

Nonseparable, Stationary Covariance Functions for Space-Time Data



NRCSE

Technical Report Series

NRCSE-TRS No. 063

May 23, 2001

The NRCSE was established in 1997 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.



Nonseparable, Stationary Covariance Functions for Space-Time Data

Tilmann Gneiting

Department of Statistics, University of Washington
Seattle, Washington 98195-4322, USA

Abstract

Geostatistical approaches to spatio-temporal prediction in environmental science, climatology, meteorology, and related fields rely on appropriate covariance models. This paper proposes general classes of nonseparable, stationary covariance functions for spatio-temporal random processes. The constructions are directly in the space-time domain and do not require Fourier inversion. The model parameters can be associated with the data's spatial and temporal structures, respectively; and a covariance model with a readily interpretable space-time interaction parameter is fitted to wind data from Ireland.

Keywords: Completely monotone; Correlation function; Geostatistics; Kriging; Positive definite; Separable; Spatio-temporal.

1 Introduction

Random process models for space-time data play increasingly important roles in various scientific disciplines; among them are environmental science, agriculture, climatology, meteorology, and hydrology. In the statistical literature, the recent works of Handcock and Wallis (1994), Sølna and Switzer (1996), Kyriakidis and Journel (1999), Christakos (2000), Christakos, Hristopulos, and Bogaert (2000), Brown, Diggle, Lord, and Young (2001), and Chilès, Bourgine, and Watremez (2001), among others, point at the significance of the approach. To fix the idea, we consider the random process model

$$Z(\mathbf{s}; t), \quad (\mathbf{s}; t) \in \mathbb{R}^d \times \mathbb{R},$$

for a real-valued variable observed at the space-time coordinates $(\mathbf{s}_1; t_1), \dots, (\mathbf{s}_k; t_k)$. The spatio-temporal variable might stand for atmospheric pollutant concentrations, soil parameters, temperature, or wind speed, to name but a few applications. More often than not, $Z(\mathbf{s}; t)$ will denote a transformed variable or a residual field, rather than the original variable. We refer to Kyriakidis and Journel (1999) for a comprehensive review and bibliography.

A statistical analysis typically aims at the optimal prediction of an unobserved part of the space-time process. Assuming that $Z(\mathbf{s}; t)$ has finite variance at all space-time coordinates $(\mathbf{s}; t) \in \mathbb{R}^d \times \mathbb{R}$, the mean function $\mu(\mathbf{s}; t) = E(Z(\mathbf{s}; t))$ and the covariance between $Z(\mathbf{s}; t)$ and $Z(\mathbf{s} + \mathbf{h}; t + u)$ exist. The *kriging predictor* of $Z(\mathbf{s}_0; t_0)$ is the linear combination

$$Z^*(\mathbf{s}_0; t_0) = \mu(\mathbf{s}_0; t_0) + \sum_{i=1}^k a_i \left(Z(\mathbf{s}_i; t_i) - \mu(\mathbf{s}_i; t_i) \right) \quad (1)$$

of the observations which minimizes the mean squared prediction error (Cressie 1993; Cressie and Huang 1999). The covariance structure of the spatio-temporal process $Z(\mathbf{s}; t)$ determines the weights

a_1, \dots, a_k of the individual observations in the predictor. It is then frequently assumed that the covariance structure is *stationary* in space and time, so that the covariance

$$\text{Cov}\left(Z(\mathbf{s}; t), Z(\mathbf{s} + \mathbf{h}; t + u)\right) = C(\mathbf{h}; u), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R},$$

depends on the space-time lag $(\mathbf{h}; u)$ only. The function $C(\mathbf{h}; u)$ is called the *covariance function* of the process, and its restrictions $C(\mathbf{h}; 0)$ and $C(\mathbf{0}; u)$ are purely spatial and purely temporal covariance functions, respectively. The assumption of stationarity in space and time needs to be assessed from case to case. For instance, Rodriguez-Iturbe, Marani, D’Odorico, and Rinaldo (1998, p. 3462) call stationarity “an important but reasonable hypothesis in the case of rainfall,” whereas Guttorp, Meiring, and Sampson (1997, pp. 407–408) dispute the assumption of temporal stationarity for the ozone-level data of Carroll et al. (1997). Spatial nonstationarity can often be dealt with by the space deformation approach of Sampson and Guttorp (1992), and we refer to Sampson, Damian, and Guttorp (2001) for recent developments and applications.

Under the assumption of stationarity, the kriging predictor (1) has variance

$$\text{Var}\left(Z^*(\mathbf{s}_0; t_0)\right) = \sum_{i=1}^k \sum_{j=1}^k a_i a_j C(\mathbf{s}_i - \mathbf{s}_j; t_i - t_j) \geq 0. \quad (2)$$

This points at a fundamental requirement for any covariance function: given any finite system of space-time coordinates $(\mathbf{s}_1; t_1), \dots, (\mathbf{s}_k; t_k) \in \mathbb{R}^d \times \mathbb{R}$ and coefficients $a_1, \dots, a_k \in \mathbb{R}$, the double sum in (2) must be nonnegative. The property is called *positive definiteness*, and it is a necessary and sufficient condition for a covariance function. The celebrated theorem of Bochner (1955, p. 58) states that a continuous function is positive definite if and only if it is the Fourier transform of a finite, nonnegative measure.

To ensure that a valid covariance model is fitted to the data, one usually considers a parametric family whose members are known to be positive definite functions. Previous space-time models of this form involve various separability assumptions with undesirable properties. For example, the space-time covariance function might decompose into the sum or the product of a purely spatial and a purely temporal covariance function. Functions of this type do not allow for space-time interaction. We point to Kyriakidis and Journel (1999, pp. 664–666) or Cressie and Huang (1999, p. 1331) for a more detailed discussion of the shortcomings of separable models. Alternatively, if $\|\cdot\|$ denotes the Euclidean norm, models of the form

$$C(\mathbf{h}; u) = \varphi\left(a_1^2 \|\mathbf{h}\|^2 + a_2^2 |u|^2\right)$$

have been fitted. Here, a_1 and a_2 are geometric anisotropy factors between the space and time dimensions, and the covariance function is constrained to the same form in space and time. Cressie and Huang (1999) introduced classes of nonseparable, stationary covariance functions that allow for space-time interaction. Their approach is novel and powerful, but depends on Fourier transform pairs in \mathbb{R}^d . In other words, it is restricted to a comparably small class of functions for which a closed form solution to the d -variate Fourier integral is known.

In this paper we take up the approach of Cressie and Huang (1999), but avoid the aforementioned limitation, and provide very general classes of valid space-time covariance models. Section 2 reviews a necessary and sufficient condition for positive definiteness, and Section 3 gives a sufficient condition which does not rely on Fourier transforms. Specifically, let $\varphi(t)$, $t \geq 0$, be any completely monotone function, such as those given in Table 1; let $\psi(t)$, $t \geq 0$, be any positive function with a completely

Table 1: Some completely monotone functions $\varphi(t)$, $t \geq 0$.

Function ^{a,b}	Parameters
$\varphi(t) = \exp(-ct^\gamma)$	$c > 0, 0 < \gamma \leq 1$
$\varphi(t) = (2^{\nu-1}\Gamma(\nu))^{-1} (ct^{1/2})^\nu K_\nu(ct^{1/2})$	$c > 0, \nu > 0$
$\varphi(t) = (1 + ct^\gamma)^{-\nu}$	$c > 0, 0 < \gamma \leq 1, \nu > 0$
$\varphi(t) = 2^\nu (\exp(ct^{1/2}) + \exp(-ct^{1/2}))^{-\nu}$	$c > 0, \nu > 0$

^a See Gneiting (1997) and the references therein for proofs of the complete monotonicity. The functions have been standardized so that $\varphi(0) = 1$.

^b K_ν denotes a modified Bessel function of the second kind of order ν (see Abramowitz and Stegun 1972, pp. 374 ff.).

Table 2: Some positive functions $\psi(t)$, $t \geq 0$, with a completely monotone derivative.

Function ^a	Parameters
$\psi(t) = (at^\alpha + 1)^\beta$	$a > 0, 0 < \alpha \leq 1, 0 \leq \beta \leq 1$
$\psi(t) = \ln(at^\alpha + b) / \ln(b)$	$a > 0, b > 1, 0 < \alpha \leq 1$
$\psi(t) = (at^\alpha + b) / (b(at^\alpha + 1))$	$a > 0, 0 < b \leq 1, 0 < \alpha \leq 1$

^a The functions have been standardized so that $\psi(0) = 1$.

monotone derivative, such as those given in Table 2; and let $\sigma^2 > 0$. Then

$$C(\mathbf{h}; u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}} \varphi\left(\frac{\|\mathbf{h}\|^2}{\psi(|u|^2)}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (3)$$

is a valid space-time covariance function. For instance, if we choose the first entry in Table 1 and the first function in Table 2, then (3) provides the family

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\beta d/2}} \exp\left(-\frac{c\|\mathbf{h}\|^{2\gamma}}{(a|u|^{2\alpha} + 1)^{\beta\gamma}}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (4)$$

of space-time covariance functions, where a and c are nonnegative scaling parameters of time and space, respectively. The smoothness parameters α and γ take values in $(0, 1]$, and σ^2 is the variance of the spatio-temporal process. Figure 1 illustrates the covariance function (4) for various values of α and γ , where $d = 2$, $a = 1$, $c = 1$, $\beta = 1$, and $\sigma^2 = 1$.

