CHAPTER 4

Geostatistical Space-Time Models, Stationarity, Separability, and Full Symmetry

Tilmann Gneiting, Marc G. Genton, and Peter Guttorp

Contents

4.1 Introduction

Environmental and geophysical processes such as atmospheric pollutant concentrations, precipitation fields and surface winds are characterized by spatial and temporal variability. In view of the prohibitive costs of spatially and temporally dense monitoring networks, one often aims to develop a statistical model in continuous space and time, based on observations at a limited number of monitoring stations. Examples include environmental monitoring and model assessment for surface ozone levels (Guttorp et al., 1994; Carroll et al., 1997; Meiring et al., 1998; Huang and Hsu, 2004), precipitation forecasts (Amani and Lebel, 1997) and the assessment of wind energy resources (Haslett and Raftery, 1989). Geostatistical approaches model the observations as a partial realization of a spatio-temporal, typically Gaussian random function

$$Z(\boldsymbol{s}, t), \quad (\boldsymbol{s}, t) \in \mathrm{R}^d \times \mathrm{R},$$

which is indexed in space by $\boldsymbol{s} \in \mathrm{R}^d$ and in time by $t \in \mathrm{R}$. Henceforth, we assume that second moments for the random function exist and are finite. Optimal least-squares prediction, or kriging, then relies on the appropriate specification of the space-time covariance structure. Generally, the covariance between $Z(\boldsymbol{s}_1, t_1)$ and $Z(\boldsymbol{s}_2, t_2)$ depends on the space-time coordinates $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$, and no further structure may exist. In practice, however, estimation and modeling call for simplifying assumptions, such as stationarity, separability, and full symmetry.

Specifically, the random field Z is said to have a *separable* covariance if there exist purely spatial and purely temporal covariance functions C_{S} and C_{T}, respectively, such that

$$\operatorname{cov}\{Z(\boldsymbol{s}_1, t_1), Z(\boldsymbol{s}_2, t_2)\} = C_{\mathrm{S}}(\boldsymbol{s}_1, \boldsymbol{s}_2) \cdot C_{\mathrm{T}}(t_1, t_2) \tag{4.1}$$

for all space-time coordinates $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ in $\mathrm{R}^d \times \mathrm{R}$. The spatio-temporal covariance structure factors into a purely spatial and a purely temporal component, which allows for computationally efficient estimation and inference. Consequently, separable covariance models have been popular even in situations in which they are not physically justifiable. Many statistical tests for separability have been proposed recently and are based on parametric models (Shitan and Brockwell, 1995; Guo and Billard, 1998; Brown et al., 2000), likelihood ratio tests and subsampling (Mitchell et al., 2005), or spectral methods (Scaccia and Martin, 2005; Fuentes, 2006).

A related notion is that of full symmetry (Gneiting, 2002a; Stein, 2004). The space-time process Z has *fully symmetric* covariance if

$$\operatorname{cov}\{Z(\boldsymbol{s}_1, t_1), Z(\boldsymbol{s}_2, t_2)\} = \operatorname{cov}\{Z(\boldsymbol{s}_1, t_2), Z(\boldsymbol{s}_2, t_1)\} \tag{4.2}$$

for all space-time coordinates $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ in $\mathrm{R}^d \times \mathrm{R}$. Atmospheric, environmental, and geophysical processes are often under the influence of prevailing air or water flows, resulting in a lack of full symmetry. Transport effects of this type are well-known in the meteorological and hydrological literature and have recently been described by Gneiting (2002a), Stein (2005a), and de Luna and Genton (2005), who considered the Irish wind data of Haslett and Raftery (1989), by Wan et al. (2003) for wind power data, by Huang and Hsu (2004) for surface ozone levels, and by Jun and Stein (2004a, 2004b) for atmospheric sulfate concentrations. Separability forms a special case of full symmetry. Hence, covariance structures that are not fully symmetric are nonseparable, and tests for full symmetry (Scaccia and Martin, 2005; Lu and Zimmerman, 2005) can be used to reject separability.

Frequently, trend removal and space deformation techniques (Haslett and Raftery, 1989; Sampson and Guttorp, 1992) allow for a reduction to a stationary covariance structure. The spatio-temporal random function has *spatially stationary* covariance if $\text{cov}\{Z(\boldsymbol{s}_1, t_1),\, Z(\boldsymbol{s}_2, t_2)\}$ depends on the observation sites $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ only through the spatial separation vector, $\boldsymbol{s}_1 - \boldsymbol{s}_2$. It has *temporally stationary* covariance if $\text{cov}\{Z(\boldsymbol{s}_1, t_1),\, Z(\boldsymbol{s}_2, t_2)\}$ depends on the observation times t_1 and t_2 only through the temporal lag, $t_1 - t_2$. If a spatio-temporal process has both spatially and temporally stationary covariance, we say that the process has *stationary* covariance. Under this assumption, there exists a function C defined on $\mathrm{R}^d \times \mathrm{R}$ such that

$$\text{cov}\{Z(\boldsymbol{s}_1, t_1), Z(\boldsymbol{s}_2, t_2)\} = C(\boldsymbol{s}_1 - \boldsymbol{s}_2, t_1 - t_2) \tag{4.3}$$

for all $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ in $\mathrm{R}^d \times \mathrm{R}$. We call C the space-time *covariance function* of the process, and its restrictions $C(\cdot, 0)$ and $C(\mathbf{0}, \cdot)$ are purely spatial and purely temporal covariance functions, respectively. For tests of stationarity we point to Fuentes (2005) and references therein.

The remainder of the chapter is organized as follows. Section 4.2 reviews the geostatistical approach to space-time modeling and returns to the aforementioned notions of stationarity, separability, and full symmetry. Clearly, these notions apply to correlation structures also and in many cases, such as the case study below, a discussion in terms of correlations is preferable. Section 4.3 turns to recent advances in the literature on stationary space-time covariance functions. This material is largely expository, and much progress has been made since the reviews of Kyriakidis and Journel (1999) and Gneiting and Schlather (2002). Section 4.4 provides a case study based on the Irish wind data of Haslett and Raftery (1989). This rich and well-known data set continues to inspire methodological advances, and recent analyses include the papers by Gneiting (2002a), Stein (2005a, 2005b), and de Luna and Genton (2005). We consider time-forward kriging predictions based on increasingly complex spatio-temporal correlation models, and our experiments suggest that the use of the richer, more realistic models results in improved predictive performance.

4.2 Geostatistical space-time models

4.2.1 Spatio-temporal domains

Geostatistical approaches have been developed to fit random function models in continuous space and time, based on a limited number of spatially and/or temporally dispersed observations. Hence, the natural domain for a geostatistical space-time model is $\mathrm{R}^d \times \mathrm{R}$, where R^d stands for space and R for time. Physically, there is clear-cut separation between the spatial and time dimensions, and a realistic statistical model will take account thereof. This contrasts with a purely mathematical perspective in which $\mathrm{R}^d \times \mathrm{R} = \mathrm{R}^{d+1}$ with no differences between the coordinates. While the latter equality may seem (and indeed is) trivial, it has important implications. In particular, all technical results on spatial covariance functions or on least-squares prediction, or kriging, in Euclidean spaces apply directly to space-time problems, simply by separating a vector into its spatial and temporal components.

Henceforth, we focus on covariance structures for spatio-temporal random functions $Z(\boldsymbol{s}, t)$ where $(\boldsymbol{s}, t) \in \mathrm{R}^d \times \mathrm{R}$. Other spatio-temporal domains are relevant in practice as well. Monitoring data are frequently observed at fixed temporal lags, and it may suffice to model a random function on $\mathrm{R}^d \times \mathrm{Z}$, with time considered discrete. The time autoregressive Gaussian models of Storvik et al. (2002) form a promising tool in this direction. A related result of Stein (2005a) characterizes space-time covariance functions that correspond to a temporal Markov structure. In atmospheric and geophysical applications, the spatial domain of interest is frequently expansive or global and the curvature of the earth needs to be taken into account (Banerjee, 2005). In this type of situation, random functions defined on $\mathrm{S} \times \mathrm{R}$ or $\mathrm{S} \times \mathrm{Z}$ become crucial, where S denotes a sphere in three-dimensional space. Perhaps the simplest way of defining a random field on $\mathrm{S} \times \mathrm{R}$ or $\mathrm{S} \times \mathrm{Z}$ is by defining a random function on $\mathrm{R}^3 \times \mathrm{R}$ and restricting it to the desired domain. Gneiting (1999), Stein (2005a, 2005b), and particularly Jun and Stein (2004b) discuss these and other ways of defining suitable parametric covariance models on global spatial or spatio-temporal domains.

