. Jun and Stein (2007) developed a space–time asymmetric covariance function for modeling sulfate concentration levels to respect the apparent space–time asymmetry displayed in their data. Cressie and Huang (1999) suggested a nonseparable spatio-temporal covariance underlying a data set of tropical winds in the Pacific Ocean based on a plot of the empirical space–time variogram. The diagnostics used in those works are useful but difficult to assess and are open to interpretation. This points to the need for a unified approach for assessing properties of space–time covariance functions.

Let $\{Z(\mathbf{s}, t) : \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{R}\}$ be a strictly stationary space–time random field with covariance function $C(\mathbf{h}, u) = \mathrm{cov}\{Z(\mathbf{s}, t), Z(\mathbf{s} + \mathbf{h}, t + u)\}$, where $\mathbf{h}$ and u denote an arbitrary spatial lag and time lag. The random field $Z(\mathbf{s}, t)$ has a fully symmetric covariance function if $C(\mathbf{h}, u) = C(\mathbf{h}, -u)$ or if $C(\mathbf{h}, u) = C(-\mathbf{h}, u)$. Among the class of fully symmetric covariances, a covariance is separable if and only if $C(\mathbf{h}, u)/C(\mathbf{h}, 0) = C(\mathbf{0}, u)/C(\mathbf{0}, 0)$, that is, the space–time covariance can be factored into the product of a purely spatial covariance and a purely temporal covariance. Note that a separable covariance function must be fully symmetric, but full symmetry does not imply separability. Spatial isotropy restricts $C(\mathbf{h}, u)$ to be a function depending only on the Euclidean norm of $\mathbf{h}$ rather than on the lag vector $\mathbf{h}$ itself. The relationships among several assumptions on space–time covariance functions have been illustrated by Gneiting et al. (2007).

This article provides a framework to assess all of the foregoing assumptions at once, based on the asymptotic joint normality of sample space–time covariance estimators given in Section 2.2. Let $\mathbf{\Lambda}$ be a set of space–time lags, and let m denote the cardinality of $\mathbf{\Lambda}$. Let D be the domain of observations, and let $\widehat{C}(\mathbf{h}, u)$ denote an estimator of $C(\mathbf{h}, u)$ over D. Let $\mathbf{G} = \{C(\mathbf{h}, u), (\mathbf{h}, u) \in \mathbf{\Lambda}\}$, and let $\widehat{\mathbf{G}} = \{\widehat{C}(\mathbf{h}, u), (\mathbf{h}, u) \in \mathbf{\Lambda}\}$ denote the estimator of $\mathbf{G}$ over D. We derive that the appropriately standardized $\widehat{\mathbf{G}}$ has an asymptotic multivariate normal distribution for a random field with a fixed spatial domain and an increasing temporal domain. The derivation does not require any marginal or joint distributional assumptions other than mild moment and mixing conditions on the strictly stationary random field.

Many approaches have been proposed for testing specific properties on covariance functions under various assumptions. Mitchell, Genton, and Gumpertz (2006) proposed a likelihood ratio test for separability of covariance models in the context of multivariate repeated measures assuming the multinormality of observations. Mitchell et al. (2005) implemented this test in the spatio-temporal context, although in a somewhat ad hoc fashion. Scaccia and Martin (2002, 2005) presented a spectral method for testing the symmetry and separability for spatial lattice processes. Fuentes (2006) proposed a nonparametric test for separability of a spatio-temporal process also based on a spectral method. Lu and Zimmerman (2001) and Guan, Sherman, and Calvin (2004) developed nonparametric tests for spatial isotropy based on the asymptotic joint distribution of sample variograms. Lu and Zimmerman (2005) proposed some diagnostic tests for reflection symmetry and complete symmetry in spatial dependence based on the two-dimensional periodogram.

We develop a unified methodology for all these assumptions through contrasts of elements in $\widehat{\mathbf{G}}$ in a very general setting. For each hypothesis, we choose an appropriate contrast. Our testing approach can be widely applied and is easy to implement due to the nonparametric basis from which the test is derived and the simple form of the test statistics.

Bo Li is a Postgraduate Scientist, Geophysical Statistical Project, National Center for Atmospheric Research, Boulder, CO 80307 (E-mail: *boli@ucar.edu*). Marc G. Genton is Professor, Department of Econometrics, University of Geneva, CH-1211 Geneva 4, Switzerland (E-mail: *Marc.Genton@metri.unige.ch*), and Associate Professor, Department of Statistics, Texas A&M University, College Station, TX 77843 (E-mail: *genton@stat.tamu.edu*). Michael Sherman is Associate Professor, Department of Statistics, Texas A&M University, College Station, TX 77843 (E-mail: *sherman@stat.tamu.edu*). The National Center for Atmospheric Research is sponsored by the National Science Foundation. Genton acknowledges partial support from National Science Foundation grants DMS-0504896 and CMG ATM-0620624. The authors thank the joint editor, the associate editor, and the referees for constructive suggestions that have improved the content and presentation of this article. The authors also thank Dr. Christopher K. Wikle for providing the Pacific Ocean wind data.

Journal of the American Statistical Association
June 2007, Vol. 102, No. 478, Theory and Methods
DOI 10.1198/016214507000000202

The rest of the article is organized as follows. Section 2 describes our asymptotic regime and our test statistics and their large-sample distributions, and discusses several related issues. Section 3 applies our test for full symmetry and separability to the Irish wind and Pacific Ocean data sets mentioned earlier. Section 4 evaluates our tests by performing simulation experiments. Section 5 discusses our testing approach and extensions of our current examples and the Appendix gives a proof of Theorem 1.

2. ASSESSMENT OF VARIOUS COVARIANCE ASSUMPTIONS

2.1 Hypotheses

A very general form of hypothesis applied to many assumptions made for covariances is

$$H_0:\quad \mathbf{Af}(\mathbf{G}) = \mathbf{0}, \tag{1}$$

where $\mathbf{A}$ is a contrast matrix of row rank q and $\mathbf{f} = (f_1, \dots, f_r)^{\mathrm{T}}$ are real-valued functions that are differentiable at $\mathbf{G}$. For example, the null hypothesis of full symmetry and separability exactly follows this form. We choose these two specific assumptions to exemplify our methodology in testing a general class of hypotheses.

According to the definitions of full symmetry and separability, we give the null hypothesis for full symmetry, denoted by H_0^1, and the null hypothesis for separability, denoted by H_0^2:

$$H_0^1:\quad C(\mathbf{h}, u) - C(\mathbf{h}, -u) = 0, \qquad (\mathbf{h}, u) \in \mathbf{\Lambda},$$

$$H_0^2:\quad \frac{C(\mathbf{h}, u)}{C(\mathbf{h}, 0)} - \frac{C(\mathbf{0}, u)}{C(\mathbf{0}, 0)} = 0, \qquad (\mathbf{h}, u) \in \mathbf{\Lambda}.$$