In Section 4 we use the Irish wind data of Haslett and Raftery (1989) in order to illustrate strategies for physically meaningful choices of $\varphi(t)$ and $\psi(t)$ functions in a given situation. Though the model (3) is in general nonseparable, we can associate $\varphi(t)$ and $\psi(t)$ with the data's spatial structure and temporal structure, respectively. We develop a correlation model which derives from (4) and a nugget effect,

$$C(\mathbf{h}; u | \beta) = \begin{cases} (0.901|u|^{1.544} + 1)^{-1}, & \text{if } \mathbf{h} = \mathbf{0}, \\ 0.968 (0.901|u|^{1.544} + 1)^{-1} \exp\left(-\frac{0.00134\|\mathbf{h}\|}{(0.901|u|^{1.544} + 1)^{\beta/2}}\right), & \text{otherwise,} \end{cases}$$

and depends on a readily interpretable space-time interaction parameter $\beta \in [0, 1]$. The case $\beta = 0$ corresponds to a separable model, in which the spatial correlations at different temporal lags u

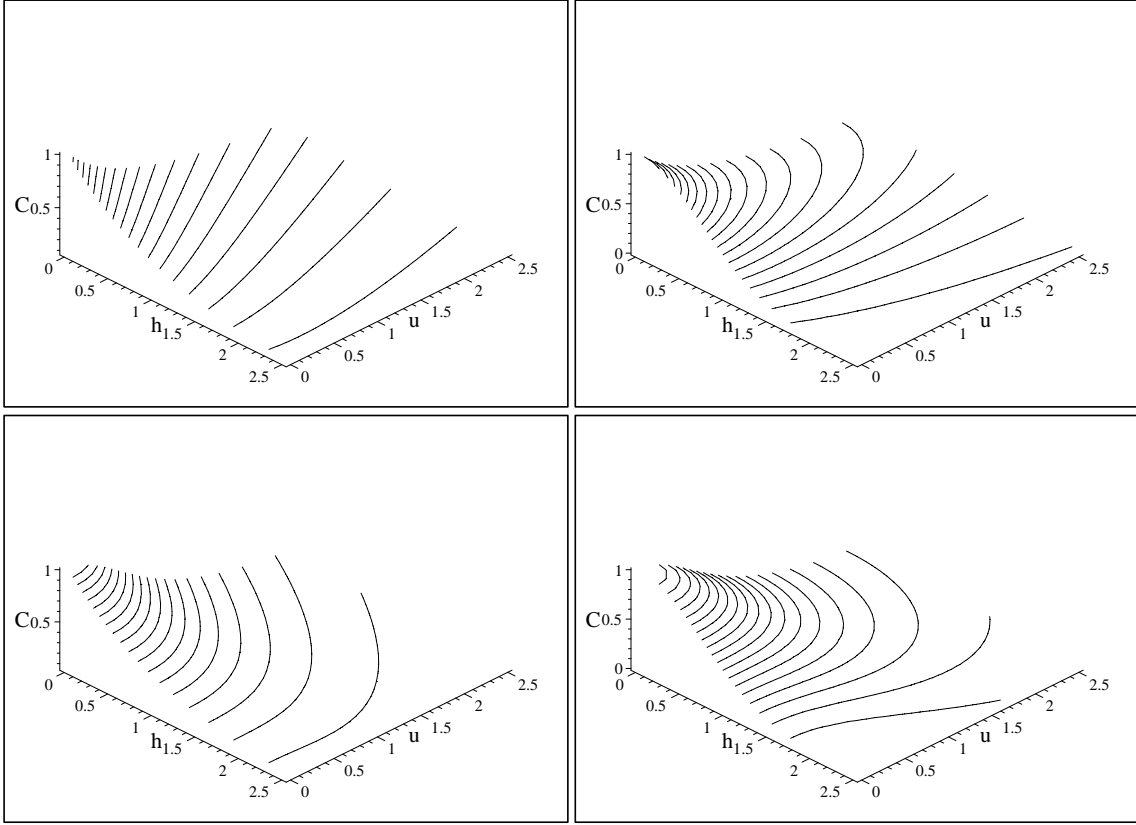


Figure 1: Contour plots of the space-time covariance function (13) versus the modulus of the spatial lag, $\|\mathbf{h}\|$, and the temporal lag, $|u|$. The functions attain their maximum, $C(0; 0) = 1$, at the origin, and the contour lines are equidistant at $0.95, 0.9, \dots, 0.05$. Upper left: $\alpha = 1/2$, $\gamma = 1/2$. Upper right: $\alpha = 1/2$, $\gamma = 1$. Lower left: $\alpha = 1$, $\gamma = 1/2$. Lower right: $\alpha = 1$, $\gamma = 1$.

are proportional to each other. As β increases, space-time interaction strengthens, and the spatial correlations at nonzero temporal lags fall off less and less rapidly. The weighted least squares estimate $\hat{\beta} = 0.61$ for the Irish wind data falls well into the nonseparable range.

Section 5 returns to the theoretical discussion. We apply the criterion of Section 2 to space-time covariance models proposed by Carroll et al. (1997) and Cressie and Huang (1999), and it will be seen that some of these are not valid covariance functions. The paper closes with a discussion of challenges in geostatistical space-time analysis in Section 6. We take another look at the Irish wind data and address the modeling of covariance structures which are not *fully symmetric*, the latter meaning that

$$C(\mathbf{h}; u) = C(-\mathbf{h}; u) = C(\mathbf{h}; -u) = C(-\mathbf{h}; -u), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}. \quad (5)$$

The assumption (5) of full symmetry is often violated when environmental, atmospheric, or oceanographic data are influenced by dynamic processes such as prevailing winds or ocean currents. In this type of situation, physically meaningful covariance models derive from the general idea of a Lagrangian reference frame, which can be thought of as attached to and moving with the center of an air or water mass.

2 A Criterion for Positive Definiteness

In this section, we discuss conditions for the validity of space-time covariance functions. The terms *valid covariance model*, *covariance function*, *stationary covariance function*, and *positive definite function* will be used interchangeably. From a mathematical perspective, there is no distinction between the space-time domain $\mathbb{R}^d \times \mathbb{R}$ and the purely spatial domain \mathbb{R}^{d+1} . In other words, the class of space-time covariance functions in $\mathbb{R}^d \times \mathbb{R}$ coincides with the class of spatial covariance functions in \mathbb{R}^{d+1} . Thus, the fundamental physical difference between space and time dimensions needs to be acknowledged through our notation and through the specific constructions below.

The well-known theorem of Bochner (1955, p. 58) states that a continuous function C on $\mathbb{R}^d \times \mathbb{R}$ is positive definite if and only if it is of the form

$$C(\mathbf{h}; u) = \iint e^{i\boldsymbol{\omega}'\mathbf{h} + i\tau u} dF(\boldsymbol{\omega}; \tau), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (6)$$

with a *spectral distribution* function F . In other words, F is the distribution function of a nonnegative, finite measure on $\mathbb{R}^d \times \mathbb{R}$. An immediate consequence of the representation is the inequality

$$|C(\mathbf{h}; u)| \leq C(\mathbf{0}; 0), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (7)$$

and we will frequently return to (7) and its analogue for purely spatial or purely temporal covariance functions,

$$|C(\mathbf{h})| \leq C(\mathbf{0}), \quad \mathbf{h} \in \mathbb{R}^d. \quad (8)$$

If $C(\mathbf{h}; u)$ is integrable, the spectral distribution function F is absolutely continuous, and Bochner's representation (6) simplifies to

$$C(\mathbf{h}; u) = \iint e^{i\boldsymbol{\omega}'\mathbf{h} + i\tau u} f(\boldsymbol{\omega}; \tau) d\boldsymbol{\omega} d\tau, \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R},$$

where f is a continuous, nonnegative, and integrable function. The covariance function C and the *spectral density* function f then form a Fourier transform pair, and

$$f(\boldsymbol{\omega}; \tau) = (2\pi)^{-d-1} \iint e^{-i\mathbf{h}'\boldsymbol{\omega} - i\tau u} C(\mathbf{h}; u) d\mathbf{h} du.$$

The following criterion is based on these results and Fubini's theorem. It is due to Cressie and Huang (1999), where it is given in a slightly different but equivalent form.

Criterion 1 (Cressie and Huang). A continuous, bounded, symmetric, and integrable function $C(\mathbf{h}; u)$, defined on $\mathbb{R}^d \times \mathbb{R}$, is a space-time covariance function if and only if

$$C_{\boldsymbol{\omega}}(u) = \int e^{-i\mathbf{h}'\boldsymbol{\omega}} C(\mathbf{h}; u) d\mathbf{h}, \quad u \in \mathbb{R},$$

is a covariance function for almost all $\boldsymbol{\omega} \in \mathbb{R}^d$.

The proof of a generalized version of Criterion 1 is given in the appendix. Integrability is not an overly restrictive assumption, since a continuous, bounded, and symmetric function $C(\mathbf{h}; u)$ is positive definite if and only if, for every $a > 0$ and $b > 0$, the integrable function $\exp(-a\|\mathbf{h}\| - b|u|) C(\mathbf{h}; u)$ is positive definite. If the assumption (5) of full symmetry holds, then all the functions

$C_{\boldsymbol{\omega}}(u)$ are real-valued and symmetric. The criterion remains valid for symmetric but not necessarily fully symmetric functions, although $C_{\boldsymbol{\omega}}(u)$ will be complex-valued for some, or all, $\boldsymbol{\omega} \in \mathbb{R}^d$.

Cressie and Huang (1999) used Criterion 1 to construct valid space-time covariance functions through analytic Fourier inversion of $C_{\boldsymbol{\omega}}(u)$ with respect to $\boldsymbol{\omega} \in \mathbb{R}^d$. In the following section, we give a Fourier-free criterion which is based on their approach; and in Section 5 we use Criterion 1 to disprove the validity of previously proposed space-time models.