4.2.2 Stationarity, separability, and full symmetry

We consider a generic, typically Gaussian spatio-temporal random function $Z(\boldsymbol{s}, t)$ where $(\boldsymbol{s}, t) \in \mathrm{R}^d \times \mathrm{R}$. The covariance between $Z(\boldsymbol{s}_1, t_1)$ and $Z(\boldsymbol{s}_2, t_2)$ generally depends on the space-time coordinates $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$, and no further structure may exist. In practice, simplifying assumptions are required, such as the aforementioned notions of stationarity, separability, and full symmetry, which are defined in (4.1), (4.2), and (4.3), respectively. In addition, it is sometimes desirable that the space-time process Z has *compactly supported* covariance, so that $\mathrm{cov}\{Z(\boldsymbol{s}_1, t_1), Z(\boldsymbol{s}_2, t_2)\} = 0$ whenever $\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ and/or $|t_1 - t_2|$ are sufficiently large. Gneiting (2002b) reviews parametric models of compactly supported, stationary, and isotropic covariance functions in a spatial setting. Unfortunately, the straightforward idea of thresholding covariances to zero using the product with an indicator function, such as the truncation device of Haas (2002, p. 320), yields invalid covariance models. That said, compactly supported covariances are attractive in that they allow for computationally efficient spatio-temporal prediction and simulation. Furrer et al. (2006), for instance, proposed their use for kriging large spatial datasets.

Figure 4.1 summarizes the relationships between the various notions in terms of classes of space-time covariance functions, and an analogous scheme applies to correlation structures. The largest class is that of general, stationary or non-stationary covariance functions. A separable covariance can be stationary or non-stationary, and similarly for fully symmetric covariances. However,

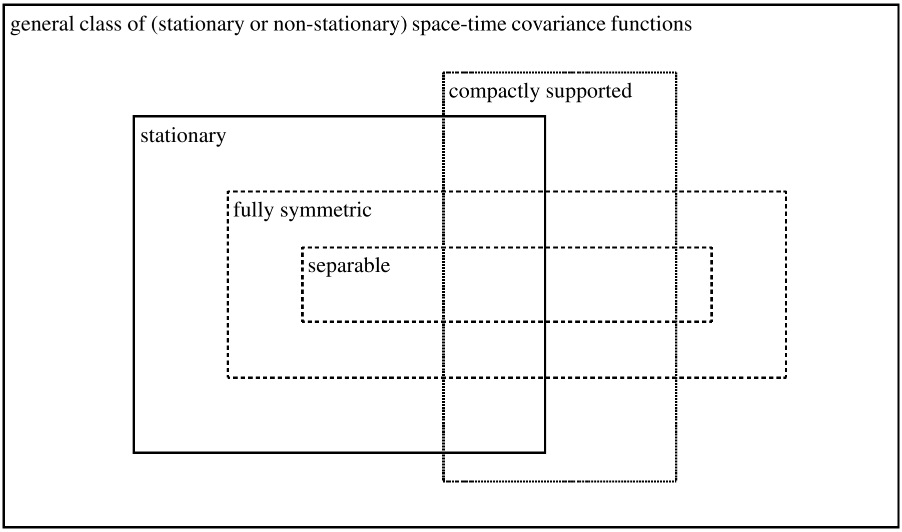

Figure 4.1 Schematic illustration of the relationships between separable, fully symmetric, stationary, and compactly supported covariances within the general class of (stationary or non-stationary) space-time covariance functions. An analogous scheme applies to correlation structures.

a separable covariance function is always fully symmetric, but not vice versa, and this has implications in testing and model fitting. In particular, to reject separability it suffices to reject full symmetry. We return to these issues in Section 4.4 below, when we discuss modeling strategies using the example of the Irish wind data.

Occasionally, the second-order structure of a spatio-temporal random function is modeled based on variances rather than covariances. Specifically, when viewed as a function of space-time coordinates $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ in $\mathrm{R}^d \times \mathrm{R}$, the quantity

$$\frac{1}{2}\,\mathrm{var}(Z(\boldsymbol{s}_1, t_1) - Z(\boldsymbol{s}_2, t_2)) \tag{4.4}$$

is called a *non-stationary variogram.* Separability is not a meaningful assumption for variograms, but we discuss full symmetry and stationarity. The space-time process has a *fully symmetric* variogram structure if

$$\mathrm{var}(Z(\boldsymbol{s}_1, t_1) - Z(\boldsymbol{s}_2, t_2)) = \mathrm{var}(Z(\boldsymbol{s}_1, t_2) - Z(\boldsymbol{s}_2, t_1))$$

for all $(\boldsymbol{s}_1, t_1)$ and $(\boldsymbol{s}_2, t_2)$ in $\mathrm{R}^d \times \mathrm{R}$. The space-time random function has a *spatially intrinsically stationary* variogram if the non-stationary variogram (4.4) depends on the observation sites $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ only through the spatial separation vector, $\boldsymbol{s}_1 - \boldsymbol{s}_2$. Similarly, it has *temporally intrinsically stationary* variogram if (4.4) depends on the observation times t_1 and t_2 only through the temporal lag, $t_1 - t_2$. The process has *intrinsically stationary* variogram if it has both spatially intrinsically stationary and temporally intrinsically stationary variogram. Under this latter assumption, there exists a function γ defined on $\mathrm{R}^d \times \mathrm{R}$ such that

$$\frac{1}{2}\,\mathrm{var}(Z(\boldsymbol{s}_1,t_1) - Z(\boldsymbol{s}_2,t_2)) = \gamma(\boldsymbol{s}_1 - \boldsymbol{s}_2, t_1 - t_2)$$

for all $(\boldsymbol{s}_1,t_1)$ and $(\boldsymbol{s}_2,t_2)$ in $\mathrm{R}^d \times \mathrm{R}$, and we call γ the *stationary variogram* of the process Z. Variograms exist under slightly weaker assumptions than covariances, and we refer to Gneiting et al. (2001) and Ma (2003c) for the various analogies and correspondences between the two classes of dependence measures. The use of variograms has been particularly popular in purely spatial problems. In discussing geostatistical space-time models, we follow the literature and focus our attention on covariances and correlations.

4.2.3 Positive definiteness

A crucial notion for stationary covariance functions is that of positive definiteness. Specifically, if $Z(\boldsymbol{s},t)$ is a random function in $\mathrm{R}^d \times \mathrm{R}$, k is a positive integer, and $(\boldsymbol{s}_1,t_1),\ldots,(\boldsymbol{s}_k,t_k)$ are space-time coordinates in $\mathrm{R}^d \times \mathrm{R}$, the covariance matrix of the random vector $(Z(\boldsymbol{s}_1,t_1),\ldots,Z(\boldsymbol{s}_k,t_k))'$ is nonnegative definite. If the process is stationary with covariance function C on $\mathrm{R}^d \times \mathrm{R}$, this matrix can be written as

$$(C(\boldsymbol{s}_i - \boldsymbol{s}_j, t_i - t_j))_{i,j=1,\ldots,k}. \tag{4.5}$$

A complex-valued function C on $\mathrm{R}^d \times \mathrm{R}$ is called *positive definite* if the matrix in (4.5) is nonnegative definite for all finite collections of space-time coordinates $(\boldsymbol{s}_1,t_1),\ldots,(\boldsymbol{s}_k,t_k)$ in $\mathrm{R}^d \times \mathrm{R}$. Any positive definite function is Hermitian, and a positive definite function is real-valued if and only if it is symmetric. It is well known that the class of stationary covariance functions is identical to the class of symmetric positive definite functions, and we use the two terms interchangeably. When we talk of a positive definite function, we explicitly allow for complex-valued functions.

Unfortunately, it can be quite difficult in general to check whether a function is positive definite, and this forms one of the key difficulties in the construction of parametric space-time covariance models. Until a few years ago, geostatistical space-time models were largely based on stationary, separable covariance functions of the form

$$C(\boldsymbol{h},u) = C_{\mathrm{S}}(\boldsymbol{h}) \cdot C_{\mathrm{T}}(u), \quad (\boldsymbol{h},u) \in \mathrm{R}^d \times \mathrm{R} \tag{4.6}$$

where $C_{\mathrm{S}}(\boldsymbol{h})$ and $C_{\mathrm{T}}(u)$ are stationary, purely spatial and purely temporal covariance functions, respectively. A convenient choice is an isotropic model, $C_{\mathrm{S}}(\boldsymbol{h}) = c_0(\|\boldsymbol{h}\|)$, $\boldsymbol{h} \in \mathrm{R}^d$, where c_0 is one of the standard models in geostatistics, such as the powered exponential class (see, for instance, Diggle

Table 4.1 Some parametric classes of isotropic covariance functions. The Whittle-Matérn covariance is defined in terms of the modified Bessel function, K_ν.