Observe that H_0^1 and H_0^2 are contrasts of covariances and contrasts of ratios of covariances. Thus H_0^1 can be rewritten in the form of $\mathbf{A}_1\mathbf{G} = \mathbf{0}$, whereas H_0^2 can be rewritten as $\mathbf{A}_2\mathbf{f}(\mathbf{G}) = \mathbf{0}$ for a specified $\mathbf{\Lambda}$ and matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, where $\mathbf{f}$ takes pairwise ratios of elements in $\mathbf{G}$. For example, if

$$\mathbf{\Lambda} = \{(\mathbf{0}, 0), (\mathbf{h}_1, u_1), (\mathbf{h}_1, -u_1), (\mathbf{h}_2, u_2), (\mathbf{h}_2, -u_2), (\mathbf{h}_1, 0), (\mathbf{0}, u_1), (\mathbf{h}_2, 0), (\mathbf{0}, u_2)\},$$

that is,

$$\mathbf{G} = \big(C(\mathbf{0}, 0), C(\mathbf{h}_1, u_1), C(\mathbf{h}_1, -u_1), C(\mathbf{h}_2, u_2), C(\mathbf{h}_2, -u_2), C(\mathbf{h}_1, 0), C(\mathbf{0}, u_1), C(\mathbf{h}_2, 0), C(\mathbf{0}, u_2)\big)^{\mathrm{T}},$$

then

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{f}(\mathbf{G}) = \left(\frac{C(\mathbf{h}_1, u_1)}{C(\mathbf{h}_1, 0)}, \frac{C(\mathbf{h}_1, -u_1)}{C(\mathbf{h}_1, 0)}, \frac{C(\mathbf{h}_2, u_2)}{C(\mathbf{h}_2, 0)}, \frac{C(\mathbf{h}_2, -u_2)}{C(\mathbf{h}_2, 0)}, \frac{C(\mathbf{0}, u_1)}{C(\mathbf{0}, 0)}, \frac{C(\mathbf{0}, u_2)}{C(\mathbf{0}, 0)} \right)^{\mathrm{T}},$$

and

$$\mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

Under the corresponding null hypotheses, we have that $\mathbf{A}_1\mathbf{G} = \mathbf{0}$ and $\mathbf{A}_2\mathbf{f}(\mathbf{G}) = \mathbf{0}$. It is easily verified that H_0^1 and H_0^2 are simply two special cases of (1). In particular, for full symmetry and separability, it is noteworthy that once H_0^1 is rejected, H_0^2 is automatically rejected as well.

2.2 Test Statistics and Asymptotic Theory

Let $\widehat{C}_n(\mathbf{h}, u)$ denote a sample-based estimator of $C(\mathbf{h}, u)$ based on observations in a sequence of increasing index sets D_n, and let $\widehat{\mathbf{G}}_n = \{\widehat{C}_n(\mathbf{h}, u) : (\mathbf{h}, u) \in \mathbf{\Lambda}\}$ denote the estimator of $\mathbf{G}$ computed over D_n. We decompose D_n into $D_n = \mathcal{F} \times \mathcal{I}_n$, where $\mathcal{F} \subset \mathbb{R}^{d_1}$ and $\mathcal{I}_n \subset \mathbb{R}^{d_2}$. Suppose that $\mathcal{F}$ is a fixed space in the sense that finitely many observations are located within this space, and that $\mathcal{I}_n$ is an increasing space. We account for the shape of the space–time domain in which we observe data as done by Sherman (1996). Let $\mathcal{A}$ denote the interior of a closed surface contained in a d_2-dimensional cube with edge length 1, and let $\mathcal{A}_n$ denote the inflation of $\mathcal{A}$ by a factor n. We define $\mathcal{I}_n = \mathbb{Z}^{d_2} \cap \mathcal{A}_n$ if the observations are regularly spaced and $\mathcal{I}_n = \mathcal{A}_n$ otherwise. This formulation allows for a wide variety of space–time domains.

In many situations, the observations are taken from a fixed space $S \subset \mathbb{R}^d$ at regularly spaced times $T_n = \{1, \dots, n\}$. In this particular case, $d_1 = d$ and $d_2 = 1$ and we define the mixing coefficient (e.g., Ibragimov and Linnik 1971, p. 306)

$$\alpha(u) = \sup_{A,B} \big\{ |P(A \cap B) - P(A)P(B)|, A \in \mathfrak{F}_{-\infty}^0, B \in \mathfrak{F}_u^\infty \big\},$$

where $\mathfrak{F}_{-\infty}^0$ is the σ-algebra generated by the past time process until $t = 0$ and $\mathfrak{F}_u^\infty$ is the σ-algebra generated by the future time process from $t = u$. We assume that the mixing coefficient $\alpha(u)$ satisfies the strong mixing condition

$$\alpha(u) = \mathrm{O}(u^{-\epsilon}) \quad \text{for some } \epsilon > 0. \tag{2}$$

Let $S(\mathbf{h}) = \{\mathbf{s} : \mathbf{s} \in S, \mathbf{s} + \mathbf{h} \in S\}$, and let $|S(\mathbf{h})|$ be the number of elements in $S(\mathbf{h})$. We assume that the mean of Z is known and equal to 0 in the following theorem, to make both the theorem and proof concise. If we remove this assumption, we then let $\widehat{C}_n^*(\mathbf{h}, u)$ and $\widehat{\mathbf{G}}_n^*$ denote the mean-corrected estimators of $C(\mathbf{h}, u)$ and $\mathbf{G}$. It is easy to show that $\widehat{\mathbf{G}}_n^*$ and $\widehat{\mathbf{G}}_n$ have the same asymptotic properties. We define the estimator of C under the mean-0 assumption as

$$\widehat{C}_n(\mathbf{h}, u) = \frac{1}{|S(\mathbf{h})||T_n|} \sum_{\mathbf{s} \in S(\mathbf{h})} \sum_{t=1}^{n-u} Z(\mathbf{s}, t) Z(\mathbf{s} + \mathbf{h}, t + u),$$

and assume the moment condition for $\widehat{C}_n(\mathbf{h}, u)$,

$$\sup_n \mathrm{E}\big\{ \big| \sqrt{|T_n|} \{\widehat{C}_n(\mathbf{h}, u) - C(\mathbf{h}, u)\} \big|^{2+\delta} \big\} \le C_\delta \quad \text{for some } \delta > 0, C_\delta < \infty. \tag{3}$$

Theorem 1. Let $\{Z(\mathbf{s}, t), \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ be a strictly stationary spatio-temporal random field observed in $D_n = S \times T_n$, where $S \subset \mathbb{R}^d$ and $T_n = \{1, \dots, n\}$. Assume that

$$\sum_{t \in \mathbb{Z}} \big| \mathrm{cov}\{Z(\mathbf{0}, 0) Z(\mathbf{h}_1, u_1), Z(\mathbf{s}, t) Z(\mathbf{s} + \mathbf{h}_2, t + u_2)\} \big| < \infty \tag{4}$$

for all $\mathbf{h}_1, \mathbf{h}_2 \in S, \mathbf{s} \in S(\mathbf{h}_2)$ and all finite u_1 and u_2. Then $\boldsymbol{\Sigma} = \lim_{n\to\infty} |T_n| \mathrm{cov}(\widehat{\mathbf{G}}_n, \widehat{\mathbf{G}}_n)$ exists, the (i,j)th element of which is

$$\frac{1}{|S(\mathbf{h}_i)||S(\mathbf{h}_j)|} \sum_{\mathbf{s}_1 \in S(\mathbf{h}_i)} \sum_{\mathbf{s}_2 \in S(\mathbf{h}_j)} \sum_{t\in\mathbb{Z}} \mathrm{cov}\{Z(\mathbf{s}_1, 0)Z(\mathbf{s}_1 + \mathbf{h}_i, u_i), \\ Z(\mathbf{s}_2, t)Z(\mathbf{s}_2 + \mathbf{h}_j, t + u_j)\}.$$

If we further assume that $\boldsymbol{\Sigma}$ is positive definite and that conditions (2) and (3) hold, then $\sqrt{|T_n|}(\widehat{\mathbf{G}}_n - \mathbf{G}) \xrightarrow{d} \mathrm{N}_m(\mathbf{0}, \boldsymbol{\Sigma})$ as $n \to \infty$.

For the proof see the Appendix.

The conditions assumed for Theorem 1 to hold are relatively mild. The assumption on the temporal correlation given by (2) holds for a large class of temporal processes, for example, AR(1) processes with normal, double-exponential, or Cauchy errors. The moment condition, (3), is only slightly stronger than the existence of an asymptotic variance of the covariance estimator given in the theorem by $\boldsymbol{\Sigma}$.

In this theorem we allow the observations to be either regularly spaced or irregularly spaced in S. However, even for irregularly spaced observations, we consider only the covariances of observed spatial lags due to the often-limited number of observations in S. Note that in this section we require that the observations be taken at the same spatial locations over time, which is very common for monitoring stations; for example, this is the case for the Irish wind data and the Pacific Ocean wind data that we analyze in Section 3.