3 Direct Constructions in the Space-Time Domain

This section provides a simple sufficient condition for positive definiteness in $\mathbb{R}^d \times \mathbb{R}$. The criterion allows us to construct parametric families of spatio-temporal covariance functions directly in the space-time domain. Any appeal to Fourier analysis is avoided.

Recall that a continuous function $\varphi(t)$, defined for $t > 0$ or $t \geq 0$, is said to be *completely monotone* if it possesses derivatives $\varphi^{(n)}$ of all orders and

$$(-1)^n \varphi^{(n)}(t) \geq 0, \quad (t > 0, n = 0, 1, 2, \dots).$$

From Bernstein's theorem (Feller 1966, p. 439), the general form of a completely monotone function $\varphi(t)$, $t > 0$, is

$$\varphi(t) = \int_{[0, \infty)} \exp(-rt) dF(r), \quad t > 0, \quad (9)$$

where F is nondecreasing. Isotropic covariance functions and completely monotone functions are closely related. Specifically, the isotropic function

$$C(\mathbf{h}) = \varphi(\|\mathbf{h}\|^2), \quad \mathbf{h} \in \mathbb{R}^d,$$

is a spatial covariance function for all dimensions d if and only if $\varphi(t)$, $t \geq 0$, is completely monotone (Schoenberg 1938; Cressie 1993, p. 86). Table 1 gives some completely monotone functions, and the first two entries lead to the *powered exponential* class,

$$C(\mathbf{h}) = \sigma^2 \exp(-c\|\mathbf{h}\|^{2\gamma}),$$

and the *Whittle-Matérn* family,

$$C(\mathbf{h}) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (c\|\mathbf{h}\|)^\nu K_\nu(c\|\mathbf{h}\|), \quad (10)$$

of isotropic covariance functions. Here, c is a nonnegative scaling parameter, $\gamma \in (0, 1]$ and $\nu > 0$ are smoothness parameters, and σ^2 is the variance of the process.

We can now formulate our key result. Its proof is based on Criterion 1 and deferred to the appendix.

Criterion 2. Let $\varphi(t)$, $t \geq 0$, be a completely monotone function, and let $\psi(t)$, $t \geq 0$, be a positive function with a completely monotone derivative. Then

$$C(\mathbf{h}; u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}} \varphi\left(\frac{\|\mathbf{h}\|^2}{\psi(|u|^2)}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (11)$$

is a space-time covariance function.

All the examples of Cressie and Huang (1999) can be written in the form (11), except for the case $c < 1$ in their Example 7, and their Examples 5 and 6 which are shown to be wrong in Section 5 below. Though the models are nonseparable in general, we can associate $\varphi(t)$ and $\psi(t)$ with the data's spatial and temporal structure, respectively. Table 1 provides a range of possible choices of completely monotone functions $\varphi(t)$. Further examples and a discussion of the associated spatial covariance functions can be found in Gneiting (1999). The entries in Table 2 are obviously positive functions with a completely monotone derivative if $\alpha = 1$. If $\alpha \in (0, 1]$, the complete monotonicity of the derivative follows from the chain rule for differentiation together with two criteria of Feller (1966, p. 441).

The following examples illustrate the breadth and simplicity of our approach. Strategies for selecting appropriate $\varphi(t)$ and $\psi(t)$ functions in order to construct a meaningful parametric family for a given situation will be discussed in Section 4.

Example 1. Putting $\varphi(t) = \exp(-ct^\gamma)$ and $\psi(t) = (at^\alpha + 1)^\beta$ in (11) leads to the previously introduced parametric family (4). The product with the purely temporal covariance function $(a|u|^{2\alpha} + 1)^{-\delta}$, $u \in \mathbb{R}$, then gives the class

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\delta + \beta d/2}} \exp\left(-\frac{c\|\mathbf{h}\|^{2\gamma}}{(a|u|^{2\alpha} + 1)^{\beta\gamma}}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (12)$$

where a and c are nonnegative scaling parameters of time and space, respectively; the smoothness parameter α and γ take values in $(0, 1]$; $\beta \in [0, 1]$, $\delta \geq 0$, and $\sigma^2 > 0$. A separable covariance function is obtained when $\beta = 0$. In practice, parameter values will frequently be fixed. For instance, Figure 1 illustrates members of the class

$$C(\mathbf{h}; u) = (|u|^{2\alpha} + 1)^{-1} \exp\left(-\frac{\|\mathbf{h}\|^{2\gamma}}{(|u|^{2\alpha} + 1)^\gamma}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^2 \times \mathbb{R}, \quad (13)$$

to which the family (12) reduces when $d = 2$; $a = 1$, $c = 1$, $\beta = 1$, $\delta = 0$, and $\sigma^2 = 1$. The parameters $\alpha \in (0, 1]$ and $\gamma \in (0, 1]$ govern the smoothness of the purely temporal and purely spatial covariance. Specifically, the spatial sections of the associated space-time process have fractal (Hausdorff) dimension $d - \gamma$, and the temporal sections have fractal dimension $1 - \alpha$ (see, for example, Chapter 8 of Adler 1981).

In many instances, a reparameterization of (12) is useful. Specifically, replacing the exponent $\delta + \beta d/2$ in (12) by $\tau \geq \beta d/2$ leads to the parametric family

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^\tau} \exp\left(-\frac{c\|\mathbf{h}\|^{2\gamma}}{(a|u|^{2\alpha} + 1)^{\beta\gamma}}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}. \quad (14)$$

If we fix $\tau \geq d/2$, we obtain a parametric family $C(\mathbf{h}; u | \beta)$ with an easily interpretable space-time interaction parameter $\beta \in [0, 1]$. The purely spatial covariance function $C(\mathbf{h}; 0 | \beta)$ and the purely temporal covariance function $C(\mathbf{0}; u | \beta)$ are independent of β . However, the case $\beta = 0$ corresponds to a separable model, in which the spatial correlations for different values of the temporal lag u are proportional to each other. As β increases, space-time interaction strengthens, and the correlations at nonzero temporal lags fall off less and less rapidly, as compared to the separable model. Figure 2 illustrates the covariances structures associated with the extremal cases $\beta = 0$ and $\beta = 1$ in the family

$$C(\mathbf{h}; u | \beta) = (|u| + 1)^{-1} \exp\left(-\frac{\|\mathbf{h}\|}{(|u| + 1)^{\beta/2}}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^2 \times \mathbb{R}, \quad (15)$$

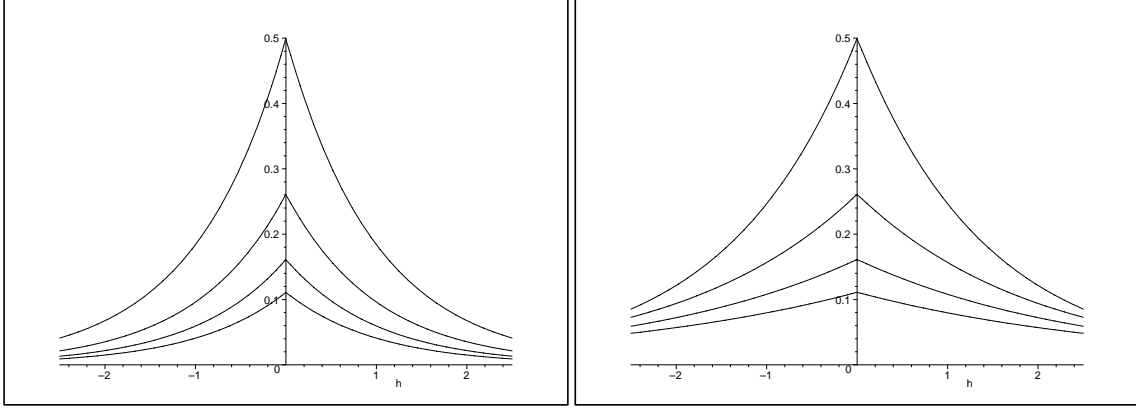


Figure 2: The space-time covariance model (15) in the extremal cases $\beta = 0$ (separable case, left) and $\beta = 1$ (right). The horizontal axis corresponds to oriented spatial distance along a line transect, and the graphs represent $C(\mathbf{h}; u)$ with temporal lags u equal to 1, 2, 3, and 4 (from top to bottom).

to which (14) reduces when $d = 2$; $a = 1$, $c = 1$, $\alpha = 1/2$, $\gamma = 1/2$, $\tau = 1$, and $\sigma^2 = 1$. The effect of the space-time interaction parameter $\beta \in [0, 1]$ is clearly visible.