Class	Functional form	Parameters
Powered exponential	$c_0(r) = \sigma^2 \exp(-(\theta r)^\gamma)$	$0 < \gamma \leq 2;\ \theta > 0;\ \sigma > 0$
Whittle-Matérn	$c_0(r) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\theta r)^\nu K_\nu(\theta r)$	$\nu > 0;\ \theta > 0;\ \sigma > 0$
Cauchy	$c_0(r) = \sigma^2 (1 + (\theta r)^\gamma)^{-\nu}$	$0 < \gamma \leq 2;\ \nu > 0;\ \theta > 0;\ \sigma > 0$

et al., 1998); the Whittle-Matérn class (Whittle, 1954; Matérn, 1986); and the Cauchy class (Gneiting and Schlather, 2004). These parametric models are listed in Table 4.1. Similarly, the temporal covariance function can conveniently be chosen as $C_{\mathrm{T}}(u) = c_0(|u|)$, $u \in \mathrm{R}$, where c_0 is another, possibly distinct standard model that guarantees positive definiteness. The product (4.6) yields a positive definite function whenever C_{S} and C_{T} are positive definite on R^d and R, respectively, because products — and also sums, convex combinations and limits — of positive definite functions are positive definite.

4.3 Stationary space-time covariance functions

4.3.1 Bochner's Theorem

The celebrated theorem of Bochner (1955) states that a continuous function is positive definite if and only if it is the Fourier transform of a finite nonnegative measure. This allows for the following characterization of stationary space-time covariance functions.

Theorem 4.3.1 **(Bochner)**
Suppose that C is a continuous and symmetric function on $R^d \times R$. Then C is a covariance function if and only if it is of the form

$$C(\boldsymbol{h}, u) = \iint e^{\,i(\boldsymbol{h}'\boldsymbol{\omega} + u\tau)}\, \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\boldsymbol{h}, u) \in R^d \times R, \tag{4.7}$$

where F is a finite, non-negative and symmetric measure on $R^d \times R$.

In other words, the class of stationary space-time covariance functions on $\mathrm{R}^d \times \mathrm{R}$ is identical to the class of the Fourier transforms of finite, non-negative and symmetric measures on this domain. The measure F in the representation

(4.7) is often called the *spectral measure.* If C is integrable, the spectral measure is absolutely continuous with Lebesgue density

$$f(\boldsymbol{\omega}, \tau) = (2\pi)^{-(d+1)} \iint e^{-i(\boldsymbol{h}'\boldsymbol{\omega}+u\tau)}\, C(\boldsymbol{h}, u)\, \mathrm{d}\boldsymbol{h}\, \mathrm{d}u, \quad (\boldsymbol{\omega}, \tau) \in \mathrm{R}^d \times \mathrm{R},$$

and f is called the *spectral density.* If the spectral density exists, the representation (4.7) in Bochner's theorem reduces to

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

and C and f can be obtained from each other via the Fourier transform.

A stationary space-time covariance function is separable if there exist stationary, purely spatial, and purely temporal covariance functions C_{S} and C_{T}, respectively, such that (4.6) holds or, equivalently, if we can factor the space-time covariance function as

$$C(\boldsymbol{h}, u) = \frac{C(\boldsymbol{h}, 0) \cdot C(\mathbf{0}, u)}{C(\mathbf{0}, 0)}$$

for all $(\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}$ (Mitchell et al., 2005). In spectral terms, a stationary covariance function is separable if and only if the spectral measure factors as a product measure over the spatial and temporal domains, respectively. In particular, if the spectral density exists, it factors as a product over the domains. A stationary space-time covariance function is fully symmetric if

$$C(\boldsymbol{h}, u) = C(\boldsymbol{h}, -u) = C(-\boldsymbol{h}, u) = C(-\boldsymbol{h}, -u)$$

for all $(\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}$. In the purely spatial context, this property is also known as axial symmetry (Scaccia and Martin, 2005) or reflection symmetry (Lu and Zimmerman, 2005). For fully symmetric covariances, Bochner's theorem can be specialized as follows.

Theorem 4.3.2
Suppose that C is a continuous function on $R^d \times R$. Then C is a stationary, fully symmetric covariance if and only if it is of the form

$$C(\boldsymbol{h}, u) = \iint \cos(\boldsymbol{h}'\boldsymbol{\omega}) \cos(u\tau)\, \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\boldsymbol{h}, u) \in R^d \times R, \tag{4.8}$$

where F is a finite, non-negative measure on $R^d \times R$.

The proof of this result is given in the Appendix. If C is fully symmetric and the spectral density f exists, then f can be chosen as fully symmetric, too, that is,

$$f(\boldsymbol{\omega}, \tau) = f(\boldsymbol{\omega}, -\tau) = f(-\boldsymbol{\omega}, \tau) = f(-\boldsymbol{\omega}, -\tau)$$

for all $(\boldsymbol{\omega}, \tau) \in \mathrm{R}^d \times \mathrm{R}$. A similar characterization applies in terms of general spectral measures. If the space-time covariance function has additional structure, such as spherical symmetry with respect to $\boldsymbol{h} \in \mathrm{R}^d$ for each $u \in \mathrm{R}$, the representation (4.8) can be further specialized. Theorem 2 of Ma (2003a) is a result of this type.

The above results apply to continuous functions. In practice, fitted stationary space-time covariance functions often involve a *nugget effect*, that is, a discontinuity at the origin. In the spatio-temporal context, the nugget effect could be purely spatial, purely temporal or spatio-temporal and takes the general form

$$C(\boldsymbol{h}, u) = a\,\delta_{(\boldsymbol{h},u)=(\boldsymbol{0},0)} + b\,\delta_{\boldsymbol{h}=0} + c\,\delta_{u=0}, \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}, \tag{4.9}$$

where a, b, and c are nonnegative constants and δ denotes an indicator function. Products and sums of continuous covariances and functions of the form (4.9) are valid covariances, and for all practical purposes functions of this type exhaust the class of valid stationary covariance functions (Gneiting and Sasvári, 1999). For the Irish wind data which we discuss in Section 4.4 below, we fit a purely spatial nugget effect.

4.3.2 Cressie-Huang representation

The following result of Cressie and Huang (1999) characterizes the class of stationary space-time covariance functions under the additional assumption of integrability.

Theorem 4.3.3 **(Cressie and Huang)**
Suppose that C is a continuous, bounded, integrable, and symmetric function on $R^d \times R$. Then C is a stationary covariance if and only if

$$\rho(\boldsymbol{\omega}, u) = \int e^{-i\boldsymbol{h}'\boldsymbol{\omega}}\, C(\boldsymbol{h}, u)\, \mathrm{d}\boldsymbol{h}, \quad u \in R, \tag{4.10}$$

is positive definite for almost all $\boldsymbol{\omega} \in R^d$.

Cressie and Huang (1999) used Theorem 4.3.3 to construct stationary space-time covariance functions through closed form Fourier inversion in R^d. Specifically, they considered functions of the form

$$C(\boldsymbol{h}, u) = \int e^{\,i\boldsymbol{h}'\boldsymbol{\omega}} \rho(\boldsymbol{\omega}, u)\, \mathrm{d}\boldsymbol{\omega}, \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R},$$

where $\rho(\boldsymbol{\omega}, u)$, $u \in \mathrm{R}$, is a continuous positive definite function for all $\boldsymbol{\omega} \in \mathrm{R}^d$. Gneiting (2002a) gave a criterion that is based on this construction but does not depend on closed form Fourier inversion and does not require integrability. Recall that a continuous function $\varphi(r)$ defined for $r > 0$ or $r \geq 0$ is *completely* *monotone* if it possesses derivatives $\varphi^{(n)}$ of all orders and $(-1)^n\varphi^{(n)}(r) \geq 0$ for $r > 0$ and $n = 0, 1, 2, \ldots$. Gneiting (2002a) and Ma (2003b) gave various examples of completely monotone functions. In particular, if c_0 is any of the functions listed in Table 4.1, then $\varphi(r) = c_0(r^{1/2})$, $r \geq 0$, is a completely monotone function. Examples of positive functions with a completely monotone derivative include $\psi(r) = (ar^\alpha + 1)^\beta$ and $\psi(r) = \ln(ar^\alpha + 1)$, where $\alpha \in (0, 1]$, $\beta \in (0, 1]$ and $a > 0$.

Theorem 4.3.4 **(Gneiting)**
Suppose that $\varphi(r)$, $r \geq 0$, is a completely monotone function, and that $\psi(r)$, $r \geq 0$, is a positive function with a completely monotone derivative. Then

$$C(\boldsymbol{h}, u) = \frac{1}{\psi(u^2)^{d/2}}\,\varphi\!\left(\frac{\|\boldsymbol{h}\|^2}{\psi(u^2)}\right), \quad (\boldsymbol{h}, u) \in R^d \times R, \tag{4.11}$$

is a stationary covariance function on $R^d \times R$.

The specific choices $\varphi(r) = \sigma^2\exp(-cr^\gamma)$ and $\psi(r) = (1 + ar^\alpha)^\beta$ recover Equation (14) of Gneiting (2002a) and yield the parametric family

$$C(\boldsymbol{h}, u) = \frac{\sigma^2}{(1 + a|u|^{2\alpha})^{\beta d/2}}\,\exp\!\left(-\frac{c\|\boldsymbol{h}\|^{2\gamma}}{(1 + a|u|^{2\alpha})^{\beta\gamma}}\right), \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}, \tag{4.12}$$

of stationary space-time covariance functions. Here, a and c are nonnegative scale parameters of time and space, respectively. The smoothness parameters α and γ and the space-time interaction parameter β take values in $(0, 1]$, and σ^2 is the variance of the spatio-temporal process. The purely spatial covariance function, $C(\boldsymbol{h}, 0)$, is of the powered exponential form, and the purely temporal covariance function, $C(\mathbf{0}, u)$, belongs to the Cauchy class.