The asymptotic distribution of $\widehat{\mathbf{G}}_n$ in other space–time regimes can be stated in a similar fashion. For example, regimes can be regularly spaced observations with an increasing spatio-temporal domain, spatially irregularly spaced observations with an increasing spatio-temporal domain, or irregularly spaced observations with an increasing spatio-temporal domain. Under appropriate conditions, the asymptotic joint normality of $\widehat{\mathbf{G}}_n$ holds in all of these types of data structures. This guarantees the validity of our test over a wide variety of data structures under mild assumptions on the domain shape and the strength of dependence in the underlying space–time random field.

Replacing $\mathbf{G}$ with an estimator $\widehat{\mathbf{G}}_n$ in (1), we obtain a contrast vector for testing H_0 as the estimated left side of (1), $\mathcal{C} = \mathbf{Af}(\widehat{\mathbf{G}}_n)$. Specifically, the contrasts for testing H_0^1 and H_0^2 are given by

$$\mathcal{C}_1 = \widehat{C}_n(\mathbf{h}, u) - \widehat{C}_n(\mathbf{h}, -u), \qquad (\mathbf{h}, u) \in \boldsymbol{\Lambda},$$

and

$$\mathcal{C}_2 = \frac{\widehat{C}_n(\mathbf{h}, u)}{\widehat{C}_n(\mathbf{h}, 0)} - \frac{\widehat{C}_n(\mathbf{0}, u)}{\widehat{C}_n(\mathbf{0}, 0)}, \qquad (\mathbf{h}, u) \in \boldsymbol{\Lambda}.$$

Apparently, $\mathcal{C}_1$ and $\mathcal{C}_2$ can be rewritten in the form of $\mathbf{A}_1\widehat{\mathbf{G}}_n$ and $\mathbf{A}_2\mathbf{f}(\widehat{\mathbf{G}}_n)$.

It is then straightforward to obtain the asymptotic distribution of the test statistics based on the asymptotic joint normality of $\widehat{\mathbf{G}}_n$. By the multivariate delta theorem (e.g., Mardia, Kent, and Bibby 1979, p. 52), we have that

$$\sqrt{|T_n|}\{\mathbf{f}(\widehat{\mathbf{G}}_n) - \mathbf{f}(\mathbf{G})\} \xrightarrow{d} \mathrm{N}_r(\mathbf{0}, \mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{B}), \tag{5}$$

where $\mathbf{B}_{ij} = \partial f_j/\partial \mathbf{G}_i$, $i = 1, \dots, m, j = 1, \dots, r$. We form the test statistic (TS) based on the contrasts of $\mathbf{f}(\widehat{\mathbf{G}}_n)$ and obtain the distribution of TS under the null hypothesis as

$$\mathrm{TS} = |T_n|\{\mathbf{Af}(\widehat{\mathbf{G}}_n)\}^{\mathrm{T}}(\mathbf{AB}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{BA}^{\mathrm{T}})^{-1}\{\mathbf{Af}(\widehat{\mathbf{G}}_n)\} \xrightarrow{d} \chi_q^2 \tag{6}$$

for a matrix $\mathbf{A}$ with row rank q. The idea of using a quadratic form to assess the discrepancy between two vectors emerged from the work of Lu and Zimmerman (2001) and was later used by Guan et al. (2004) to test isotropy of a spatial random field. This idea lends itself naturally to testing general hypotheses of the form (1) for spatio-temporal random fields.

By (6), we have that under H_0^1, the contrast $\mathcal{C}_1$ yields the test statistic for full symmetry,

$$\mathrm{TS1} = |T_n|(\mathbf{A}_1\widehat{\mathbf{G}}_n)^{\mathrm{T}}(\mathbf{A}_1\boldsymbol{\Sigma}\mathbf{A}_1^{\mathrm{T}})^{-1}(\mathbf{A}_1\widehat{\mathbf{G}}_n) \xrightarrow{d} \chi_{q_1}^2,$$

and under H_0^2, the contrast $\mathcal{C}_2$ yields the test statistic for separability,

$$\mathrm{TS2} = |T_n|\big[\mathbf{A}_2\{\mathbf{f}(\widehat{\mathbf{G}}_n)\}\big]^{\mathrm{T}}(\mathbf{A}_2\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{BA}_2^{\mathrm{T}})^{-1}\big[\mathbf{A}_2\{\mathbf{f}(\widehat{\mathbf{G}}_n)\}\big] \\ \xrightarrow{d} \chi_{q_2}^2,$$

for matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ with row ranks q_1 and q_2. We estimate the matrix $\mathbf{B}$ empirically by replacing $\mathbf{G}$ with $\widehat{\mathbf{G}}_n$ in the estimation. For the other types of domain, D_n, the distribution of the test statistic can be derived in an analogous manner with the normalizing sequence $|\mathcal{I}_n|$ replacing $|T_n|$ and noting the corresponding change in $\boldsymbol{\Sigma}$.

2.3 Estimating $\boldsymbol{\Sigma}$

The covariance matrix $\boldsymbol{\Sigma}$ defined in Theorem 1 is usually unknown and thus must be estimated. Noting the large number of elements in $\boldsymbol{\Sigma}$, we apply a subsampling technique for this estimation. Specifically, if the data are observed over a fixed spatial domain S and an increasing time domain T_n as in Sections 3 and 4, then we form overlapping $S \times l(n)$ subblocks using a moving subblock window along time. Much research has addressed the choice of the optimal block length, $l(n)$, in the sense of minimizing the mean squared error (MSE) of estimators in various contexts (see, e.g., Lahiri 2003). For simplicity, we follow the approach of Carlstein (1986) to determine the block length $l(n)$. Although Carlstein gave a formula based on nonoverlapping blocks, we use all overlapping blocks. This reduces the variance of the variance estimator by a factor of 2/3 (Künsch 1989) and allows the use of slightly longer subblocks by a factor of $(3/2)^{1/3}$. The block length for a series of length n is then $l(n) = (\frac{2\gamma}{1-\gamma^2})^{2/3}(\frac{3n}{2})^{1/3}$, where we estimate γ by $\widehat{\gamma}_n = \widehat{C}_n(\mathbf{0}, 1)/\widehat{C}_n(\mathbf{0}, 0)$. This approach assumes that the statistic of interest is the sample mean and that the temporal correlation follows an AR(1) process with parameter γ. This procedure often works well in practice. For more fully model-free approaches, see the methods of, for instance, Lahiri (2003).

2.4 Choice of Contrast Matrices

In Section 2.1 we gave examples of the matrices used in testing the appropriate linear hypothesis. The choices are not unique, however, because we can choose a subset of rows from $\mathbf{A}_1$ or $\mathbf{A}_2$ to obtain new test statistics for their corresponding tests. For example, we can pick only the first two rows of $\mathbf{A}_2$ to form a new test statistic with an asymptotic distribution that follows χ_2^2. Although these tests will have approximately the same sizes, the power depends on the specific choice.

Generally, it is preferable to use lags combining small spatial and temporal lags, because typically covariance estimators of smaller lags are obtained over more observations than larger lags, and they play a more important role in making predictions over the random field (see, e.g., Stein 1999). Given sufficient data, however, we can also include a larger variety of lags in terms of both space and time to ensure that we are assessing the characteristics of the whole space–time random field. In addition, contingent on the understanding of the physical process, we should definitely take the features of the physical process into consideration while choosing testing lags. For example, if the random field is related to the wind or precipitation, then it is more appropriate to use the dominant wind direction as a guide for choosing space–time lags with strong correlation, because the wind plays an important role in governing the structure of this type of random field.

3. DATA ANALYSIS

Starting in this section, we focus our research on the assessment of full symmetry and separability. We apply our testing procedures to the Irish wind data and the Pacific Ocean wind data to illustrate how our methodology functions as a guide in choosing covariance models for the data.