Example 2. If we choose the second entry in Table 1 and the first entry in Table 2, Eq. (11) leads to the parametric family

$$C(\mathbf{h}; u) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu) (a|u|^{2\alpha} + 1)^{\delta+\beta d/2}} \left(\frac{c\|\mathbf{h}\|}{(a|u|^{2\alpha} + 1)^{\beta/2}} \right)^\nu K_\nu \left(\frac{c\|\mathbf{h}\|}{(a|u|^{2\alpha} + 1)^{\beta/2}} \right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (16)$$

of space-time covariance functions. Here a and c are nonnegative scaling parameters of time and space, respectively; $\alpha \in (0, 1]$ is the smoothness parameter of time, $\nu > 0$ is the smoothness parameter of space; $\beta \in [0, 1]$, $\delta \geq 0$, $\sigma^2 > 0$; and K_ν is the modified Bessel function of the second kind of order ν (see, for example, Abramowitz and Stegun 1972, pp. 374ff). The purely temporal covariance is the corresponding limit as $\|\mathbf{h}\| \rightarrow 0$,

$$C(\mathbf{0}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\delta+\beta d/2}}, \quad u \in \mathbb{R},$$

and the purely spatial covariance $C(\mathbf{h}; 0)$ is the Whittle-Matérn class (10). If $\nu = 1/2$ the space-time covariance function (16) reduces to

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\delta+\beta d/2}} \exp\left(-\frac{c\|\mathbf{h}\|}{(a|u|^{2\alpha} + 1)^{\beta/2}}\right),$$

which is the same as (12) with $\gamma = 1/2$; and if $\nu = 3/2$ we get

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\delta+\beta d/2}} \left(1 + \frac{c\|\mathbf{h}\|}{(a|u|^{2\alpha} + 1)^{\beta/2}}\right) \exp\left(-\frac{c\|\mathbf{h}\|}{(a|u|^{2\alpha} + 1)^{\beta/2}}\right).$$

A separable covariance function is obtained when $\beta = 0$. Again, a subset of the parameters is usually held fixed. Figure 3, for example, illustrates the effect of the parameters $\alpha \in (0, 1]$ and $\nu > 0$ in the

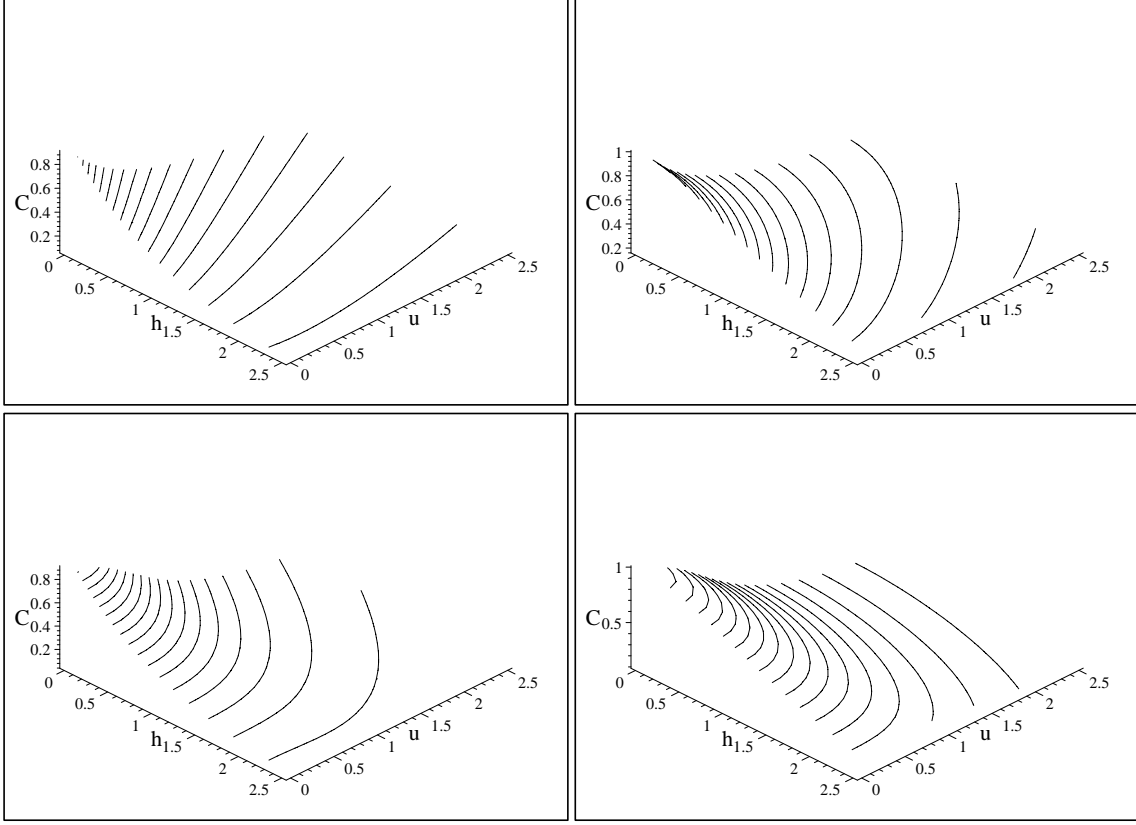


Figure 3: Contour plots of the space-time covariance function (17) versus the modulus of the spatial lag, $\|\mathbf{h}\|$, and the temporal lag, $|u|$. The functions attain their maximum, $C(0; 0) = 1$, at the origin, and the contour lines are equidistant at $0.95, 0.9, \dots, 0.05$. Upper left: $\alpha = 1/2, \nu = 1/2$. Upper right: $\alpha = 1/2, \nu = 3/2$. Lower left: $\alpha = 1, \nu = 1/2$. Lower right: $\alpha = 1, \nu = 3/2$.

family

$$C(\mathbf{h}; u) = \frac{1}{2^{\nu-1}\Gamma(\nu)(|u|^{2\alpha}+1)} \left(\frac{\|\mathbf{h}\|}{(|u|^{2\alpha}+1)^{1/2}} \right)^{\nu} K_{\nu} \left(\frac{\|\mathbf{h}\|}{(|u|^{2\alpha}+1)^{1/2}} \right), \quad (\mathbf{h}; u) \in \mathbb{R}^2 \times \mathbb{R}, \quad (17)$$

to which the covariance model (16) reduces when $d = 2$; $a = 1$, $c = 1$, $\beta = 1$, $\delta = 0$, and $\sigma^2 = 1$. Then $\alpha \in (0, 1]$ is a temporal smoothness parameter, and the spatial smoothness parameter $\nu > 0$ governs the differentiability of the purely spatial covariance and spatial sections of the space-time process. See Hancock and Wallis (1994) and Gneiting (1999) for further comments on the Whittle-Matérn class. As in the previous example, replacing the exponent $\delta + \beta d/2$ by $\tau \geq d/2$ in (16) leads to parametric covariance models with a meaningful and easily interpretable space-time interaction parameter β .

Table 3: The 11 synoptic meteorological stations in the Irish wind data set.

Station ^a	Latitude ^b	Longitude ^b	Elevation ^c	Mean ^d
Roche’s Point	51 48’ N	8 15’ W	41	2.46
Valentia	51 56’ N	10 15’ W	14	2.26
Kilkenny	52 40’ N	7 16’ W	64	1.73
Shannon	52 42’ N	8 55’ W	20	2.25
Birr	53 05’ N	7 53’ W	72	1.82
Dublin	53 26’ N	6 15’ W	85	2.17
Mullingar	53 32’ N	7 22’ W	104	2.02
Claremorris	53 43’ N	8 59’ W	69	2.01
Clones	54 11’ N	7 14’ W	89	2.04
Belmullet	54 14’ N	10 00’ W	10	2.53
Malin Head	55 22’ N	7 20’ W	25	2.76

^a Latitude, longitude, and elevation as posted by the Naval Atlantic Meteorology & Oceanography Detachment at <http://205.67.212.10/station.htm>

^b in degrees and minutes

^c in meters

^d mean of the square roots of daily mean wind speeds in meters per seconds

4 Irish Wind Data

This section illustrates strategies for selecting appropriate $\varphi(t)$ and $\psi(t)$ functions in the general model (11) in order to construct a physically meaningful, parametric family of space-time covariance functions for a given situation.

We consider the Irish wind data of Haslett and Raftery (1989), which consist of daily averages of wind speeds at 11 synoptic meteorological stations in Ireland during the period 1961–1978. The data are available at Statlib, <http://lib.stat.cmu.edu/datasets/>. Following Haslett and Raftery (1989), we take a square root transformation to stabilize the variance over both stations and time periods and to make the marginal distributions approximately normal. Table 3 summarizes latitude, longitude, elevation, and the mean of the square roots of daily average wind speeds for the 11 meteorological stations. Generally, wind speeds decrease with distance from the coastline. Figure 4 shows time series plots of the square roots of daily mean wind speeds at Kilkenny and Malin Head in 1961. Kilkenny and Malin Head are the stations with the lowest and highest mean wind speeds, respectively. Spatial and temporal dependencies are clearly recognizable. As in Haslett and Raftery (1989), we estimate the seasonal effect by calculating the average of the square roots of the daily means over all years and stations for each day of the year, and then regress the result on a set of annual harmonics. Subtraction of the estimated seasonal effect and the estimated spatially varying mean, as given in the right-hand column of Table 3, results in data hereinafter referred to as *velocity measures*. Haslett and Raftery argue convincingly that a stationary model for the velocity measures is an appropriate approximation. In their paper, the goal was estimation of the spatially varying mean at a new site, where only a short run of data is available. This required a careful and innovative modeling of temporal long-memory dependence, which was achieved through ARMA modelling and fractional differencing.