Clearly, any stationary covariance of the form (4.11) is fully symmetric. Furthermore, under the assumption of full symmetry the test functions $\rho(\boldsymbol{\omega}, u)$ of Theorem 4.3.3 are real-valued and symmetric functions of $u \in \mathrm{R}$. If C is not fully symmetric then $\rho(\boldsymbol{\omega}, u)$ is generally complex-valued. For instance, the function $C(h, u) = \exp(-h^2 + hu - u^2)$ has Fourier transform proportional to $\exp(-\frac{1}{3}(\omega^2 + \omega\tau + \tau^2))$ and therefore is a stationary covariance on $\mathrm{R} \times \mathrm{R}$. The associated test function $\rho(\omega, u)$, $u \in \mathrm{R}$, is proportional to $\exp(-\frac{1}{4}(3u^2 + 2i\omega u))$ and positive definite yet generally complex-valued.

4.3.3 Fully symmetric, stationary covariance functions

Non-separable, fully symmetric stationary space-time covariance functions can be constructed as mixtures of separable covariances. The following theorem summarizes relevant results by De Iaco et al. (2002) and Ma (2003b). In view of Theorem 4.3.2, the construction is completely general.

Theorem 4.3.5
Let μ be a finite, nonnegative measure on a non-empty set Θ. Suppose that for each $\theta \in \Theta$, C_{S}^{θ} and C_{T}^{θ} are stationary covariances on R^d and R, respectively, and suppose that $C_{\mathrm{S}}^{\theta}(\mathbf{0})\, C_{\mathrm{T}}^{\theta}(0)$ has finite integral over Θ. Then

$$C(\boldsymbol{h}, u) = \int C_{\mathrm{S}}^{\theta}(\boldsymbol{h})\, C_{\mathrm{T}}^{\theta}(u)\, \mathrm{d}\mu(\theta), \quad (\boldsymbol{h}, u) \in R^d \times R, \tag{4.13}$$

is a stationary covariance function on $R^d \times R$.

Explicit constructions of the form (4.13) have been reported by various authors. Perhaps the simplest special case is the product-sum model of De Iaco et al. (2001),

$$C(\boldsymbol{h}, u) = a_0 C_{\mathrm{S}}^{0}(\boldsymbol{h}) C_{\mathrm{T}}^{0}(u) + a_1 C_{\mathrm{S}}^{1}(\boldsymbol{h}) + a_2 C_{\mathrm{T}}^{2}(u), \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R},$$

where a_0, a_1 and a_2 are nonnegative coefficients and C_{S}^{0}, C_{S}^{1} and C_{T}^{0}, C_{T}^{2} are stationary, purely spatial and purely temporal covariance functions, respectively. Examples 1 and 2 of De Iaco et al. (2002) consider the particular case of (4.13) in which μ is a gamma distribution on $\Theta = [0, \infty)$ and both C_{S}^{θ} and C_{T}^{θ} are of powered exponential type. This construction yields the parametric family

$$C(\boldsymbol{h}, u) = \sigma^2 \left(1 + \left\| \frac{\boldsymbol{h}}{a} \right\|^{\alpha} + \left| \frac{u}{b} \right|^{\beta} \right)^{-\gamma}, \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}, \tag{4.14}$$

of stationary space-time covariance functions, where $\alpha \in (0, 2]$, $\beta \in (0, 2]$, $\gamma > 0$, $a > 0$, $b > 0$, and $\sigma > 0$. De Iaco et al. (2002) gave more stringent parameter ranges which are unnecessarily restrictive. Ma (2003a) reported various interesting examples of parametric space-time covariance functions that are also based on the mixture representation (4.13).

4.3.4 Stationary covariance functions that are not fully symmetric

Environmental, atmospheric, and geophysical processes are often influenced by prevailing winds or ocean currents, which are incompatible with the assumption of full symmetry. In this type of situation, the general idea of a Lagrangian reference frame applies, which can be thought of as attached to and moving with the center of an air or water mass. This approach has been studied in fluid dynamics, meteorology, and hydrology, and we refer to May and Julien (1998) for a comparison of empirical correlations observed in the classical (fixed) Eulerian and in the (moving) Lagrangian reference frame. Specifically, consider a purely spatial random field with stationary covariance function C_{S} on R^d, and suppose that the entire field moves time-forward with random velocity vector $\boldsymbol{V} \in \mathrm{R}^d$. The resulting spatio-temporal random field has stationary covariance

$$C(\boldsymbol{h}, u) = \mathrm{E}\, C_{\mathrm{S}}(\boldsymbol{h} - \boldsymbol{V} u), \quad (\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}, \tag{4.15}$$

where the expectation is taken with respect to the random vector $\boldsymbol{V}$. Similar constructions of stationary space-time covariance functions have been reported by Cox and Isham (1988) and Ma (2003b), among others. Stationary space-time covariance functions that are not fully symmetric can also be constructed on the basis of diffusion equations or stochastic partial differential equations. These and related approaches have been discussed by Jones and Zhang (1997), Christakos (2000), Brown et al. (2000), Kovolos et al. (2004), Jun and Stein (2004b), and Stein (2005a, 2005b).

For the specification of the random velocity vector $\boldsymbol{V}$ in the Lagrangian covariance (4.15), various choices can be physically motivated and justified. We discuss these in the context of atmospheric transport effects driven by prevailing winds. The simplest case occurs when $\boldsymbol{V} = \boldsymbol{v}$ is constant and represents the mean or prevailing wind. Gupta and Waymire (1987) referred to this as the *frozen field* model. For the Irish wind data of Haslett and Raftery (1989), for instance, $\boldsymbol{v}$ can be identified with the prevailing westerly wind, and we give details in Section 4.4 below. Alternatively, $\boldsymbol{V}$ might attain a small number of distinct values that represent wind regimes, or we might identify the distribution of $\boldsymbol{V}$ with the empirical distribution of velocity vectors, or a smoothed version thereof, as inferred from meteorological records. Finally, the distribution of $\boldsymbol{V}$ could be updated dynamically according to the current state of the atmosphere. This option yields nonstationary, flow-dependent covariance structures similar to those posited by Riishøgaard (1998) and Huang and Hsu (2004). A related approach has been studied under the heading of Lagrangian kriging (Amani and Lebel, 1997), which operates directly within the (moving) Lagrangian reference frame. The recent advent of ensemble Kalman filter techniques in atmospheric and oceanic data assimilation (Evensen, 1994; Houtekamer and Mitchell, 1998; Hamill and Snyder, 2000) marks another promising avenue to nonstationary space-time covariance modeling based on multiple realizations of an underlying spatio-temporal random field.

4.3.5 Taylor's hypothesis

A stationary space-time covariance function C on $\mathbb{R}^d \times \mathbb{R}$ satisfies *Taylor's hypothesis* (Taylor, 1938, p. 478) if there exists a velocity vector $\boldsymbol{v} \in \mathbb{R}^d$ such that

$$C(\mathbf{0}, u) = C(\boldsymbol{v}u, 0), \quad u \in \mathbb{R}. \tag{4.16}$$

Taylor's hypothesis concerns the relationships between the purely spatial and the purely temporal covariance functions only and has found widespread interest in fluid dynamics, meteorology and hydrology. Zawadzki (1973) argued on the basis of empirical correlations that Taylor's hypothesis is plausible for precipitation data and temporal lags less than 40 minutes. Gupta and Waymire (1987) and Cox and Isham (1988) studied the approximate validity of the hypothesis for various space-time covariance models. The following covariance functions, among others, admit Taylor's hypothesis exactly.