3.1 Irish Wind Data

The Irish wind data described by Haslett and Raftery (1989) consist of time series of daily average wind speed at 11 synoptic meteorological stations in Ireland during the period 1961–1978. To normalize the data, we follow Haslett and Raftery (1989) and Gneiting (2002) by taking a square root transformation and subtracting the seasonal effects and the spatially varying mean from the wind speed to obtain velocity measures before performing our test.

Stein (2005) noted an apparent asymmetric property of the covariance function by viewing variogram plots. Gneiting et al. (2007) explored the validity of the assumptions of full symmetry and separability of the covariance function and then fitted a separable model, a fully symmetric model, and a general stationary covariance model on training data 1961–1970. They assessed the prediction performance using these three fitted models on the test period of 1971–1978. Their findings suggest that the data violate the assumptions of full symmetry and separability. However, their parameter estimate $\hat{\lambda} = .0573$ seems reasonably close to $\lambda = 0$, indicating that full symmetry holds. (See Sec. 4.1 for the formula involving λ and for its interpretation.)

To formally test the full symmetry and separability using the data, we choose 5 pairs of stations among the 55 pairs, and choose time lags $u = 1$ and 2 days, because the correlations for the velocity measures decay rapidly in time (Gneiting 2002). An apparently natural choice of the station pairs is the 5 pairs with the shortest spatial distance $\|\mathbf{h}\|$ among the 55 pairs; however, the prevailing westerly wind (e.g., Gneiting et al. 2007) suggests choosing the 5 pairs of stations with the smallest ratio of h_2/h_1, where h_1 and h_2 are the east–west component and the north–south component of the spatial lag $\mathbf{h}$. We adopt the latter choice in our test. To show the effects of testing lags so as to provide guidance on how to choose them, we compare the test based on the five station pairs with the smallest ratio to the tests based on the five pairs with the largest ratio and with the shortest $\|\mathbf{h}\|$. For each station pair, we choose time lags $u = 1$ and 2 days with the west station leading the east station, because the wind propagates from west to east. The three sets of station pairs are shown in Figure 1.

The set $\mathbf{\Lambda}$ contains 10 lags of the combination of 5 $\mathbf{h}$'s and 2 u's and the other lags introduced by these 10 lags in the test. Thus $m = r = 20$ for the full symmetry test and $m = 18$, $r = 12$ for the separability test, whereas $q = 10$ for both tests. The test statistics and p values for the full symmetry and separability tests based on the five station pairs with the smallest ratio are TS1 $= 262.7$ ($p = 0$) and TS2 $= 445.2$ ($p = 0$). Therefore, our test results thoroughly reject the assumptions of full symmetry and separability. This provides a theoretical basis for the statement of asymmetry of Stein (2005) and gives strength to the suggested model of Gneiting et al. (2007) that allows for a non–fully symmetric (and thus nonseparable) covariance function. Compared with the test based on the five station pairs with the largest ratio, which gives TS1 $= 20.2$ and TS2 $= 103.6$, and the test based on the five pairs with the smallest $\|\mathbf{h}\|$, which gives TS1 $= 132.9$ and TS2 $= 202.5$, the test based on our choice detects further departure from full symmetry and separability by

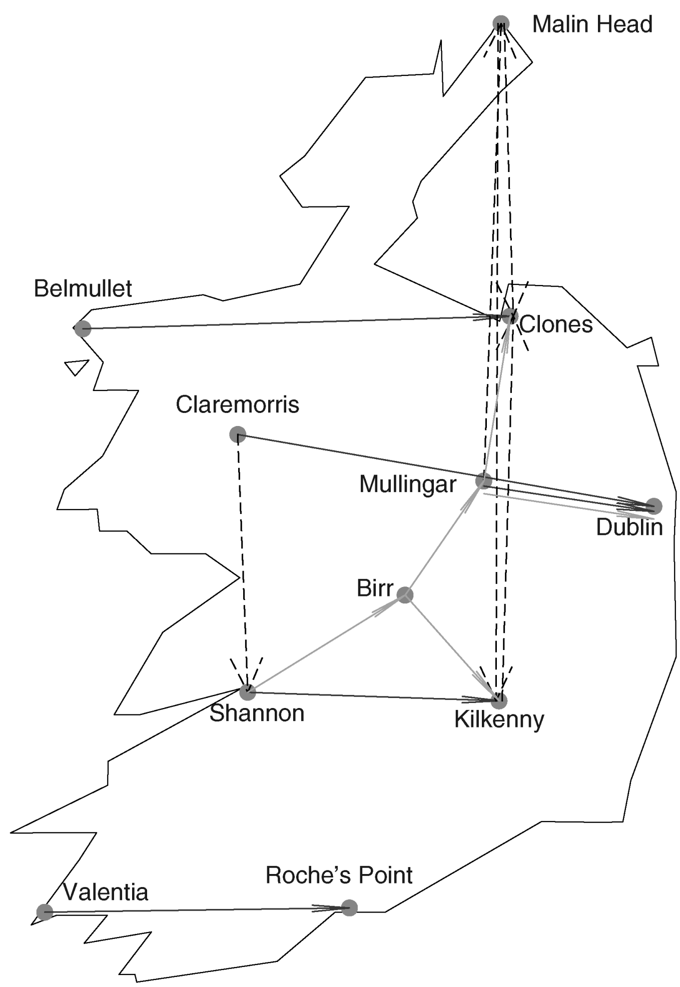

Figure 1. Irish Wind Stations and the Five Pairs Selected for the Test (— smallest h_2/h_1; – – largest h_2/h_1; — shortest $\|h\|$).

taking the wind direction into account. We note that if the east station is leading the west station in time for each pair, then the value of TS2 drops significantly. (The value of TS1 remains the same because H_0^1 is invariant under time reversal.) This implies that the power of the separability test is weakened by this choice of testing lags.

3.2 Pacific Ocean Wind Data

The Pacific Ocean wind data consist of the east–west component of the wind velocity vector from a region over the tropical western Pacific Ocean for the period from November 1992 to February 1993. The winds are given every 6 hours and at a 17×17 grid with grid interval of about 210 km. (See Wikle and Cressie 1999 for a more detailed description of the data.) Cressie and Huang (1999) graphically showed an apparent nonseparability of the spatio-temporal covariance through an examination of the empirical space–time variogram plot. Based on this, they fit several nonseparable covariance models to the data. To verify the need for fitting a nonseparable model, we performed our tests for full symmetry and separability to this wind data.

We remove the time-averaged mean for each grid location to create the mean-0 data set, as done by Wikle and Cressie (1999). We choose three sets of east–west spatial lags, each consisting of three distinct $\|\mathbf{h}\|$'s, and two sets of time lags, each consisting of five distinct u's. Because neither east wind nor west wind is clearly predominant (e.g., Cressie and Huang 1999), we simply choose u with the east grid location leading the west grid location in time. The degrees of freedom are $q = 15$ for all of the tests. The test results, given in Table 1, indicate that our tests provide strong evidence against separability for all testing lags and against full symmetry for small spatial lags. This corroborates the need for Cressie and Huang's nonseparable models. This table also clearly illustrates that different lags can lead to quite different p values for the test; for example, it is not easy to detect the asymmetric property of the covariance if we use large $\|\mathbf{h}\|$ rather than small $\|\mathbf{h}\|$. This is in accordance with the empirical spatio-temporal variogram of Cressie and Huang (1999). If we choose u with the opposite leading direction (i.e., the west grid location leading the east grid location in time), then the results are very similar to those in Table 1; however, the results change if we choose $\mathbf{h}$ in the north–south direction.

4. SIMULATION

We rejected full symmetry and separability in Sections 3.1 and 3.2. These conclusions will be strengthened if we assess the size and power of our testing procedures. To do this, we perform two simulations. In the first, we assess the test for full symmetry and separability for the data structure in Section 3.1. In the second, we study our test for separability for data on a grid as in Section 3.2 over a range of grid sizes, temporal lengths, and temporal correlations. Each situation is analyzed over 1,000 simulated space–time data sets.