Here, our goals in analyzing the velocity measures differ. Long-memory effects are irrelevant in short-term prediction problems, and we restrict our attention to the spatio-temporal covariance

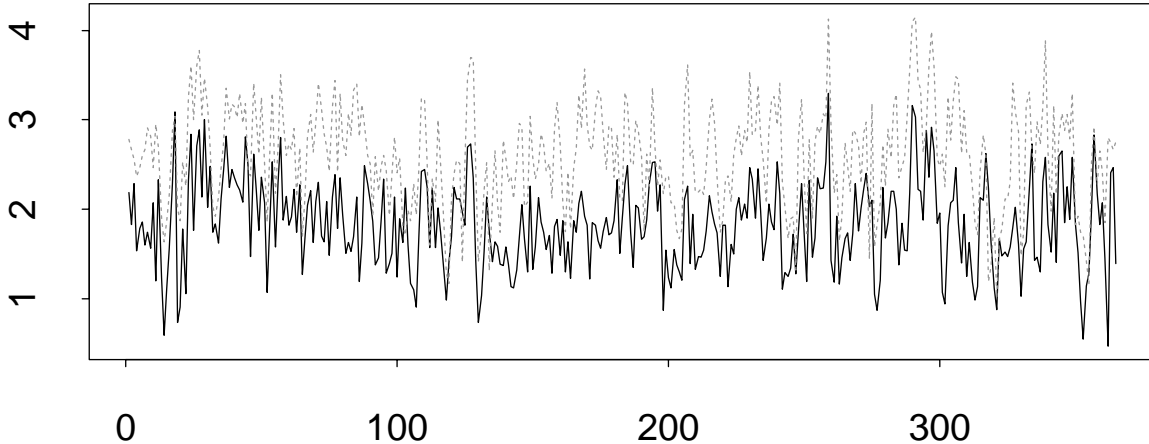


Figure 4: Time series plot of the square roots of daily mean wind speeds at Kilkenny (solid line) and Malin Head (broken line) in 1961. Kilkenny and Malin Head are the stations with the lowest and highest mean wind speeds, respectively.

structure for temporal lags up to 3 days, a range which is crucial in many environmental applications. The correlations for the velocity measures fall off rapidly in time, and a meaningful distinction between separable and nonseparable covariance structures may not be feasible at higher lags. Figure 5 illustrates the empirical space-time correlations at temporal lags less than or equal to three days. Raftery, Haslett, and McColl (1982) introduced this type of graph as a *distance-time autocorrelation plot*. The upper left display shows the purely spatial correlations for the 55 pairs of meteorological stations as a function of distance in km, along with the spatial correlation model fitted by Haslett and Raftery (1989),

$$C(\mathbf{h}; 0) = \begin{cases} 1, & \text{if } \mathbf{h} = \mathbf{0}, \\ 0.968 \exp(-0.00134\|\mathbf{h}\|), & \text{otherwise,} \end{cases} \quad (18)$$

where the usable range of spatial lags is $\|\mathbf{h}\| \leq 450$ km. This can be written as a convex combination of a continuous, exponential model and a *nugget effect*,

$$C_0(\mathbf{h}) = \begin{cases} 1, & \text{if } \mathbf{h} = \mathbf{0}. \\ 0, & \text{otherwise.} \end{cases}$$

The nugget effect allows for a discontinuity at the origin and corresponds to measurement error and/or small scale spatial variability (see, for example, Cressie 1993, p. 58). For the Irish wind data, either explanation is likely to apply, due to possible instrument variations and the highly irregular nature of wind speeds.

For the purely temporal covariance structure, a continuous model with limited smoothness at the origin is a physically reasonable compromise. On the one hand, wind speeds are highly irregular and measurement error may be nonnegligible, suggesting the presence of a nugget effect. On the other hand, the data were obtained by temporal aggregation of a continuous trace over 24 hours, which tends to smooth out the discontinuities. Here we choose the correlation function

$$C(\mathbf{0}; u) = (0.901|u|^{1.544} + 1)^{-1} \quad (19)$$

which fits the empirical temporal correlations observed at the 11 stations well; these average to .526, .267, and .179 at lags u equal to 1, 2, and 3 days, respectively. Evidently, a wide range of models

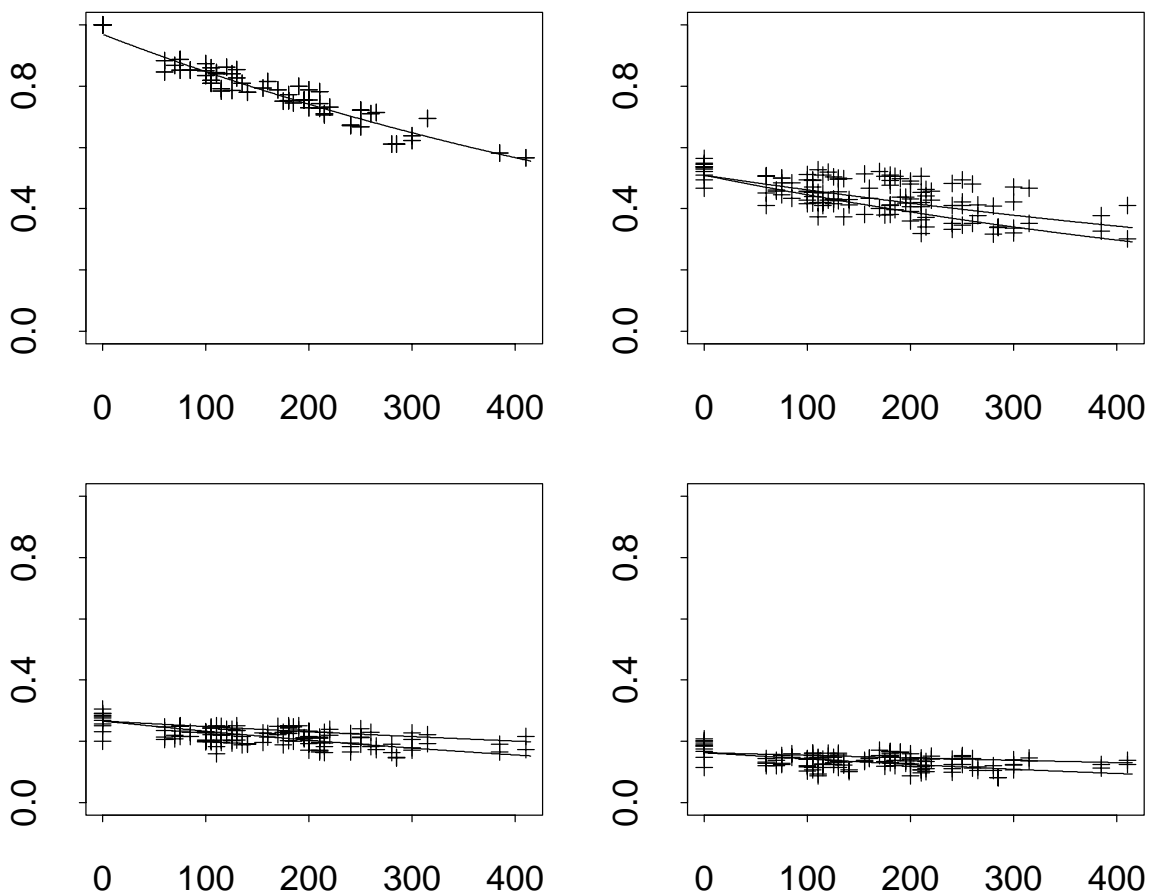


Figure 5: Empirical space-time correlations for time lag 0 (upper left), 1 (upper right), 2 (lower left), and 3 days (lower right) as a function of spatial distance in km. The solid lines illustrate the extremal cases $\beta = 0$ (lower line) and $\beta = 1$ (upper line) for the covariance function (21).

of the form $C(\mathbf{0}; u) = (a|u|^{2\alpha} + 1)^{-\tau}$ is physically justifiable and fits the empirical correlations. We choose (19) because it is easily embedded into a rich, parametric family which includes both separable and nonseparable space-time covariance functions. Specifically, larger values of the parameter a in (14) allow for stronger space-time interaction effects.

The product of the purely spatial correlation function (18) and the purely temporal correlation function (19) is

$$C(\mathbf{h}; u) = \begin{cases} (0.901|u|^{1.544} + 1)^{-1}, & \text{if } \mathbf{h} = \mathbf{0}. \\ 0.968 (0.901|u|^{1.544} + 1)^{-1} \exp(-0.00134\|\mathbf{h}\|), & \text{otherwise.} \end{cases} \quad (20)$$

The separable model (20) corresponds to the case $\beta = 0$ in the parametric family

$$C(\mathbf{h}; u | \beta) = \begin{cases} (0.901|u|^{1.544} + 1)^{-1}, & \text{if } \mathbf{h} = \mathbf{0}, \\ 0.968 (0.901|u|^{1.544} + 1)^{-1} \exp\left(-\frac{0.00134\|\mathbf{h}\|}{(0.901|u|^{1.544} + 1)^{\beta/2}}\right), & \text{otherwise,} \end{cases} \quad (21)$$

where the usable range of space-time lags is given by $\|\mathbf{h}\| \leq 450$ km and $|u| \leq 3$ days. Note that (21) can be written as a convex combination of two permissible space-time covariance functions and is therefore itself a permissible, positive definite function. The first component is the product of the purely temporal covariance function (19) and a purely spatial nugget effect; the second component is the continuous space-time covariance function (14) with $d = 2$; $a = 0.901$, $c = 0.00134$, $\alpha = 0.772$, $\gamma = 1/2$, and $\tau = 1$. Evidently, (21) is a permissible covariance model in $\mathbb{R}^2 \times \mathbb{R}$, although lags larger than 450 km or 3 days will not be used in the fitting procedure and are not required in typical prediction problems. Figure 5 illustrates the empirical spatio-temporal correlations along with the extremal members of the family (21), corresponding to $\beta = 0$ and $\beta = 1$. The case $\beta = 0$ gives a separable model, in which the spatial correlations for different values of the temporal lag u are proportional to each other. As β increases, correlations at nonzero temporal lags fall off less and less rapidly than under the separable model.

Similar to the technique proposed by Cressie (1993, p. 96) and Cressie and Huang (1999), we use a weighted-least-squares method to estimate the space-time interaction parameter β by minimizing

$$W(\beta) = \sum_{i,j} \sum_{u=1}^3 \left(\frac{\widehat{C}(\mathbf{h}_{ij}; u) - C(\mathbf{h}_{ij}; u | \beta)}{1 - C(\mathbf{h}_{ij}; u | \beta)} \right)^2 \quad (22)$$

over $\beta \in [0, 1]$. Here, \mathbf{h}_{ij} is the spatial lag between stations i and j , $\widehat{C}(\mathbf{h}_{ij}; u)$ is the empirical correlation between the velocity measures at stations i and j and temporal lag u , and the summation is over all ordered pairs of meteorological stations. The weighted-least-squares estimate is $\hat{\beta} = 0.61$, indicating a nonseparable covariance structure.