1. The covariance function (4.15) for the frozen field model, that is, $C(\boldsymbol{h}, u) = C_S(\boldsymbol{h} - \boldsymbol{v}u)$ where C_S is a stationary covariance on R^d and $\boldsymbol{v} \in \mathrm{R}^d$ is a nonzero velocity vector, satisfies (4.16).
2. Geometrically anisotropic models of the form $C(\boldsymbol{h}, u) = c_0((a^2\|\boldsymbol{h}\|^2 + b^2u^2)^{1/2})$, where a and b are positive constants and c_0 is any of the parametric models in Table 4.1, admit Taylor's relationship (4.16) with $\boldsymbol{v} = (b/a, 0, \ldots, 0)' \in \mathrm{R}^d$.
3. Separable models of the form $C(\boldsymbol{h}, u) = c_0(\|\boldsymbol{h}\|)\, c_0(|u|)$, where c_0 is any of the parametric models in Table 4.1, satisfy (4.16) with $\boldsymbol{v} = (1, 0, \ldots, 0)' \in \mathrm{R}^d$.
4. Putting $d = 2$, $\varphi(r) = (1 + cr^\alpha)^{-\nu}$ and $\psi(r) = (1 + ar^\alpha)^\nu$ in (4.11), where $a > 0$, $c > 0$, $\alpha \in (0, 1]$, and $\nu \in (0, 1]$, yields

$$C(\boldsymbol{h}, u) = \sigma^2(1 + a|u|^{2\alpha})^{-\nu}\left(1 + \frac{c\|\boldsymbol{h}\|^{2\alpha}}{1 + (a|u|^{2\alpha})^{\alpha\nu}}\right)^{-\nu}, \quad (\boldsymbol{h}, u) \in \mathrm{R}^2 \times \mathrm{R}.$$

   This function admits Taylor's hypothesis with $\boldsymbol{v} = ((a/c)^{1/(2\alpha)}, 0)'$. Generally, if $d = 2$ the covariance function (4.11) satisfies (4.16) if and only if there exists a nonnegative number b such that the product $\varphi(r)\,\psi(br)$ does not depend on r.
5. If $\alpha = \beta \in (0, 2]$, the stationary space-time covariance function (4.14) satisfies Taylor's relationship (4.16) with $\boldsymbol{v} = (a/b, 0, \ldots, 0)' \in \mathrm{R}^d$.

4.4 Irish wind data

We turn to a case study that considers the Irish wind data of Haslett and Raftery (1989). The dataset consists of time series of daily average wind speed at eleven synoptic meteorological stations in Ireland during the period 1961–1978, as described in Table 4.2. The observations for February 29 were removed in the interest of a convenient handling of the seasonal trend component. We refer to Haslett and Raftery (1989) for background information and to Gneiting (2002a), Stein (2005a, 2005b), and de Luna and Genton (2005) for subsequent analyses of the Irish wind data. To allow for an out-of-sample evaluation of predictive performance, we split the data set into a training period consisting of years 1961–1970 and a test period, comprising 1971–1978. Following the aforementioned authors, we apply the square root transform to the time series of daily average wind speed, fit and extract a common seasonal trend component, and remove the station-specific means, resulting in time series of *velocity measures* at the eleven meteorological stations. We apply the same set of transformations to the training data and to the test data, with the common seasonal trend component and the station-specific means estimated from the training data. The square root transform stabilizes the variance over both stations and time periods and makes the marginal distributions approximately normal.

Table 4.2 Latitude, longitude, and mean wind speed (in $\mathrm{m} \cdot \mathrm{s}^{-1}$) at eleven meteorological stations in Ireland, as observed in 1961–1978 and described by Haslett and Raftery (1989)

Station	Latitude	Longitude	Mean wind
Valentia (Val)	51° 56′ N	10° 15′ W	5.48
Belmullet (Bel)	54° 14′ N	10° 00′ W	6.75
Claremorris (Cla)	53° 43′ N	8° 59′ W	4.32
Shannon (Sha)	52° 42′ N	8° 55′ W	5.38
Roche's Point (Roc)	51° 48′ N	8° 15′ W	6.36
Birr (Birr)	53° 05′ N	7° 53′ W	3.65
Mullingar (Mul)	53° 32′ N	7° 22′ W	4.38
Malin Head (Mal)	55° 22′ N	7° 20′ W	8.03
Kilkenny (Kil)	52° 40′ N	7° 16′ W	3.25
Clones (Clo)	54° 11′ N	7° 14′ W	4.48
Dublin (Dub)	53° 26′ N	6° 15′ W	5.05

4.4.1 Exploratory analysis

This section provides an initial study of the spatio-temporal correlation structure for the velocity measures during the training period (1961–1970). Figure 4.2 shows the empirical correlations of contemporary velocity measures at the 55 pairs of distinct meteorological stations as a function of the Euclidean distance between the sites. The correlations decay with distance, and we fit a stationary spatial correlation function of the form

$$C_{\mathrm{S}}(\boldsymbol{h}) = (1-\nu)\exp(-c\|\boldsymbol{h}\|) + \nu\,\delta_{\boldsymbol{h}=\boldsymbol{0}}, \tag{4.17}$$

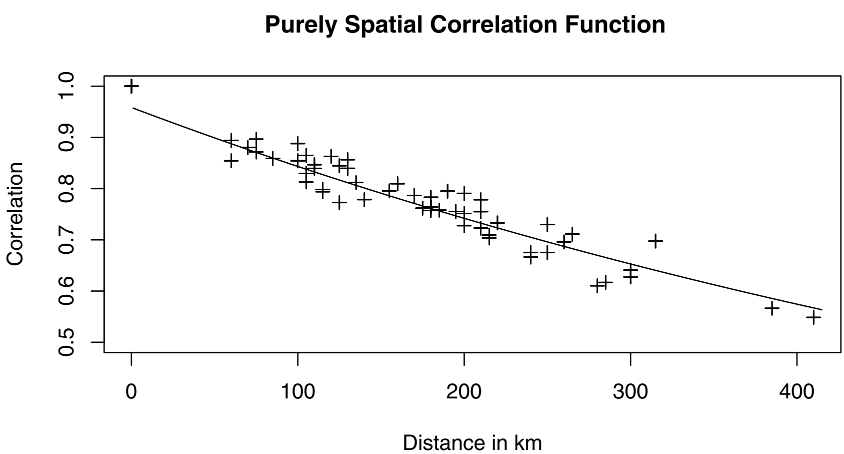

Figure 4.2 Empirical correlations and the fitted spatial correlation function (4.17) for the velocity measures during the training period (1961–1970).

Table 4.3 Empirical correlations between velocity measures at selected meteorological stations in Ireland, as observed in the training period (1961–1970). The table shows the correlations between velocity measures at a temporal lag of one day, with the westerly station leading one day (WE) or lagging one day (EW), respectively.

Westerly station	Easterly station	WE	EW
Valentia	Roche's Point	.48	.35
Belmullet	Clones	.52	.39
Claremorris	Mullingar	.51	.41
Claremorris	Dublin	.50	.36
Shannon	Kilkenny	.51	.39
Mullingar	Dublin	.49	.45

that is, a convex combination of an exponential model and a nugget effect. The weighted least-squares (WLS) estimates for the spatial nugget effect and the scale parameter are $\widehat{\nu} = 0.0415$ and $\widehat{c} = 0.00128$, with distances measured in kilometers. Furthermore, we fit a stationary, purely temporal correlation function of Cauchy type,

$$C_{\mathrm{T}}(u) = (1 + a|u|^{2\alpha})^{-1}, \tag{4.18}$$

for temporal lags $|u| \leq 3$ days; higher temporal lags are irrelevant given our goal of one-day ahead forecasts at the stations. The WLS estimates for the parameters in (4.18) are $\widehat{a} = 0.972$ and $\widehat{\alpha} = 0.834$, respectively. Plots of the empirical purely temporal correlation functions at the meteorological stations are shown in Figure 4 of Haslett and Raftery (1989).

We now describe some of the interactions between the spatial and the temporal correlation structures. Winds in Ireland are predominantly westerly, so that the velocity measures propagate from west to east. Hence, we expect the empirical correlation between a westerly station today and an easterly station later on to be higher than vice versa. Indeed, Table 4.3 and Figure 4.3(a) illustrate the lack of full symmetry — and thereby the lack of separability — in the correlation structure of the velocity measures. Table 4.3 compares the west-to-east and east-to-west correlations at a temporal lag of one day for the six pairs of meteorological stations with the most dominant longitudinal (east-west) component of the separation vector between the stations. The violation of full symmetry is most pronounced for Valentia and Kilkenny, though, with lag one west-to-east and east-to-west correlations reaching .50 and .30, respectively.