4.1 Full Symmetry and Separability in the Irish Wind Data

We use the correlation model fitted to the Irish wind data of Gneiting et al. (2007) to simulate a space–time random field, $\mathbf{Z}$, of size $11 \times |T_n|$ at the 11 stations. Let $\mathbf{\Sigma}_0$ denote the spatio-temporal correlation matrix of the vectorized $\mathbf{Z}$, and let $\mathbf{Z}_0$ denote a vector of independent mean 0 standard normal variables. In this case $\mathbf{\Sigma}_0$ is of dimension $11n \times 11n$ and $\mathbf{Z}_0$ is of length $11n$, where $n = |T_n| = 3{,}650$. An apparent way to simulate $\mathbf{Z}$ is through $\mathbf{\Sigma}_0^{1/2}\mathbf{Z}_0$, where $\mathbf{\Sigma}_0^{1/2}$ can be computed using an eigenvalue decomposition. But, it is not practical to generate $\mathbf{Z}$ in this manner because of the large dimension of $\mathbf{\Sigma}_0$. To bypass this difficulty, we split the whole random field, $\mathbf{Z}$, into $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_{\lceil n/k \rceil}$, each of size $11 \times k$. We assume that temporal correlation exists only between the nearest neighbors of $\mathbf{Z}_i$, $i = 1, 2, \ldots, \lceil n/k \rceil$. This assumption is justified by the negligible temporal correlation when the time lag exceeds a certain value k. Here $k = 15$ suffices. Then the explicit form of the conditional distribution of two jointly normal random vectors allows us to generate the data consecutively. Specifically, we generate an initial 11×15 Gaussian random field $\mathbf{Z}_1$, and then, given $\mathbf{Z}_1 = \mathbf{z}_1$ generate $\mathbf{Z}_2|(\mathbf{Z}_1 = \mathbf{z}_1)$, $\mathbf{Z}_3|(\mathbf{Z}_2 = \mathbf{z}_2)$, and so on. This procedure frees us from the restriction of the dimension of $\mathbf{\Sigma}_0$, enabling generation of the space–time random field of a size similar to the training data set on which the model is fitted. The training period is 3,650 days, so we choose the space–time random field of size $11 \times 3{,}650$.

The spatio-temporal correlation model used by Gneiting et al. (2007) is given by

$$C(\mathbf{h}, u) = \frac{(1-\nu)(1-\lambda)}{1+a|u|^{2\alpha}}\left[\exp\left\{-\frac{c\|\mathbf{h}\|}{(1+a|u|^{2\alpha})^{\beta/2}}\right\} + \frac{\nu}{1-\nu}\delta_{\mathbf{h}=\mathbf{0}}\right] + \lambda\left(1 - \frac{1}{2v}|h_1 - \tilde{v}u|\right)_+,$$

where $(\cdot)_+ = \max(\cdot, 0)$ and the constants a and c are nonnegative scale parameters of time and space. The smoothness parameter α, the space–time interaction parameter β, the nugget parameter ν, and the symmetry parameter λ all take values in $[0, 1]$. The vector $\mathbf{h} = (h_1, h_2)^{\mathrm{T}}$ comprises an east–west component h_1 and a north–south component h_2. The scalar $\tilde{v} \in \mathbb{R}$ is an east–west velocity. When $\lambda = 0$, this model simplifies to a fully symmetric model. Further assuming that $\beta = 0$, this model reduces to a separable model. Note that β controls the degree of nonseparability.

We vary the parameters λ and β to study the behavior of our test procedure under different settings. Specifically, we determine the size and power of our tests by setting λ and β at the boundary values, and we also obtain the power of the test in Section 3.1 by setting $\lambda = .0573$ and $\beta = .681$, which are the

Table 1. Pacific Ocean Wind Data: Testing Full Symmetry and Separability

$\|\mathbf{h}\|$	u	Full symmetry		Separability	
		TS1	p value	TS2	p value
1, 2, 3	1, 3, 5, 7, 9	152.7	<1.0e−16	459.7	<1.0e−16
1, 5, 10	1, 3, 5, 7, 9	85.6	6.5e−12	205.5	<1.0e−16
10, 11, 12	1, 3, 5, 7, 9	27.6	.025	40.4	.0004
1, 2, 3	1, 2, 3, 4, 5	96.8	5.1e−14	458.1	<1.0e−16
1, 5, 10	1, 2, 3, 4, 5	87.5	2.9e−12	130.3	<1.0e−16
10, 11, 12	1, 2, 3, 4, 5	24.7	.054	49.7	1.4e−5

NOTE: Units of $\|\mathbf{h}\|$: grid interval; units of u: 6 hours.

Table 2. Empirical Sizes and Powers for Testing Full Symmetry and Separability in the Irish Wind Speed Data Based on Five Station Pairs With the Smallest Ratio h_2/h_1

λ	β	*Full symmetry*	*Separability*
0	0	Size = .074 [.067] (.083)	Size = .084 [.073] (.091)
0	1	Size = .070 [.081] (.092)	Power = .945 [1.000] (.503)
.2	0	Power = 1.000 [.141] (1.000)	Power = 1.000 [.510] (1.000)
.4	0	Power = 1.000 [.309] (1.000)	Power = 1.000 [.959] (1.000)
.0573	0	Power = .900 [.095] (.529)	Power = .985 [.125] (.445)
0	.681	Size = .095 [.075] (.078)	Power = .717 [.929] (.306)
.0573	.681	Power = .876 [.083] (.478)	Power = 1.000 [.941] (.867)

NOTE: The nominal level is .05. Sizes/powers in brackets are obtained by using the five station pairs with the largest ratio of h_2/h_1, and sizes/powers in parentheses are obtained by using the five station pairs with the shortest $\|\mathbf{h}\|$.

estimates of those two parameters using the training data set. We choose the same three sets of testing lags as in Section 3.1. The empirical sizes with respect to the nominal level .05 and powers are given in Table 2.

We first look at the results based only on our choice of testing lags. Table 2 shows that the size of the test is close to the nominal level and the power approaches 1 as λ or β increases. In particular, when λ and β are set to the fitted values from the Irish wind data, we have approximately 88% power against full symmetry ($\lambda = 0$) and 100% power against separability ($\lambda = 0$ and $\beta = 0$) under these given parameter values. It makes sense that we have greater power against the more restrictive hypothesis of separability, but we have strong confidence in our conclusion of rejecting full symmetry as well. Next, we compare the results from different sets of testing lags. The sizes from all three sets of testing lags are not appreciably different; however, the powers from the five station pairs with the largest ratio or the shortest $\|\mathbf{h}\|$ drop sharply in many scenarios. This implies that the physically motivated choice of space–time lags increases the reliability of the test.

4.2 Separability for Gridded Data

This simulation experiment is focused on evaluating our separability test for gridded data, such as the Pacific Ocean wind data. The Pacific Ocean wind data are collected over a medium-sized spatial grid; however, we are interested in studying small spatial grid sizes, which presents a more challenging setting for our test. Along with evaluating our test in terms of size and power, we assess the effect of estimation of the matrix $\boldsymbol{\Sigma}$ by $\widehat{\boldsymbol{\Sigma}}$ using subsampling, the grid size, the temporal size $|T_n|$, and the temporal correlation on our test. We also compare our test with the likelihood ratio test proposed by Mitchell et al. (2005). To implement all of these tasks, we choose a first-order vector autoregressive model, VAR(1), to simulate the random field, as done by de Luna and Genton (2005). This model allows us to obtain the asymptotic covariance matrix $\boldsymbol{\Sigma}$ explicitly (Priestley 1981, p. 693). In the VAR(1) process, $\mathbf{Z}_t = \mathbf{R}\mathbf{Z}_{t-1} + \boldsymbol{\epsilon}_t$, where $\mathbf{Z}_t = (Z(\mathbf{s}_1, t), Z(\mathbf{s}_2, t), \ldots, Z(\mathbf{s}_K, t))^{\mathrm{T}}$, K is the cardinality of S, $\boldsymbol{\epsilon}_t$ is a Gaussian multivariate white noise process with a spatially stationary and isotropic exponential correlation function, and $\mathbf{R}$ is a matrix of coefficients that determines the dependency between $\mathbf{Z}_t$ and $\mathbf{Z}_{t-1}$.