5 Previously Proposed Models

In this section, we apply Criterion 1 to covariance models proposed by Cressie and Huang (1999) and Carroll et al. (1997).

Example 3. Cressie and Huang (1999) propose space-time covariance models of the form

$$C(\mathbf{h}; u) = \sigma^2 \exp(-a^\delta |u|^\delta - b^2 \|\mathbf{h}\|^2 - c|u|^\delta \|\mathbf{h}\|^2) \quad (23)$$

where \mathbf{h} is the spatial lag in \mathbb{R}^d , u is the temporal lag, a and b are nonnegative scaling parameters, $c \geq 0$ is a space-time interaction parameter, and $\sigma^2 > 0$. Examples 5 and 6 of Cressie and Huang correspond to the specific choices $\delta = 2$ and $\delta = 1$, respectively. We consider the general case with a positive shape parameter δ . If $c = 0$, the model is separable, and it is valid if and only if $\delta \leq 2$ or $a = 0$. In the nonseparable case, $c > 0$, we proceed to prove that (23) is not a valid covariance model for any $\delta > 0$. In particular, the graphs in Figure 3(b), (c), and (d) of Cressie and Huang (1999) do not show space-time covariance functions.

We may assume that $\sigma^2 = 1$, and it suffices to consider the case when \mathbf{h} is scalar, because it corresponds to the restriction of a space-time process in $\mathbb{R}^d \times \mathbb{R}$ to a space-time process in $\mathbb{R} \times \mathbb{R}$. By Criterion 1, (23) is a covariance function if and only if, for almost all $\omega \in \mathbb{R}$,

$$\begin{aligned} C_\omega(u) &= \int_{-\infty}^{\infty} e^{-i\mathbf{h}\omega} C(\mathbf{h}; u) d\mathbf{h} \\ &= \pi^{1/2} (b^2 + c|u|^\delta)^{-1/2} \exp\left(-a^\delta |u|^\delta - \frac{\omega^2}{4(b^2 + c|u|^\delta)}\right) \end{aligned}$$

is a covariance function. Straightforward calculations show that if $c > 0$ and $\omega^2 > 2b^2(1 + 2a^\delta b^2 c^{-1})$, then $C_\omega(u)$ has precisely three extremal points at $u = 0$ and $u = \pm u_0$, where

$$u_0^\delta = \frac{\omega^2 - 2b^2}{2c}$$

if $a = 0$, and

$$u_0^\delta = \frac{1}{4a^\delta c} \left((c^2 + 4a^\delta c \omega^2)^{1/2} - (c + 4a^\delta b^2) \right)$$

if $a > 0$. Since $C_\omega(u)$ is an even, continuous, and positive function with $\lim_{u \rightarrow \pm\infty} C_\omega(u) = 0$, the extremal points at $\pm u_0$ are maxima. Thus $C_\omega(\pm u_0) > C_\omega(0)$, contrary to inequality (8). We conclude that (23) is not a valid space-time covariance function. The problem stems from an erroneous claim of monotonicity and convexity for the function

$$\rho(\omega; u) = \frac{c_0^{d/2}}{(|u|^\delta + c_0)^{d/2}} \exp\left(-\frac{\|\omega\|^2}{4(|u|^\delta + c_0)} + \frac{\|\omega\|^2}{4c_0}\right), \quad u > 0,$$

in Cressie and Huang (1999, pp. 1333 and 1334).

Before proceeding, we note that Example 7 of Cressie and Huang (1999) involves a similar, erroneous claim of convexity for the function

$$\rho(\omega; u) = (u^2 + 1 + (u^2 + c)\|\omega\|^2)^{-\nu-d/2} (1 + c\|\omega\|^2)^{\nu+d/2}, \quad u > 0.$$

However, it is easy to establish directly that if $c \geq 0$ and $\nu > 0$ then $\rho(\omega; u)$ is a covariance function in $u \in \mathbb{R}$, which is the desired conclusion. Thus, in this case the model proposed by Cressie and Huang remains valid.

Example 4. Carroll et al. (1997) consider correlation models of the form

$$C(\mathbf{h}; u) = \exp(-a_1|u| - a_2u^2) \exp((-b_0 - b_1|u| - b_2u^2)\|\mathbf{h}\|) \quad (24)$$

for space-time data on ozone levels in Harris county, Texas. Here, \mathbf{h} is the spatial lag in \mathbb{R}^2 , u is the temporal lag, and a_1 , a_2 , b_0 , b_1 , and b_2 are parameters to be fitted from the data. Concerns about the validity of the model were raised in comments by Cressie (1997) and Guttorp, Meiring, and Sampson (1997).

Inequality (7) supplies necessary conditions on the parameters, because it holds for the correlation model (24) if and only if $a_1 \geq 0$, $a_2 \geq 0$, $b_0 \geq 0$, $b_2 \geq 0$, and $b_1 \geq -2(b_0b_2)^{1/2}$. The parameter estimates in Table 1 of Carroll et al. (1997) satisfy these constraints except for the years 1981, 1982, and 1987. The violations for 1981 and 1987 were noted in Cressie's comment and in the reply by Carroll et al. (1997), respectively. We return to this point below. If the necessary conditions hold and the inequalities are strict, then (24) is an integrable function and Criterion 1 applies. Thus, (24) is a covariance function if and only if, for every $\omega \in \mathbb{R}^2$,

$$\begin{aligned} C_\omega(u) &= \int e^{-i\mathbf{h}'\omega} C(\mathbf{h}; u) d\mathbf{h} \\ &= 2\pi \left(1 + \frac{\|\omega\|^2}{(b_0 + b_1|u| + b_2u^2)^2} \right)^{-3/2} (b_0 + b_1|u| + b_2u^2)^{-2} \exp(-a_1|u| - a_2u^2) \end{aligned}$$

is a covariance function in $u \in \mathbb{R}$. The parameter estimates for 1980 in Table 1 of Carroll et al. (1997) are $a_1 = .1608$, $a_2 = .0051$, $b_0 = 1.8354$, $b_1 = -.2942$, and $b_2 = .0205$, and satisfy the

aforementioned necessary conditions. It is easily verified that for the fitted values of the parameters, and in a neighborhood of $\|\omega\| = 1$, $C_\omega(u)$ is not a covariance function, because inequality (8) is violated. Thus, the fitted correlation model is not positive definite on the space-time domain $\mathbb{R}^2 \times \mathbb{R}$.

Two observations are relevant here. First, Carroll et al. (1997) do not fit the correlation model (24) itself, but a convex combination of the continuous covariance function (24) and a nugget effect. A fundamental decomposition theorem for positive definite functions (Sasvári 1994, Theorem 3.1.2) implies that any practically relevant covariance function on the Euclidean space $\mathbb{R}^d \times \mathbb{R}$ can be written as a convex combination of a valid continuous covariance and a nugget effect. In particular, the sum of a continuous function and a nugget effect is positive definite if and only if the continuous part is such.

Our second observation continues the discussion on the paper by Carroll et al. (1997). In their reply, the authors point out that the covariance model is not fitted over $\mathbb{R}^d \times \mathbb{R}$, but over a bounded domain $S \times T$, where $S \subset \mathbb{R}^2$ corresponds to the spatial lags in Harris county, Texas, and $T \subset \mathbb{R}$ is a bounded and discrete set of temporal lags. Carroll et al. (1997, p. 415) wonder whether the covariance model “is positive definite over the usable range of distances and time lags.” The question relates to extension problems for positive definite functions, which are discussed in Chapter 4 of Sasvári (1994). Nonetheless, results are sparse and not readily applicable, unless the covariance model is isotropic (Gneiting and Sasvári 1999) or the usable range of space-time lags is purely discrete. Otherwise, the only approach to ensure that a valid space-time covariance is fitted, is to use known classes of positive definite functions in $\mathbb{R}^d \times \mathbb{R}$ and restrict these to the spatio-temporal lags of interest. We saw an example of such a strategy in Section 4, when fitting the parametric model (21) to the Irish wind data of Haslett and Raftery (1989).

6 Discussion

Until recently, valid space-time covariance models were mostly subject to separability assumptions or constrained to the same parametric form in space and time. Cressie and Huang (1999) introduced general classes of nonseparable, stationary covariance functions that allow for space-time interaction and include separable models as a special case. The present work provides a Fourier-free implementation of their approach and enlarges the class of valid space-time covariance functions at the modeler’s disposal. Our constructions provide flexible models in closed form and with parameters which have clear-cut interpretations. Using the Irish wind data in Section 4 as an example, we showed how to develop covariance models with a readily interpretable space-time interaction parameter. We identified and estimated a nonseparable covariance structure, in which the spatial correlations at nonzero temporal lags decay slower than would be expected under a separable model.

Physically based approaches might be crucial for further progress in geostatistical space-time analysis. Christakos (2000, p. 18), for example, argues that “in modern spatiotemporal geostatistics, the rational approach for choosing the appropriate model from the data is by means of a theory that represents the physical knowledge available. . .” Dynamic geophysical processes such as wind patterns or ocean currents play key roles here.