Figure 4.3(a) shows the difference between the empirical west-to-east and east-to-west correlations for all 55 pairs of distinct stations as a function of the longitudinal (east-west) component of the spatial separation vector, for temporal lags of one day (red), two days (green), and three days (blue). Similar graphical displays can be found in de Luna and Genton (2005) and

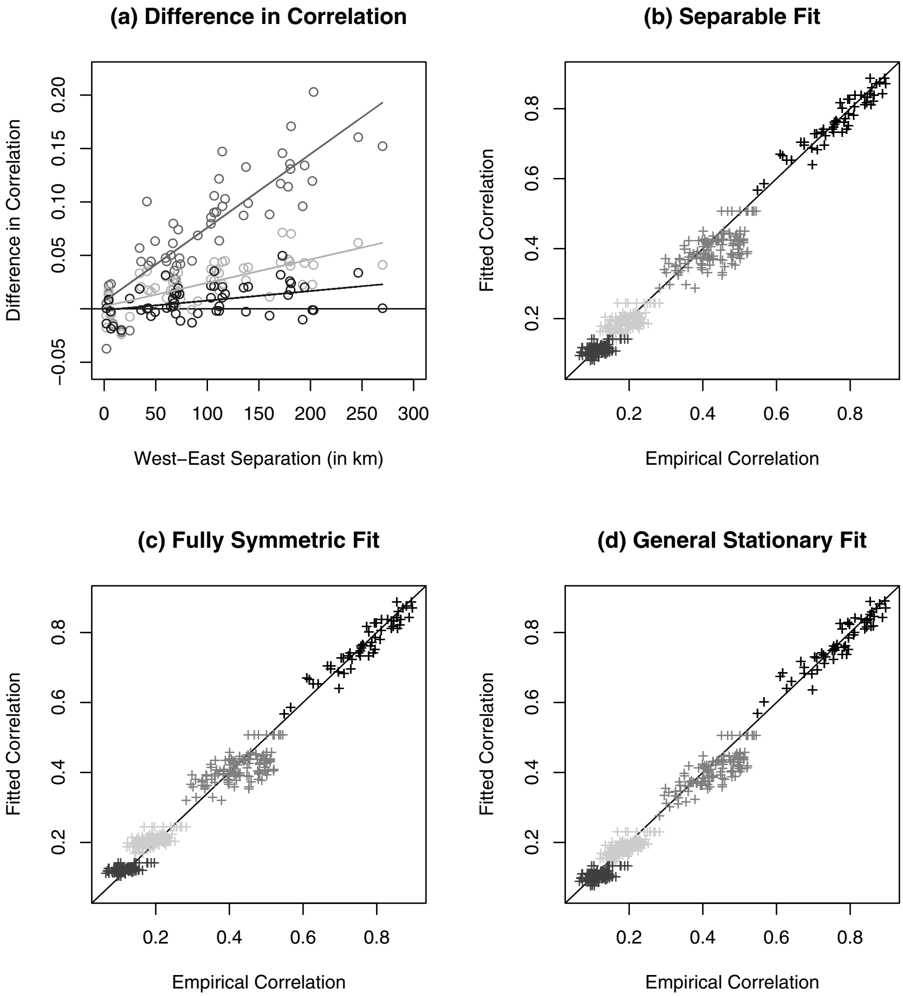

Figure 4.3 (SEE COLOR INSERT FOLLOWING PAGE 142) (a) Difference between the empirical west-to-east and east-to-west correlations for the 55 pairs of distinct stations and temporal lag one day (red), two days (green), and three days (blue), in dependence on the longitudinal (east-west) distance between the stations (in km). Linear least-squares fits are shown as well. (b), (c) and (d) Empirical versus fitted correlations at temporal lag zero (black), one day (red), two days (green), and three days (blue), for the separable model (4.19), the fully symmetric model (4.20), and the general stationary model (4.21), respectively. All displays are based on the training period (1961–1970).

Stein (2005a). The marked difference between the west-to-east and east-to-west correlations persists for temporal lags up to three days. For each lag, the difference increases linearly with the longitudinal distance between the stations.

4.4.2 Fitting a parametric, stationary space-time correlation model

A heuristic approach to fitting a parametric, stationary space-time correlation function can be described as follows. Suppose that a stationary correlation structure is a reasonable approximation for the spatio-temporal data set at hand. To fit a parametric stationary correlation model, we first check (or test) for full symmetry. If full symmetry is not a justifiable assumption, we fit a general stationary model in ways exemplified below. If full symmetry is an appropriate assumption, we check (or test) for separability. If the use of a separable model cannot be justified, we fit a fully symmetric but non-separable model. Otherwise, a separable model is adequate. Note that we fit and discuss correlation structures rather than covariance models.

In the case of the Irish wind data, Table 4.3 and Figure 4.3(a) show that the assumptions of separability and full symmetry are clearly violated. We fit a general stationary correlation model to the velocity measures, based on initial separable and fully symmetric fits. Generally, we consider only those spatio-temporal lags $(\boldsymbol{h}, u)$ in the WLS fits that are relevant to the subsequent prediction experiment; that is, we require that $\|\boldsymbol{h}\| \leq 450$ km and $|u| \leq 3$ days. All fits are based on the training period (1961–1970). The fitted separable model,

$$C_{\mathrm{SEP}}(\boldsymbol{h}, u) = C_{\mathrm{S}}(\boldsymbol{h}) \cdot C_{\mathrm{T}}(u), \tag{4.19}$$

is simply the product of the fitted purely spatial correlation function (4.17) and the fitted purely temporal correlation (4.18), respectively. This separable model can be embedded into the fully symmetric but generally non-separable correlation function

$$C_{\mathrm{FS}}(\boldsymbol{h}, u) = \frac{1-\nu}{1+a|u|^{2\alpha}} \left(\exp\!\left(-\frac{c\|\boldsymbol{h}\|}{(1+a|u|^{2\alpha})^{\beta/2}} \right) + \frac{\nu}{1-\nu}\, \delta_{\boldsymbol{h}=\boldsymbol{0}} \right) \tag{4.20}$$

which derives from (4.12) and reduces to the separable model when $\beta = 0$. With all other parameters fixed at the previous estimates, the WLS estimate of the space-time interaction parameter, $\beta \in [0, 1]$, is $\widehat{\beta} = 0.681$.

To fit a general stationary but not necessarily fully symmetric correlation model, we consider convex combinations,

$$C_{\mathrm{STAT}}(\boldsymbol{h}, u) = (1-\lambda)\, C_{\mathrm{FS}}(\boldsymbol{h}, u) + \lambda\, C_{\mathrm{LGR}}(\boldsymbol{h}, u), \tag{4.21}$$

of the fitted fully symmetric model (4.20) and the compactly supported, Lagrangian correlation function

$$C_{\mathrm{LGR}}(\boldsymbol{h}, u) = \left(1 - \frac{1}{2v}\, |h_1 - vu| \right)_{+}. \tag{4.22}$$

Here, the spatial separation vector $\boldsymbol{h} = (h_1, h_2)'$ has longitudinal (east-west) component h_1 and latitudinal (north-south) component h_2, $v \in \mathrm{R}$ is a longitudinal velocity, and we write $p_+ = \max(p, 0)$. Note that (4.22) considers the longitudinal distance between the stations only and forms a special case of the Lagrangian model (4.15). The particular parametric form is motivated by the prevalence of westerly winds over Ireland and the linear relationships in Figure 4.3(a). The random field model underlying (4.21) can be thought of as a superposition of two independent processes with correlation functions C_{FS} and C_{LGR}, respectively, and the Lagrangian component can be interpreted in term of the frozen field model. A less ambitious interpretation is simply in terms of increasing the flexibility of the correlation structure. All other parameters held constant, the WLS estimates in (4.22) are $\widehat{\lambda} = 0.0573$ and $\widehat{v} = 234\ \mathrm{km} \cdot \mathrm{d}^{-1}$ or $2.71\ \mathrm{m} \cdot \mathrm{s}^{-1}$. To summarize, the fitted general stationary model is

$$C_{\mathrm{STAT}}(\boldsymbol{h}, u) = \frac{(1-\nu)(1-\lambda)}{1 + a|u|^{2\alpha}} \left(\exp\!\left(- \frac{c\|\boldsymbol{h}\|}{(1 + a|u|^{2\alpha})^{\beta/2}} \right) + \frac{\nu}{1-\nu}\, \delta_{\boldsymbol{h}=\boldsymbol{0}} \right) + \lambda \left(1 - \frac{1}{2v}\, |h_1 - vu| \right)_+$$

with parameter estimates $\widehat{\nu} = 0.0415$, $\widehat{\lambda} = 0.0573$, $\widehat{a} = 0.972$, $\widehat{\alpha} = 0.834$, $\widehat{c} = 0.00128$, $\widehat{\beta} = 0.681$ and $\widehat{v} = 234$, where spatio-temporal lags are measured in kilometers and days, respectively. Full WLS optimization over all correlation parameters simultaneously yields essentially identical estimates. That said, our approach should be understood as illustrative. Ramifications in various directions remain to be explored, such as a more efficient and more appropriate handling of the nugget effect in the general stationary correlation model.

The graphs in Figure 4.3(b)–(d) illustrate the goodness of fit for the separable correlation model (4.19), the fully symmetric model (4.20), and the general stationary model (4.21), respectively. In each of the graphs, the fitted correlations are plotted vs. the empirical ones, for all pairs of stations and for temporal lag zero (black), one day (red), two days (green), and three days (blue). The general stationary correlation model fits better than the fully symmetric model, and the fully symmetric yet non-separable model provides a better fit than the separable model, particularly at temporal lags of one day.

4.4.3 Predictive performance

To assess the performance of predictions based on increasingly complex correlation structures, we consider time-forward predictions of the velocity measures during the test period (1971–1978), based on space-time correlation models fitted to the training data (1961–1970).