We first follow the setting of Mitchell et al. (2005) by choosing the spatial covariance $C(\mathbf{h}, 0) = \exp(-\frac{\|\mathbf{h}\|}{\phi})$, $\phi = 3.476$, and $\mathbf{R} = \rho\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. This produces random fields with a separable space–time covariance. We vary grid size, temporal size $|T_n|$, and temporal correlation parameter ρ in the simulation to assess their effect on the size of the test. We choose two lags ($\|\mathbf{h}\| = 1, u = 1$) and ($\|\mathbf{h}\| = 1, u = 2$) in the test, so the degrees of freedom of the test statistic are $q = 2$. Let p_1 denote the empirical size using the asymptotic covariance matrix $\boldsymbol{\Sigma}$, and let p_2 denote the empirical size using $\widehat{\boldsymbol{\Sigma}}$ estimated using the subsampling technique described in Section 2.3. The nominal level is set to be .05, and the results are summarized in Table 3. This table shows that the grid size does not appreciably affect the size of the test and that the test size is around .05 even with $|T_n| = 200$, whereas the parameter ρ seems to bring the size upward slightly as it increases. Figure 2 displays the distribution of p values for a 3×3 grid when $|T_n| = 1{,}000$. The boxplots indicate that the distributions of p_1 and p_2 are quite similar for each ρ, implying that estimating $\boldsymbol{\Sigma}$ does not strongly affect the test. Comparing the boxplots for each ρ shows that the performance of the separability test does not strongly depend on ρ.

To assess power in this setting, we set the coefficients in $\mathbf{R}$ as follows. For each $(\mathbf{s}_i, t)$, the coefficient is ρ for $(\mathbf{s}_i, t-1)$, whereas it is .05 for $\{(\mathbf{s}_j, t-1) : \|\mathbf{s}_j - \mathbf{s}_i\| = 1\}$, and 0 for the remaining $(\mathbf{s}, t-1)$'s. This produces nonseparable space–time covariances. The powers of the tests for a 3×3 grid are given in Table 4. As before, p_1 and p_2 denote powers under known $\boldsymbol{\Sigma}$ and estimated $\widehat{\boldsymbol{\Sigma}}$. Table 4 shows that although the power increases mildly as ρ increases, the power increases significantly as temporal size $|T_n|$ increases. Yet $\widehat{\boldsymbol{\Sigma}}$ has little effect on the power, especially when $|T_n|$ reaches 1,000.

To compare the size of our test to that of the likelihood ratio test of Mitchell et al. (2005), we further obtain $p_2 = .059$ for $\rho = .7$ and $p_2 = .080$ for $\rho = .9$ when $|T_n| = 200$ and the grid

Table 3. Empirical Sizes of the Separability Test for Gridded Data

		$\lvert T_n\rvert = 200$		$\lvert T_n\rvert = 500$		$\lvert T_n\rvert = 1{,}000$	
Grid size	ρ	p_1	p_2	p_1	p_2	p_1	p_2
3×3	.4	.033	.030	.046	.030	.052	.038
	.6	.036	.049	.046	.045	.049	.044
	.8	.040	.076	.043	.060	.053	.060
5×5	.4	.044	.033	.036	.028	.048	.032
	.6	.040	.044	.035	.042	.044	.038
	.8	.035	.057	.038	.053	.044	.048
7×7	.4	.045	.032	.046	.033	.048	.033
	.6	.042	.055	.052	.048	.045	.039
	.8	.042	.062	.058	.058	.046	.036

NOTE: Nominal level is .05. p_1 represents sizes based on asymptotic covariance matrix; p_2, sizes based on subsampling estimation of the covariance matrix.

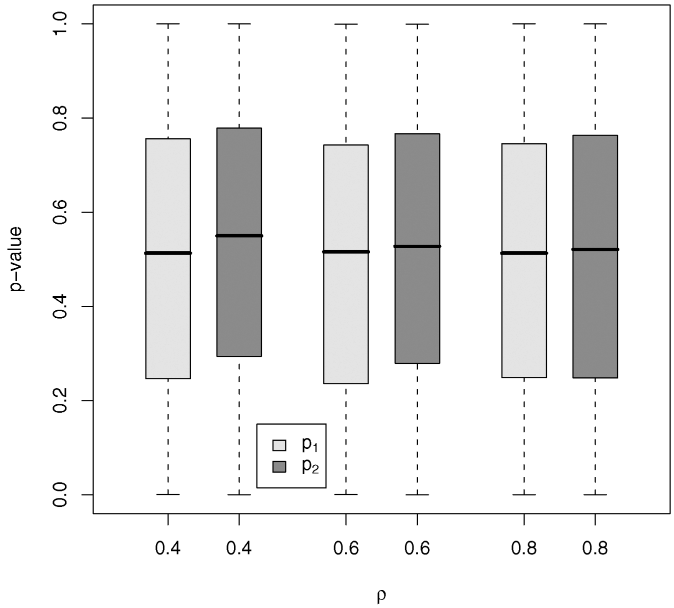

Figure 2. Distribution of p Values for the Separability Test for a 3 × 3 Grid, $|T_n| = 1{,}000$ (p_1: p values based on the asymptotic covariance matrix; p_2: p values based on the subsampling estimation of the covariance matrix).

size is 3 × 3. Comparing our p_2's with the values in Mitchell et al. (2005) shows that the size of our test is much closer to the nominal level for large ρ (i.e., for strong temporal correlation) than the test of Mitchell et al. (2005).

5. DISCUSSION

Working with an appropriate covariance model is of great importance in the analysis of space–time data and in the ability to make effective predictions. However, it is often not easy to decide which class of models we should choose due to the complex structure of the data. We have proposed a convenient method applicable for testing the full symmetry and separability of space–time covariance functions. Both the data analysis and simulation results demonstrated the reliability and accuracy of our method. Note that we discussed all of the tests in terms of covariance estimators. But tests retain the same properties if we replace the covariance estimators with correlation estimators, because all asymptotic distributions are in the unit-free chi-squared family of distributions. In addition, there are alternative covariance estimators to the moment estimator used in this article, for example, the kernel smoothed covariance estimator described by Hall, Fisher, and Hoffmann (1994).

Both H_0^1 and H_0^2 have no specific requirements for signs of covariances, so they can be applied to random fields with either positive or negative covariances, the latter are often exhibited in hydrological and meteorological applications (e.g., Stol 1983). Once the data are detected to be incompatible with the assumptions of full symmetry or separability, it is more appropriate to model the data using asymmetric or nonseparable covariance functions. Cressie and Huang (1999), Gneiting (2002), Stein (2005), and Gneiting et al. (2007) proposed rich classes of asymmetric or nonseparable covariance functions. They also presented details of a strategy for fitting an appropriate model by examining the data from different perspectives.

Table 4. Empirical Powers of the Separability Test for a 3 × 3 Grid

	$\lvert T_n\rvert = 200$		$\lvert T_n\rvert = 500$		$\lvert T_n\rvert = 1{,}000$	
ρ	p_1	p_2	p_1	p_2	p_1	p_2
.4	.468	.382	.829	.797	.984	.977
.6	.565	.512	.916	.904	.997	.997
.8	.697	.658	.992	.984	1.000	1.000

NOTE: p_1 represents powers based on asymptotic covariance matrix; p_2, powers based on subsampling estimation of the covariance matrix.