To illustrate this type of situation, we return to the discussion in Section 4 and to the Irish wind data of Haslett and Raftery (1989). To fix the idea, we consider the correlation coefficients between the velocity measures at Kilkenny and Shannon, and Clones and Belmullet, respectively. For these pairs, either station has basically the same latitude (see Table 3), so that the north-south component of the spatial separation vector is negligible. The empirical correlations between the western station at a given day and the eastern station one day later are .53 and .52, respectively. However, the correlations between the western station at a given day and the eastern station one day earlier are

.42 and .40, respectively. The deviation from the fully symmetric assumption (5) is not surprising. Winds over Ireland are predominantly westerly, so that velocity measures propagate from west to east. Similar features might well occur in other geophysical or environmental data sets, such as wind speeds over the tropical western Pacific Ocean, as analyzed by Cressie and Huang (1999), or atmospheric pollutant in the Milan district, Italy, as recently modeled by De Cesare, Myers, and Posa (2001).

The covariance models proposed by Cressie and Huang (1999) and in the present paper cannot capture features of this type, since they are fully symmetric as defined in (5). The recent approach of Brown, Kåresen, Roberts, and Tonellato (2000) allows for covariance structures which are not fully symmetric, but the resulting covariance models do not have closed form expressions, and it is not obvious how to proceed in a given situation. Another approach to model dynamic environmental and atmospheric processes builds on the general idea of a Lagrangian reference frame, which can be thought of as attached to and moving with the center of an air or water mass. Lagrangian covariance structures have indeed been discussed in the meteorological and hydrological literature, and we refer to Cox and Isham (1988), Bouttier (1993), Desroziers and Lafore (1993), Fischer, Joly, and Lalaurette (1998), and May and Julien (1998), among others. Cox and Isham (1988) show that if \mathbf{V} is a random vector in \mathbb{R}^2 , and $G(r)$ denotes the area of intersection of two disks of common unit radius whose centers are a distance r apart, then

$$C(\mathbf{h}; u) = E_{\mathbf{V}} G(\|\mathbf{h} - \mathbf{V}u\|), \quad (\mathbf{h}; u) \in \mathbb{R}^2 \times \mathbb{R}, \quad (25)$$

is a valid space-time covariance function. Evidently, (25) is in general not fully symmetric. This model is easily extended to the Euclidean space \mathbb{R}^d and general functions G , of which Christakos (2000, p. 227) gives further examples. Conceptually, we think of (25) as the covariance function of a spatio-temporal random field, in which fixed air masses move with random velocity \mathbf{V} . Convex combinations of fully symmetric space-time covariance models and models of the form (25) might well provide improved fits and improved prediction skill for atmospheric and environmental space-time datasets. The general idea is to perturb a fully symmetric model, say of the form (11), so that the dynamic features are captured, too. For the specification of the random velocity \mathbf{V} , various choices are physically reasonable. The simplest case is a constant $\mathbf{V} = \mathbf{v}$, which represents the mean wind vector as determined from synoptic or local wind patterns, such as a westerly wind in the case of Ireland. Research along these lines is presently under development, and we hope to report on the details elsewhere. Well-founded strategies for spatio-temporal modeling remain in great demand.

Appendix

In this appendix, we state and prove generalized versions of Criterion 1 and Criterion 2, which apply to covariance functions defined on the Euclidean space $\mathbb{R}^k \times \mathbb{R}^l$. We do so because the proofs are identical to those in the special case when $k = d$ and $l = 1$, and because the generalizations might lead to further applications. For instance, a promising approach to the statistical analysis of deterministic simulation experiments (Sacks, Welch, Mitchell, and Wynn 1989; Currin, Mitchell, Morris, and Ylvisaker 1991) relies on analytical covariance models in \mathbb{R}^n where n , the number of parameters in a simulation experiment, is often large. In this situation, the parameter set might split into two groups of size k and l , respectively, calling for a covariance model in $\mathbb{R}^k \times \mathbb{R}^l$. Returning to space-time problems, we see from Eq. (28) with $k = 1$ and $l = d$ that under the conditions of Criterion 2

$$C(\mathbf{h}; u) = \frac{\sigma^2}{\psi(\|\mathbf{h}\|^2)^{1/2}} \varphi\left(\frac{|u|^2}{\psi(\|\mathbf{h}\|^2)}\right), \quad (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R}, \quad (26)$$

is a valid space-time covariance function. Note the symmetry between (11) and (26): now $\psi(t)$ and $\varphi(t)$ are associated the data's spatial structure and temporal structure, respectively.

Criterion 1 (generalized). A continuous, bounded, symmetric, and integrable function $C(\mathbf{h}; \mathbf{u})$, defined on $\mathbb{R}^k \times \mathbb{R}^l$, is a covariance function if and only if

$$C_{\boldsymbol{\omega}}(\mathbf{u}) = \int e^{-i\mathbf{h}'\boldsymbol{\omega}} C(\mathbf{h}; \mathbf{u}) d\mathbf{h}, \quad \mathbf{u} \in \mathbb{R}^l, \quad (27)$$

is a covariance function for almost all $\boldsymbol{\omega} \in \mathbb{R}^k$.

Proof. Notice that $C(\mathbf{h}; \mathbf{u})$ is square-integrable over $\mathbb{R}^k \times \mathbb{R}^l$, and that its Fourier transform $f(\boldsymbol{\omega}; \boldsymbol{\tau})$ is a real-valued, continuous, and symmetric function. Furthermore, for all $\boldsymbol{\omega} \in \mathbb{R}^k$, $C_{\boldsymbol{\omega}}(\mathbf{u})$ is continuous and integrable, because $C(\mathbf{h}; \mathbf{u})$ is integrable and uniformly continuous on compact sets, and

$$\left| \int C_{\boldsymbol{\omega}}(\mathbf{u}) d\mathbf{u} \right| \leq \iint |C(\mathbf{h}; \mathbf{u})| d\mathbf{h} d\mathbf{u} < \infty.$$

From Bochner's theorem and Fourier inversion, $C(\mathbf{h}; \mathbf{u})$ is positive definite if and only if

$$\begin{aligned} f(\boldsymbol{\omega}; \boldsymbol{\tau}) &= (2\pi)^{-k-l} \iint e^{-i\mathbf{h}'\boldsymbol{\omega} - i\boldsymbol{\tau}'\mathbf{u}} C(\mathbf{h}; \mathbf{u}) d\mathbf{h} d\mathbf{u} \\ &= (2\pi)^{-k-l} \int e^{-i\boldsymbol{\tau}'\mathbf{u}} C_{\boldsymbol{\omega}}(\mathbf{u}) d\mathbf{u} \end{aligned}$$

is nonnegative everywhere.

Now suppose that $C_{\boldsymbol{\omega}}(\mathbf{u})$ is a covariance function for almost all $\boldsymbol{\omega} \in \mathbb{R}^k$. Since $C_{\boldsymbol{\omega}}(\mathbf{u})$ is continuous and integrable, we find that $f(\boldsymbol{\omega}; \boldsymbol{\tau}) \geq 0$ almost everywhere on $\mathbb{R}^k \times \mathbb{R}^l$. Thus, the continuous function $f(\boldsymbol{\omega}; \boldsymbol{\tau})$ is nonnegative everywhere. Conversely, if $C(\mathbf{h}; \mathbf{u})$ is a covariance function, then $f(\boldsymbol{\omega}; \boldsymbol{\tau})$ is nonnegative and integrable, by Bochner's theorem applied in $\mathbb{R}^k \times \mathbb{R}^l$. By Fubini's theorem, $f(\boldsymbol{\omega}; \boldsymbol{\tau})$ is also integrable as a function of $\boldsymbol{\tau} \in \mathbb{R}^l$, for almost all $\boldsymbol{\omega} \in \mathbb{R}^k$. Thus, $C_{\boldsymbol{\omega}}(\mathbf{u})$ is a covariance function for almost all $\boldsymbol{\omega} \in \mathbb{R}^k$, by Bochner's theorem applied in \mathbb{R}^l . The proof is complete.

Criterion 2 (generalized). Let k and l be nonnegative integers, and let $\sigma^2 > 0$. Suppose that $\varphi(t)$, $t \geq 0$, is a completely monotone function, and let $\psi(t)$, $t \geq 0$, be a positive function with a completely monotone derivative. Then

$$C(\mathbf{h}; \mathbf{u}) = \frac{\sigma^2}{\psi(\|\mathbf{u}\|^2)^{k/2}} \varphi\left(\frac{\|\mathbf{h}\|^2}{\psi(\|\mathbf{u}\|^2)}\right), \quad (\mathbf{h}; \mathbf{u}) \in \mathbb{R}^k \times \mathbb{R}^l, \quad (28)$$

is a covariance function.