Specifically, we consider one-day ahead simple kriging predictions for the velocity measures at the eleven meteorological stations in Ireland. For each station, we obtain $365 \times 9 = 3285$ forecasts during the test period (1971–1978). The predictor variables are the 33 velocity measures observed during the past three days at the eleven stations. At each station, the simple kriging point predictor or point forecast, μ_t, for the velocity measure at time t is given by

$$\mu_t = \mathbf{c}_0{}' \, \mathbf{C}^{-1} \boldsymbol{z}_t \tag{4.23}$$

where $\mathbf{C}$ denotes the 33×33 variance-covariance matrix of the predictor variables, $\mathbf{c}_0$ denotes a vector with the covariances between the predictand and the predictor variables, and $\boldsymbol{z}_t$ is a vector with the realized values of the predictor variables. The associated simple kriging variance is

$$\sigma_t^2 = \sigma_0^2 - \mathbf{c}_0{}'\mathbf{C}^{-1}\mathbf{c}_0 \tag{4.24}$$

where σ_0^2 denotes the unconditional variance of the velocity measure at the station at hand (Cressie, 1993, pp. 109–110 and 359; Chilès and Delfiner, 1999, pp. 154–164). Note that σ_t^2 in fact is constant and does not change in time. In addition to the point forecast (4.23), simple kriging provides a probabilistic forecast in the form of a Gaussian predictive distribution, $F_t = \mathcal{N}(\mu_t, \sigma_t^2)$, with predictive mean μ_t and predictive variance σ_t^2, respectively.

In the following we compare out-of-sample simple kriging predictions based on the fitted separable correlation model (4.19), the fully symmetric model (4.20), and the general stationary model (4.21). Specifically, we build $\mathbf{C}$ and $\mathbf{c}_0$ on the basis of the respective correlation model and the station-specific empirical variances of the velocity measures during the training period. Furthermore, we consider simple kriging predictions with $\mathbf{C}$ and $\mathbf{c}_0$ based on the empirical correlations in the training period.

To assess and rank the point forecasts, μ_t, for the velocity measures, x_t, we use the root-mean-square error or RMSE, defined as

$$\text{RMSE} = \left(\frac{1}{3285} \sum_{t=1}^{3285} (\mu_t - x_t)^2 \right)^{1/2} \tag{4.25}$$

and the mean absolute error or MAE, given by

$$\text{MAE} = \frac{1}{3285} \sum_{t=1}^{3285} |\mu_t - x_t| \,. \tag{4.26}$$

To evaluate the predictive distributions, $F_t = \mathcal{N}(\mu_t, \sigma_t^2)$, we consider the logarithmic score and the continuous ranked probability score. The logarithmic score is simply the negative of the logarithm of the predictive density evaluated at the observation. Let $\mathbf{1}(y \geq x)$ denote the function that attains the value 1 when $y \geq x$ and the value 0 otherwise. If F denotes the predictive cumulative distribution function and x materializes, the continuous ranked probability score is defined as

$$\text{crps}(F, x) = \int_{-\infty}^{\infty} (F(y) - \mathbf{1}(y \geq x))^2 \, \mathrm{d}y.$$

For point forecasts the continuous ranked probability score reduces to the absolute error. If the predictive distribution is normal with mean μ and variance σ^2, the integral can be evaluated as

$$\mathrm{crps}\Big(\mathcal{N}(\mu,\sigma^2),x\Big) = \sigma\left(\frac{x-\mu}{\sigma}\left(2\,\Phi\Big(\frac{x-\mu}{\sigma}\Big)-1\right)+2\,\phi\Big(\frac{x-\mu}{\sigma}\Big)-\frac{1}{\sqrt{\pi}}\right)$$

where ϕ and Φ denote the probability density function and the cumulative distribution function of the standard normal distribution, respectively. We refer to Gneiting and Raftery (2005) and Gneiting et al. (2006) for detailed discussions of the continuous ranked probability score and other devices for the evaluation of probabilistic forecasts. Here, we assess the predictive distributions based on the various correlation structures by comparing the associated LogS and CRPS values, where

$$\mathrm{LogS} = \frac{1}{3285}\sum_{t=1}^{3285}\left(\frac{1}{2}\ln(2\pi\sigma_t^2)+\frac{(x_t-\mu_t)^2}{2\sigma_t^2}\right) \tag{4.27}$$

and

$$\mathrm{CRPS} = \frac{1}{3285}\sum_{t=1}^{3285}\mathrm{crps}\big(\mathcal{N}(\mu_t,\sigma_t^2),x_t\big) \tag{4.28}$$

denote the mean of the logarithmic score and the continuous ranked probability score over the test period, respectively. Clearly, the smaller these values the better.

Tables 4.4, 4.5, 4.6, and 4.7 compare the RMSE, MAE, LogS and CRPS values for the kriging predictions based on the separable, fully symmetric, general stationary, and empirical correlation structures, respectively. In terms of all four performance measures, the simple kriging predictions based on the general stationary correlation model performed better than those based on the fully symmetric model, and the predictions based on the fully symmetric model outperformed those based on the separable model. This suggests that the use of richer, more complex, and more physically realistic correlation and covariance models results in improved predictive performance. Perhaps surprisingly, the predictions based on the empirical space-time correlations performed best. The empirical correlations are computed on 10 years' worth of replicated data and need to be taken at face value — in stark contrast to the classical geostatistical, purely spatial one realization scenario, in which reliable estimates of the correlation and covariance structure are unlikely to be available. However, the simple kriging approach based on the empirical correlations does not allow for predictions of velocity measures away from the meteorological stations. The model-based approaches admit such an extension when combined with a statistical model for the spatially varying variance of the velocity measures.

Table 4.4 Root-mean-square error (RMSE) for one-day ahead simple kriging predictions of the velocity measures at the eleven meteorological stations during the test period (1971–1978)

	Val	Bel	Clo	Sha	Roc	Birr	Mul	Mal	Kil	Clo	Dub
separable	.501	.495	.491	.468	.483	.477	.427	.496	.439	.486	.450
fully symmetric	.501	.495	.492	.468	.479	.476	.424	.492	.436	.484	.445
general stationary	.499	.495	.490	.466	.474	.472	.419	.488	.429	.479	.440
empirical	.500	.494	.486	.454	.465	.462	.414	.479	.414	.466	.427

Table 4.5 Mean absolute error (MAE) for one-day ahead simple kriging predictions of the velocity measures at the eleven meteorological stations during the test period (1971–1978)

	Val	Bel	Clo	Sha	Roc	Birr	Mul	Mal	Kil	Clo	Dub
separable	.398	.395	.389	.372	.387	.375	.340	.399	.347	.385	.359
fully symmetric	.399	.396	.389	.372	.384	.373	.338	.396	.344	.382	.356
general stationary	.397	.395	.387	.369	.379	.370	.334	.393	.339	.377	.351
empirical	.394	.392	.384	.359	.369	.362	.327	.385	.325	.367	.339

Table 4.6 Logarithmic score (LogS) for one-day ahead simple kriging predictions of the velocity measures at the eleven meteorological stations during the test period (1971–1978)

	Val	Bel	Clo	Sha	Roc	Birr	Mul	Mal	Kil	Clo	Dub
separable	.727	.716	.707	.659	.692	.680	.577	.720	.596	.699	.626
fully symmetric	.728	.716	.709	.661	.682	.677	.570	.712	.589	.694	.617
general stationary	.724	.715	.705	.655	.672	.670	.560	.704	.574	.683	.606
empirical	.726	.714	.698	.630	.654	.648	.542	.684	.538	.658	.571

Table 4.7 Continuous ranked probability score (CRPS) for one-day ahead simple kriging predictions of the velocity measures at the eleven meteorological stations during the test period (1971–1978)

	Val	Bel	Clo	Sha	Roc	Birr	Mul	Mal	Kil	Clo	Dub
separable	.282	.279	.276	.264	.273	.268	.241	.281	.247	.273	.254
fully symmetric	.282	.279	.277	.264	.271	.267	.240	.279	.245	.272	.252
general stationary	.281	.279	.275	.262	.267	.265	.237	.276	.241	.269	.249
empirical	.280	.278	.273	.255	.262	.259	.233	.271	.232	.261	.240

Appendix

Proof of Theorem 4.3.2 Suppose that C is a stationary, fully symmetric covariance function. By Bochner's theorem, the representation (4.7) holds. Expand the exponential term under the integral in (4.7) as the product of $\cos(\boldsymbol{h}'\boldsymbol{\omega}) + i\sin(\boldsymbol{h}'\boldsymbol{\omega})$ and $\cos(u\tau) + i\sin(u\tau)$. Since C is real-valued, we can write (4.7) as

$$C(\boldsymbol{h}, u) = \iint \Big(\cos(\boldsymbol{h}'\boldsymbol{\omega})\cos(u\tau) - \sin(\boldsymbol{h}'\boldsymbol{\omega})\sin(u\tau)\Big)\,\mathrm{d}F(\boldsymbol{\omega}, \tau),$$
$$(\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}. \tag{4.29}$$

Similarly,

$$C(\boldsymbol{h}, -u) = \iint \Big(\cos(\boldsymbol{h}'\boldsymbol{\omega})\cos(u\tau) + \sin(\boldsymbol{h}'\boldsymbol{\omega})\sin(u\tau)\Big)\,\mathrm{d}F(\boldsymbol{\omega}, \tau),$$
$$(\boldsymbol{h}, u) \in \mathrm{R}^d \times \mathrm{R}. \tag{4.30}$$

Full symmetry implies that $C(\boldsymbol{h}, u) = C(\boldsymbol{h}, -u)$ and therefore (4.29) and (4.30) yield the representation (4.8), as desired. Conversely, any function C of the form (4.8) is fully symmetric and admits Bochner's representation (4.7).