Our method can be naturally generalized to other tests, for example, testing for spatial isotropy at a specific time point. The null hypothesis in this case is

$$H_0^3: \quad C(\mathbf{h}_1, u) - C(\mathbf{h}_2, u) = 0, \quad (\mathbf{h}_1, u), (\mathbf{h}_2, u) \in \boldsymbol{\Lambda}, \mathbf{h}_1 \neq \mathbf{h}_2, \text{ but } \|\mathbf{h}_1\| = \|\mathbf{h}_2\|,$$

with the corresponding contrast

$$\mathcal{C}_3 = \widehat{C}_n(\mathbf{h}_1, u) - \widehat{C}_n(\mathbf{h}_2, u), \quad (\mathbf{h}_1, u), (\mathbf{h}_2, u) \in \boldsymbol{\Lambda}, \mathbf{h}_1 \neq \mathbf{h}_2, \text{ but } \|\mathbf{h}_1\| = \|\mathbf{h}_2\|.$$

Consider, for example,

$$\mathbf{A}_3 = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix},$$

in the context of $\boldsymbol{\Lambda}$ in Section 2.1. Then $H_0^3 : \mathbf{A}_3\mathbf{G} = \mathbf{0}$, and $\mathcal{C}_3$ can be rewritten in the form of $\mathbf{A}_3\widehat{\mathbf{G}}_n$, where $\mathbf{h}_1 \neq \mathbf{h}_2$, with $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$ and $u_1 = u_2$. Guan et al. (2004) used this setup to test for isotropy for spatial data using the variogram in place of the covariance function that we have considered here. The variogram has some natural advantages for purely spatial data (see, e.g., Cressie 1993), but the notion of separability for space–time processes is defined naturally in terms of the covariance function rather than the variogram function.

As another example, we have Taylor (1938)'s hypothesis, which addresses the relationship between the purely spatial and purely temporal covariances by examining whether there exists a velocity vector $\mathbf{v} \in \mathbb{R}^d$ such that $C(\mathbf{0}, u) = C(\mathbf{v}u, 0)$ for all u (see Gneiting et al. 2007 for a recent account). Cox and Isham (1988) discussed some restrictions for a specific covariance function to satisfy Taylor's hypothesis. If the vector $\mathbf{v}$ is known, then, analogously to Section 2.2, we write the null hypothesis as $H_0^4 : C(\mathbf{0}, u) - C(\mathbf{v}u, 0) = 0$ and the contrast as $\mathcal{C}_4 = \widehat{C}_n(\mathbf{0}, u) - \widehat{C}_n(\mathbf{v}u, 0)$. Given $\boldsymbol{\Lambda}$ including the involved lags, we can find a matrix, say $\mathbf{A}_4$, to make $\mathbf{A}_4\mathbf{G} = \mathbf{0}$ under the null hypothesis. The rest of the test is completely analogous to the previous tests. If the vector $\mathbf{v}$ is unknown, then the test could be used in an exploratory fashion by viewing the testing results over a range of vectors $\mathbf{v}$.

APPENDIX: PROOF OF THEOREM 1

Let $T_n(u) = \{t : t \in T_n, t + u \in T_n\}$. First, we have

$$\operatorname{cov}\{\widehat{C}_n(\mathbf{h}_i, u_i), \widehat{C}_n(\mathbf{h}_j, u_j)\} = \frac{1}{|S(\mathbf{h}_i)||T_n|}\frac{1}{|S(\mathbf{h}_j)||T_n|} \times \sum_{\mathbf{s}_1 \in S(\mathbf{h}_i)} \sum_{t_1 \in T_n(u_i)} \sum_{\mathbf{s}_2 \in S(\mathbf{h}_j)} \sum_{t_2 \in T_n(u_j)} \operatorname{cov}\{Z(\mathbf{s}_1, t_1)$$

$$\times Z(\mathbf{s}_1+\mathbf{h}_i, t_1+u_i), Z(\mathbf{s}_2, t_2)Z(\mathbf{s}_2+\mathbf{h}_j, t_2+u_j)\big\}$$

$$= \frac{1}{|S(\mathbf{h}_i)||S(\mathbf{h}_j)|}$$

$$\times \sum_{S(\mathbf{h}_i)}\sum_{S(\mathbf{h}_j)}\sum_{T_n(u_j)-T_n(u_i)} \operatorname{cov}\big\{Z(\mathbf{s}_1, 0)Z(\mathbf{s}_1+\mathbf{h}_i, u_i),$$

$$Z(\mathbf{s}_2, t)Z(\mathbf{s}_2+\mathbf{h}_j, t+u_j)\big\}\frac{|T_n(u_i)\cap\{T_n(u_j)-t\}|}{|T_n|^2}.$$

Applying condition (4) and Kronecker's lemma, we conclude that

$$|T_n|\times \operatorname{cov}\{\widehat{C}_n(\mathbf{h}_i, u_i), \widehat{C}_n(\mathbf{h}_j, u_j)\}$$

$$\to \frac{1}{|S(\mathbf{h}_i)||S(\mathbf{h}_j)|}\sum_{S(\mathbf{h}_i)}\sum_{S(\mathbf{h}_j)}\sum_{t\in\mathbb{Z}} \operatorname{cov}\big\{Z(\mathbf{s}_1, 0)Z(\mathbf{s}_1+\mathbf{h}_i, u_i),$$

$$Z(\mathbf{s}_2, t)Z(\mathbf{s}_2+\mathbf{h}_j, t+u_j)\big\}.$$

Letting $A_n = \sqrt{|T_n|}\{\widehat{C}_n(\mathbf{h}, u) - C_n(\mathbf{h}, u)\}$, we prove $A_n \xrightarrow{d} \mathrm{N}(0, \sigma^2)$ by applying a blocking technique and telescope arguments. Specifically, we divide the whole field into blocks along time; that is, there is no division in space S. Figure A.1 shows a schematic illustration when $d = 2$. Let $l(n) = n^\alpha$ and let $m(n) = n^\alpha - n^\eta$, for some $1/(1+\epsilon) < \eta < \alpha < 1$. Divide the original field D_n into nonoverlapping cylinders, $D^i_{l(n)} = S \times T^i_{l(n)}$, $i = 1, \dots, k_n$, where $|T^i_{l(n)}| = l(n)$; within each cylinder, further obtain $D^i_{m(n)} = S \times T^i_{m(n)}$, which shares the same center as $D^i_{l(n)}$. Thus $\operatorname{dist}(D^i_{m(n)}, D^{i'}_{m(n)}) \ge n^\eta$ for $i \ne i'$. Let $a_n = \sum_{i=1}^{k_n} a^i_n/\sqrt{k_n}$, $a'_n = \sum_{i=1}^{k_n}(a^i_n)'/\sqrt{k_n}$, where $a^i_n = \sqrt{m(n)}\{\widehat{C}^i_n - C\}$ and $(a^i_n)'$ have the same marginal distributions as a^i_n but are independent. Let $\phi'_n(x)$ and $\phi_n(x)$ be the characteristic functions of a'_n and a_n. The proof comprises the following three steps:

S1: $A_n - a_n \xrightarrow{p} 0$.
S2: $\phi'_n(x) - \phi_n(x) \to 0$.
S3: $a'_n \xrightarrow{d} \mathrm{N}(0, \sigma^2)$.

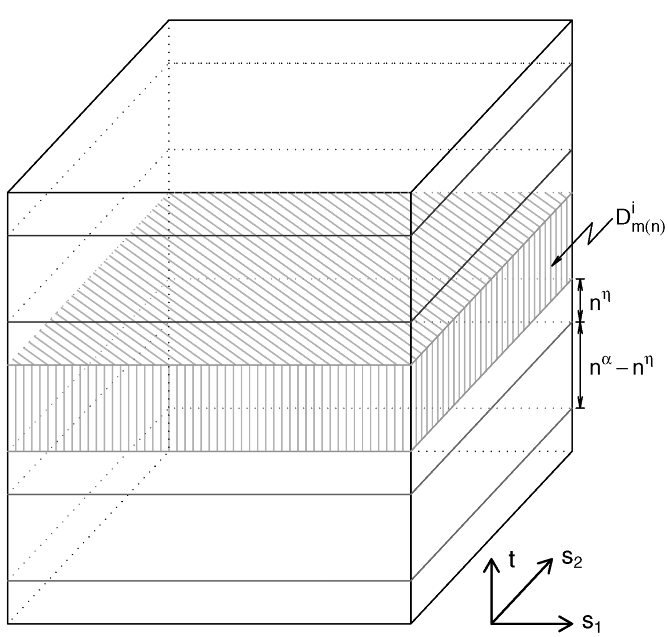

Figure A.1. Partition of the Space–Time Random Field for the Proof of Theorem 1.