Proof. We assume initially that the isotropic function $\varphi(\|\mathbf{h}\|^2)$, $\mathbf{h} \in \mathbb{R}^k$, is integrable. Then

$$C(\mathbf{h}; \mathbf{u}) = \exp(-a\|\mathbf{u}\|^2) \frac{\sigma^2}{\psi(\|\mathbf{u}\|^2)^{k/2}} \varphi\left(\frac{\|\mathbf{h}\|^2}{\psi(\|\mathbf{u}\|^2)}\right), \quad (29)$$

which differs from (28) by the extra first factor, is integrable over $(\mathbf{h}; \mathbf{u}) \in \mathbb{R}^k \times \mathbb{R}^l$, for all $a > 0$. We may apply Criterion 1, and (29) is a covariance function if and only if the associated function (27) is a covariance function, for almost all $\boldsymbol{\omega} \in \mathbb{R}^k$. Notice that the nondecreasing function F in

Bernstein's representation (9) for φ is bounded and continuous at zero, because φ is bounded and $\lim_{t \rightarrow \infty} \varphi(t) = 0$ by the integrability assumption. From (9) and Fubini's theorem,

$$\begin{aligned} C_{\boldsymbol{\omega}}(\mathbf{u}) &= \int e^{-i\mathbf{h}'\boldsymbol{\omega}} C(\mathbf{h}; \mathbf{u}) \, d\mathbf{h} \\ &= \exp(-a\|\mathbf{u}\|^2) \frac{\sigma^2}{\psi(\|\mathbf{u}\|^2)^{k/2}} \int \int_{(0,\infty)} e^{-i\mathbf{h}'\boldsymbol{\omega}} \exp\left(-\frac{r}{\psi(\|\mathbf{u}\|^2)}\|\mathbf{h}\|^2\right) \, dF(r) \, d\mathbf{h} \\ &= \sigma^2 \pi^{k/2} \exp(-a\|\mathbf{u}\|^2) \int_{(0,\infty)} \exp\left(-\frac{\|\boldsymbol{\omega}\|^2}{4r} \psi(\|\mathbf{u}\|^2)\right) \frac{1}{r^{k/2}} \, dF(r). \end{aligned}$$

Putting $\boldsymbol{\omega} = \mathbf{0}$ and $\mathbf{u} = \mathbf{0}$, we see that

$$\int_{(0,\infty)} \frac{1}{r^{k/2}} \, dF(r) = \frac{1}{\sigma^2 \pi^{k/2}} C_{\mathbf{0}}(\mathbf{0})$$

is finite, because $C_{\mathbf{0}}(\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^l$, is a continuous function. We may therefore write

$$C_{\boldsymbol{\omega}}(\mathbf{u}) = \varphi_{\boldsymbol{\omega}}(\|\mathbf{u}\|^2), \quad \mathbf{u} \in \mathbb{R}^k,$$

where

$$\varphi_{\boldsymbol{\omega}}(t) = \sigma^2 \pi^{k/2} \exp(-at) \int_{(0,\infty)} \exp(-s\psi(t)) \, dG_{\boldsymbol{\omega}}(s), \quad t \geq 0,$$

with a certain nondecreasing, bounded function $G_{\boldsymbol{\omega}}$. From Bernstein's theorem and the two criteria for complete monotonicity on p. 441 of Feller (1966), $\varphi_{\boldsymbol{\omega}}(t)$, $t \geq 0$, is a completely monotone function, for all $\boldsymbol{\omega} \in \mathbb{R}^k$. By Schoenberg's theorem (Schoenberg 1938; Cressie 1993, p. 86), $C_{\boldsymbol{\omega}}(\mathbf{u})$ is a covariance function for all $\boldsymbol{\omega} \in \mathbb{R}^k$. We conclude from Criterion 1 that (29) is a covariance function. Now (29) converges to (28) as $a \rightarrow 0$, and since limits of covariance functions are covariance functions, (28) is a covariance function.

Another approximation argument is needed to dispose of our initial assumption of integrability. Given a completely monotone function $\varphi(t)$, $t \geq 0$, and a positive number b , the product $\exp(-bt)\varphi(t)$, $t \geq 0$, is completely monotone, and $\exp(-b\|\mathbf{h}^2\|)\varphi(\|\mathbf{h}^2\|)$ is integrable over $\mathbf{h} \in \mathbb{R}^k$. Thus,

$$C(\mathbf{h}; \mathbf{u}) = \frac{\sigma^2}{\psi(\|\mathbf{u}\|^2)^{k/2}} \exp\left(-\frac{b\|\mathbf{h}\|^2}{\psi(\|\mathbf{u}\|^2)}\right) \varphi\left(\frac{\|\mathbf{h}\|^2}{\psi(\|\mathbf{u}\|^2)}\right) \quad (30)$$

is a covariance function on $\mathbb{R}^k \times \mathbb{R}^l$ by the above. Since (28) is the limit of (30) as $b \rightarrow 0$, the proof is complete.

Acknowledgements

I am grateful to Adrian Raftery for information on the Irish wind data, to Peter Guttorp, Donald Richards, and Paul Sampson for discussions on space-time covariance models, and to Martin Schlather for detailed comments on an earlier version of the manuscript and a simplified proof of Criterion 1.

References

- Abramowitz, M., and Stegun, I. A. (1972), *Handbook of Mathematical Functions*, New York: Dover.
- Adler, R. J. (1981), *The Geometry of Random Fields*, Chichester: Wiley.
- Bochner, S. (1955), *Harmonic Analysis and the Theory of Probability*, Berkeley and Los Angeles: University of California Press.
- Bouttier, F. (1993), “The dynamics of error covariances in a barotropic model,” *Tellus A*, 45, pp. 408–423.
- Brown, P. E., Kåresen, K. F., Roberts, G. O., and Tonellato, S. (2000), “Blur-generated non-separable space-time models,” *Journal of the Royal Statistical Society Ser. B*, 62, pp. 847–860.
- Brown, P. E., Diggle, P. J., Lord, M. E., and Young, P. C. (2001), “Space-time calibration of radar-rainfall data,” *Applied Statistics*, 50, pp. 221–242.
- Carroll, R. J., Chen, R., George, E. I., Li, T. H., Newton, H. J., Schmiediche, H., and Wang, N. (1997), “Ozone exposure and population density in Harris county, Texas” (with discussion), *Journal of the American Statistical Association*, 92, pp. 392–415.
- Chilès, J.-P., Bourguin, B., and Watremez, P. (2001), “Space-time modelling of sand beach data: a geostatistical approach,” to appear in *GeoENV2000 Proceedings*, eds. D. Allard and P. Monestiez.
- Christakos, G. (2000), *Modern Spatiotemporal Geostatistics*, New York: Oxford University Press.
- Christakos, G., Hristopulos, D. T., and Bogaert, P. (2000), “On the physical geometry concept at the basis of space/time geostatistical hydrology,” *Advances in Water Resources*, 23, pp. 799–810.
- Cox, D. R., and Isham, V. (1988), “A simple spatial-temporal model of rainfall,” *Proceedings of the Royal Society of London, Ser. A*, 415, pp. 317–328.
- Cressie, N. (1993), *Statistics for Spatial Data*, revised ed., New York: Wiley.
- Cressie, N. (1997), “Comment,” *Journal of the American Statistical Association*, 92, pp. 411–413.
- Cressie, N., and Huang, H.-C. (1999), “Classes of nonseparable, spatio-temporal stationary covariance functions,” *Journal of the American Statistical Association*, 94, pp. 1330–1340.
- Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1991), “Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments,” *Journal of the American Statistical Association*, 86, pp. 953–963.
- De Cesare, L., Myers, D. E., and Posa, D. (2001), “Product-sum covariance for space-time modeling: an environmental application,” *Environmetrics*, 12, pp. 11–23.
- Desroziers, G., and Lafore, J.-P. (1993), “A coordinate transformation for objective frontal analysis,” *Monthly Weather Review*, 121, pp. 1531–1553.
- Feller, W. (1966), *An Introduction to Probability Theory and Its Applications, Vol. II*. New York: Wiley.
- Fischer, C., Joly, A., and Lalaurette, F. (1998), “Error growth and Kalman filtering within an idealized baroclinic flow,” *Tellus A*, 50, pp. 596–615.
- Gneiting, T. (1997), “Normal scale mixtures and dual probability densities,” *Journal of Statistical Computation and Simulation*, 59, pp. 375–384.
- Gneiting, T. (1999), “Correlation functions for atmospheric data analysis,” *Quarterly Journal of the Royal Meteorological Society*, 125, pp. 2449–2464.
- Gneiting, T., and Sasvári, Z. (1999), “The characterization problem for isotropic covariance functions,” *Mathematical Geology*, 31, pp. 105–111.
- Guttorp, P., Meiring, W., and Sampson, P. D. (1997), “Comment,” *Journal of the American Statistical Association*, 92, pp. 405–408.

Handcock, M. S., and Wallis, J. R. (1994), "An approach to spatial-temporal modeling of meteorological fields" (with discussion), *Journal of the American Statistical Association*, 92, pp. 368–390.

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, pp. 1–50.

Kyriakidis, P. C., and Journel, A. G. (1999), "Geostatistical space-time models: a review," *Mathematical Geology*, 31, pp. 651–684.

May, D. R., and Julien, P. Y. (1998), "Eulerian and Lagrangian correlation structures of convective rainstorms," *Water Resources Research*, 34, pp. 2671–2683.

Raftery, A. E., Haslett, J., and McColl, E. (1982), "Wind power: a space-time process?," in *Time Series Analysis: Theory and Practice 2*, ed. O. D. Anderson, Amsterdam: North-Holland, pp. 191–202.

Rodriguez-Iturbe, I., Marani, M., D'Odorico, P., and Rinaldo, A. (1998), "On space-time scaling of cumulated rainfall fields," *Water Resources Research*, 34, pp. 3461–3469.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989), "Design and analysis of computer experiments" (with discussion), *Statistical Science*, 4, pp. 409–435.

Sampson, P. D., and Guttorp, P. (1992), "Nonparametric estimation of nonstationary spatial covariance structure," *Journal of the American Statistical Association*, 87, pp. 108–119.

Sampson, P. D., Damian, D., and Guttorp, P. (2001), "Advances in modeling and inference for environmental processes with nonstationary spatial covariance," to appear in *GeoENV2000 Proceedings*, eds. D. Allard and P. Monestiez.

Sasvári, Z. (1994), *Positive Definite and Definitizable Functions*, Berlin: Akademie Verlag.

Schoenberg, I. J. (1938), "Metric spaces and completely monotone functions," *Annals of Mathematics*, 39, pp. 811–841.

Sølna, K., and Switzer, P. (1996), "Time trend estimation for a geographic region," *Journal of the American Statistical Association*, 91, pp. 577–589.