Acknowledgments

While working on this chapter, Tilmann Gneiting was on sabbatical leave at the Soil Physics Group, Universität Bayreuth, Universitätsstr. 30, 95440 Bayreuth, Germany, and supported in part by Office of Naval Research Grant N00014-01-10745 and National Science Foundation Award 0134264.

References

Amani, A. and Lebel, T. (1997), Lagrangian kriging for the estimation of Sahelian rainfall at small time steps, *Journal of Hydrology*, 192, 125–157.

Banerjee, S. (2005), On geodetic distance computations in spatial modeling, *Biometrics*, 61, 617–625.

Bochner, S. (1955), *Harmonic Analysis and the Theory of Probability*, University of California Press, Berkeley and Los Angeles.

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, 847–860.

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, 392–415.

Chilès, J.P. and Delfiner, P. (1999), *Geostatistics. Modeling Spatial Uncertainty*, Wiley, New York.

Christakos, A.G. (2000), *Modern Spatiotemporal Geostatistics*, Oxford University Press, Oxford.

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*, revised ed., Wiley, New York.

Cressie, N. and Huang, H.-C. (1999), Classes of nonseparable, spatio-temporal stationary covariance functions, *Journal of the American Statistical Association*, 94, 1330–1340; 96, 784.

De Iaco, S., Myers, D.E., and Posa, T. (2001), Space-time analysis using a general product-sum model, *Statistics & Probability Letters*, 52, 21–28.

De Iaco, S., Myers, D.E., and Posa, T. (2002), Nonseparable space-time covariance models: some parametric families, *Mathematical Geology*, 34, 23–42.

de Luna, X. and Genton, M.G. (2005), Predictive spatio-temporal models for spatially sparse environmental data, *Statistica Sinica*, 15, 547–568.

Diggle, P.J., Tawn, J.A., and Moyeed, R.A. (1998), Model-based geostatistics (with discussion), *Applied Statistics*, 47, 299–350.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research – Oceans*, 99, 10143–10162.

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

Fuentes, M. (2006), Testing for separability of spatial-temporal covariance functions, *Journal of Statistical Planning and Inference*, 136, 447–466.

Furrer, R., Genton, M.G., and Nychka, D. (2006), Covariance tapering for interpolation of large spatial datasets, *Journal of Computational and Graphical Statistics*, in press.

Gneiting, T. (1999), Correlation functions for atmospheric data analysis, *Quarterly Journal of the Royal Meteorological Society*, 125, 2449–2464.

Gneiting, T. (2002a), Nonseparable, stationary covariance functions for space-time data, *Journal of the American Statistical Association*, 97, 590–600.

Gneiting, T. (2002b), Compactly supported correlation functions, *Journal of Multivariate Analysis*, 83, 493–508.

Gneiting, T. and Raftery, A.E. (2005), Strictly proper scoring rules, prediction and estimation, University of Washington, Department of Statistics, Technical Report no. 463R.

Gneiting, T. and Sasvári, Z. (1999), The characterization problem for isotropic covariance functions, *Mathematical Geology*, 31, 105–111.

Gneiting, T. and Schlather, M. (2002), Space-time covariance models, in *Encyclopedia of Environmetrics*, El-Shaarawi, A.H. and Piegorsch, W.W., Eds., Vol. 4, Wiley, Chichester, pp. 2041–2045.

Gneiting, T. and Schlather, M. (2004), Stochastic models that separate fractal dimension and the Hurst effect, *SIAM Review*, 46, 269–282.

Gneiting, T., Sasvári, Z., and Schlather, M. (2001), Analogies and correspondences between variograms and covariance functions, *Advances in Applied Probability*, 33, 617–630.

Gneiting, T., Larson, K., Westrick, K., Genton, M.G., and Aldrich, E. (2006), Calibrated probabilistic forecasting at the Stateline Wind Energy Center: the regime-switching space-time (RST) method, *Journal of the American Statistical Association*, in press.

Guo, J.H. and Billard, L. (1998), Some inference results for causal autoregressive processes on a plane, *Journal of Time Series Analysis*, 19, 681–691.

Gupta, V.K. and Waymire, E. (1987), On Taylor's hypothesis and dissipation in rainfall, *Journal of Geophysical Research*, 92, 9657–9660.

Guttorp, P., Meiring, W., and Sampson, P.D. (1994), A space-time analysis of ground-level ozone, *Environmetrics*, 5, 241–254.

Haas, T.C. (2002), New systems for modeling, estimating, and predicting a multivariate spatiotemporal process, *Environmetrics*, 13, 311–332.

Hamill, T.M. and Snyder, C. (2000), A hybrid ensemble Kalman filter-3D variational analysis scheme, *Monthly Weather Review*, 128, 2905–2919.

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.

Houtekamer, P.L. and Mitchell, H.L. (1998), Data assimilation using an ensemble Kalman filter technique, *Monthly Weather Review*, 126, 796–811.

Huang, H.-C. and Hsu, N.-J. (2004), Modeling transport effects on ground-level ozone using a non-stationary space-time model, *Environmetrics*, 15, 251–268.

Jones, R. and Zhang, Y. (1997), Models for continuous stationary space-time processes, in *Modelling Longitudinal and Spatially Correlated Data*, Lecture Notes in Statistics, 122, Gregoire, T.G., Brillinger, D.R., Diggle, P.J., Russek-Cohen, E., Warren, W.G., and Wolfinger, R.D., Eds., Springer, New York, pp. 289–298.

Jun, M. and Stein, M.L. (2004a), Statistical comparison of observed and CMAQ modeled daily sulfate levels, *Atmospheric Environment*, 38, 4427–4436.

Jun, M. and Stein, M.L. (2004b), An approach to producing space-time covariance functions on spheres, University of Chicago, Center for Integrating Statistical and Environmental Science, Technical Report No. 18.

Kolovos, A., Christakos, G., Hristopulos, D.T., and Serre, M.L. (2004), Methods for generating non-separable spatiotemporal covariance models with potential environmental applications, *Advances in Water Resources*, 27, 815–830.

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

Lu, N. and Zimmerman, D.L. (2005), Testing for directional symmetry in spatial dependence using the periodogram, *Journal of Statistical Planning and Inference*, 129, 369–385.

Ma, C. (2002), Spatio-temporal covariance functions generated by mixtures, *Mathematical Geology*, 34, 965–975.

Ma, C. (2003a), Spatio-temporal stationary covariance models, *Journal of Multivariate Analysis*, 86, 97–107.

Ma, C. (2003b), Families of spatio-temporal stationary covariance models, *Journal of Statistical Planning and Inference*, 116, 489–501.

Ma, C. (2003c), Nonstationary covariance functions that model space-time interactions, *Statistics & Probability Letters*, 61, 411–419.

Matérn, B. (1986), *Spatial Variation*, 2nd edition, Springer, Berlin.

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

Meiring, W., Guttorp, P., and Sampson, P.D. (1998), Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model, *Environmental and Ecological Statistics*, 5, 197–222.

Mitchell, M., Genton, M.G., and Gumpertz, M. (2005), Testing for separability of space-time covariances, *Environmetrics*, 16, 819–831.

Riishøjgaard, L.P. (1998), A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus Ser. A*, 50, 42–57.

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

Scaccia, L. and Martin, R.J. (2005), Testing axial symmetry and separability of lattice processes, *Journal of Statistical Planning and Inference*, 131, 19–39.

Shitan, M. and Brockwell, P. (1995), An asymptotic test for separability of a spatial autoregressive model, *Communications in Statistics, Theory and Methods*, 24, 2027–2040.

Stein, M.L. (2005a), Space-time covariance functions, *Journal of the American Statistical Association*, 100, 310–321.

Stein, M.L. (2005b), Statistical methods for regular monitoring data, *Journal of the Royal Statistical Society Ser. B,* 67, 667–687.

Storvik, G., Frigessi, A., and Hirst, D. (2002), Stationary space-time Gaussian fields and their time autoregressive representation, *Stochastic Modelling*, 2, 139–161.

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

Wan, Y., Milligan, M., and Parsons, B. (2003), Output power correlation between adjacent wind power plants, *Journal of Solar Energy Engineering*, 125, 551–555.

Whittle, P. (1954), On stationary processes in the plane, *Biometrika*, 41, 434–449.

Zawadzki, I.I. (1973), Statistical properties of precipitation patterns, *Journal of Applied Meteorology*, 12, 459–472.