Proof of S1

Let $D^{m(n)}$ denote the union of all $D^i_{m(n)}$, and let $T^{m(n)}$ denote the union of all $T^i_{m(n)}$. Specifically, $|T^{m(n)}| = k_n m(n)$ and $|D^{m(n)}| = |S||T^{m(n)}|$. Observe that

$$a_n = \frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n} a^i_n = \sqrt{|T^{m(n)}|}\big(\widehat{C}_{D^{m(n)}} - C\big)$$

and $\frac{|T_n|}{|T^{m(n)}|} \to 1$, $\frac{|D_n|}{|D^{m(n)}|} \to 1$. Thus we get $\operatorname{cov}(A_n, a_n) \to \sigma^2$ and $\operatorname{var}(A_n - a_n) \to 0$.

Proof of S2

We use telescope arguments here. Let ι denote the imaginary number. We define $U_i = \exp(\iota x \frac{a^i_n}{\sqrt{k_n}})$, $X_j = \prod_{i=1}^j U_i$ and $Y_j = U_{j+1}$. Extending theorem 17.2.1 (p. 306) and the telescope argument (p. 338) of Ibragimov and Linnik (1971) to our space–time context, we have

$$\operatorname{cov}(X_j, Y_j) \le 16\alpha(n^\eta) = \mathrm{O}(n^{-\epsilon\eta})$$

by (2) and

$$|\phi'_n(x) - \phi_n(x)| \le 16 k_n \mathrm{O}(n^{-\epsilon\eta}) = \mathrm{O}(n^{1-\alpha-\epsilon\eta}).$$

The last equality in the foregoing expression follows from $\mathrm{O}(k_n) = \mathrm{O}(\frac{n}{n^\alpha}) = \mathrm{O}(n^{(1-\alpha)})$. Because $1/(1+\epsilon) < \eta < \alpha < 1$, $1 - \alpha - \epsilon\eta < 1 - \eta - \epsilon\eta < 0$. Then $|\phi'_n(x) - \phi_n(x)| \to 0$.

Proof of S3

Observe that $\mathrm{E}(|(a^i_n)'|^{2+\delta}) < C_\delta$ for some constant C_δ. Because $(a^i_n)'$ are iid,

$$\operatorname{var}\left\{\sum_{i=1}^{k_n}(a^i_n)'\right\} = k_n \operatorname{var}\{(a^i_n)'\}.$$

Defining $\sigma^2_n = \operatorname{var}\{(a^i_n)'\}$, we have $\sigma^2_n \to \sigma^2$ from the proof of S1. Thus

$$\lim_{n\to\infty}\sum_{i=1}^{k_n}\frac{\mathrm{E}(|(a^i_n)'|^{2+\delta})}{\sqrt{[\operatorname{var}\{\sum_{i=1}^{k_n}(a^i_n)'\}]^{2+\delta}}} \le \lim_{n\to\infty} C_\delta \frac{k_n}{(k_n\sigma^2_n)^{(2+\delta)/2}} = 0.$$

Therefore, applying Lyapounov's theorem, we have

$$\frac{1}{\sqrt{k_n}}\sum_{i=1}^{k_n}(a^i_n)' \xrightarrow{d} \mathrm{N}(0, \sigma^2).$$

The Cramér–Wold device proves the joint normality.

[Received July 2006. Revised January 2007.]

REFERENCES

Carlstein, E. (1986), "The Use of Subseries Values for Estimating the Variance of a General Statistic From a Stationary Sequence," *The Annals of Statistics*, 14, 1171–1179.
Cox, D. R., and Isham, V. (1988), "A Simple Spatial-Temporal Model of Rainfall," *Proceedings of the Royal Society of London*, Ser. A, 415, 317–328.
Cressie, N. (1993), *Statistics for Spatial Data*, New York: Wiley.
Cressie, N., and Huang, H.-C. (1999), "Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions," *Journal of the American Statistical Association*, 94, 1330–1340.
de Luna, X., and Genton, M. G. (2005), "Predictive Spatio-Temporal Models for Spatially Sparse Environmental Data," *Statistica Sinica*, 15, 547–568.
Fuentes, M. (2006), "Testing for Separability of Spatial-Temporal Covariance Functions," *Journal of Statistical Planning & Inference*, 136, 447–466.
Gneiting, T. (2002), "Nonseparable, Stationary Covariance Functions for Space–Time Data," *Journal of the American Statistical Association*, 97, 590–600.
Gneiting, T., Genton, M. G., and Guttorp, P. (2007), "Geostatistical Space–Time Models, Stationarity, Separability and Full Symmetry," in *Statistics of Spatio-Temporal Systems*, Vol. 107, eds. B. Finkenstaedt, L. Held, and V. Isham, Boca Raton, FL: Chapman & Hall/CRC Press, pp. 151–175.

Guan, Y., Sherman, M., and Calvin, J. A. (2004), "A Nonparametric Test for Spatial Isotropy Using Subsampling," *Journal of the American Statistical Association*, 99, 810–821.

Hall, P., Fisher, N. I., and Hoffmann, B. (1994), "On the Nonparametric Estimation of Covariance Functions," *The Annals of Statistics*, 22, 2115–2134.

Haslett, J., and Raftery, A. E. (1989), "Space–Time Modelling With Long-Memory Dependence: Assessing Ireland's Wind-Power Resource" (with discussion), *Applied Statistics*, 38, 1–50.

Ibragimov, I. A., and Linnik, Y. V. (1971), *Independent and Stationary Sequences of Random Variables*, Gronigen, The Netherlands: Wolters-Noordhoff Publishing.

Jun, M., and Stein, M. L. (2007), "An Approach to Producing Space–Time Covariance Functions on Spheres," *Technometrics*, in press.

Künsch, H. (1989), "The Jackknife and Bootstrap for General Stationary Observations," *The Annals of Statistics*, 17, 1217–1241.

Lahiri, S. N. (2003), *Resampling Methods for Dependent Data*, New York: Springer-Verlag.

Lu, H., and Zimmerman, D. L. (2001), "Testing for Isotropy and Other Directional Symmetry Properties of Spatial Correlation," preprint.

——— (2005), "Testing for Directional Symmmetry in Spatial Dependence Using the Periodogram," *Journal of Statistical Planning & Inference*, 129, 369–385.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, New York: Academic Press.

Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2005), "Testing for Separability of Space–Time Covariances," *Environmetrics*, 16, 819–831.

——— (2006), "A Likelihood Ratio Test for Separability of Covariances," *Journal of Multivariate Analysis*, 97, 1025–1043.

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

Scaccia, L., and Martin, R. J. (2002), "Testing for Simplification in Spatial Models, in *COMPSTAT 2002*, Heidelberg: Physica, pp. 581–586.

——— (2005), "Testing Axial Symmetry and Separability of Lattice Processes," *Journal of Statistical Planning & Inference*, 131, 19–39.

Sherman, M. (1996), "Variance Estimation for Statistics Computed From Spatial Lattice Data," *Journal of the Royal Statistical Society*, Ser. B, 58, 509–523.

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

——— (2005), "Space–Time Covariance Functions," *Journal of the American Statistical Association*, 100, 310–321.

Stol, P. T. (1983), "Rainfall Interstation Correlation Functions: VII. On Non-Monotonousness," *Journal of Hydrology*, 64, 69–92.

Taylor, G. I. (1938), "The Spectrum of Turbulence," *Proceedings of the Royal Society of London*, Ser. A, 164, 476–490.

Wikle, C. K., and Cressie, N. (1999), "A Dimension-Reduction Approach to Space–Time Kalman Filtering," *Biometrika*, 86, 815–829.